

SI: SCA Measures – Fuzzy rough set features of cognitive computations in the visual system

Andrzej W. Przybyszewski^{a,b,*}

^aPolish Japanese Academy of Information Technology, Koszykowa, Warsaw, Poland

^bDepartment Neurology, UMass Medical School, Worcester, MA, USA

Abstract. Object recognition is a complex neuronal process determined by interactions between many visual areas: from the retina, thalamus to the ventral visual pathway. These structures transform variable, single pixel signal in photoreceptors to a stable object representation. Neurons in visual area V4, midway in ventral stream, represent such stable shape detector. A feed forward hierarchy of increasing in size and complexity receptive fields (RF) leads to *grand mother* cell concept. Our question is how these processes might identify an object or its elements in order to recognize it in new, unseen conditions? We propose a new approach to this problem by extending the classical definition of the RF to a fuzzy detector. RF properties are also determined by the computational properties of the bottom-up and top-down pathways comparing stimulus with many predictions. The “driver-type” logic (DTL) of bottom-up computations looks for large number of possible object parts (hypotheses – rough set (RS) upper approximation), as object’s elements are similar to RF properties. The optimal combination is chosen, in unsupervised, parallel, multi-hierarchical pathways by the “modulator-type” logic (MTL) of top-down computations (RS lower approximation). Interactions between DTL (hypotheses) and MTL (predictions) terminates when RS boundary became small - the object is recognized.

Keywords: Brain’s logic, ascending, descending pathways, object categorization, predictive coding, inconsistent rules

1. Introduction

How slow and noisy brain’s computations make our recognition so effective that it outperforms many times faster artificial intelligent (AI) systems? Can we at least find out what differences are in computations between these systems?

In our everyday life we actively perceive only a small part of our environment. This part depends on our interest, which determines where we direct our eyes. This paper describes neurological mechanisms that determine how different brain structures may anticipate what we actually see and mechanisms

how to recognize familiar objects in new positions, light conditions, perspective, and environment. There are two anatomically different pathways that interact in order to focus our attention on a specific object. One pathway has specific sub-cortical ascending input (core cells) whereas another pathway has diffused sub-cortical inputs (matrix cells) related to descending pathways from the higher visual areas. The first pathway classifies objects on the basis of their pure visual attributes in an ambiguous way as it is influenced by many sources of noise. In contrast to the second pathway is based on predictions and diverse classifications that may be related to different motor activities like eye movement or possibility to grasp an object, anticipation of obtaining higher value food reward, avoidance of the danger or

*Corresponding author. Andrzej W. Przybyszewski. E-mail: Andrzej.Prybyszewski@umassmed.edu.

obtaining of the pleasure. Generally, there are many still not well-described factors that make our visual recognition powerful and universal. For example: we have recently demonstrated that we (monkey) perceive differently exactly the same object (stimulus) in different positions (related the eye positions) [1]. As images in the retina of this object are identical it means that the descending pathways related to the eye positions are changing the meaning of this stimulus (object).

As the precision and meanings of these two pathways are different, in order to classify objects, we have already proposed that the visual system is using the principles [2] based on rough set theory [3]. Our model is also using the fuzzy similarities relations between objects and RF attributes takes into account differences between different anatomical pathways. On the basis of the fuzzy similarity relation definition [4], we propose to classify objects by assembly of RF related granules that differentiate our method from that used in most AI applications.

A popular understanding on how neurological processes in the visual system lead to objects classification is based on generalization of simple and complex cell properties from visual area V1 as described by Hubel and Wiesel [5]. They proposed that an array of spatially aligned receptive fields (RFs) of LGN cells might give orientation sensitivity to V1 simple cell (SC), and that several phase (or position) shifted SCs with similar orientations convergent on the complex cell (CC). Such convergence might give spatial invariance in complex cells. On the basis of simple and complex cells properties, Fukushima [6] made simulation of a self-organizing network: cognition, and later introduced improved model with a position invariant property [7]. Networks with similar principles are still used nowadays in most models of the visual system. There are based on a first-order description of primary visual cortex V1 that consist of a collection of locally-normalized, threshold Gabor wavelet functions spanning a range of orientations and spatial frequencies [8]. More complex cells' properties arise in such linear models as summation of simple/complex cells from V1. There are many dissertations using this approach in different visual areas from the thalamus to the inferotemporal cortex.

The linear combination of simple and complex cell RF attributes from areas V1, V2 may explain selectivity and position invariance properties of cells in area V4 [9, 10]. The main assumptions of above models are that simple units in higher areas (V4)

generate selectivity for complex features or shapes by summation of units selective to different orientations and different receptive field sizes. Such linear, feed forward models can simulate certain sensitivity of V4 cells to complex object but cannot explain universality of higher brain areas to recognize complex objects in unseen conditions. Another problem with these models is that they do not take into account nonlinear properties of the complex cells such as, for example, overlapping of the on and off subfields [11]. Also some basic experimental findings of cell properties in area V4 like nonlinear interactions between subfields are not taken into account in above models [12].

2. Methods

2.1. Theoretical basis

Our data mining analysis is based rough set theory (RST) proposed by Pawlak [3]. Our data is converted to the decision table where rows were related to different measurements and columns represent different attributes. An information system [3] a pair $S = (U, A)$, where U, A are nonempty finite sets called the universe of objects U and the set of attributes A . If $a \in A$ and $u \in U$, the value $a(u)$ is a unique element of V (where V is a value set).

We define as in [3] for RST the *indiscernibility relation* of any subset B of A or $IND(B)$ as: $(x, y) \in IND(B)$ or $xI(B)y$ iff $a(x) = a(y)$ for every $a \in B$ where the value of $a(x) \in V$. It is an equivalence *crisp* relation $[u]_B$ that we understand as a *B-elementary granule*. The family of $[u]_B$ gives the partition U/B containing u will be denoted by $B(u)$. The set $B \subset A$ of information system S is a reduct $IND(B) = IND(A)$ and no proper subset of B has this property [13]. In most cases, we are only interested in such reducts that are leading to expected rules (classifications). On the basis of the reduct we have generated rules using four different ML methods (RSES 2.2): exhaustive algorithm, genetic algorithm [14], covering algorithm, or LEM2 algorithm [15].

A *lower approximation* of set $X \subseteq U$ in relation to an attribute B is defined as all elements have B attribute: $\underline{B}X = \{u \in U : [u]_B \subseteq X\}$. The *upper approximation* of X is defined as some elements have B attribute: $\overline{B}X = \{u \in U : [u]_B \cap X \neq \emptyset\}$. The difference of $\overline{B}X$ and $\underline{B}X$ is the boundary region of X

that we denote as $BN_B(X)$. If $BN_B(X)$ is empty then set than X is *exact* with respect to B ; otherwise if $BN_B(X)$ is not empty and X is not *rough* with respect to B .

