

The Story of A Single Cell: Peeking into the Semantics of Spikes

Roi Kliper*[†], Thomas Serre[§], Daphna Weinshall*[†] and Israel Nelken^{‡†}

*The Interdisciplinary Center for Neural Computation

[†]School of Computer Science and Engineering

[‡]Department of Neurobiology, The Alexander Silberman Institute of Life Sciences,
The Hebrew University of Jerusalem, Jerusalem, Israel 91904

[§]Brown University Department of Cognitive and Linguistic Sciences

Abstract—Traditionally, the modeling of sensory neurons has focused on the characterization and/or the learning of input-output relations. Motivated by the view that different neurons impose different partitions on the stimulus space, we propose instead to learn the structure of the stimulus space, as imposed by the cell, by learning a cell specific *distance function* or kernel. Metaphorically speaking, this direction attempts to bypass the syntactic question of “how the cell speaks”, by focusing instead on the semantic and fundamental question of “what the cell says”.

Here we consider neural data from both the inferotemporal cortex (ITC) and the prefrontal cortex (PFC) of macaque monkeys. We learn a cell-specific distance function over the stimulus space as induced by the cell response; the goal is to learn a function such that the distance between stimuli is large when the responses they evoke are very different, and small when the responses they evoke are similar.

Our main result shows that after training, when given new stimuli, our ability to predict their similarity to previously seen stimuli is significantly improved. We attempt to exploit this ability to predict the response of the cell to a novel stimuli using KNN over the learnt distances. Furthermore, using our learned kernel we obtain a partitioning of the stimulus space which is more similar to the partition induced by the cell responses as revealed by low dimension embedding, and thus, are able in some of the cases to peek at the semantic partition induced by the cell.

I. INTRODUCTION

A. Functional characterization of single cells

Transformations of representations are an immanent part of our cognitive processes. Objects and sensations (with distinct physical properties) are analyzed, synthesized and transformed through the brain’s different pathways, taking detailed forms of membrane potentials and action potentials that eventually compose a larger mental representation¹. The question of what exactly is being represented in different brain regions has attracted great interest and some remarkable work, but while the picture is fairly clear in some domains, most remain only partially understood.

A case of special interest is the characterization of sensory neurons: one of its main goals is to characterize dimensions in stimulus space to which the neurons are highly sensitive (causing large gradients in the neural responses), or alternatively dimensions in stimulus space to which the neuronal

response is invariant (defining iso-response manifolds). This challenge is especially pronounced when trying to learn the representation of visual objects in higher brain areas, where simple features representations (and models) are neglected in favor of complex, non-trivial and possibly semantic ones.

Visual processing in the cortex is classically considered to be hierarchical, with simple feature representations being gradually neglected in favor of distributed complex object representations at the level of the inferotemporal cortex (ITC) [3], [4]. Consistent with this idea, a recent study [5] showed that it was possible to reliably extract position and scale invariant object category information from a small population of neurons (~ 200) in IT cortex. A second study pointed out a possible functional shift that might be taking place between the closely connected ITC and prefrontal cortex (PFC) regions [6], [7]. Others [8] point out the manifestation of non-linear representation already in the V4 regions and offered a model that may account for it. Understanding the actual manifestation of this shift and characterizing the role of single cells in such schemes as well as gathering information to support such models of representation has proved to be a major challenge. In this work we focus our efforts on the task of single cell characterization.

The dominant approach to the functional characterization of sensory neurons attempts to learn the input-output relations $f : S \rightarrow R$ of a cortical neuron, and to predict the response of the neuron to novel stimuli. Prediction is often estimated from a set of known responses of a given neuron to a set of stimuli, by modeling some linear filter over the stimuli. One typical method of building such predictors uses linear models and their second order variants, in order to approximate the function that is assumed to generate the response [9], [10]. However, these models may fail when responses are highly non-linear or when the smoothness of the response as dependent on the stimulus space is lost [11] - a process which is hypothesized to occur as one moves along the Visual pathways from V4 to the ITC and further to the PFC [7].

