# Feature selection and dimensionality reduction of neural data to model adaptation in the retina

Neda Nategh
Electrical Engineering Department, Stanford University
Stanford, CA 94305
nnategh@stanford.edu

## Abstract

*Retinal ganglion cells are most sensitive to the visual feature defined by the linear spatio-temporal receptive field. They encode this feature according to a nonlinear sensitivity curve that often has a threshold and saturation. Both the linear receptive field and nonlinearity are adaptive, in that these parameters change depending on the recent statistics of the stimulus.*

*In the context of motion processing, changes in gain are important for a cell to detect textures of low contrast or luminance, but not be saturated by high contrast or fast motion. In order to make the comparison of trajectory in the center and background regions, both sites should avoid saturation, yet detect all available weak motion signals.*

*One potentially rich source to generate adaptation is the diverse population of inhibitory amacrine cells, which comprise about thirty types. Amacrine transmission is thought to play a role in retinal adaptation to more complex stimulus statistics (Hosoya et al., 2005), but not for simple statistics such as luminance and contrast.*

*To understand how the circuit transforms the visual scene, we will identify components of retinal image processing using a novel combined experimental and theoretical approach that includes intracellular recording, simultaneous current injection and multielectrode recording, and computational modeling. Here, we carry this analysis further and divide the population of ganglion cells into functional classes using quantitative clustering algorithms that combine several response characteristics. We first used the dimensionality reduction methods to extract the visual features encoded by the amacrine and ganglion cells that best describe their response properties. Using these features to classify the interactions between these two cells revealed seven types of transmissions, in agreement with the types of modulations in the response properties of ganglion cells derived by the amacrine cell's output.*

## 1. Introduction

### *Object Motion Sensitivity in the retina*

One of the important circuitries in the retina that contains multiple sites of adaptation is the object motion sensitivity circuitry, in which amacrine cells play an important role.

Recently, it was discovered that segmentation of moving objects, and rejection of background motion begins in the retina [1]. A subset of retinal ganglion cells responds to motion in the receptive field center, but only if the motion trajectory is different from that of the surrounding region (Figure 1).

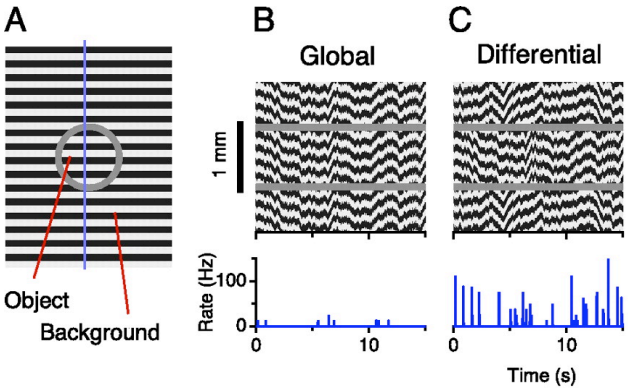



Figure 1: Objection motion sensitivity in a retina ganglion cell. Electrical activity in OMS ganglion cells was recorded from the isolated retina of a salamander. A video monitor was projected onto the retina, and extracellular electrical impulses were recorded with an array of electrodes. ***A***, Diagram of object and background regions in the stimulus display. ***B,C***, First row, Space–time plot of a vertical cross section through the center of the stimulus (line in ***A***), showing trajectories for global motion and differential motion. Global Motion represents fixational eye movements with no object motion. Differential Motion represents object motion in the presence of eye movements. Motion in the object region is identical in both cases. Second row, Average firing rate of an OMS ganglion cell in response to 10 repeats of each stimulus sequence. The cell is nearly silent in Global Motion, but fires precise bursts of activity during Differential Motion. (From Baccus et al., 2008).

These cells are termed "Object Motion Sensitive" (OMS) cells. Like many retinal ganglion cells, the

responses of OMS neurons are highly precise [2]. When the same visual stimulus is repeated, the action potentials are highly reproducible in time, in some cases to within less than a millisecond. This has enabled mathematical models to capture the responses of OMS cells, involving the interaction of OMS cells and other interneurons in the retinal circuitry (Figure 2).

