Approximation and Radial-Basis-Function Networks

Jooyoung Park
Irwin W. Sandberg
Department of Electrical and Computer Engineering,
The University of Texas at Austin, Austin, TX 78712 USA

This paper concerns conditions for the approximation of functions in certain general spaces using radial-basis-function networks. It has been shown in recent papers that certain classes of radial-basis-function networks are broad enough for universal approximation. In this paper these results are considerably extended and sharpened.

1 Introduction

This paper concerns the approximation capabilities of radial-basis-function (RBF) networks. It has been shown in recent papers that certain classes of RBF networks are broad enough for universal approximation (Park and Sandberg 1991; Cybenko 1989). In this paper these results are considerably extended and sharpened.

Throughout this paper, we use the following definitions and notation, in which $N$ and $\mathbb{R}$ denote the natural numbers and the set of real numbers, respectively, and, for any positive integer $r$, $\mathbb{R}^r$ denotes the normed linear space of real $r$-vectors with norm $\| \cdot \|_r$; $L^r$ denotes the standard inner product in $\mathbb{R}^r$, $L^r(N)$, and $C_r^{\infty}(\mathbb{R})$; respectively, denote the usual spaces of $r$-valued maps $f$ defined on $\mathbb{R}^r$ such that $f$ is $j$th power integrable, essentially bounded, and continuous with compact support. With $W \subset \mathbb{R}^r$, $C(W)$ denotes the space of continuous $r$-valued maps defined on $W$. The usual $L^r$ and uniform norms are denoted by $\| \cdot \|_r$, and $\| \cdot \|_{\infty}$, respectively. The characteristic function of a Lebesgue measurable subset $A$ of $\mathbb{R}^r$ is denoted by $1_A$. The convolution operation is denoted by $\ast$ and the Fourier transform (Stein and Weiss 1971) of a Fourier-transformable function $f$ is written as $\hat{f}$. By a cone in $\mathbb{R}^r$ we mean a set $C \subset \mathbb{R}^r$ such that $\alpha x \in C$ for all $x \in C$ and $\alpha > 0$. By a proper cone we mean a cone that is neither empty nor the singleton $\{0\}$.

The block diagram of a typical RBF network with one hidden layer is shown in Figure 1. Each unit in the hidden layer of this RBF network has its own centroid, and for each input $x = (x_1, x_2, \ldots, x_r)$, it computes the distance between $x$ and its centroid. Its output (the output signal at one
A Simple Coding Procedure: Enhances a Neuron’s Information Capacity

Simon Laughlin

Department of Neurobiology, Research School of Biological Sciences, P. O. Box 473, Canberra City, A. C. T. 2601, Australia


Neural Coding, Contrast, Fly Eye, Information Theory

The contrast-response function of a class of first-order interneurons in the fly’s compound eye approximates to the cumulative probability distribution of contrast levels in natural scenes. Elementary information theory shows that this matching enables the neurons to encode contrast fluctuations most efficiently.

Neurons carry and process information and there ought to be situations in which neural coding follows the dictates of information theory [1]. Large monopolar cells (LMCs) are first-order interneurons of the insect compound eye. Like the analogous bipolar cells of the vertebrate retina, their graded responses are driven, by small groups of receptors with the same field of view [2]. The compressive intensity-response function of the receptors, combined with lateral and self-inhibition, adjusts the LMC sensitivity to the background intensity so that their responses code contrast fluctuations rather than absolute intensity [3]. I show here that the interneuron’s contrast-response function matches the range of contrasts encountered in natural scenes so as to increase the efficiency with which information is encoded.

A fundamental limitation upon neural coding is the restricted range of responses with which a neuron can represent the states of its inputs. For a graded potential cell like the LMC, the response range is ultimately limited by reversal potentials. How should a neuron weigh its inputs so as to best represent their states? If sensitivities are set too high then inputs will often saturate the response, and information will be lost through clipping. Conversely, when sensitivities are set too low, large parts of the response range are underutilised because they correspond to exceptionally large excursions of input. Information theory [4] suggests an efficient means of approximating the neuron’s linear response range: the inputs should be encoded so that all response levels are used with equal frequency. Under this condition the information carried by the response can be maximised because information carried achieves its maximum entropy [3, 4].

