Neurocomputing 199 (2016) 185–196

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A neural model for straight line detection in the human visual cortex

Theju Jacob [a,*], Wesley Snyder [a], Jing Feng [b], HeeSun Choi [b]

[a] Department of Electrical and Computer Engineering, NC State University, Raleigh, USA
[b] Department of Psychology, NC State University, Raleigh, USA

ARTICLE INFO

Article history:
Received 11 August 2015
Received in revised form
1 February 2016
Accepted 7 March 2016
Communicated by Bin Fan
Available online 1 April 2016

Keywords:
Hough transform
Straight line detection
Primary visual cortex
Oja's rule
Orientation columns

ABSTRACT

Building a computational model for how the visual cortex identifies objects is a problem that has attracted much attention over the years. Generally, the interest has been in creating models that are translation, rotation, and luminance invariant. In this paper, we utilize the philosophy of Hough Transform to create a model for detecting straight lines under conditions of discontinuity and noise. A neural network that can learn to perform a Hough Transform-like computation in an unsupervised manner is the main takeaway from this work. Performance of the network when presented with straight lines is compared with that of human subjects. Optical illusions like the Poggendorff illusion could potentially find an explanation in the framework of our model.

1. Introduction

The human visual system is vastly superior to man-made vision systems in its ability to identify objects. Building a model for how the visual cortex identifies different objects is a problem that has attracted much attention over the years. Both top-down and bottom-up [1,42,4] approaches have been proposed towards this end. Top-down models start at a high-level representation of the incoming visual input, while bottom-up models start with simple features within the input and then move to more complex features.

Oriented lines are one of the first features detected by the visual cortex in a bottom-up approach, see [18,19]. In this paper, we utilize the philosophy of the Hough Transform to explain how the primary visual cortex could possibly detect straight lines. We describe in detail a neural model that can learn a Hough Transform-like structure in an unsupervised manner. Our focus is on the learning principles and patterns of connectivity of the network of artificial neurons, without getting into creating a detailed biologically realistic network like in [33]. Experiments involving human subjects that add support to our hypothesis are also outlined.

The Hough Transform is a popular feature detection technique in computer vision applications, see [32,7,39]. The Hough domain, when used to detect straight lines, is characterized by two parameters: (1) the orientation of the line and (2) the shortest distance from the origin to the line, Fig. 1a. The underlying parametric transformation from Cartesian coordinates is represented by the following equation:

$$\rho = x\ \cos(\theta) + y\ \sin(\theta) \tag{1}$$

A straight line in the Cartesian domain is thus represented by a single ρ, θ point in the Hough domain.

Practical implementation of the Hough Transform algorithm relies on an accumulator array $\mathcal{A}(\rho,\theta)$ with all possible orientations and distances. The parameters of the line are obtained by looking at the ρ and θ that reflect the maximum increments. The transform is insensitive to clutter and partial occlusion, as we rely on a voting process to determine the parameters of the line.

Neural networks that can detect straight lines have appeared in literature. Ref. [34] discusses a neural network structure for detecting straight lines of various orientations. However, no learning happens in the network. The network replicates input lines during detection, and no accumulation is involved. A spiking neural network model that implements the Hough Transform is discussed in [43]. Once again, no learning is involved as the weights for the network connections are derived directly from the Hough Transform formulation.

A neural network capable of achieving a Hough Transform parameterization is discussed in [2]; however, the accumulator

* Corresponding author. Current address: Massachusetts General Hospital/ Harvard Medical School, Boston, USA.
E-mail address: theju.jacob@mgh.harvard.edu (T. Jacob).
URL: http://www.tjacob.org (T. Jacob).

http://dx.doi.org/10.1016/j.neucom.2016.03.023
0925-2312/

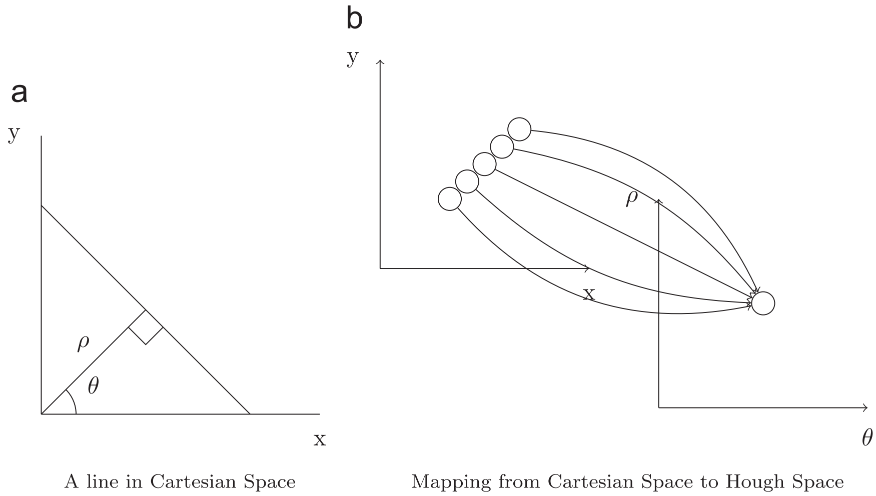

Fig. 1. Parameters of a straight line in the Hough space.

