Universality and individuality in a neural code

Elad Schneidman,[1,2] **Naama Brenner,**[3] **Naftali Tishby,**[1,3]
Rob R. de Ruyter van Steveninck,[3] **William Bialek**[3]

[1]School of Computer Science and Engineering, Center for Neural Computation and
[2]Department of Neurobiology, Hebrew University, Jerusalem 91904, Israel
[3]NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540, USA
{elads,tishby}@cs.huji.ac.il, {bialek,ruyter,naama}@research.nj.nec.com

Abstract

The problem of neural coding is to understand how sequences of action potentials (spikes) are related to sensory stimuli, motor outputs, or (ultimately) thoughts and intentions. One clear question is whether the same coding rules are used by different neurons, or by corresponding neurons in different individuals. We present a quantitative formulation of this problem using ideas from information theory, and apply this approach to the analysis of experiments in the fly visual system. We find significant individual differences in the structure of the code, particularly in the way that temporal patterns of spikes are used to convey information beyond that available from variations in spike rate. On the other hand, all the flies in our ensemble exhibit a high coding efficiency, so that every spike carries the same amount of information in all the individuals. Thus the neural code has a quantifiable mixture of individuality and universality.

1 Introduction

When two people look at the same scene, do they see the same things? This basic question in the theory of knowledge seems to be beyond the scope of experimental investigation. An accessible version of this question is whether different observers of the same sense data have the same neural representation of these data: how much of the neural code is universal, and how much is individual?

Differences in the neural codes of different individuals may arise from various sources: First, different individuals may use different 'vocabularies' of coding symbols. Second, they may use the same symbols to encode different stimulus features. Third, they may have different latencies, so they 'say' the same things at slightly different times. Finally, perhaps the most interesting possibility is that different individuals might encode different features of the stimulus, so that they 'talk about different things'.

If we are to compare neural codes we must give a quantitative definition of similarity or divergence among neural responses. We shall use ideas from information theory

[1, 2] to quantify the the notions of distinguishability, functional equivalence and content in the neural code. This approach does not require a metric either on the space of stimuli or on the space of neural responses (but see [3]); all notions of similarity emerge from the statistical structure of the neural responses.

We apply these methods to analyze experiments on an identified motion sensitive neuron in the fly's visual system, the cell H1 [4]. Many invertebrate nervous systems have cells that can be named and numbered [5]; in many cases, including the motion sensitive cells in the fly's lobula plate, a small number of neurons is involved in representing a similarly identifiable portion of the sensory world. It might seem that in these cases the question of whether different individuals share the same neural representation of the visual world would have a trivial answer. Far from trivial, we shall see that the neural code even for identified neurons in flies has components which are common among flies and significant components which are individual to each fly.

2   Distinguishing flies according to their spike patterns

Nine different flies are shown precisely the same movie, which is repeated many times for each fly (Figure 1a). As we show the movie we record the action potentials from the H1 neuron.[1] The details of the stimulus movie should not have a qualitative impact on the results, provided that the movie is sufficiently long and rich to drive the system through a reasonable and natural range of responses. Figure 1b shows a portion of the responses of the different flies to the visual stimulus − the qualitative features of the neural response on long time scales (~ 100 ms) are common to almost all the flies, and some aspects of the response are reproducible on a (few) millisecond time scale across multiple presentations of the movie to each fly. Nonetheless the responses are not identical in the different flies, nor are they perfectly reproduced from trial to trial in the same fly.

To analyze similarities and differences among the neural codes, we begin by discretizing the neural response into time bins of size $\Delta t = 2$ ms. At this resolution there are almost never two spikes in a single bin, so we can think of the neural response as a binary string, as in Fig. 1c-d. We examine the response in blocks or windows of time having length T, so that an individual neural response becomes a binary 'word' W with $T/\Delta t$ 'letters'. Clearly, any fixed choice of T and Δt is arbitrary, and so we explore a range of these parameters.