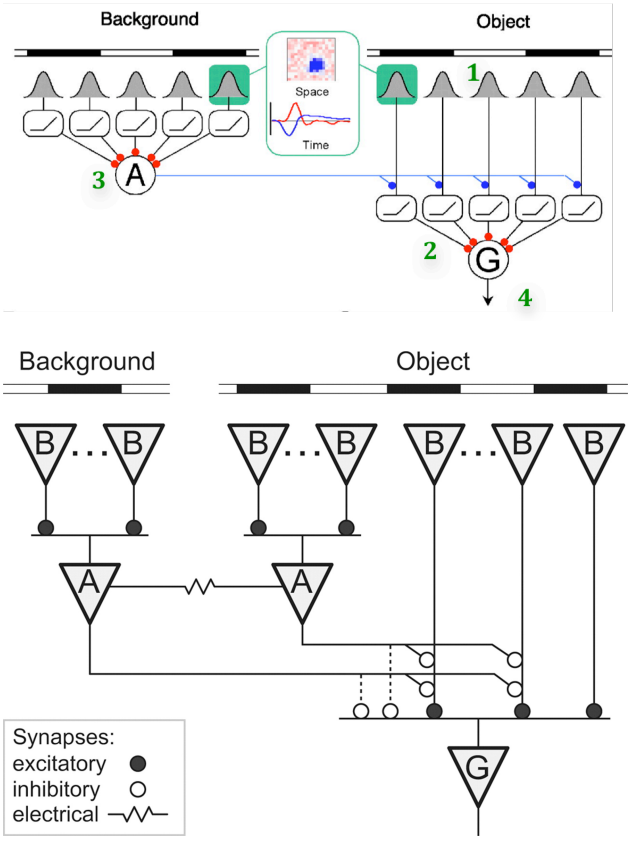



Figure 2: Model and neural circuit for object motion sensitivity. (Top). The visual stimulus is first processed by linear subunits with a small receptive field and transient dynamics. An OMS ganglion cell (G) receives excitatory input in the object region from multiple small subunits. Each subunit applies a linear spatiotemporal filter to the stimulus in its receptive field. Amacrine and ganglion cells each sum the rectified output of many linear bipolar cells. To predict the response of an OMS ganglion cell to differential motion, amacrine inhibition from the background was combined with bipolar input from the object region before the stage of rectification. (Bottom) The neural circuit that produces object motion sensitivity. Bipolar cells perform the spatio-temporal linear filtering. Rectification is at the output synapse of bipolar cells. A specific subtype of amacrine cell with long axons produces background suppression. (From Baccus et al., 2008).

Although this block diagram describes the response of the OMS ganglion cell accurately, it does not specify what neural circuit and what algorithm implements the computation. To flesh out the schematic with actual processing units and building blocks of an object motion algorithm, one needs to answer the following questions: (1) What is the identity and properties of the subunits and inhibitory cell? (2) How is the output of these subunits integrated? (3) At what level is the signal from background motion combined with that from the object region?

In our quest to answer these questions, we measured how the signals transmitted through individual amacrine cells contribute to the ganglion cell response by recording intracellularly from single amacrine cells while simultaneously recording spiking activity from the ganglion cell population using a multielectrode array. We presented a randomly flickering visual stimulus drawn from a Gaussian distribution while injecting Gaussian white-noise current into the amacrine cell. By this direct perturbation of the circuit we measured how the interneuron generates adaptation of the ganglion cell visual response.

To model the contribution of each amacrine cell to each ganglion cell's visual response, we combined elements of a linear-nonlinear (LN) model, consisting of a linear temporal or spatio-temporal filter followed by a static nonlinearity. The model consisted of the linear receptive field and nonlinearity of the ganglion cell, a modulatory pathway containing the LN model of the amacrine cell, and a transmission filter linking the two pathways.

We found that amacrine transmission scales the ganglion cell nonlinear response function by a gain factor, and in some cases, also modulates the linear receptive field of the ganglion cell, changing it from being more integrating to more differentiating. This modulation is driven by the preferred feature of the amacrine cell, even if this feature is different from that of the ganglion cell.