A decision table (training sample in ML) for S is the triplet: $S = (U, C, D)$ where: C, D are condition and decision attributes [3]. Each row of the information table gives a particular rule that connects condition and decision attributes for a single measurements of a particular receptive field. As there are many rows related to different cells and stimuli, they gave many particular rules. Rough set approach allows generalizing these rules into universal hypotheses that may determine optimal classification for different objects. The decision attribute D is related to neuron classification defined as normalized (nominative) response to the stimulus or stimuli.

Dubois and Prade [16] has been generalized RST to FRTS (fuzzy rough set theory) by extending RST indiscernibility with concepts of tolerance after Zadeh's membership degrees in fuzzy sets [17].

As the effect, 'crisp' dependences were replaced by a fuzzy tolerance relation $R_a(x, y)$ as a value between two observations x and y . As $R_a(x, y)$ is a similarity relation, it must be reflexive, symmetric and transitive. As summarized in [18] there are several tolerance relationships such as the normalized difference (so-called 'Equation 1') or Gaussian or exponential differences [18]. There are also formulas related to normalized differences between pairs of attributes. The most common are *Lukasiewicz* and *t.cos t-norms* - τ [18]. As decision attributes are nominative we used crisp relations between them.

We define *B-lower* and *B-upper* approximations for each observation x in FRST as following: *B-lower* approximation as: $(R_B \downarrow X)(x) = \inf_{y \in U} I(R_B(x, y), X(y))$, where I is an *implicator* [18]. The *B-lower* approximation for the observation x is then the set of observations, which are the most similar to observation x and it can predict the decision attribute with the highest confidence, based on conditional attributes B .

The *B-upper* approximation is defined by $(R_B \uparrow X)(x) = \sup_{y \in U} \tau(R_B(x, y), X(y))$ where τ is the *t-norm*. The *B-upper* approximation is a set of observations for which the prediction of decision attribute has the smallest confidence [18].

Also rules in FRST have different construction than in RST. They are based on the tolerance classes

and appropriate decision concepts. The *fuzzy rule* is a triple (B, C, D) , where B is a set of conditional attributes that appear in the rule, C stands for fuzzy tolerance class of object and D stands for decision class of object.

2.2. Objects' attributes

We will represent experimental data ([12]) in the following table. In the first column are neural measurements. Neurons are identified using numbers related to a collection of figures in the previous paper [12]. Stimuli typically used in neuroscience have the following properties:

1. Orientation in degrees appears in the column labeled $o1$.
2. Spatial frequency is denoted as $sf1$.
3. X-axis stimulus size is denoted by $xs1$.
4. Y-axis stimulus size is denoted by $ys1$.
5. X-axis position is denoted by $xp1$.
6. Y-axis position is denoted by $yp1$.
7. Stimulus contrast $c1$.

Similar attributes are for the second stimulus define as '*2'.

Decision attributes are divided into three classes determined by the strength of the neural responses. Small cell responses r are classified as class r_0 with value 0, medium to strong responses are classified as class r_1 with value 1, and the strongest cell responses are classified as class r_2 with value 2. Therefore each cell divides stimuli into its own family of equivalent objects. It is similar approach to popular used in neuroscience normalization of neuronal responses from 0 to 100% (here 0 to 2). The full set of stimulus attributes is expressed as $B = \{o1, sf1, xs1, ys1, xp1, yp1, c1, o2, sf2, xs2, ys2, xp2, yp2, c2\}$.

In this work we are looking into single cell responses only in one area - V4 that will divide all patterns into equivalent (or at least similar to a degree of the fuzzy tolerance) classes of V4-elementary granules. Neurons in V4 are sensitive only to the certain attributes of the stimulus, like for example space localization, and they are insensitive to other stimulus attribute like e.g. contrast changes, but when put value of the contrast equal 0 it means that there is no stimulus). Different V4 cells have different receptive field properties, which mean that one object (B-elementary granule) can be classified in many ways by different cells (V4-elementary granules).

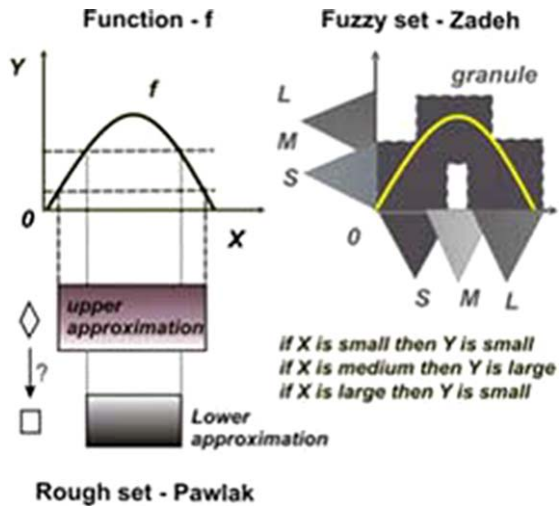


Fig. 1. Two models of the function approximation: the rough set [3] on the left, and on the right fuzzy set [27, 28] approaches.

2.3. Receptive field as a computation unit that determines similarities between objects

Kuffler [19] first defined the receptive field as antagonistic circular center-surround filter in the output of the retina. Hubel and Wiesel [5] found elongated orientation-sensitive ON and OFF subfields in the cat primary visual cortex (V1).

Receptive field properties in the early stages of the visual pathway have been explained in terms of many different models generally as linear filters (Gaussian, Gabor or wavelets) parameterized by temporal and spatial frequencies, orientation, phase and position [5, 11]. Even if such local filters are well suited for the effective and sparse encoding of natural images, none of the computational vision systems that use them have managed to achieve robust recognition performance. It is appropriate, therefore, to consider different strategies for image processing that assist recognition.

We have assumed that generally stronger neuronal responses measured in spikes/sec better classify stimulus attributes related to RF properties than weaker neuronal responses. In other words, higher response means that a certain attribute(s) of the object and RF are more similar than for smaller responses. It is in agreement with the standard RF properties understanding.

However, we will make following modification to the classical view: we divide neuronal activity into several ranges: below a certain threshold we assume

that very weak activity is not related to the stimulus (a classical approach); for activity above the threshold, as an example, we will discuss medium and strong responses in different ranges of spike frequencies (see Fig. 1).