Figure 1f shows that different flies 'speak' with similar but distinct vocabularies. We quantify the divergence among vocabularies by asking how much information the observation of a single word W provides about the identity of its source, that is about the identity of the fly which generates this word:

$$I(W \to \text{identity}; T) = \sum_{i=1}^{N} P_i \sum_{W} P^i(W) \log_2 \left[\frac{P^i(W)}{P^{\text{ens}}(W)} \right] \text{ bits,} \qquad (1)$$

---

[1] The stimulus presented to the flies is a rigidly moving pattern of vertical bars, randomly dark or bright, with average intensity $\bar{I} \approx 100$ mW/(m$^2 \cdot$ sr). The pattern position was defined by a pseudorandom sequence, simulating a diffusive motion or random walk. Recordings were made from the H1 neuron of immobilized flies, using standard methods. We draw attention to three points relevant for the present analysis: (1) The flies are freshly caught female *Calliphora*, so that our 'ensemble of flies' approaches a natural ensemble. (2) In each fly we identify the H1 cell as the unique spiking neuron in the lobula plate that has a combination of wide field sensitivity, inward directional selectivity for horizontal motion, and contralateral projection. (3) Recordings are rejected only if raw electrode signals are excessively noisy or unstable.
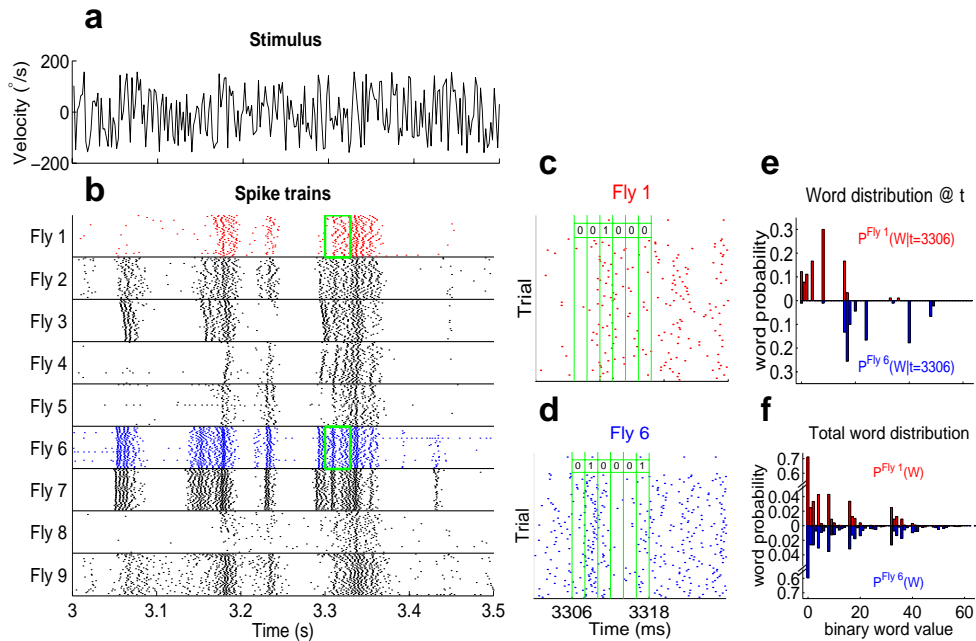
Figure 1: **Different flies' spike trains and word statistics**. **(a)** All flies view the same random vertical bar pattern moving across their visual field with a time dependent velocity, part of which is shown. In the experiment, a 40 sec waveform is presented repeatedly, 90 times. **(b)** A set of 45 response traces to the part of the stimulus shown in (a) from each of the 9 flies. The traces are taken from the segment of the experiment where the transient responses have decayed. **(c)** Example of construction of the local word distributions. Zooming in on a segment of the repeated responses of fly 1 to the visual stimuli, the fly's spike trains are divided into contiguous 2 ms bins, and the spikes in each of the bins are counted. For example, we get the 6 letter words that the fly used at time 3306 ms into the input trace. **(d)** Similar as (c) for fly 6. **(e)** The distributions of words that flies 1 and 6 used at time $t = 3306$ ms from the beginning of the stimulus. The time dependent distributions, $P^1(W|t = 3306\,\text{ms})$ and $P^6(W|t = 3306\,\text{ms})$ are presented as a function of the binary value of the actual 'word', e.g., binary word value $'17'$ stands for the word $'010001'$. **(f)** Collecting the words that each of the flies used through all of the visual stimulus presentations, we get the total word distributions for flies 1 and 6, $P^1(W)$ and $P^6(W)$.

where $P_\text{i} = 1/N$ is the a priori probability that we are recording from fly i, $P^\text{i}(W)$ is the probability that fly i will generate (at some time) the word W in response to the stimulus movie, and $P^\text{ens}(W)$ is the probability that any fly in the whole ensemble of flies would generate this word,

$$P^\text{ens}(W) = \sum_{\text{i}=1}^{N} P_\text{i} P^\text{i}(W). \qquad (2)$$

The measure $I(W \rightarrow \text{identity}; T)$ has been discussed by Lin [11] as the 'Jensen–Shannon divergence' D_JS among the distributions $P^\text{i}(W)$.[2]

---

[2]Unlike the Kullback–Leibler divergence [2] (the 'standard' choice for measuring dissimilarity among distributions), the Jensen–Shannon divergence is symmetric, and bounded (see also [12]). Moreover, D_JS can be used to bound other measures of similarity, such as the optimal or Bayesian probability of identifying correctly the origin of a sample.