Even at a fixed luminance and contrast, retinal ganglion cells adapt at a fast timescale. For this type of adaptation, an amacrine cell provides contextual information that modulates the ganglion cell visual response. Thus, the space of visual features encoded by the diverse population of amacrine cells defines a multidimensional context that gates and modifies a different space of visual feature encoded by the population of ganglion cells.

These modulations vary in their selected features, patterns, model parameters, and strength. This project studies the diversity of these modulations, potential sources of this variability, and their possible functional contributions to the retinal processing. To understand the contribution of amacrine transmission to fast adaptation of retinal ganglion cells, we will identify and characterize the functional role of an amacrine cell using the following procedure.

First we divide the population of ganglion cells and

amacrine cells into functional classes using quantitative clustering algorithms that combine several response characteristics. Since the response characteristics of the neurons occupy a high dimensional space of features, we use the dimensionality reduction techniques to find a low dimensional representation of the feature space that is being used for clustering the neurons. Then we divide the modulations into functional classes using clustering algorithms, based on changes in the response variables that are reflected in our model's parameters.

Finally we explore the correspondences of these two clustering schemes. If members of each modulation cluster corresponded to the distinct functional clusters of neurons, we could understand which properties or components of ganglion cells' or amacrine cells' response are contributing to or shaping the adaptation components.

## 2. Methods

*Stimulation*

A uniform field randomly flickering visual stimulus, drawn from a Gaussian distribution, is projected from a video monitor onto the intact, isolated salamander retina and an array of extracellular electrodes is used to record the light responses of many ganglion cells at once. Simultaneously, an intracellular recording monitors the visual responses of an amacrine cell. Then, white noise current is injected into the amacrine cell to measure the kinetics and nonlinear properties of how the cell's output modifies the circuit's behavior. Thus, many simultaneous paired recordings are performed (Figure 3).

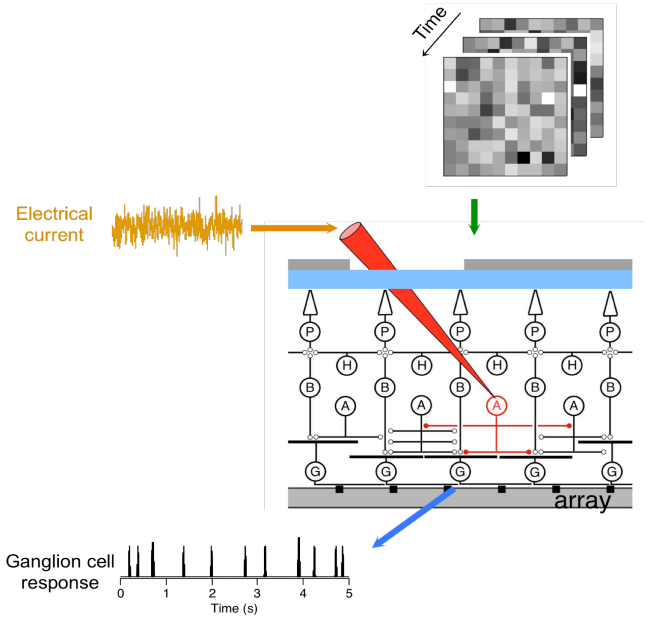



Figure 3: Schematic diagram of simultaneous intracellular and multielectrode recording preparation.

*Analysis: modeling*

A simple model that has been used to approximate the behavior of amacrine and ganglion cells is a linear-nonlinear (LN) model (Figure 4) [3]. The LN model describes the average behavior of the circuit, but does not account for many of the nonlinear contributions of individual amacrine circuits.

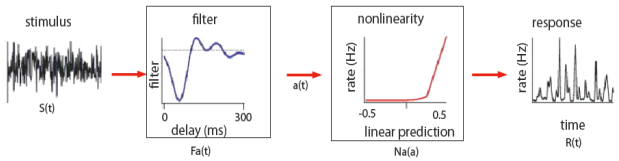



Figure 4. Linear-Nonlinear (LN) model. The LN model consists of the stimulus intensity weighted over time by a linear temporal filter or spatio-temporal filter, followed by a static nonlinearity.