array concept is not utilized and the network needs to search for the weight vectors representing straight lines over multiple iterations, every time. By comparison, after training, our network detects the lines in a two-step computation. Ref. [28] also discusses a neural network for learning the Hough Transform parameter space. Their training data includes parameter values and input image values, and the network is trained using back-propagation. Our network is unsupervised, and uses only the incoming input image as training data. The problem of finding the maxima in a Hough Transform parameter space using biologically inspired ideas as opposed to computing the Hough Transform itself is addressed in [5].

Neural activity models that lead to orientation-filter like characteristics have also appeared in the literature previously. For example, learning a sparse code for images as a model for neural activity is discussed in [36,9]. Predictive coding as an explanation for visual processing in the cortex is discussed in [38]. Information maximization as a goal of sensory coding is discussed in [3,31,30]. A self-organizing map architecture capable of extracting features similar to that extracted in the early stages of visual processing is discussed in [25].

Our network differs from previous networks in that we explain not only how neurons can learn to be orientation sensitive, but also how the neurons in subsequent layers can pool information from orientation sensitive cells and learn to detect entire lines even in the presence of discontinuity and noise. Orientation sensitivity becomes evident in two aspects of our network after training: (1) In the connections developed by neurons in the same layer and (2) in connections between neurons in different layers. The former contributes to an associative memory like behavior [26,27], and a derivation of our learning rule that enables it is discussed in our work [24].

The latter demonstrates itself in the receptive fields of neurons in the receiving layer. This in turn causes a single neuron or a small group of neurons in the later layer to represent an entire line in the previous layer, much like a mapping from Cartesian space to the Hough Space. The idea of pooling of neuronal outputs in subsequent layers has been explored previously [29,12]. Further, the local nature of connections employed by our network leads to a topological mapping of data between subsequent layers—that is, neighboring lines in the input space would be represented by neighboring neurons in our network. Local connections between neurons and topological mapping are features that are present in the cortex as well.

After training, the computation done by our network is comparable to that of the standard model of object recognition [40]. In the standard model, a battery of filters is fixed as S1 cells, and their outputs are combined in C1 cells. After training, the receptive fields of our network neurons become Gabor-like, like those of S1 cells, while the winner-neuron computation performed by our network becomes a computation comparable to that of a C1 cell operation.

Parallels may be drawn between our network architecture and that of Boltzmann/Restricted-Boltzmann machines [16,17]. However, the learning algorithms are very different—for Boltzmann machines, the parameter of interest is the global energy of the network. In our network, the interest of the learning rule is in the normalization of incoming weights to individual neurons. Boltzmann machines aim to model the input distribution, while our network learns input features and pool them, retaining spatial relationships.

It is well established from experiments by Hubel and Wiesel, see [18,19], that the primary visual cortex contains cells that are sensitive to the orientation of the visual stimuli. It is further known that the orientation sensitive cells aggregate into columns called orientation columns, and that the sensitivity of the columns themselves vary in a sequential manner. We also know that the cells of the visual cortex are organized into retinotopic maps. That is, neighboring areas in the visual fields map to neighboring areas in the cortex itself [20–23]. Parallels may thus be drawn between architectural features of Hubel and Wiesel's 'ice-cube model' of the cortex and that of the accumulator array.

An accumulator like structure could potentially explain the phenomenon of log polar transformation observed in the cortex [41,6]. For this discussion, assume that the center of an input image acts as the origin, and that the origin corresponds to the fovea. Suppose there is a straight line in the input image. Every point along that line will increment a set of cells in the accumulator array, and a maximum is obtained in a single cell or set of cells. The cell or group of cells with the maximum value would indicate the position, or the distance of the line from the origin, ρ.

Now consider a particular point p in the input image. Let the point be at a distance d from the origin. Any line which passes through p will have a ρ which is less than or equal to d, i.e., $\rho \le d$. Since p will contribute to the accumulator array every time a line passes through p, the cells excited by point p will indicate lines with ρ from 0 to d. That is, p should excite a total of d number of cells in the accumulator array. Since a point closer to the origin will have a d that is smaller than that of a point which is farther