The approach taken in this work avoids learning the input-output relation f . Instead, we attempt to learn the specific “geometric” structure induced by a neuron on the visual stimuli. In other words, we try to learn the non-linear partition of stimulus space as induced by the neuronal responses.

¹The dual representation terminology is in itself deeply embedded in the brain research discourse, see for example [1], [2]

B. Learning a single cell's invariant space

Motivated by the view that different neurons impose different partitions of stimulus space which are not necessarily simply related to the simple feature structure of the stimuli [12], we attempt instead to learn the structure of the stimulus space by learning a *distance function*. Specifically, we characterize a neuron by learning a pairwise distance function over the stimulus space that is consistent with the similarities between the responses to different stimuli. A distance function is a function defined over pairs of data-points $D : S \times S \rightarrow \mathbb{R}$, which assigns a real (and possibly bounded) valued number to any pair of points from the input space $\{s_i, s_j\} \in S$. The assigned number measures the distance between pairs of points, which reflects the similarities between them.

Intuitively, a good distance function would capture the desired structure and assign small distance values to pairs of stimuli that elicit a similar neuronal response, and large values to pairs of stimuli that elicit different neuronal responses. Knowledge of the structure is in itself valuable, as it can be used to understand the kind of classification that a single cell performs on the stimulus space. Interestingly, it could also be used for prediction, using some variant of K Nearest Neighbors (KNN) with the learnt distance function.

Our approach offers several advantages: first, it allows us to aggregate information from a number of neurons and reach a good hypothesis even when the number of known stimuli responses per neuron is small, which is a typical concern in the domain of neuronal characterization. Second, unlike most functional characterizations that are limited to linear or weakly non-linear models, distance learning can approximate functions that are highly non-linear.

Metaphorically speaking, learning a cell specific distance function allows the investigator to bypass the question of "how the cell speaks", or "how many spikes are fired to a given stimuli?". Instead, we attempt to touch upon the more fundamental question of what exactly the cell is saying, or "what partition does the cell induce on stimulus space?".

II. RELATED WORK

A. Neurophysiological background

The common view of sensory systems considers neurons as feature detectors arranged in an anatomical and functional hierarchy. Visual processing in cortex is classically modeled as a hierarchy of increasingly sophisticated representations, naturally extending the model of simple to complex cells of Hubel and Wiesel. In this view of the visual system, information from simple feature detectors in the retina converges at the level of the primary visual cortex (V1) to represent elaborate features of direction and orientation. This hierarchical view continues to dominate models characterizing both the ventral pathway, associated with object recognition and form representation, and the dorsal pathway associated with determining objects' position in space.

Several models have been proposed to account for the impressive human ability to locate, identify, recognize and track objects in a visual scenery e.g. [13], [14], [15]. Little quantitative modeling has been done to explore the biological

feasibility of this class of models to explain aspects of higher-level visual processing such as object recognition. The role of a single cell in such models is, in particular, a subject of great debate and while researchers have acknowledged the need to account for properties of invariance and specificity, the prediction requirement of a single cell model is still phrased in terms of response predictions.

In addition, physiological evidence accumulated over the past decade remains controversial, particularly because models of object recognition in the cortex have been mostly applied to tasks involving the recognition of isolated objects presented on blank backgrounds. Ultimately models of the visual system have to prove themselves in real world object recognition tasks, such as face detection in cluttered scenes, a standard computer vision benchmark task. Understanding the role of a single cell in such complex task places a very difficult challenge on the experimental setting: that of revealing the invariant response space of the cell.

The accepted approach to functional single cell characterization has focused on characterizing the Spatio-Temporal Receptive Fields (STRFs) of a neuron: a function that maps time-varying visual inputs to neural responses [9], [11]. Classically, STRFs have been used to implement linear and sometimes second-order models of neural responses [10]; ultimately, evaluating the success of such a model by measuring the model's response prediction against the real responses. This approach has frequently been used to study the visual system and has been proven quite efficient for early stages of the visual system; however, using these models to characterize neurons in higher stages of the visual system has not been proven as efficient.