As it is explained in Fig. 1, lower approximation (strong) neural response is related to certainty (belief) in the classification of object attributes, whereas upper approximation (weaker) response is related to the possibility (plausibility) that an object may have detected attributes. Therefore our hypothesis is that by studying the strength of single cell responses to different stimulus attributes, we can find ranges of “similarities” between stimulus and RF properties. In this paper we are looking for the basis of how the brain changes the precision of object classification from uncertain to confident. Let us take a simple example like the RF of ON-center retinal ganglion cell (GC) approximated by the DOG function (like in Fig. 2). We say that the RF better fits (more similar) to the larger spot (size attribute of the object) when GC gives stronger responses (Fig. 1: lower vs. upper approximation). Another possibility is a fuzzy set approximation (Fig. 2 right side). In this model we have three granules: small spot size give small responses, larger spot size (near size of the RF center) gives large responses, even larger spot size (that also partly covers RF surround) gives smaller

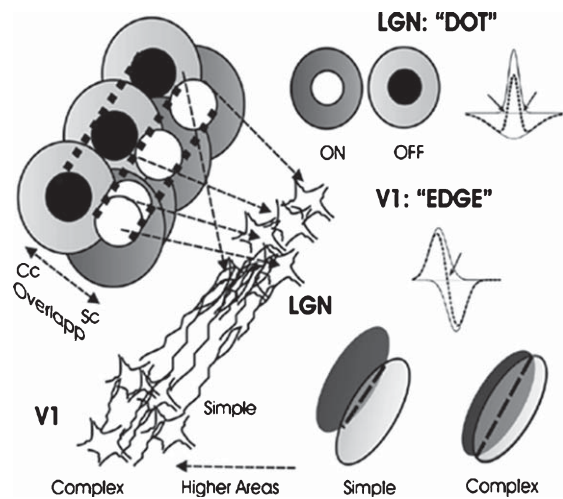


Fig. 2. Modified schematic shows RF of LGN, simple and complex V1 cells. ON- OFF-center LGN RF is well described by DOG (difference of Gaussian) functions. Aligned LGN RF may give orientation properties of V1 simple cells. V1 complex cells may arise from overlapping V1 simple cells or by higher area modulations. On-and off-subfields of V1 cells can be approximated by shifted Gaussian functions (see text).

responses. These two models are to some extent interchangeable but can also be fused to one fuzzy-rough set approach that we are proposing in this work. But we have to mention that they are related to the first order approach, responses measured by mean spike frequency or by a first harmonic if the stimulus changes its intensity in time. In this case the second stimulus attribute is the optimal frequency (spatial vs. temporal frequency tuning). However, even if this example is limited to the retinal output that is not influenced by the feedback from higher areas, retinal classification are not well described except the first order approximation and probably more complex [22]. A more careful analysis of the spike train and its frequencies in response to change of the light spot diameter and frequency shows a wide range of different oscillatory responses [22, 23]. We have revealed (also in the intracellular recordings) that synchronization of certain oscillations with the stimulus might code certain stimulus attributes [23, 24]. More generally the retina (and the brain) may be seen as a system of coupled nonlinear oscillators, which synchronizations might be related to cognition [25, 26].

2.4. Decision rules for a single neuron

A classical approach to neuro computation is a modification of the linear threshold summation. Neurons obtain excitatory and inhibitory inputs (via synapses) and sum of inputs exceed the threshold and to generate an action potential (the first order approach). Each neuron receives thousands inputs (the most synapses with different weights are on the dendritic tree) and has a single axon as the output. As the single action potential is the result of many inputs, it can be seen as a result of analog computation. We still have only weak understanding how these analog computers are working. Therefore we will mainly approximate neuronal activity as spikes mean frequency in a certain time (so-called the first order approach). We will consider only neurons in the visual system with activity dependent from the visual input from the retina. Each neuron in this sensory system is characterized by its receptive field (RF) properties. RF is small part of the visual field in which neuron is sensitive to the luminance or color changes. In the following, we study RF properties of neurons in different parts of the visual brain.

An extension of this approach will be to take into account membrane properties as assembly of ion channels with different dynamic. In this case the

membrane can sense different frequencies in assemble of input (synaptic) signals and generate spikes with complex frequency patterns. It is the basis of the oscillatory theory of the cognition. In the retina, ganglion cells show intracellular oscillations that for certain parameters of the stimulus that can lock (see above) to the input giving appropriate burst of spikes [25]. Then the decision become more complex as the mean spike frequency give information about stimulus attributes that fit (to a certain extend) to RF properties, their frequency can give additional information about other stimulus attributes. Therefore, oscillations can be seen as a higher order decisions related to object's attributes.

2.5. Decision rules for thalamus - LGN

The LGN neurons RFs have the concentric center-surround shapes that are similar to the retinal ganglion cells RFs [19]. In our model, we take into account only on- and off type RFs. The ON-type neurons increase their activity by an increase of the light luminance in their RF center and/or decrease of the light luminance in the RF surround (Fig. 2). The OFF-type neurons act in the inverse direction.

An example below consists of two equations describing rules for on- off- center in the LGN neuron. The RF of this neuron has the position: xp_0 , yp_0 , and RF size is $x_s = 0.5 \text{ deg}$, $y_s = 0.5 \text{ deg}$ and RF center size is $x_s = 0.2 \text{ deg}$, $y_s = 0.2 \text{ deg}$. There is no positive feedback from higher areas therefore the maximum response is r_1 .

$$\text{DRLGN1} : xp_0 \wedge yp_0 \wedge xs_{0,2} \wedge ys_{0,2} \wedge c_1 \rightarrow r_1 \quad (1)$$

$$\text{DRLGN2} : xp_0 \wedge yp_0 \wedge xs_{0,5} \wedge ys_{0,5} \wedge c_{-1} \rightarrow r_0 \quad (2)$$

In (1) change in the luminance that cover the RF center gives response r_1 . When change in the luminance covers the whole RF it gives response r_0 as the sums of excitation and inhibitions are equal. It is an example of very simple stimulus like a short light flash. We can also stimulate separately RF center and RF surround and by changing contrast and frequency of luminance changes in order to obtain more complex responses that better characterize a particular neuron. Another simpler method is to replace RF center and surround stimulations by the drifting grating that covers the whole RF. As the RF – center is small then it is stimulated the high spatial frequency

of the grating. The low spatial frequencies stimulate center and surround. By findings differences in these frequencies we can characterize sizes of RF center and surround. By changing contrast and color one can find other properties of the particular RF and formulate them as rules [29]. These rules will represent different LGN-elementary granules.

2.6. Decision rules for area V1

In the area V1 (so-called primary visual cortex) neurons by aligned LGN RF get a new, orientation sensitivity property (Fig. 2 left). It is in contrast to the lower areas: retina or LGN where RF have circular on-off shapes (Fig. 2 top, right). There are generally two cell types related to their RF properties in the primary visual cortex: simple and complex RF cells (Fig. 2 bottom, right). Both cell types have incremental (On - responding to white bars) and decremental (Off - responding to black bars) subfields. The major difference is that in simple cells On- and Off- subfields are separated, and in complex cells they are overlapping (see rules in [29]).

The classical concept related to the difference between simple and complex cells is that complex cells RFs are effect of the convergence of several simple cells [5] (Fig. 2). However, some experiments suggest that the nonlinearity of the complex cells RFs [11] might be related to the properties of the feedback and/or horizontal connections [30].

2.7. Decision rules for area V4

In the higher areas such as the area V4, RFs became larger (till several degrees) and more complex than in the lower areas such as V1 or LGN. As V4 RFs are also nonlinear, it is very difficult to find optimal stimulus for some of these cells. In our experiments [12], we found that V4 RFs are consistent of the interactions between many subfields related to the lower areas RFs [12]. Examples of such RFs properties and related decision rules are in [12, 29].