To capture the effect of the amacrine cell, we have modified the direct, single pathway LN model of the ganglion cell's firing rate by adding an indirect parallel pathway that includes the effect of amacrine cell on the direct visual pathway (Figure 5). Our modified model consists of the following components, all computed from a single experiment:

- Direct input. A spatio-temporal LN model of each ganglion cell's firing rate is computed. This pathway represents inputs to the ganglion cell other than the amacrine cell.
- Amacrine input. A spatio-temporal LN model of the amacrine cell's membrane potential is computed by correlating the visual stimulus with the cell's response (Figure 5).
- Amacrine transmission kinetics. By correlating the Gaussian white-noise current injected into the amacrine cell with each ganglion cell's firing rate, a linear temporal filter is computed representing the average kinetics of transmission for each cell pair (Figure 5).
- Amacrine effect on the ganglion cell nonlinearity. We characterize the transmission of the cell by computing a two-dimensional nonlinear function that combines the amacrine transmission with the direct pathway. We found that the amacrine transmission changes ganglion cells' nonlinear characteristics by scaling its nonlinear response function by a gain factor (Figure 5).
- Amacrine effect on ganglion cell temporal processing. We have found that this amacrine cell strongly modulates the temporal response of ganglion cells. When the amacrine cell is more hyperpolarized, the ganglion cell speeds up and becomes more differentiating, thus encoding changes in light intensity more than the absolute light intensity (Figure 5).

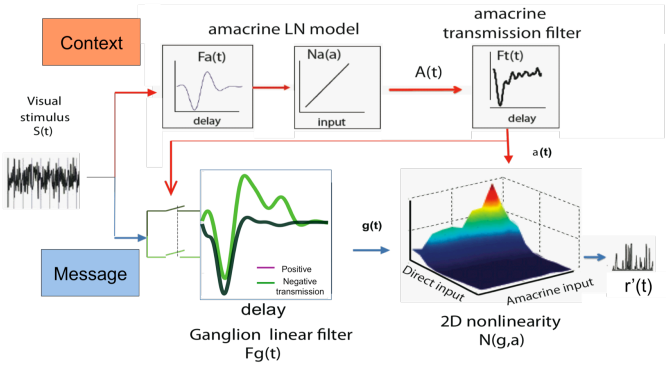



Figure 5. Model of amacrine cell transmission acting on a ganglion cell. The visual stimulus *s(t)* passes through *Fa(t)* and *Na(a)*, the linear filter and nonlinearity of the amacrine cell, followed by *Ft(t)*, the amacrine transmission filter. The amacrine pathway modulates both the kinetics and the nonlinearity of the ganglion cell. The ganglion cell visual linear filter, *Fg(t),* was calculated by averaging visual stimuli preceding the time of a ganglion cell spike.

## *Analysis: Classification*

We used a broad parameter ensemble-linear temporal filters, and nonlinearity functions of the two-pathway LN model-to classify cells or cell pairs into types having systematically different temporal responses and nonlinear sensitivities. Each linear temporal filter is characterized by a high dimensional visual response vector as a function of time, with the size in the order of hundreds of milliseconds. A nonlinearity function is characterized by three parameters $a_1, a_2, a_3$ of its exponential fit of the form $a_1^{erf(x+a_2)+1} + a_3$.

The functional similarity between two cells or two cell pairs is initially quantified by computing the mean-squared difference between their temporal dynamics and nonlinearity parameters. However, since we have hundreds of cells and cell pairs, each characterized by a high dimensional feature vector consisting of its temporal dynamics and static nonlinearities, we need to construct a low dimensional representation of the response characteristics to find a reasonable quantitative measure of functional similarity used to feed our clustering algorithms.

Because there exists no a priori method of functional classification, we have made several choices of dimensionality reduction and clustering methods to divide the ganglion cells, amacrine cells, and ganglion-amacrine cell pairs into broader or finer groupings.

We initially used the temporal dynamics of the receptive field to classify ganglion cells into functional types. We used linear dimensionality reduction methods including principal component analysis (PCA), and multidimensional scaling (MDS) (Table 1) [4] to find a low dimensional representation of the high dimensional visual response vectors- namely the linear temporal filters and static nonlinearity functions of ganglion and amacrine cells.