These and other findings seem to suggest that highly complex representations of the environment cannot be accounted for by the mere use of simple linear models, nor can its success be measured purely by its ability to predict neuronal response, and that a new approach to the functional characterization of these cells may be in order. A first step in this direction was taken by [16], introducing the learning of distance functions into this domain and showing some preliminary and promising results. Here we pursue this research direction further, trying to identify and characterize complex cells in the visual system. Our results may help to make this approach an accepted one in experimental neuroscience.

B. Distance function learning

While *distance function learning* is a somewhat new area of research, the concept of a distance function is well known and is widely used in various applications and for various computational tasks. Recent years have seen a lot of interest in distance function learning algorithms (e.g. [17], [18], [19]) These algorithms aim at automatically incorporating domain specific knowledge and/or side information into the applied distance function. In the general setting of such an algorithm, side information (typically in the form of equivalence constraints) is used to learn a pairwise function that adequately captures the structure of the space and the relations between the data-points.

Unlike the classical learning scenario, where one attempts to learn some function $f : X \rightarrow Y$ that approximates the relation between a given input and its resulting output using a training sample $S = \{(s_1, r_1), (s_2, r_2), \dots, (s_N, r_N)\}$, distance functions provide information about the similarity of pairs of points - essentially capturing relations within the input data-points themselves $D : S \times S \rightarrow \mathbb{R}$ rather than an input-output relation. In some cases capturing the relations between data-points can provide information which cannot be easily extracted from directly estimating input-output relations. Such cases occur when the input-output function is highly complex, multi staged, or just difficult to estimate. In other cases, the structure of the space is what one is interested in (rather than the transformation).

In the case of the characterization of single neurons, both incentives for using distance function learning appeal: on the one hand estimating the actual transformation is hard and poses many physiological and technical limitations. On the other hand, we argue, information about the partition of space may in many cases be much more interesting than predicting the actual response. Ideally, learning an adequate distance function may render the task of response prediction redundant.

III. EXPERIMENTAL SETTING

A. Problem formulation

Our approach is based on the idea of learning a distance function over the stimuli space, using side-information extracted from the response space. The initial data consists of stimulus-response paired representations. To generate the side-information, we use a similarity measure over pairs of points in the response space. These are used in turn to generate equivalence constraints on pairs of stimuli: two stimuli are related by a positive equivalence constraint if their paired responses are highly similar; they are related by a negative equivalence constraint if their paired responses are highly dissimilar.

Next, this side information is used to train an algorithm that learns a distance function between pairs of stimuli points, thus capturing implicitly the structure of the stimuli space as induced by the cell. In our framework, the cell becomes a teacher, specifying similarities between stimuli using its own language of action potentials. These similarities are then used to learn a cell-specific distance function over the space of all possible stimuli. This learned distance function should reveal what exactly is represented by the changes in the response of the specific cell.

We can formally define the computational task as follows:

Input: A set of stimuli-response pairs $\{s_i, r_i\}_{i=1}^N$

- 1) Represent the responses and stimuli in their own 'natural' feature space, along with a predefined similarity measure in the responses space.
- 2) Use the responses to extract equivalence constraints on stimuli, as described above.
- 3) Learn a distance function over the stimuli space $D(s_i, s_j) \rightarrow R$ using these constraints.

- 4) Use the generated distance function to understand and predict the nature of stimuli space.

In the remainder of this section we present the details of our suggested scheme and how it is used for the characterization of visual neurons.

B. Neural computational setting implementation

The data was collected by Freedman et al [6], [7] and consisted of stimuli-response pairs of data recorded in the ITC and PFC of macaque monkeys. The stimuli used consisted of a continuous set of cat and dog stimuli constructed from six prototypes with a three-dimensional morphing system. The stimuli were generated by morphing different amounts of the prototypes (see Fig. 1).

This allowed to continuously vary the stimulus shape, and precisely define a category boundary. The category of a stimulus was defined by whichever category contributed more (50%) to a given morph. The behavioral paradigm required monkeys to release a lever if two stimuli (separated by a 1 sec delay) were from the same category (a category match), see [6], [7] for details. The continuous nature of the stimuli in combination with this behavioral task presumably broke smoothness of response over stimuli space making the task difficult for simple models.