We find that information about identity is accumulating at more or less constant rate well before the undersampling limits of the experiment are reached (Fig. 2a). Thus $I(W \to \text{identity}; T) \approx R(W \to \text{identity}) \cdot T$; $R(W \to \text{identity}) \approx 5$ bits/s, with a very weak dependence on the time resolution Δt. Since the mean spike rate can be measured by counting the number of 1s in each word W, this information includes the differences in firing rate among the different flies.

Even if flies use very similar vocabularies, they may differ substantially in the way that they associate words with particular stimulus features. Since we present the stimulus repeatedly to each fly, we can specify the stimulus precisely by noting the time relative to the beginning of the stimulus. We can therefore consider the word W that the ith fly will generate at time t. This word is drawn from the distribution $P^i(W|t)$ which we can sample, as in Fig. 1c–e, by looking across multiple presentations of the same stimulus movie. In parallel with the discussion above, we can measure the information that the word W observed at known t gives us about the identity of the fly,

$$I(W \to \text{identity}|t; T) = \sum_{i=1}^{N} P_i \sum_{W} P^i(W|t) \log_2 \left[\frac{P^i(W|t)}{P^{\text{ens}}(W|t)} \right] \;, \qquad (3)$$

where the distribution of words used at time t by the whole ensemble of flies is

$$P^{\text{ens}}(W|t) = \sum_{i=1}^{N} P_i P^i(W|t). \qquad (4)$$

The natural quantity is an average over all times t,

$$I(\{W, t\} \to \text{identity}; T) = \langle I(W \to \text{identity}|t; T) \rangle_t \text{ bits}, \qquad (5)$$

where $\langle \cdots \rangle_t$ denotes an average over t.

Figure 2b shows a plot of $I(\{W, t\} \to \text{identity}; T)/T$ as a function of the observation time window of size T. Observing both the spike train and the stimulus together provides 32 ± 1 bits/s about the identity of the fly. This is more than six times as much information as we can gain by observing the spike train alone, and corresponds to gaining one bit in ~ 30 ms; correspondingly, a typical pair of flies in our ensemble can be distinguished reliably in ~ 30 ms. This is the time scale on which flies actually use their estimates of visual motion to guide their flight during chasing behavior [6], so that the neural codes of different individuals are distinguishable on the time scales relevant to behavior.

3 Different flies encode different amounts of information about the same stimulus

Having seen that we can distinguish reliably among individual flies using relatively short samples of the neural response, we turn to ask whether these substantial differences among codes have an impact on the ability of these cells to convey information about the visual stimulus. As discussed in Refs. [7, 8], the information which the neural response of the ith fly provides about the stimulus, $I^i(W \to s(t); T)$, is determined by the same probability distributions defined above: [3]

$$I^i(W \to s(t); T) = \left\langle \sum_{W} P^i(W|t) \log_2 \left[\frac{P^i(W|t)}{P^i(W)} \right] \right\rangle_t . \qquad (6)$$

---

[3] Again we note that our estimate of the information rate itself is independent of any metric in the space of stimuli, nor does it depend on assumptions about which stimulus features are most important in the code.
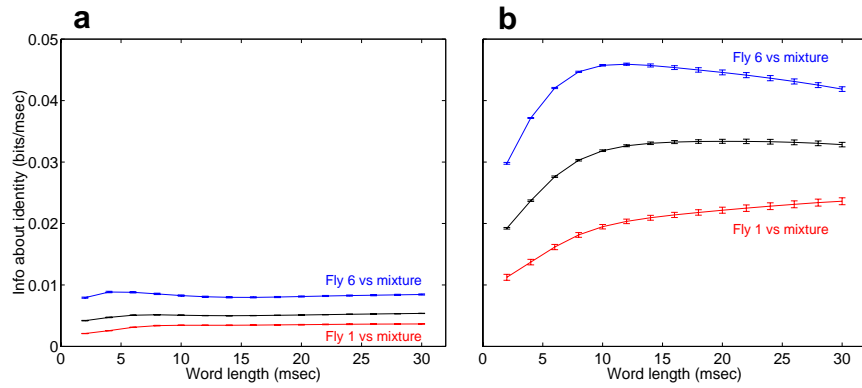
Figure 2: **Distinguishing one fly from others based on spike trains**. (a) The average rate of information gained about the identity of a fly from its word distribution, as a function of the word size used (middle curve). The information rate is saturated even before we reach the maximal word length used. Also shown is the average rate of information that the word distribution of fly 1 (and 6) gives about its identity, compared with the word distribution mixture of all of the flies. The connecting line is used for clarification only. (b) Similar to (a), we compute the average amount of information that the distribution of words the fly used at a specific point in time gives about its identity. Averaging over all times, we show the amount of information gained about the identity of fly 1 (and 6) based on its time dependent word distributions, and the average over the 9 flies (middle curve). Error bars were calculated as in [8]. A "baseline calculation", where we subdivided the spike trains of one fly into artificial new individuals, and compared their spike trains, gave significantly smaller values (not shown).

