Available online at www.sciencedirect.com
SCIENCE DIRECT
Neural Networks 18 (2005) 33–43
Neural Networks
www.elsevier.com/locate/neunet

Restoring partly occluded patterns: a neural network model

Kunihiko Fukushima*

School of Media Science, Tokyo University of Technology, 1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan

Abstract

This paper proposes a neural network model that has an ability to restore missing portions of partly occluded patterns. It is a multi-layered hierarchical neural network, in which visual information is processed by interaction of bottom-up and top-down signals. Memories of learned patterns are stored in the connections between cells. Occluded parts of a pattern are reconstructed mainly by top-down signals from higher stages of the network, while the unoccluded parts are reproduced mainly by signals from lower stages. The restoration progresses successfully, even if the occluded pattern is a deformed version of a learned pattern. The model tries to complete even an unlearned pattern by interpolating and extrapolating visible edges. Resemblance of local features to other learned patterns are also utilized for the restoration.

Keywords: Vision; Partly occluded pattern; Pattern recognition; Pattern restoration; Neural network model; Bottom-up and top-down; Completion of familiar patterns; Completion of unlearned patterns

1. Introduction

When we human beings watch a pattern that is partly occluded by other objects, we often can perceive the original shape of the occluded pattern. If the occluded pattern is already familiar to us, like an alphabetical character, we can perceive a complete shape of the original pattern even if the occluded portion of the pattern has a complicated shape.

If the pattern is unfamiliar to us, we imagine the original shape using the geometry of the visible parts of the occluded pattern.

The perception sometimes changes depending on the shape and location of occluding objects. In Fig. 1, for example, pattern (a) is perceived as 'R' while pattern (b) is perceived as 'B', where the black parts of the patterns are actually identical in shape between (a) and (b). We feel as though different black patterns are occluded by gray objects. This paper proposes a neural network model that shows such a human-like response.

Various theories and neural network models have been proposed so far for recognition and restoration of partly occluded patterns (Grossberg & Mingolla, 1985; Peterhans & von der Heydt, 1991; Sajda & Finkel, 1995).

Among them, there is a neural network model proposed previously by the author (Fukushima, 2001). The model can recognize partly occluded patterns correctly, but lacks an ability to restore missing portions of the occluded pattern.

The author also proposed previously another neural network model called *Selective Attention Model* (Fukushima, 1987), which can emulate the mechanism of visual attention. The model has a function of restoring damaged patterns. It is a multilayered hierarchical network having not only forward (bottom-up) but also backward (top-down) signal paths. In this previous model, however, the top-down signals are fed back only from the recognition cells at the highest stage of the hierarchical network. If an unlearned pattern is presented to the input layer of the network, no recognition cell will respond at the highest stage, hence top-down signals cannot start flowing.

This paper proposes an extended neural network model, combining these two models. It is a multi-layered hierarchical neural network and has forward (bottom-up) and backward (top-down) signal paths. The backward signals come down to lower stages, not only from the highest stage, but also from every stages of the hierarchy.

The new model can restore a complete shape of a partly occluded pattern, if the pattern has been shown to the model during the learning phase. It should be noted here that the model does not use a simple template matching method to recognize and restore an occluded pattern. The model can accept some amount of deformation of the input pattern, and can restore the occluded portion of the pattern even if the pattern is a deformed version of a learned pattern. Occluded parts of a pattern are restored mainly by feedback signals

* Tel.: +81-426-37-2469; fax: +81-426-37-2790.
E-mail address: fukushima@media.teu.ac.jp (K. Fukushima).

0893-6080/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2004.05.001

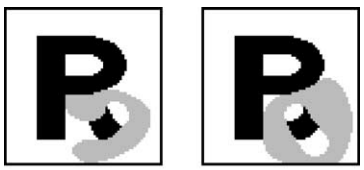

Fig. 1. An identical pattern is perceived differently by the placement of different gray objects.

from the higher stages of the network, while the unoccluded parts are reproduced mainly by signals from lower stages. The model tries to complete even an unlearned pattern by interpolating and extrapolating visible edges. Resemblance of local features to other learned patterns are also utilized for the restoration.