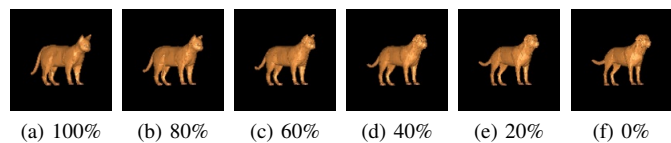


Fig. 1: A morph line between 'Cat I' (left) and 'Dog I' (right). Three Prototypes of dogs and three prototypes of cats were morphed in six levels of inter-species. 100%, 80%, 60%, 40%, 20%, 0% and 4 levels of within species morphing 100%, 60% 40% 0% yielding 42 Cat-Dog morphs and 12 within species morphs.

1) *Data representation:* As input for our learning algorithm we used the first 20 principal components of a gray scale representation of the images. The neuronal response for each stimulus was represented as a vector containing in each of its entries the spike rate of one out of multiple stimulus presentations (trials), where the number of trials per stimulus varied in the range 7-13.

C. Obtaining equivalence constraints from neural data

For the sake of simplicity, the distance between responses was measured using the Cohen's d which is defined as the difference between two means divided by the standard deviation for the data, thus creating a distance matrix over all pairs of responses. This is very similar to taking the F-statistic of a 1-way ANOVA between vectors of responses over multiple trials. By choosing Cohen's d as our response distance measure, we implicitly assume that responses of the cells to different stimuli, are normally distributed across trial repetitions to same stimuli and are generally of equal variance (homoscedasticity assumption), but may be characterized by different means.

We then used the complete linkage algorithm to cluster the data into 8 clusters,². All of the points in each cluster were marked as similar to one another, thus providing positive equivalence constraints. Negative constraints were determined to exist between points in the 4 furthest clusters.

D. Training a distance function

We took the scheme described above and implemented it using the Kernel-Boost distance learning algorithm described in [20],³ Kernel-Boost is a variant of the Dist-Boost algorithm [21] which was used in [16] and showed promising results. Kernel-Boost is a semi-supervised distance learning algorithm that learns distance functions using unlabeled data-points and equivalence constraints. While the Dist-Boost algorithm has been shown to enhance clustering and retrieval performance, it was never used in the context of classification mainly due to the fact that the learnt distance function is not a kernel (and is not necessarily metric). Therefore it cannot be used by the large variety of kernel based classifiers that have shown to be highly successful in fully labeled classification scenarios. Kernel-Boost alleviates this problem by modifying the weak learner of Dist-Boost to produce a 'weak' kernel function. The 'weak' kernel has an intuitive probabilistic interpretation - the similarity between two points is defined by the probability that they both belong to the same Gaussian component within the constrained Gaussian Mixture Model (cGMM) learned by the weak learner.

An additional important advantage of Kernel-Boost over Dist-Boost is that it is not restricted to model each class at each round using a single Gaussian model, therefore removing the assumption that classes are convex. This restriction is dealt with by using an adaptive label dissolve mechanism, which splits the labeled points from each class into several local subsets. An important inherited feature of Kernel Boost is that it is semi-supervised, and can naturally accommodate unlabeled data in the learning process.

For each neuron, a subset of all pairs of stimuli was selected such that the responses of the two stimuli in a pair were either very similar or very dissimilar. The distance function was trained using a cross validation scheme to fit these constraints. The resulting distance functions generalized to predict the distances between the responses of a test stimulus and all trained stimuli.

Evaluation: We used a number of ways to evaluate the quality of the learned distance function. First, we evaluated the learned distance function by rank correlating (Spearman) the learned distances to the actual distances as measured over the cell responses and measured using Cohen's d. In a second evaluation step, the distances were used to cluster the stimulus data, which was then compared with the clustering induced by the cell responses; the accordance between the two clusterings was measured using the Rand index. Finally, the distance function was used along with a KNN classifier to generate predictions for novel samples in a cross validation scheme.