Figure 3a shows that the flies in our ensemble span a range of information rates from ≈ 50 to ≈ 150 bits/s. This threefold range of information rates is correlated with the range of spike rates, so that each of the cells transmits nearly a constant amount of information per spike, 2.39 ± 0.24 bits/spike. This universal efficiency (10% variance over the population, despite three fold variations in total spike rate), reflects that cells with higher firing rates are not generating extra spikes at random, but rather each extra spike is equally informative about the stimulus.

Although information rates are correlated with spike rates, this *does not* mean that information is carried by a "rate code" alone. To address the rate/timing distinction we compare the total information rate in Fig. 3a, which includes the detailed structure of the spike train, with the information carried in the temporal modulations of the spike rate. As explained in Ref. [10], the information carried by the arrival time of a single spike can be written as an integral over the time variations of the spike rate, and multiplying by the number of spikes gives us the expected information rate if spikes contribute independently; information rates larger than this represent synergy among spikes, or extra information in the temporal patterns of spike. For all the flies in our ensemble, the total rate at which the spike train carries information is substantially larger than the 'single spike' information—2.39 vs. 1.64 bits/spike, on average. This extra information is carried in the temporal patterns of spikes (Fig. 3b).

4 A universal codebook?

Even though flies differ in the structures of their neural responses, distinguishable responses could be functionally equivalent. Thus it might be that all flies could be
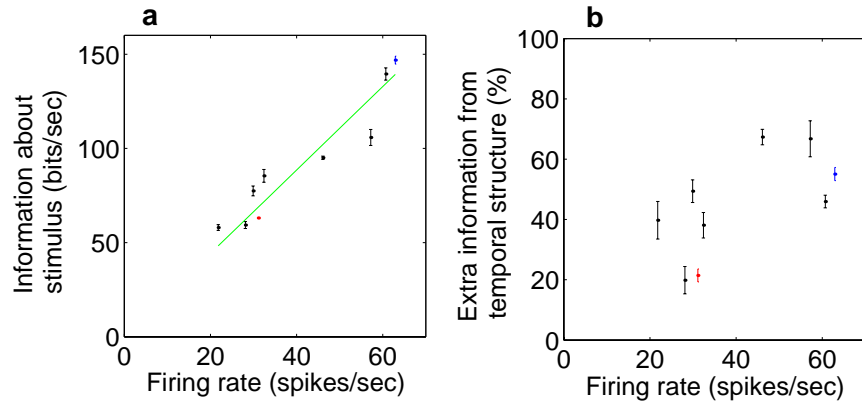
Figure 3: **The information about the stimulus that a fly's spike train carries is correlated with firing rate, and yet a significant part is in the temporal structure**. **(a)** The rate at the H1 spike train provides information about the visual stimulus is shown as a function of the average spike rate, with each fly providing a single data point The linear fit of the data points for the 9 flies corresponds to a universal rate of 2.39 ± 0.24 bits/spike, as noted in the text. **(b)** The extra amount of information that the temporal structure of the spike train of each of the flies carry about the stimulus, as a function of the average firing rate of the fly (see [10]). The average amount of additional information that is carried by the temporal structure of the spike trains, over the population is $45 \pm 17\%$. Error bars were calculated as in [8]

endowed (genetically?) with a universal or consensus codebook that allows each individual to make sense of her own spike trains, despite the differences from her conspecifics. Thus we want to ask how much information we lose if the identity of the flies is hidden from us, or equivalently how much each fly can gain by knowing its own individual code.