2. Network architecture

Fig. 2 illustrates the network architecture of the proposed model, showing how the information flows between layers. The architecture resembles that of the Selective Attention Model (Fukushima, 1987). U represents a layer of cells in the forward paths, and W in the backward paths. Each layer consists of cell-planes, in a similar way as in the neocognitron.

2.1. Forward paths

The architecture and the function of the forward paths (from U_0 through U_{C4}) is almost the same as that of the neocognitron (Fukushima, 1988, 2003). Masker layer U_M is added to the neocognitron-type network and inhibits irrelevant features extracted in U_{S1} (S-cell layer of the first stage). Layer U_{S1} receive signals, not only from the forward paths, but also from the backward paths.

2.1.1. Contrast extraction

Layer U_0 is the input layer consisting of photoreceptor cells, to which visual pattern is presented. Layer of contrast-extracting cells (U_G) follows U_0. The layer consists of two cell-planes: one with concentric on-center receptive fields, and the other with off-center receptive fields.

The response of U_G can be expressed mathematically as follows. Let the output of a photoreceptor cell of input layer U_0 at time t be $u_0^t(\mathbf{n})$, where $\mathbf{n}$ represent the location (two-dimensional coordinates) of the cell. The output of a contrast-extracting cell of layer U_G, whose receptive field center is located at $\mathbf{n}$, is given by

$$u_G^t(\mathbf{n},k)=\varphi\left[(-1)^k\sum_{|\boldsymbol{\nu}|<A_G}a_G(\boldsymbol{\nu})\cdot u_0^t(\mathbf{n}+\boldsymbol{\nu})\right],\qquad (k=1,2), \tag{1}$$

where $\varphi[\,]$ is a function defined by $\varphi[x]=\max(x,0)$. Parameter $a_G(\boldsymbol{\nu})$ represents the strength of fixed connections to the cell and takes the shape of a Mexican hat. Layer U_G has two cell-planes: one consisting of on-center cells ($k=2$) and one of off-center cells ($k=1$). A_G denotes the radius of summation range of $\boldsymbol{\nu}$, that is, the size of spatial spread of the input connections to a cell.

The input connections to a single cell of layer U_G are designed in such a way that their total sum is equal to zero. This means that the dc component of spatial frequency of the input pattern is eliminated in the contrast-extracting layer U_G. As a result, the output from layer U_G is zero in the area where the brightness of the input pattern is flat.

2.1.2. Edge extraction

The output of U_G is sent to the S-cell layer of the first stage (U_{S1}). Cells of U_{S1} correspond to simple cells in the primary visual cortex. They have been trained using supervised learning to extract edge components of various orientations from the input image.

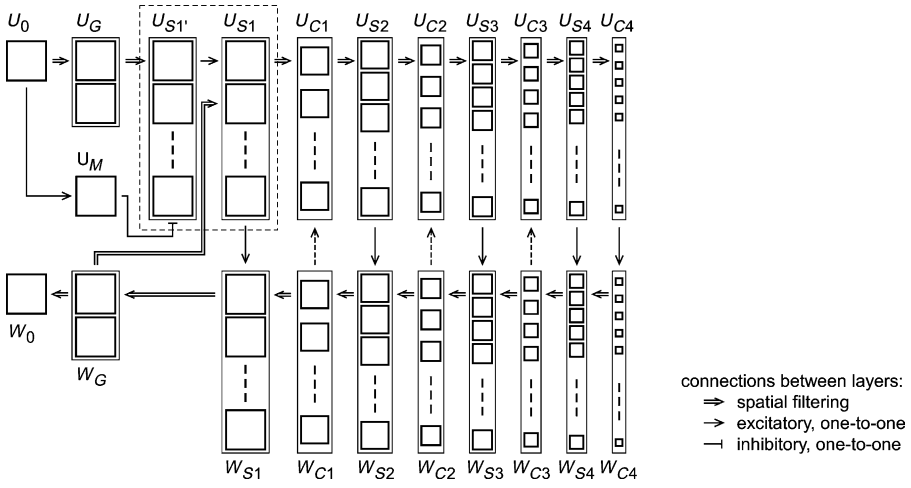

Fig. 2. The architecture of the proposed network.