²We found that choosing this number of clusters does not over constrain the problem not positively nor negatively

³In our comparative study Kernel-Boost performed extremely well, especially when given only a small amount of data.

IV. RESULTS AND EVALUATION

We narrowed our analysis to neurons which displayed some stimulus selectivity (not necessarily category selectivity). Such selectivity was established by performing 1-way ANOVA with each of the 54 samples as grouping factors at $p < 0.01$. This analysis resulted in 162 ITC neurons and 61 PFC neurons. However, while this procedure filters out non discriminative cells it still keeps cells that are selective to a very small number of stimulus (1-3 stimulus). Since this analysis is a demonstration of a new technique we limit it to the ten most selective cells in each setting.⁴ We start the results analysis by measuring the success of the learner to fit and generalize the distances as defined by the responses (Sec. IV-1). We then continue to show how this knowledge can be used for response prediction (Sec. IV-2) and stimulus classification (Sec. IV-3).

1) *Fitting Power and Generalization:* As a first step we examined the ability of our algorithm to fit the distances as induced by the cell responses. We evaluated this by measuring the mean Spearman rank correlation between the distances computed by our distance learning algorithm and those induced by the cell as measured by the 'actual' distances over the responses: we use our distance function to calculate the distances between the stimuli and check whether close pairs of stimuli (as defined by a cell) are indeed close as measured by our cell specific distance function. The results as seen in Fig. 2 (left most bar) show dramatic improvement after learning.

After establishing the improvement in fitting power across both data sets, we turned to evaluate the generalization properties of the algorithm. We tested three scenarios: Leave-One-Out (LOO) where at each simulation one stimulus was left out of the training set and used for testing, Leave-Five-Out (LFO) and Leave-Ten-Out (LTO) where a random sample of five and ten (in accordance) stimuli were left out of the training set for later test (20 repetitions). For each stimulus that was tested in one of this manners, we measured its distance to all other stimuli using the learned distance function. We then computed the rank order correlation coefficient between the learned distances in the stimulus domain, and Cohen's d between the corresponding response vectors. This procedure yielded a single correlation coefficient for each of the simulations using the stimuli which were left out. To measure performance, we took the average of the correlation coefficient over all runs for each cell.

The results (Fig. 2) show persistent improvement in test correlation in all scenarios, as compared to baseline (naive) correlation. These results highlight the fact that the algorithm can generalize well, and some aspects of topology of the stimulus space have indeed been captured. Test correlation scores are generally reduced as the size of the sample used for training is reduced, but the reduction is minor. Finally, individual cells also displayed very strong correlation between training performance and test performance. (Not shown)

⁴Cells that did not display stimulus selectivity, or only displayed marginal selectivity during the recording setting cannot be used as 'teachers' when data is so sparse (54 stimulus) and are thus omitted from the analysis. We claim that using our technique all cells will ultimately be guided to areas of stimulus space with high selectivity and will thus allow learning. All reported results pertain to this subset of the data.

