Neural Feature Layers Can Establish Correspondences In Physiological Time

Jörg Lücke\textsuperscript{1}, Christian Keck\textsuperscript{2}, Philipp Wolfrum\textsuperscript{3}, Christian Wolff\textsuperscript{3}, Jan D. Bouecke\textsuperscript{2}, and Christoph von der Malsburg\textsuperscript{3,4}

\textsuperscript{1}Gatsby Unit, UCL, UK; \textsuperscript{2}Inst. f. Neuroinformatik, Ruhr-Univ. Bochum, Germany; \textsuperscript{3}FIAS, Frankfurt University, Germany; \textsuperscript{4}Computer Science Dept., USC, USA

We describe a neural network model able to rapidly establish correspondences between neural feature layers. There are good reasons to assume that the brain performs correspondence-based invariant object recognition [1]. In the technical domain, this represents state-of-the-art object and face recognition technology (e.g. [2]). The inherent task of finding corresponding points in different images of the same object has to be solved in very short time intervals in order to account for physiological object recognition times below $100$ms [3].

As input we use natural images which are overlaid by a regular grid. The surrounding of a grid-point excites subpopulations of a cortical column model [4,5] via Gabor-like receptive fields (RFs) (compare [1,6]). Thus, the neural activity distribution within such a feature column encodes the image texture at a grid-point, and a layer of feature columns encodes the whole image. We use two layers of feature columns to represent two different images of the same object. To find corresponding points, each feature column is paired with a model column whose neural subpopulations control the connections to the other layer, i.e., they realize dynamic links [6]. The subpopulations of these control columns are excited by both, similarities between activity distributions of pairs of feature columns, and by the activities in neighboring control columns. The latter implements a topological constraint for potential connectivities.

Given a pair of images, the system converges from a state of all-to-all connectivity between the layers to a state of neighborhood preserving one-to-one connectivity. In simulations convergence is shown to be very fast, corresponding to times of about $50$ms and below if translated to physiological times. Using different pairs of natural grey-level images, we verify that the final one-to-one connectivity states connect corresponding points.

We have extended the model to a full object recognition system by adding a third gallery layer which stores representations of a multitude of memorized images. Recognition is fast (about $100$ms) and, e.g., images of faces can be recognized on the basis of more than $1000$ stored identities. A previous neural model of correspondence-based recognition [6] had problems with the neural evaluation of feature similarity and with speed. The model [7] is fast but did not attempt to cope with different feature types.

Acknowledgments: We gratefully acknowledge funding by the Gatsby Charitable Foundation, the Hertie Foundation, the Volkswagen Foundation, and the EU project FP6-2005-015803.

References
\begin{enumerate}
\end{enumerate}