2.8. Complex cell properties determine local computations

As mentioned above, the default strategy for many recognition systems based on the image encoding approach is to use local filters for the transformation of image information in terms of local (Gaussian-like) gradients. These image compressions and reconstruction strategies have had such limited success in the

task of the natural object recognition that it is difficult to compare them to the recognition capabilities of primates. We suggest that it may be related to different principles: primate's image recognition strategy is different from direct image encoding by band of linear filters.

Therefore, we will analyze the receptive field (RF) properties of thalamic (LGN) and cortical cells in order to compare them to linear filters used in artificial systems. At first, we will show how RF properties of simple and complex cells in V1 may emerge from the LGN RFs.

The schematic in Fig. 2 demonstrates convergence of the LGN cells into V1 cells. An array of spatially aligned RFs of LGN cells may give orientation sensitivity to a V1 simple cell (SC) [4] (Fig. 2 left side). However, the origin of the area V1 complex cell (CC) RF is less clear and several hypotheses are still under debate today: 1) there is synaptic convergence of several (phase shifted) SCs on one CC [4]; 2) CC properties are an effect of LGN RFs overlap [1] (Fig. 2); 3) feedback from the higher areas can change RF properties of V1 cells from simple to complex [29].

The most popular model approximates the LGN RF by the Difference of the Gaussian (DOG) function, which linearly transforms local properties of visual images (Fig. 2 right side). As mentioned above, a popular model of V1 SC and CC RFs are Gabor or Gaussian functions, which transform image linearly, whereas the electrophysiology shows that CC RFs in V1 and higher areas are nonlinear. Intracellular recordings demonstrate that there are several distinct nonlinear processes between membrane modulation and the spike generation mechanism; therefore linearity of SC RF is an exception, which depends on stimulus parameters [31]. The simple/complex cell dichotomy is also characterized by overlap between ON and OFF RF sub-regions. More precisely, ON/OFF activating regions (ARs) can be mapped with light increment/decrement (INC/DEC) bars and described as INC/DEC ARs. Recently, it has been shown that in awake monkeys, SCs are characterized by minimal overlapping (less than 30%) of the ARs, but larger group of CCs have strongly overlapping (over 50%) ARs [31]. The response of each elongated AR can be approximate by the Gaussian function [23]. If overlap is less than 30% then we can still estimate if an INC or DEC AR was stimulated and recover the input image. However, for CC with ARs overlapping more than 50%, it is not even possible to say what the stimulus polarity in the overlapping

region was. Even if Shams and von der Malsburg [32] suggested that CC population responses contain sufficient information to recover the essence of images, we will concentrate on individual cells as feedback loops act on them non-uniformly [33]. Our complex cells are from the second cortical stage (layer 2+3) and not in input layer 4, which mainly integrate lower area (thalamic) input [34]. Therefore, mentioned above properties of CCs eliminate them as encoders, and they only can be detectors. As shown schematically in Fig. 2, larger overlap in CC RFs make CCs better edge detectors than SCs. In addition their nonlinearities help in sharpening edge detections. Moreover, the higher areas may influence the overlap of INC/DEC ARs in V1 RFs [40], as well as other RF attributes like e.g. orientation [32]. Therefore, the region of the edge detections may become variable within the RF; we call this effect the tuning of the lower areas properties to the higher areas predictions. In addition, positive feedback from higher to lower areas may regulate edge detection sensitivity [35].

In summary, CCs even from early visual areas (V1) do not encode local image features but detect attributes to which they are tuned. In consequence higher areas can only access encoded information about images in lower areas with the help of feedback pathways.

We will divide information transformation in the brain into bottom-up (BUCs) and top-down computations (TDCs). The BUCs are determined by anatomical and physiological properties of ascending pathways, whereas TDCs are related to descending pathways.

2.9. Local vs. global computations: simplified connections from thalamus to area V4; core vs. matrix projections

We will demonstrate an anatomical basis of network computation that generally suggest that there are local in each area as well as global - between areas computations with different properties (see below). This schematic is giving evidence that a popular view of a serial computations going from lower to upper anatomical areas has to be modified. There are no pure feed forward computations as all areas are strongly interconnected.

We suggest that the retina is responsible for creating preliminary hypotheses about certain features of perceived objects [35]. In one part of the Thalamus: in the Lateral Geniculate Nucleus (LGN), each hypothesis is compared with the prediction from the higher

visual areas [35, 36]. If prediction and hypothesis are in agreement the decision signal is sent to the motor system to perform action [37]. This process of predictions and hypotheses is repeated in different levels of higher visual areas. In this project, we will limit our model to three hierarchical levels: LGN, V1, and V4.

The feedback interactions with horizontal connections are anatomically complex and still not fully clarify functionally [37]. Cortico-thalamic cells with somas in layer V have far more extensive axonal ramifications in the cortex and thalamus. They have dendrites in the layer I and their axons give off a number of horizontal collaterals in layers III and V and then descend to the thalamus and to other sub-cortical structures such as the tectum, other parts of the brain stem, or the spinal cord. Unlike the axons of a layer VI cells, axons of layer V cells do not give off collaterals to the reticular nucleus and they are not restricted to the nucleus from which their parent cortical area receives inputs (like for a layer VI neurons). Their axons extend into one or more adjacent nuclei, although in each nucleus the terminals can be more focused than those of the axons of layer VI cells. The focusing of the layer V projection in comparison with layer VI projection does not imply a greater degree of topographic specificity because their intracortical projections are widespread in comparison to highly columnar layer VI projections.

2.10. Logic of the anatomical connections

As it was mentioned above, our model consists of three interconnected visual areas. Their connections can be divided into feedforward (FF) and feedback (FB) pathways. We have proposed [35] that FF connections are related to the hypothesis about stimulus attributes and FB pathways are related to predictions. Below, we suggest that the different anatomical properties of the FB and FF pathways may determine their different logical rules.

We define LGN_i , as LGN i -cell attributes for cells $i = 1, \dots, n$, $V1_j$ as primary visual cortex j -cell attributes for cells $j = 1, \dots, m$, and $V4_k$ as area V4 attributes for cells $k = 1, \dots, l$.

The specific stimulus attributes for a single cell can be found in the neurophysiological experiment by recording cell responses to the set of various test stimuli. As we have mentioned above, cell responses are divided into several (here 3) ranges, which will define several granules for each cell. It is different

from the classical receptive field definition, which assumes that the cell responds (logical value 1) or does not respond (logical value 0) to the stimulus with certain attributes. In the classical electrophysiological approach all receptive field granules are crisp. In our approach, cell responses below the threshold – r_0 , have logical value 0, the maximum cell responses – r_2 , have a logical value 1 but we will introduce cell responses between r_0 and r_2 , in this paper only one value r_1 . The physiological interpretation of cell responses between the threshold and the maximum response may be related to the influence of the feedback, horizontal pathways or matrix projections. We assume that the tuning of each structure is different and we will look for decision rules in each level that give responses r_1 and r_2 . For example, we assume that r_1 means that the local structure is tuned to the attributes of the stimulus and such granule for j -cell in area V1 will be define as $[u]_{1V1j}$.