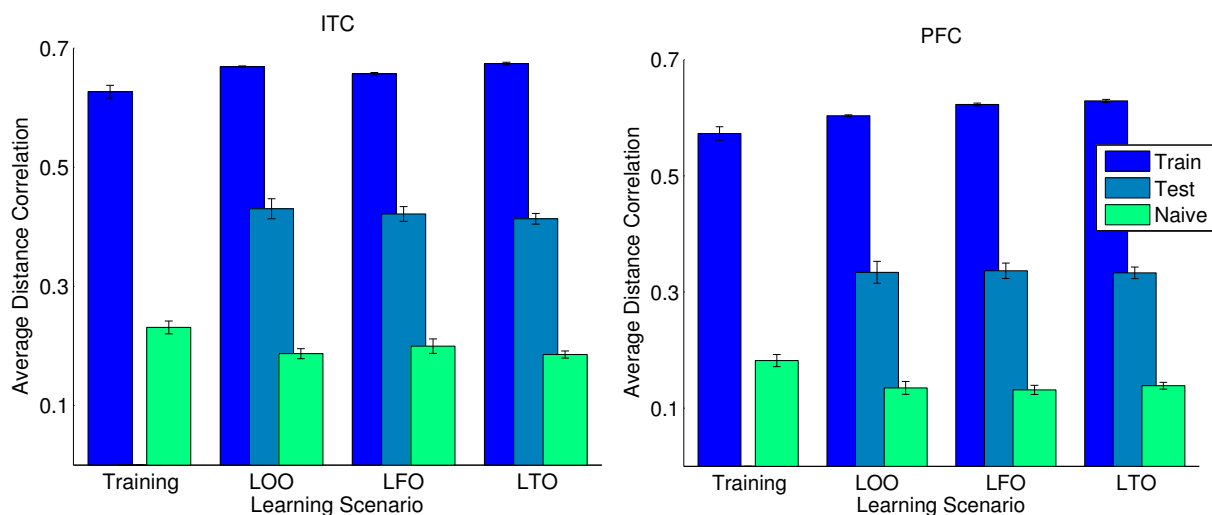


Fig. 2: **Distance correlation score.** 1) Trained vs. Test vs. Naive (distances between images) performance (upper row): The algorithm achieved a dramatic increase in average rank correlation, suggesting that some aspect of the structure of stimulus space was indeed captured. Correlation between train and test scores were in the order of 0.9 (Not shown).

Interestingly, we have observed that some of the stimuli were more easily predictable than others when left out of the training set. Presumably good results will ultimately depend on the quality of sampling of stimulus space, a task to which this algorithm can be helpful by pointing out to interesting areas in the stimulus space.

2) *Generating response predictions:* Having proved the generalization properties of the learning algorithm, we attempt to use the newly learnt distance function to generate a response prediction to novel stimuli. We do this by matching to each novel stimulus its 1st -Nearest Neighbor (1-NN) from the training sample using our learnt distance function. We then predict that the cell would respond to the novel stimulus in a similar way to its response to this matched 1-NN stimulus. We note that since the stimulus space is rather restricted in our experimental setting, responses tend to be relatively smooth locally, in the sense that similar (gray level) stimuli tend to give rise to similar responses. Thus, even the naive 1-NN predictor, based on the inter-stimuli distance in the original grey-level representation, is expected to perform well in most cases. We compared our prediction scheme to both the linear regression predictor and a naive prediction based on the 1-NN as computed using the gray levels of the stimuli. Our prediction method outperforms both competitors, as shown in Fig. 3. The success in prediction, and our advantage over the alternative predictors, is highly dependent on the size of the training sample - it increases with sample size, which is consistent with the increase in test distance correlations demonstrated earlier. Specifically, we do better or equal to other predictors in 64%, 66%, 61% for ITC cells, and 68% 65% and 61% for PFC cells, with 53, 49, 44 training samples respectively.

3) *Peeking into the semantics: Inferring iso-response manifolds:* A single cell can divide the world it sees to a few iso-response sets. This division is reflected by a distinct response pattern to each of these sets. In our setting, for the sake of

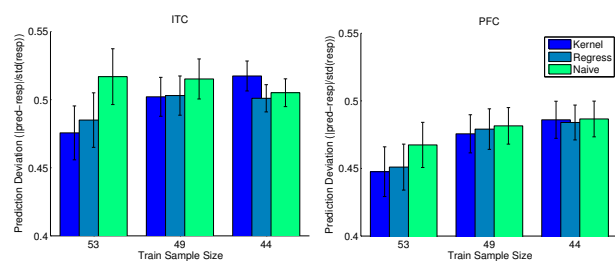


Fig. 3: **Prediction advantage:** shown above is a measure for the success of response prediction. The quality of a prediction success is calculated as the normalized absolute distance to actual response: $\frac{|prediction - response|}{std(response)}$. In other words we measure how many standard deviation our prediction is from the actual response. Our algorithm is advantageous in the majority of the scenarios both against naive KNN predictor (Naive) and against a linear regressor (regress). This advantage becomes stronger as sample size increases.

simplicity, we assume that a cell induces a binary partition on the stimulus space and thus defines two iso-response sets.