If we observe the response of a neuron but don't know the identity of the individual generating this response, then we are observing responses drawn from the ensemble distributions defined above, $P^{\mathrm{ens}}(W|t)$ and $P^{\mathrm{ens}}(W)$. The information that words provide about the visual stimulus then is

$$I^{\mathrm{mix}}(W \rightarrow s(t); T) = \left\langle \sum_W P^{\mathrm{ens}}(W|t) \log_2 \left[\frac{P^{\mathrm{ens}}(W|t)}{P^{\mathrm{ens}}(W)} \right] \right\rangle_t \text{ bits.} \qquad (7)$$

On the other hand, if we know the identity of the fly to be i, we gain the information that its spike train conveys about the stimulus, $I^{\mathrm{i}}(W \rightarrow s(t); T)$, Eq. (6). The average information loss is then

$$I^{\mathrm{avg}}_{\mathrm{loss}}(W \rightarrow s(t); T) = \sum_{\mathrm{i}=1}^{N} P_{\mathrm{i}} I^{\mathrm{i}}(W \rightarrow s(t); T) - I^{\mathrm{mix}}(W \rightarrow s(t); T). \qquad (8)$$

After some algebra it can be shown that this average information loss is related to the information that the neural responses give about the identity of the individuals, as defined above:

$$\begin{aligned} I^{\mathrm{avg}}_{\mathrm{loss}}(W \rightarrow s(t); T) &= I(\{W, t\} \rightarrow \mathrm{identity}; T) \\ &\quad - I(W \rightarrow \mathrm{identity}; T). \end{aligned} \qquad (9)$$

The result is that, on average, not knowing the identity of the fly limits us to extracting only 64 bits/s of information about the visual stimulus. This should be

compared with the average information rate of 92.3 bits/s in our ensemble of flies: knowing her own identity allows the average fly to extract 44% more information from H1. Further analysis shows that each individual fly gains approximately the same relative amount of information from knowing its personal codebook.

5   Discussion

We have found that the flies use similar yet distinct set of 'words' to encode information about the stimulus. The main source of this difference is not in the total set of words (or spike rates) but rather in how (i.e. when) these words are used to encode the stimulus; taking this into account the flies are discriminable on time scales of relevance to behavior. Using their different codes, the flies' H1 spike trains convey very different amounts of information from the same visual inputs. Nonetheless, all the flies achieve a *high* and *constant* efficiency in their encoding of this information, and the temporal structure of their spike trains adds nearly 50% more information than that carried by the rate.

So how much is universal and how much is individual? We find that each individual fly would lose $\sim 30\%$ of the visual information carried by this neuron if it 'knew' only the codebook appropriate to the whole ensemble of flies. We leave the judgment of whether this is high individuality or not to the reader, but recall that this is the individuality in an identified neuron. Hence, we should expect that all neural circuits—both vertebrate and invertebrate—express a degree of universality and a degree of individuality. We hope that the methods introduced here will help to explore this issue of individuality more generally.

References

[1] Shannon, C. E. A mathematical theory of communication, *Bell Sys. Tech. J.* **27,** 379–423, 623–656 (1948).

[2] Cover, T. & Thomas J. *Elements of information theory* (Wiley, 1991).

[3] Victor, J. D. & Purpura, K. Nature and precision of temporal coding in visual cortex: a metric–space analysis, *J. Neurophysiol.* **76,** 1310–1326 (1996).

[4] Hausen, K. The lobular complex of the fly, in *Photoreception and vision in invertebrates* (ed Ali, M.) pp. 523–559 (Plenum, 1984).

[5] Bullock, T. *Structure and Function in the Nervous Systems of Invertebrates* (W. H. Freeman, San Francisco, 1965).

[6] Land, M. F. & Collett, T. S. Chasing behavior of houseflies (Fannia canicularis). A description and analysis, *J. Comp. Physiol.* **89,** 331–357 (1974).

[7] de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P., Koberle, R. & W. Bialek, Reproducibility and variability in neural spike trains, *Science* **275,** 1805–1808, (1997).

[8] Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. & Bialek, W. Entropy and information in neural spike trains, *Phys. Rev. Lett.* **80,** 197–200 (1998).

[9] Rieke, F., Warland, D., de Ruyter van Steveninck, R. & Bialek, W. *Spikes: Exploring the Neural Code* (MIT Press, 1997).

[10] Brenner, N., Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. & Bialek, W. Synergy in a neural code, *Neural Comp.* **12,** 1531–52 (2000).

[11] Lin, J., Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Theory,* **37,** 145–151 (1991).

[12] El-Yaniv, R., Fine, S. & Tishby, N. Agnostic classification of Markovian sequences, *NIPS* **10** pp. 465–471 (MIT Press, 1997).