Combining all these linear and nonlinear receptive field's characteristics results in a nonlinear feature space. However, linear dimensionality reduction methods will fail to find any lower dimensional space that is embedded non-linearly in a higher dimension. For Euclidean manifolds, Isomap (Table 2) [5] and locally linear embedding (LLE) (Table 3) [6, 7] avoid this shortcoming of linear projection. The idea is that for a given point in a well sampled space, the point's nearest neighbors will lie only in that low dimensional. Then, if we preserve the local geometry and dimension of each neighborhood, we should be able reconstruct the manifold using only the dimension of those neighborhoods.

**Table 1. Multidimensional scaling**
Given high-dimensional points $x_i$, recovers the low-dimensional coordinates of the data that describe where the points lie on the manifold, in other words, find an embedding of the data in a low dimensional space, that preserves its essential regularities

- images ($\mathbf{y}_i$) of original points ($\mathbf{x}_i$) should have approximately the same interpoint distances as the originals
- let $\delta_{ij}$ = distance between $\mathbf{x}_i$ and $\mathbf{x}_j$, and $d_{ij}$ = distance between $\mathbf{y}_i$ and $\mathbf{y}_j$ e.g. Euclidean distance $d_{ij} = \|\mathbf{y}_i - \mathbf{y}_j\|_2$
- we want $\forall i, j$ $(\delta_{ij} \approx d_{ij})$
- exact equality generally not possible, so minize an error function $J$ of the $\mathbf{y}$'s

**Table 2. Isomap**
The Isomap algorithm takes as input the distances $d_X(i,j)$ between all pairs $i,j$ from $N$ data points in the high-dimensional input space $X$, measured either in the standard Euclidean metric or in some domain-specific metric. The algorithm outputs coordinate vectors $\mathbf{y}_i$ in a $d$-dimensional Euclidean space $Y$ that best represent the intrinsic geometry of the data.

| Step | | |
|---|---|---|
| 1 | Construct neighborhood graph | Define the graph $G$ over all data points by connecting points $i$ and $j$ if [as measured by $d_X(i,j)$] they are closer than $\epsilon$ ($\epsilon$-Isomap), or if $i$ is one of the $K$ nearest neighbors of $j$ ($K$-Isomap). Set edge lengths equal to $d_X(i,j)$. |
| 2 | Compute shortest paths | Initialize $d_G(i,j) = d_X(i,j)$ if $i,j$ are linked by an edge; $d_G(i,j) = \infty$ otherwise. Then for each value of $k = 1, 2, \ldots, N$ in turn, replace all entries $d_G(i,j)$ by $\min\{d_G(i,j), d_G(i,k) + d_G(k,j)\}$. The matrix of final values $D_G = \{d_G(i,j)\}$ will contain the shortest path distances between all pairs of points in $G$ (16, 19). |
| 3 | Construct $d$-dimensional embedding | Let $\lambda_p$ be the $p$-th eigenvalue (in decreasing order) of the matrix $\tau(D_G)$ (17), and $v_p^i$ be the $i$-th component of the $p$-th eigenvector. Then set the $p$-th component of the $d$-dimensional coordinate vector $\mathbf{y}_i$ equal to $\sqrt{\lambda_p} v_p^i$. |

**Table3. Locally linear embedding**
LLE maps a data set X, globally to a data set Y. Assuming the data lies on a nonlinear manifold which locally can be

approximated linearly, it uses two stages: (I) locally fitting hyper-planes around each sample xi, based on its k nearest neighbors, and calculating reconstruction weights, and (II) finding lower-dimensional co-ordinates yi for each xi, by minimizing a mapping function based on these weights.

1. For each $\vec{x}_i$, find the $k$ nearest neighbors.
2. Find the weight matrix $\mathbf{w}$ which minimizes the residual sum of squares for reconstructing each $\vec{x}_i$ from its neighbors,

$$RSS(\mathbf{w}) \equiv \sum_{i=1}^{n} \|\vec{x}_i - \sum_{j \neq i} w_{ij}\vec{x}_j\|^2 \quad (1)$$

where $w_{ij} = 0$ unless $\vec{x}_j$ is one of $\vec{x}_i$'s $k$-nearest neighbors, and for each $i$, $\sum_j w_{ij} = 1$. (I will come back to this constraint below.)