The learnt distance function can be thought of as somewhat analogous to a transformation of the feature space. We now ask - in the transformed space, is the division to iso-response sets more evident? does this transformation make the input space easier to analyze and understand? To answer this question, we visualize the data using 2-D embeddings. We use classic MDS to compare three embeddings: one based on the true distances in response space (presumably the goal embedding), a second based on the Euclidean distance in the original grey-level features space, and a third based on the learnt distances. Specifically, we show two examples of cells in a leave-ten-out setting in Fig. 4. Clearly the embedding based on the learnt distances is more similar to the goal embedding than the original one. Moreover, these response-induced partitions are evident in stimulus space after distance transformation, but hardly evident in the original feature space. This result was found for other cells as well. The point to appreciate

here is the fact that while creating the cell embedding (left) is illustrative in figuring out the semantics of the cell, it is depended completely on the actual recording experiment. Our distance based embedding, however is not and can be used on any number of (possibly unused) data-points.

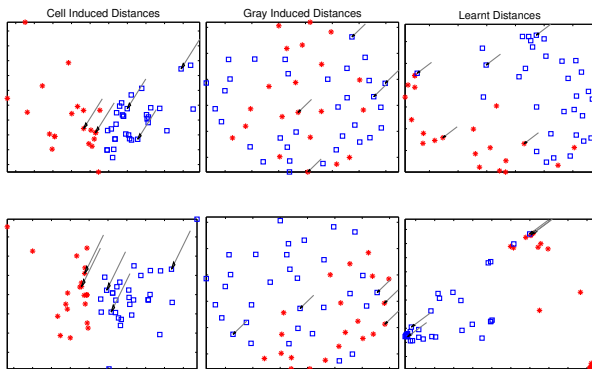


Fig. 4: Visualizing classification effect: Two cells were selected to display the effect of learning on the feature space topology, which is visualized via the embedding of points in a 2-D space. For each cell we show embeddings based on response distances (left column), on original gray levels values (middle column), and based on the learnt distance function (right column). In each case we used classic MDS to embed the data points in 2-D. Points were partitioned into two iso-response sets according to the response they elicit: red circle indicate points that elicited high response while blue square points that elicited low response. The goal of the learner was to be able to replicate the partition induced by the cell (left column). Arrows mark points that were left out during the training phase.

V. SUMMARY AND CONCLUSIONS

To begin the analysis, for each cell from the training sample, the stimuli were divided to N clusters according to the elicited cell's responses. Next, we used our algorithm to learn a distance function for each cell separately, using equivalence constraints extracted from the clustered stimuli. We then used the learnt distance matrix to generate response prediction and to re-cluster unseen stimuli. In order to visualize the results, by way of simplicity and as a good abstraction, we assumed that the cell divides the world into two iso-response manifold: 'Response' / 'No Response'. We then showed in Fig. 4 that this partition is captured almost perfectly when using the learnt distances between stimuli.

Our proposed scheme can serve as an integral part in a neuroscientist's experimental setup. An effective machinery for distance function learning can help direct the investigator towards "interesting" areas in the stimuli space as defined by the cell itself, and thus reduce the time and frustration involved in a search based on trial and error. For that to happen, we plan to develop the scheme so that it will work in an online manner, being able to handle information fast and in an accumulative manner.

ACKNOWLEDGMENT

This study was supported by the European Union under the DIRAC integrated project IST-027787.

We would like to thank David Freedman and Earl Miller for sharing their monkey electrophysiology data with us.

