Neural Networks 23 (2010) 528–540

Contents lists available at ScienceDirect

Neural Networks

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

2010 Special Issue

Neural network model for completing occluded contours

Kunihiko Fukushima [*,1]

Kansai University, Takatsuki, Osaka, Japan

ARTICLE INFO

Article history:
Received 6 July 2008
Received in revised form 18 September 2009
Accepted 12 October 2009

Keywords:
Vision
Occlusion
Amodal completion
Neural network model
Bend-extracting cell
Bottom-up and top-down

ABSTRACT

When some parts of a pattern are occluded by other objects, the visual system can often estimate the shape of occluded contours from visible parts of the contours. This paper proposes a neural network model capable of such function, which is called *amodal completion*. The model is a hierarchical multi-layered network that has bottom-up and top-down signal paths. It contains cells of area V1, which respond selectively to edges of a particular orientation, and cells of area V2, which respond selectively to a particular angle of bend. Using the responses of bend-extracting cells, the model predicts the curvature and location of the occluded contours. Missing contours are gradually extrapolated and interpolated from the visible contours. Computer simulation demonstrates that the model performs amodal completion to various stimuli in a similar way as observed by psychological experiments.

1. Introduction

Even if some parts of a pattern are occluded by other objects, the visual system can often estimate the shape of missing portions from visible parts of the contours. If we see an image shown in the left of Fig. 1, we feel as though a circular disk is occluded by a gray square and that the contour of the disk is connected behind the square. Although all patterns listed in the right of the figure look the same when occluded, we human beings usually feel that the top right is the case. Such a process of perceptually filling in parts of objects that are hidden from view is called *amodal completion*. It is closely related to the phenomena of subjective contour, figure-ground segregation, border ownership, perceptual grouping, and so on. A great deal of psychological research related to these phenomena has been performed (e.g., August, Siddiqi, & Zucker, 1999; Kellman & Shipley, 1991; Singh & Fulvio, 2005, 2007; Takeichi, Nakazawa, Murakami, & Shimojo, 1995), and a lot of models have been proposed so far (e.g., Craft, Schütze, & von der Heydt, 2007; Fantoni & Gerbino, 2003; Fowlkes, Martin, & Malik, 2007; Grossberg & Mingolla, 1985; Kikuchi, Anada, & Fukushima, 1998; Neumann & Mingolla, 2001; Sajda & Finkel, 1995; Sajda & Baek, 2004; Sakai & Nishimura, 2006).

Although the neurophysiological basis of amodal completion is not yet well understood, there are several findings that give hints for its neural mechanisms: in the primary visual cortex (area V1) of cats and monkeys, there are cells that respond selectively to oriented lines or edges, and these cells are classified into simple and complex cells (e.g., Hubel & Wiesel, 1962). Ito and Komatsu (2004) reported that, in area V2 of monkeys, there are cells that show highly selective responses to a particular angle of bend of line stimuli. We will call them *bend-extracting cells* in this paper. Bend-extracting cells are found also in V4 (Pasupathy & Connor, 1999). Sugita (1999) reported that some orientation selective cells in V1 respond to amodal contours. In V2, there are cells that respond to modal or subjective contours (Peterhans & von der Heydt, 1989; von der Heydt & Peterhans, 1989). Zhou, Friedman, and von der Heydt (2000) reported that some cells in V1, V2 and V4, which are usually thought to respond to local contrast and orientation of edges, are also sensitive to global configuration of contours and encode border ownership.

Suggested from these findings, this paper proposes a neural network model that has an ability of amodal completion. Among many neural mechanisms that work for amodal completion, we focus our attention in our model especially to the process of extrapolating smoothly curved contours.

The model is a hierarchical multi-layered neural network that has bottom-up and top-down signal paths. The network contains simple, complex and bend-extracting cells. The model estimates the curvature and location of occluded contours, and tries to extrapolate the hidden contours. Through the interaction of bottom-up and top-down signals, the occluded contours are smoothly completed, step by step. We demonstrate by computer simulation that our model can make amodal completion for various stimuli in a similar way as those observed by psychological experiments.

* Tel.: +81 44 988 5272; fax: +81 44 988 5272.
E-mail address: fukushima@m.ieice.org.
URL: http://www4.ocn.ne.jp/~fuku_k/index-e.html.

1 Permanent address: 634-3, Miwa, Machida, Tokyo 195–0054, Japan.

0893-6080/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2009.10.002

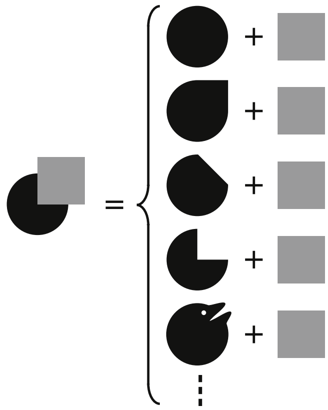

Fig. 1. Amodal completion. Although all patterns listed in the right look the same when occluded, we human beings usually feel that the top right is the case.