For the simplest case of a neuron represented by a single input parameter, this optimum can be attained when input-output function corresponds to the cumulative probability function for the input. In Figs. 1, because equal output excitations correspond to equal probabilities of input. The technique of using a cumulative probability function as a response rule is also used in the fly eye (see Fig. 2). In this example the neuron has 10 response states corresponding to 10 just noticeable differences in intensity. The intensity-response function ensures that the interval between each response level encompasses an equal area under the intensity distribution, so that each state is used with equal frequency. In the line where the states are vanishingly small this intensity-response function corresponds to the cumulative probability function for stimulus intensity.

Fig. 1. The coding strategy for maximizing a neuron’s information capacity by ensuring that all response levels are used with equal frequency. Upper curve = probability density function for stimulus intensities. Lower curve = intensity-response function that implements the strategy. In this example the neuron has 10 response states corresponding to 10 just noticeable differences in response. The intensity-response function ensures that the interval between each response level encompasses an equal area under the intensity distribution, so that each state is used with equal frequency.
characteristic is used in digital image pro-
gress and is called "histogram equalization" [5].
that this coding procedure amplifies inputs in
tion to their expected frequency of occur-
using the response range for the better resolu-
tion of common events, rather than reserving large
for the improbable.
see if the first order interneurons of the
fly's compound eye see this coding procedure I
pared their contrast-response functions with the
intensity levels measured in natural scenes, such as
ophthalmic woodland and lakeside vegetation.
were intensities measured across these
or using a detector which scanned horizontally.
the ommatidium of a turning fly. It consisted of
photodiode, operating within its linear range.
The focal plane of a quartz lens. A combination of
coated glass filters (Schott, KG 3 + BG 38) was
used to give the detector a spectral sensitivity
similar to a fly monopole cell [3]. The scans were
divided into intervals of 0.05 and convolved with a
Gaussian point spread function of half-width 1.4°,
the angular sensitivity of a fly
photoreceptor [6]. Contrast values were obtained by
dividing each scan into intervals of 10, 25 or 50.
Within each interval the mean intensity, \( \bar{I} \), was
and, subtracted from every data point to
give the fluctuation about the mean, \( \Delta I \). This
difference value was divided by the mean to give the
contrast, \( \Delta I/\bar{I} \). The cumulative probability distri-
bution of contrast levels (Fig. 2) was derived from
1000 readings. As expected, the range of contrast
countered increased with the width of the interval
used, but the difference between the two larger
intervals, 25° and 50°, was small.
The interneuron's contrast-response function was
measured using the techniques described for intra-
ular recording in the intact retinas of the blowfly,
Calliphora erythrocephala [3, 6]. Individual LMC's were light-
posed by a bright steady background and the
rines to sudden increments and decrements
of this level recorded (Fig. 2). Repeated re-
to the same stimulus were averaged to
crease the reliability of the data. The light adapted
contrast-response function approximates to the sig-
rial form of the cumulative probability function
contrast levels in natural scenes (Fig. 2). This
thing indicates that this class of neurons uses
strategy for efficient coding suggested by infor-
theory. Since small differences from re-
spone equiprobability have a marginal effect on
entropy [4], there should be little redundancy asso-
ciated with the LMC response to natural scenes.
The successful application of a central concept
from information theory, entropy, to a neuron's
transfer function validates Barlow's suggestion that
redundancy reduction is an important principle in
neural coding [1]. In this context it is interesting that
many of the bipolar (7) and ganglion cells (8) of
the vertebrate retina have intensity-response functions
similar to LMCs, suggesting that the efficient coding of the available information into neurons is an essential first step in visual processing. The strategy of matching a neuron's input-output function to the expected distribution of signals so as to increase the information capacity, is equivalent to impedance matching, ensuring that the maximum amount of information is transferred between elements. For this reason the coding procedure described here might find a wider application among nervous systems.

Acknowledgements

I would like to thank Ray Stone for building operating the light detector and Jo Howard, R. Payze, M. V. Srinivasan, and Austin Hughes for their advice.