REFERENCES

- [1] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, "An internal model for sensorimotor integration." *Science*, vol. 269, no. 5232, pp. 1880–1882, Sep 1995.
- [2] H. Imamizu, S. Miyauchi, T. Tamada, Y. Sasaki, R. Takino, B. PÄtz, T. Yoshioka, and M. Kawato, "Human cerebellar activity reflecting an acquired internal model of a new tool." *Nature*, vol. 403, no. 6766, pp. 192–195, Jan 2000. [Online]. Available: <http://dx.doi.org/10.1038/35003194>
- [3] N. K. Logothetis and D. L. Sheinberg, "Visual object recognition," *Ann. Rev. Neurosci.*, vol. 19, pp. 577–621, 1996.
- [4] K. Tanaka, "Inferotemporal cortex and object vision," *Ann. Rev. Neurosci.*, vol. 19, pp. 109–139, 1996.
- [5] C. Hung, G. Kreiman, T. Poggio, and J. DiCarlo, "Fast read-out of object identity from macaque inferior temporal cortex," *Science*, vol. 310, pp. 863–866, Nov. 2005.
- [6] D. J. Freedman, M. Riesenhuber, T. Poggio, and E. K. Miller, "Categorical representation of visual stimuli in the primate prefrontal cortex." *Science*, vol. 291, no. 5502, pp. 312–316, Jan 2001. [Online]. Available: <http://dx.doi.org/10.1126/science.291.5502.312>
- [7] —, "A comparison of primate prefrontal and inferior temporal cortices during visual categorization." *J Neurosci*, vol. 23, no. 12, pp. 5235–5246, 2003.
- [8] C. Cadieu, M. Kouh, A. Pasupathy, C. E. Connor, M. Riesenhuber, and T. Poggio, "A model of v4 shape selectivity and invariance." *J Neurophysiol*, vol. 98, no. 3, pp. 1733–1750, Sep 2007. [Online]. Available: <http://dx.doi.org/10.1152/jn.01265.2006>
- [9] F. E. Theunissen, K. Sen, and A. J. Doupe, "Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds." *J Neurosci*, vol. 20, no. 6, pp. 2315–2331, Mar 2000.
- [10] N. C. Rust, O. Schwartz, J. A. Movshon, and E. P. Simoncelli, "Spatiotemporal elements of macaque v1 receptive fields." *Neuron*, vol. 46, no. 6, pp. 945–956, Jun 2005. [Online]. Available: <http://dx.doi.org/10.1016/j.neuron.2005.05.021>
- [11] S. David and J. L. Gallant, "Predicting neuronal responses during natural vision." *Network*, vol. 16, no. 2-3, pp. 239–260, 2005.
- [12] O. Bar-Yosef, Y. Rotman, and I. Nelken, "Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context," *J Neurosci*, vol. 22, no. 19, pp. 8619–8632, Oct 2002.
- [13] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *nature neuroscience*, vol. 2, pp. 1019–1025, 1999.
- [14] T. Serre, M. Kouh., C. Cadieu, U. Knoblich, G. Kreiman, and T. Poggio, "A theory of object recognition: computations and circuits in the feedforward path of the ventral stream in primate visual cortex," MIT, Cambridge, MA, AI Memo 2005-036 / CBCL Memo 259, 2005.
- [15] T. Serre, A. Oliva, and T. Poggio, "A feedforward theory of visual cortex accounts for human performance in rapid categorization," *Proceedings of the National Academy of Science*, vol. 104, no. 15, pp. 6424–6429, 2007.
- [16] I. Weiner, T. Hertz, I. Nelken, and D. Weinshall, "Analyzing auditory neurons by learning distance functions," in *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. Platt, Eds. Cambridge, MA: MIT Press, 2006, pp. 1481–1488.
- [17] E. Xing, A. Ng, M. Jordan, and S. Russell, "Distance metric learning with application to clustering with side-information," *Advances in neural information processing systems*, pp. 521–528, 2003.
- [18] A. Bar-Hillel, T. Hertz, N. Shental, and D. Weinshall, "Learning distance functions using equivalence relations," in *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*, vol. 20, no. 1, 2003, p. 11.
- [19] J. Davis, B. Kulis, P. Jain, S. Sra, and I. Dhillon, "Information-theoretic metric learning," in *Proceedings of the 24th international conference on Machine learning*. ACM, 2007, p. 216.
- [20] T. Hertz, A. Bar-Hillel, and D. Weinshall, "Learning a kernel function for classification with small training samples." in *International Conference on Machine Learning (ICML)*, 2006.
- [21] —, "Learning distance functions for image retrieval," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2004*, vol. 2, 2004, pp. II-570–II-577 Vol.2.