2.10.1. Bottom-up computations (BUCs)

We will describe the logic of BUCs on the basis of LGN to V1 pathways, and by simplified direct and indirect influence of area V1 on area V4. Thalamic axons target specific cells in layers 4 and 6 of the primary visual cortex (V1). As Hubel and Wiesel [5] proposed, LGN cells determine orientation of SCs with their receptive fields arranged along the preferred orientation of the V1 cell (Fig. 2). There is high specificity between RF properties of the LGN cells and SC if they have monosynaptic connections [45]. The precision goes beyond simple retinotopy and includes such RF properties as RF sign, timing, subregion's strength, and size [45]. This high specificity of connections determines that V1 cell response is a result of assembly activity of several specific LGN cells "connected" by the logical "AND" as it was already discussed above. This is related to the fact that several aligned receptive fields in LGN must be simultaneously activated ("AND") in order to activate V1 cell connected to them [24]. As Sherman and Guillery [38] have proposed, we will call such inputs drivers. See formal rules in [29].

We assume that the neuron in area V4 receives driver inputs directly from cells in area V1 as well as indirectly through area V2 with highly specific RF properties (as described above for connections between LGN and V1 – equation 1). Therefore, the logical "and" has the same meaning as above: every input neuron from V1 "connected" to V4 (x_n, y_n) cell must be activated in order to activate V4 cell (more explicit formula in [29]). However, in this case

"connection" can be changed by the descending pathways (see below).

2.10.2. Top-down computations (TDCs)

The bases of TDCs are anatomical and physiological properties of descending pathways. Their function is to perform similarity verification that may lead to recognition. In the primate visual system the first descending pathway is from area V1 to the LGN.

Experimental results show that V1 feedback connections are restricted to the LGN region, visual-topically coextensive with the size of the classical RF of V1 layer 6 cells [37]. We will call after [38], feedback inputs as the modulators.

Decision Rules for TDCs from V4 to V1 or V4 to LGN will have similar syntax even if anatomical and physiological properties of the feedback pathways are different. Retrograde anatomical tracing has shown descending axons from area V4 directly to area V1 [39]. Axons of V4 cells span into area V1 in distinct clusters or in a linear array. The different semantics in decision rules are V4 cell specific and are related to the shapes of individual and variable axon branches in area V1. An axon's cluster that has terminals on V1 cells near "pinwheel centers" where cells show sub-threshold responses to all orientations will be responsible for the V4 subfield orientation tuning. If a linear array of terminals is connected to V1 neurons with similar orientation preference (narrowly tuned neurons [40]) - place tuning will take place. Retrograde tracing from area V4 showed axons projecting to different layers of the LGN with terminations in distinct clusters or in linear branches [41]. These projections will also tune orientation and place of V4 cell subfields but with different precision than V4 to V1 pathways [41]. To summarize, object recognition has two stages: at first BUCs classify all possible objects' similarities in different visual areas; in the next stage TDCs verify BUCs classification. In the following paragraph we will apply our computational model to experimental data from the area V4.

3. Results and analysis

We have analyzed the experimental data from several neurons recorded in the monkey's V4 [12]. Below we show a modified figure from the above work (Fig. 3), along with the associated decision table (Table 1). On the basis of the decision table we have made a schematic of the optimal stimulus for this cell (Fig. 4 right side). Figure 4 (left side) shows the cell's

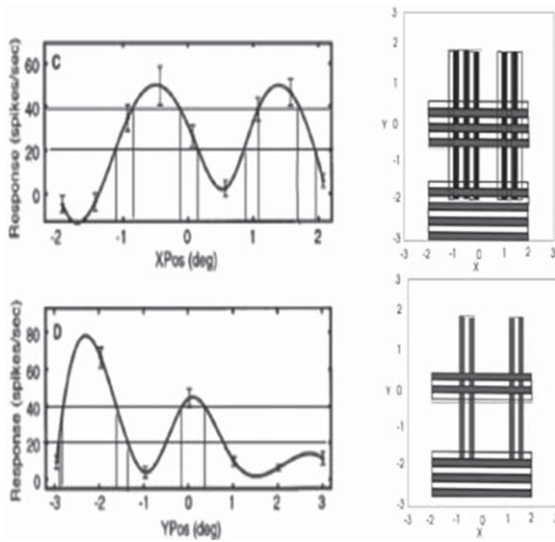


Fig. 3. Curves represent approximated responses of a cell from area V4 to vertical (C), and horizontal (D) bars. Bars change their position along x-axis (Xpos) or along y-axis (Ypos). Responses of the cell are measured in spikes/sec. Mean cell responses ± SE are marked in the figures. Cell responses are divided into three ranges (concepts) by two horizontal lines. On the right is a schematic representation of cell response on the basis of Table 1. Vertical and horizontal bars in certain x- and y-positions gave strong (r1: class 1 – upper schematic) or very strong (r2: class 2 – lower schematic) responses.

Table 1
Decision table for cell shown in Fig. 3 upper part

Cell	o1	sf1	xs1	ys1	xp1	yp1	c1	r
12	90	0.5	0.4	4	-2	0	0.9	0
12	90	0.5	0.4	4	-1.5	0	0.9	0
12	90	0.5	0.4	4	-1	0	0.9	1
12	90	0.5	0.4	4	-0.5	0	0.9	2
12	90	0.5	0.4	4	0	0	0.9	1
12	90	0.5	0.4	4	0.5	0	0.9	0
12	90	0.5	0.4	4	1	0	0.9	2
12	90	0.5	0.4	4	1.5	0	0.9	2
12	90	0.5	0.4	4	2	0	0.9	0

responses to the stimulus, which was a long narrow bar with vertical (Fig. 4C) or horizontal (Fig. 4D) orientation.

The decision table (Table 1) describes properties of stimuli and their position as a function of response strength. This table is converted into a schematic (right of Fig. 1), which shows areas of cell responses related to category 1 (upper part) and to category 2 (lower part). Strong cell responses are not symmetric along the middle of the receptive field, but divide the receptive field into several smaller subfields.

These results are the basis of the idea that the receptive field of V4 neurons can be divided into several independent parts [12]. Our results can be presented as follows:

$$o_{90} \wedge (xpr_{0.5} \vee xpr_{0.6}) \wedge xs_{0.4} \wedge ys_4 \rightarrow r_2 \quad (3)$$

$$o_0 \wedge (ypr_{1.2} \vee ypr_{0.7}) \wedge xs_4 \wedge ys_{0.4} \rightarrow r_2 \quad (4)$$

All attributes as described above. Similar table we write for the lower part of the Fig. 3. For each row we can write a rule that describes results from Fig. 3.