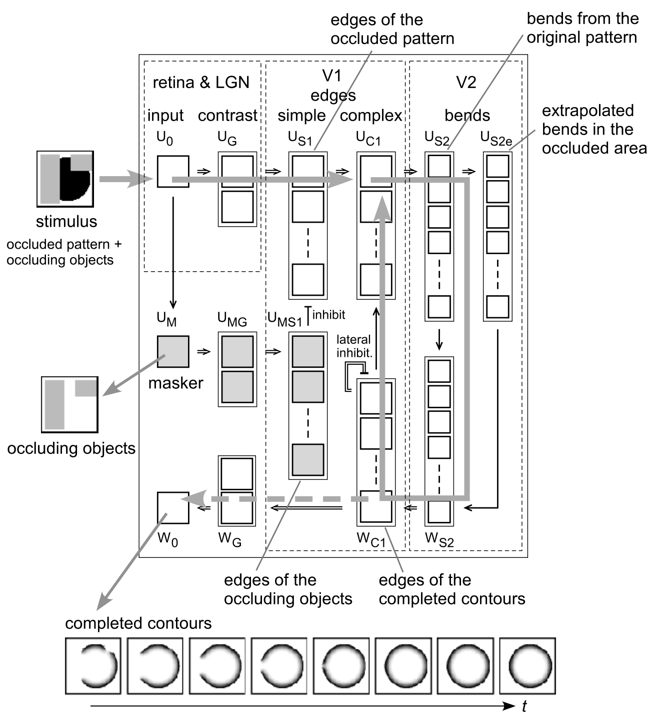

Fig. 2. Architecture of the model for amodal completion, in which an example of the responses of some layers are shown.

2. Neural network model

2.1. Outline of the model

The model is a hierarchical multi-layered network as illustrated in Fig. 2. It has bottom-up and top-down signal paths, by which a feedback loop is formed in the network. In the figure, U and W represent layers of cells in the bottom-up and top-down paths, respectively.

Each layer of the network consists of a number of *cell-planes*, depending on the difference in the features to which cells respond selectively. Incidentally, a cell-plane is a group of cells that are arranged retinotopically and share the same set of input connections (Fukushima, 1988). As a result, all cells in a cell-plane have receptive fields of an identical characteristic, but the locations of the receptive fields differ from cell to cell. For example, in layer U_{S1}, cells of the same preferred orientation constitute a cell-plane.

This section discusses the model qualitatively, and a more detailed mathematical description appears in the Appendix.

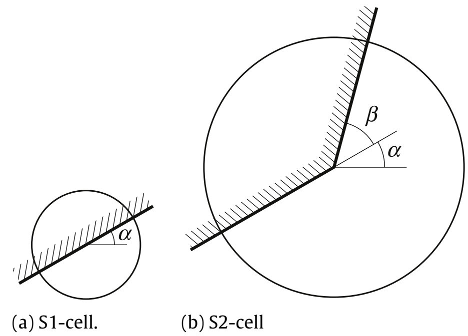

Fig. 3. Training patterns for S1- and S2-cells. These patterns become the preferred stimuli for S1- and S2-cells, respectively.

2.2. Extraction of oriented edges

The structure in $U_0 \rightarrow U_G \rightarrow U_{S1} \rightarrow U_{C1}$ is similar to that of the *neocognitron* (Fukushima, 2003; Kikuchi & Fukushima, 1996), which was suggested by the classical hypothesis of Hubel and Wiesel (1962).

The stimulus pattern is presented to input layer U_0, which consists of a two-dimensional array of photoreceptors. The output of layer U_0 is fed to contrast-extracting layer U_G, whose cells resemble retinal ganglion cells or lateral geniculate nucleus cells. Layer U_G consists of two cell-planes: one with concentric on-center receptive fields, and the other with off-center receptive fields. The former cells extract positive contrast in brightness, whereas the latter extract negative contrast from the image presented to the input layer.

The output of layer U_G is fed to edge-extracting layer U_{S1}, which consists of S1-cells. S1-cells resemble simple cells in area V1 (the primary visual cortex), and respond selectively to edges of a particular orientation. Namely, layer U_{S1} consists of K_1 cell-planes, and all cells in the kth cell-plane respond selectively to edges of orientation $2\pi k/K_1$. As a result, the contours of the input image are decomposed into edges of every orientation.

The orientation tuning curve of an S1-cell is chosen so broad that a straight edge of a particular orientation elicits responses from a couple of cell-planes of slightly different preferred orientations. Incidentally, we take $K_1 = 32$ in the computer simulation discussed later, but the exact number of the cell-planes does not matter so much.

The input connections to an S1-cell have the same spatial shape as the response of presynaptic cells to an oriented straight edge. To produce these connections, we use supervised learning, like that used for the neocognitron (Fukushima, 2003).[2]

When we train the kth cell-plane, for example, we first choose an arbitrary cell from the cell-plane. We present a training pattern like the one shown in Fig. 3(a) to input layer U_0. To be more specific, the training pattern is an edge of orientation $\alpha = 2\pi k/K_1$ that crosses the center of the receptive field of the chosen S1-cell. The input connection to the S1-cell, whose initial values are zero, are increased in proportion to the responses of the cells of layer U_G, from which the connections are leading. Thus the spatial distribution of input connections of an S1-cell comes to have the same shape as the response of on- and off-center cells to a straight edge.

The output of layer U_{S1} is fed to layer U_{C1}, which consists of C1-cells. C1-cells resemble complex cells in area V1. Similarly to

2 It is also possible to create the connections from mathematical equations instead of supervised learning, especially when the cells are located densely in each layer. The number of cells is limited in the present model, however, to reduce the computation time in computer simulation. Since the cells are arranged sparsely, the mathematical description of the connections becomes complicated because of a coarse spatial sampling of the locations of the cells. Hence we use supervised learning to create the connections.