For example for row 1 we write the rule:

$$"cell = 12" \& "o1 = 90" \& \dots \& "c1 = 0.9" \rightarrow r = 0 \quad (5)$$

We read it as, if cell number is 12 and stimulus orientation is 90 and ... and stimulus contrast is 0.9 then cell response is 0.

We would like to find how general are these rules and if we can simulate responses of another cell to different stimulus on the basis of such rules?

Two-bar experiment demonstrates responses to small bars along x-axis in the receptive field (Fig. 4).

In the next step we can use FRST hybrid rules with $t.norm = "Lukasiewicz"$, $tolerance = "Equation 1"$, $implicator = "Lukasiewicz"$ to predict results from Fig. 3 by Two-bars experiment we obtained a cross-validation Table 3.

Total accuracy is 0.66 and coverage for FRST is always 1. The best accuracy is for class 0, for class 2 there were no such predictions. We have also analyzed 'ShiftPatch' experiments where a small patch of grating was placed in different parts of the V4 receptive field [12]. These patches can have different orientations or spatial frequencies. Predictions of the experimental results from 'ShiftPatch' using RST rules from Two-bar experiments gave the total accuracy 0.494 exactly the same as prediction with FRST. But predictions with FRST rules of 'ShiftPatch' experiments where only orientation was changing, from the 'Two-bar' experiment, gave the total accuracy of 0.556. FRST rules from other experiments 'MapRF' [12] gave predictions of 'ShiftPatch' experiments with only spatial frequency changes with the total accuracy of 0.583.

We can divide all above measurement into 9 random groups and find rules from 8 groups in order to predict responses in the 9th testing group and changing learning and testing groups average results in 9-fold cross validation for 'Two-bar' experiment we have obtained total accuracy 0.907 but with total coverage 0.444 using the decision tree classification.

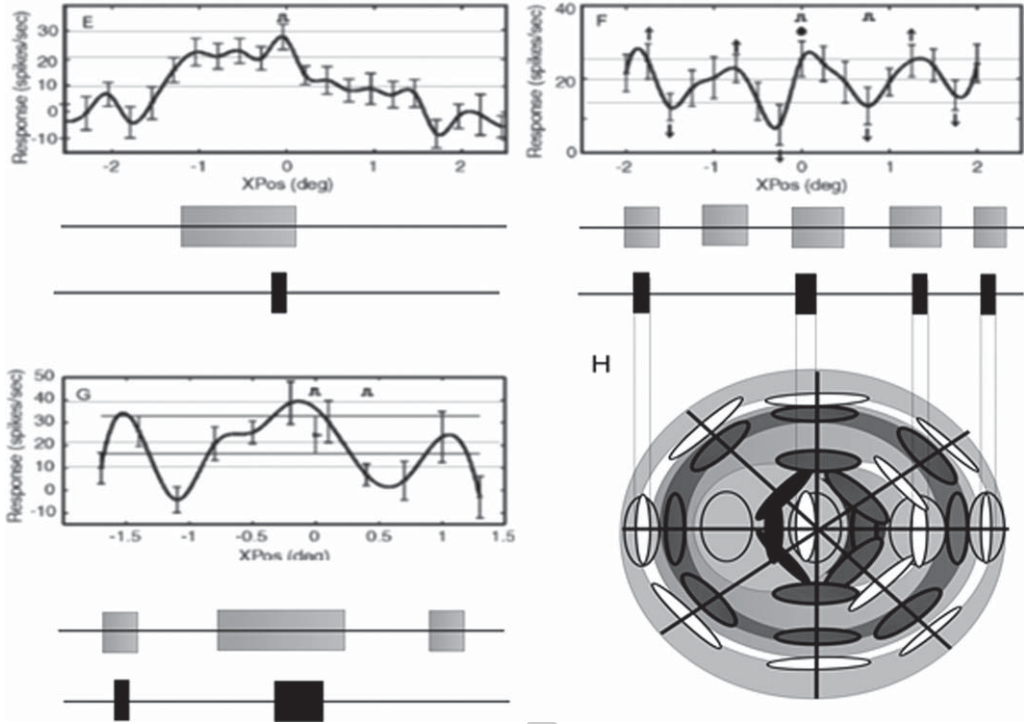


Fig. 4. Modified plots from [12]. Curves represent responses of two cells from area V4 to small single (E) and double (F, G) vertical bars. Bars change their position along x-axis (Xpos). Responses are measured in spikes/sec. Mean cell responses SE are marked in E, F, and G. Cell responses are divided into three ranges by thin horizontal lines. Below each plot are schematics showing bar positions giving r1 (gray) and r2 (black) responses; below (E) for a single bar, below (F and G) for double bars (one bar was always in position 0). (H) This schematic extends responses for horizontally placed bars (E) to the whole RF with assumption that each axis gives the same responses as x-axes: white color shows excitatory related to r2 responses, gray color is related to r1 responses and black color inhibitory interactions between bars.

Table 2
Part of the decision table for cell shown in Fig. 4E one bar

Cell	o ₁	sf ₁	xs ₁	ys ₁	xp ₁	yp ₁	c ₁	r
61	0	0.5	0.25	4	-2.25	0	1	0
61	0	0.5	0.25	4	-2.0	0	1	0
61	0	0.5	0.25	4	-1.75	0	1	0
61	0	0.5	0.25	4	-1.5	0	1	0
61	0	0.5	0.25	4	-1.0	0	1	1
61	0	0.5	0.25	4	-0.75	0	1	1
61	0	0.5	0.25	4	-0.5	0	1	1
61	0	0.5	0.25	4	-0.25	0	1	1
61	0	0.5	0.25	4	0.0	0	1	1

All together there is 39 rows in Table 2 and using rough set theory (RST) we can find 24 rules, e.g.

$$(xp1 = "(1.375, Inf)") \& (c2 = "(-Inf, 0.5)") \Rightarrow (r = 0[5]) \quad (6)$$

$$(xp1 = "(-0.775, 0.375)") \Rightarrow (r = 1[4]) \quad (7)$$

$$(xp1 = "(-1.25, 1.0)") \Rightarrow (r = 0[3]) \quad (8)$$

Table 3
Cross validation table predicts results in Fig. 3 from Fig. 4

		Predicted		
		0	1	2
Actual	0	0.44	0	0
	1	0	0.22	0
	2	0.11	0.22	0

$$(xp1 = "(-0.225, 0.05)") \& (c2 = "(0.5, -Inf)") \Rightarrow (r = 2[2]) \quad (9)$$

We can also find Two-bar experiment rules using fuzzy RST (FRST):

$$(xp1 = ' -1.1') \Rightarrow (r = 1) \quad (10)$$

$$(xp1 = ' 0') \Rightarrow (r = 0) \quad (11)$$

$$(sf2 = ' 0') \& (xp1 = ' -2') \Rightarrow (r = 0) \quad (12)$$

$$(sf2 = ' 0.5') \& (xp1 = ' 1.5') \Rightarrow (r = 1) \quad (13)$$

4. Discussion

In this paper we have considered possible mechanisms on ‘how visual system can figure out’ properties of the unseen object. We have proposed to formalize the receptive field (RF) properties with help of rough and fuzzy set theories. By using this concept and by normalization to several levels neuronal responses one can check decisions performed by each neuron in response to different stimuli. These decisions tell us how similar RF and object (stimulus) properties are.

Neurons in area V4 integrate an object’s attributes from the properties of its parts in two ways: (1) within the area via horizontal or intra-laminar local excitatory-inhibitory interactions, (2) between areas via feedback connections tuned to lower visual areas. Our research put more emphasis on feedback connections because they are probably faster than horizontal interactions [39]. Different neurons have different Subfield Interactions Rules as described in the Results section and perceive objects by way of multiple ‘fuzzy windows’. If an object’s attributes fit the fuzzy window, a neuron sends positive feedback [33] to lower areas, which as described above, use the ‘modulator-type’ logic (MTL) to sharpen the attribute-extracting window and therefore change response of the neuron from class 1 to class 2. The above analysis of our experimental data leads us to suggest that the central nervous system chiefly uses at least two different ‘logical rules’: ‘driver-types’ logical rules (DTL) and ‘modulator –type’ (MTL) rules.

The first, DTL processes data using a large number of possible algorithms (over-representation). The second, MTL supervises decisions and chooses the right algorithm. As we have described, there are experimental [42, 43] and theoretical [44] findings suggesting that properties of RF in lower areas can be tuned by descending pathways. These findings are basis for the universality of our visual system that by learning and trials can recognize unseen objects by changing hypotheses about their actual properties. It is based on the similarities.

We have demonstrated of concept on the experimental data with very limited number of cells and variety of stimuli. Advantage of this approach is that it is independent on any subjective research hypothesis.

Already in 1997 Biederman [46] suggested that we recognize objects by components and he has proposed simple geometrical elementary components as a base for human image understanding. In one of

my previous paper, I have demonstrated how Biederman’s figures could fit to V4 receptive fields (RF) [2]. Interactions between stimuli being part of the RF were used in many studies in V4 and higher areas like IT (infortemporal cortex) [47]. For example, texture segregation is related texture-defined figures with homogeneous textures. It is related to an early enhancement of the figure representation as small V4 RF subfield (as V1 RF size), and a later suppression of the background in full V4 RF [48]. Another recent finding related to subfields of V4 RF is the visual crowding. It is based on the tuning selectivity for stimuli within the receptive field (RF) of the area V4. As V4 RF are much larger than letter-like stimuli, the fusion of separate objects into a single identity have to occur within the V4 RF [49].

In the recent paper [50], authors demonstrated the first time of **blur encoding** in the V4 area of monkeys. They have measured tuning in V4 neurons for both object shape and boundary blur. Moreover, they observed that blur tuning is specific and not confused with stimulus size, intensity, or curvature [50]. These new data can be explained by changing tolerance in our fuzzy set rules.

5. Conclusion

By applying the FRST to neuro-physiological data we have demonstrated a new formalized approach: how the visual brain may perform object categorization in the psychophysical space. These processes are related to anatomical and physiological properties of the visual system: ascending and descending pathways are related to hypotheses and predictions and mirrored by different logical systems (DLT vs. MLT: driver-type vs. modulatory-type logic). These different logical rules look for similarities between properties of the object or its parts in comparison to RF properties of neurons in LGN, V1, V2, V4 and higher areas in the ventral stream. In agreement with previous experiences the right hypothesis that is the most similar to our predictions about the object is chosen. It is the basis of the cognition related to the first order processes (spike rates). Using the same fuzzy logical systems (DLT vs. MLT) one can describe higher order processes related to oscillatory processes. By extending of our retina model as the coupled nonlinear oscillatory system to higher visual areas we propose that in this case also DLT vs. MLT interactions will be the basis of cognition. The bottom-up system consists of a large number of

possible orbits and only some of them are chosen by top-down parametric control of the lower level oscillators.

Acknowledgments

This material was partly presented as an invited talk in RIKEN, Tokyo, Japan on April 2017.

References

- [1] A.W. Przybyszewski, I. Kagan and M. Snodderly, Primate area V1: Largest response gain for receptive fields in the straight-ahead direction, *NeuroReport* **25** (2014), 1109–1115.
- [2] A.W. Przybyszewski, The neurophysiological bases of cognitive computation using rough set theory, *Trans Rough Sets IX, LNCS 5390* (2008), 287–317.
- [3] Z. Pawlak, *Rough Sets—Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Boston, London, Dordrecht, 1999.
- [4] L. Polkowski and A. Skowron, Rough mereology: A new paradigm for approximate reasoning, *Int J Approx Reason* **15** (1996), 333–365.
- [5] D.H. Hubel and T.N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J Physiol* **160** (1962), 106–154.
- [6] K. Fukushima, Cognitron: A self-organizing multilayered neural network, *Biol Cyber* **20** (1975), 121–136.
- [7] K. Fukushima, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biol Cyber* **36**(4) (1980), 193–202.
- [8] J.P. Gaska, D.A. Pollen and P. Cavanagh, Diversity of complex cell responses to even- and odd-symmetric luminance profiles in the visual cortex of the cat, *Exp Brain Res* **68** (1987), 249–259.
- [9] C. Cadieu, M. Kouh, A. Pasupathy, C.E. Connor, M. Riesenhuber and T. Poggio, A model of V4 shape selectivity and invariance, *J Neurophysiol* **98**(3) (2007), 1733–1750.
- [10] M. Riesenhuber and T. Poggio, Neural mechanisms of object recognition, *Curr Opin Neurobiol* **12** (2002), 162–168.
- [11] J.P. Gaska, D.A. Pollen and P. Cavanagh, Diversity of complex cell responses to even- and odd-symmetric luminance profiles in the visual cortex of the cat, *Exp Brain Res* **68** (1987), 249–259.
- [12] D.A. Pollen, A.W. Przybyszewski, M.A. Rubin and W. Foote, Spatial receptive field organization of macaque V4 neurons, *Cerebral Cortex* **12** (2002), 601–616.
- [13] J. Bazan, Son, H. Nguyen, T. Trung, Nguyen, A. Skowron and J. Stepaniuk, Decision rules synthesis for object classification, in: E. Orłowska (ed.), *Incomplete Information: Rough Set Analysis*, Physica-Verlag, Heidelberg, 1998, pp. 23–57.
- [14] J.G. Bazan and M. Szczuka, RSES and RSESlib—A collection of tools for rough set computations, in: W. Ziarko and Y. Yao (eds.), *RSCTC 2000 LNCS(LNAI)*, 2001, vol. 2005, Springer, Heidelberg, pp. 106–113.
- [15] J.W. Grzymala-Busse, A new version of the rule induction system LERS, *Fund Inform* **31**(1) (1997), 27–39.
- [16] D. Dubois and H. Prade, Rough fuzzy sets and fuzzy rough sets, *Int J Gen Syst* **17** (1990), 91–209.
- [17] L.A. Zadeh, Fuzzy sets, *Inform Control* **8** (1965), 338–353.
- [18] L. Riza, A. Janusz, C. Bergmeir, C. Cornelis, F. Herrera, D. Slezak and J.M. Benítez, Implementing algorithms of rough set theory and fuzzy rough set theory in the R package “RoughSets”, *Inform Sci* **287** (2014), 68–89.
- [19] S.W. Kuffler, Neurons in the retina: Organization, inhibition and excitation problems, *Cold Spring Harbor Symp Quant Biol* **17** (1952), 281–292.
- [20] H.B. Barlow, The coding of sensory messages, in: W.H. Thorpe and O.L. Zangwill (eds), *Current Problems in Animal Behaviour*, Cambridge University Press, Cambridge, 1961, pp. 331–360.
- [21] J.D. Daugman, Uncertainty relation for the resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, *J Opt Sci Am A* **2** (1985), 1160–1169.
- [22] A.W. Przybyszewski, An analysis of the oscillatory patterns in the central nervous system with the wavelet method, *J Neurosci Methods* **38** (1991), 247–257.
- [23] A.W. Przybyszewski, M.J. Lankheet and W.A. van de Grind, Nonlinearity and oscillations in X-type ganglion cells of the cat retina, *Vis Res* **33** (1993), 861–875.
- [24] A.W. Przybyszewski, M.J. Lankheet and W.A. van de Grind, On the complex dynamics of intracellular ganglion cell light responses in the cat retina, *Biol Cyber* **74** (1996), 299–308.
- [25] A.W. Przybyszewski, P.S. Linsay, P. Gaudiano and C. Wilson, Basic difference between brain and computer: Integration of asynchronous processes implemented as hardware model of the retina, *IEEE Trans Neural Networks* **18** (2007), 70–85.
- [26] D. Wang and D. Terman, Image segmentation based on oscillatory correlation, *Neural Comput* **9** (1997), 805–836.
- [27] L.A. Zadeh, Toward a perception-based theory of probabilistic reasoning with imprecise probabilities, *J Statist Plan Inference* **105** (2002), 233–264.
- [28] L.A. Zadeh, From computing with numbers to computing with words—From manipulation of measurements to manipulation of perceptions, *Int J Appl Math Comput Sci* **12**(3) (2002), 307–324.
- [29] A.W. Przybyszewski, Logical rules of visual brain: From anatomy through neurophysiology to cognition, *Cogn Syst Res* **11** (2010), 53–66.
- [30] C. Bardy, J.Y. Huang, C. Wang, T. Fitz Gibbon and B. Dreher, Simplification of responses of complex cells in cat striate cortex: Suppressive surrounds and ‘feedback’ inactivation, *J Physiol* **574** (2006), 731–750.
- [31] I. Kagan, M. Gur and D.M. Snodderly, Spatial organization of receptive fields of V1 neurons of alert monkeys: Comparison with responses to gratings, *J Neurophysiol* **88** (2002), 2557–2574.
- [32] L. Shams and C. von der Malsburg, The role of complex cells in object recognition, *Vis Res* **42** (2002), 2547–2554.
- [33] A.W. Przybyszewski, J.P. Gaska, W. Foote and D.A. Pollen, Striate cortex increases contrast gain of macaque LGN neurons, *Vis Neurosci* **17** (2000), 485–494.
- [34] R.J. Douglas and K.A. Martin, Neuronal circuits of the neocortex, *Annu Rev Neurosci* **27** (2004), 419–451.
- [35] A.W. Przybyszewski and M.A. Kon, Synchronization-based model of the visual system supports recognition. Program No. 718.11. 2003 Abstract Viewer/Itinerary Planner. Society for Neuroscience, Washington, DC, Online.

- [36] A.W. Przybyszewski, Vision: Does top-down processing help us to see? *Curr Biol* **8** (1998), R135–R139.
- [37] E.G. Jones, Synchrony in the interconnected circuitry of the thalamus and cerebral cortex, *Ann NY Acad Sci* **1157** (2009), 10–23.
- [38] S.M. Sherman and R.W. Guillery, On the actions that one nerve cell can have on another: Distinguishing drivers and modulators, *Proc Natl Acad Sci USA* **95** (1998), 7121.
- [39] K.S. Rockland, K.S. Saleem and K. Tanaka, Divergent feedback connections from areas V4 and TEO in the macaque, *Vis Neurosci* **11** (1994), 579–600.
- [40] J. Schummers, J. Marino and M. Sur, Synaptic integration by V1 neurons depends on location within the orientation map, *Neuron* **36** (2002), 969–978.
- [41] A.W. Przybyszewski, D.O. Potapov and K.S. Rockland, Feedback connections from area V4 to LGN, *Ann Meet Society for Neuroscience*, San Diego, USA, 2001. <http://sfn.scholarone.com/itin2001/prog#620.9>
- [42] P. Girard, L.J.M. Hupe and J. Bullier, Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities, *J Neurophysiol* **85** (2001), 1328–1331.
- [43] C. Bardy, J.Y. Huang, C. Wang, T. Fitz Gibbon and B. Dreher, Simplification of responses of complex cells in cat striate cortex: Suppressive surrounds and ‘feedback’ inactivation, *J Physiol* **574** (2006), 731–750.
- [44] T. Poggio and S. Edelman, A network that learns to recognize three-dimensional objects, *Nature* **343** (1990), 263–266.
- [45] J.M. Alonso, W.M. Usrey and R.C. Reid, Rules of connectivity between geniculate cells and simple cells in cat primary visual cortex, *J Neurosci* **21** (2001), 4002–4015.
- [46] I. Biederman, Recognition-by-components: A theory of human image understanding, *Psychol Rev* **94** (1987), 115–147.
- [47] A.P. Sripati and C.R. Olson, Responses to compound objects in monkey inferotemporal cortex: The whole is equal to the sum of the discrete parts, *J Neurosci* **30** (2010), 7948–7960.
- [48] J. Poort, M.W. Self, B. van Vugt, H. Malkki and P.R. Roelfsema, Texture segregation causes early figure enhancement and later ground suppression in areas V1 and V4 of visual cortex, *Cereb Cortex* **26**(10) (2016), 3964–3976.
- [49] B.C. Motter, Stimulus conflation and tuning selectivity in V4 neurons: A model of visual crowding, *J Vis* **18**(1) (2018), 1–24.
- [50] T.D. Oleskiw, A. Nowack and A. Pasupathy, Joint coding of shape and blur in area V4, *Nat Commun* **9**(466) (2018), 1–13.