3. Find the coordinates $\mathbf{Y}$ which minimize the reconstruction error using the weights,

$$\Phi(\mathbf{Y}) \equiv \sum_{i=1}^{n} \|\vec{y}_i - \sum_{j \neq i} w_{ij}\vec{y}_j\|^2 \quad (2)$$

subject to the constraints that $\sum_i Y_{ij} = 0$ for each $j$, and that $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$.

Finding a low dimensional representation of the linear dynamics and nonlinear sensitivity characteristics for the cells or cell pairs, we formalize classification using k-means clustering and hierarchical clustering algorithms.

### Broad classification

Functional classification was carried out using the method of agglomerative clustering [9], an iterative algorithm, that at each step merges the most functionally similar cells into the same cluster and averages all of their properties together, weighted by the number of cells in each cluster; By examining the similarity of the clusters that are merged at each step of this algorithm, we can assess the significance of the merger, which is the functional difference between the two clusters that were merged together. By looking at the merger score as fewer and fewer clusters remain, we can find that the differences between clusters suddenly become large, which indicates that these clusters are significant.

An alternative way to set the significance threshold is by looking at the histogram of the merger score. In this manner, one can identify all of the outlier values of the merger score and set the number of clusters as the maximal number that includes all these outliers (Figure 6A).

Further discussion of the issue of choosing the number of significant clusters in a data set can be found in several interesting books and articles [9]. For broad functional types, the algorithm was applied to all of the cells recorded from multiple retinas.

### Fine classification

To split the ganglion cell population or ganglion-amacrine cell pair interactions into as many types as could be justified by the data, we used k-means clustering to define cluster boundaries, because this method is known to be biased toward forming extra clusters when used on data with relatively few examples [9].

In K-means clustering, we first decide how many clusters the data will be divided into and randomly assign one cell to each cluster. All remaining cells are assigned to the nearest cluster, based on the (normalized) mean-squared difference between cell $i$ and cluster $k$, $a_{ik}$. For this analysis, we used the visual features extracted by our dimensionality reduction algorithms. Next, the cluster waveform is computed by averaging the features of all members. At this point, a goodness-of-fit measure is obtained by calculating the total mean-squared difference between all cells and their respective clusters, $\bar{a} = \sum_i a_{ik}$.

The algorithm iterates by starting with the new cluster waveforms and reassigning all cells to the nearest cluster. This iteration is continued until the total difference between cells and clusters, $\bar{a}$, no longer decreases. Because the resulting cluster structure depends on the choice of initial clusters, we repeated this algorithm with 1000 different random choices of initial cluster definitions for each value of $K$ and selected the final cluster partition that had the smallest total difference, $\bar{a}$ [8].

The number of clusters $K$ is a parameter of this algorithm, and as more clusters are used to describe the population, the total difference $\bar{a}^{(K)}$ must decrease. To determine what value of $K$ resolves significant clusters, we plotted the decrease in the total difference as new clusters were added, $\Delta^{(K)} = \bar{a}^{(K)} - \bar{a}^{(K-1)}$. When this decrease is large, clusters are significant, and when the decrease is small, the new clusters resolve only minor details in the ganglion cell population (Figure 7A).

## 3. Results

To understand the manner in which ganglion cells and amacrine cells collectively represent a visual scene, we calculated the temporal amacrine transmission filter of every recorded ganglion cell in a small patch of the retina. By reducing the dimensionality of the feature space characterizing the interaction between response characteristics of the visual pathway consisting of a ganglion cell's LN model and response characteristics of the modulatory pathway containing the LN model of the amacrine cell, we could quantitatively cluster the interactions-namely the transmission filters linking the two pathways. Such computation will elucidate how these clusters of interactions distribute their responses to a dynamic input variable.

*Classification of ganglion cells*

Our objective here was to use quantitative clustering techniques along with a very descriptive set of features, so that our results better reflect the information encoded collaboratively by ganglion cells and amacrine cells about the visual scene. Because there exists no a priori method of functional classification, we made several choices of feature space and clustering method to divide the cells into broader or finer groupings. We found that the feature dimensions found by the PCA and MDS methods were not descriptive enough to cover all types of the observed modulations. However, we found that LLE performed as well as Isomap and both better than the linear dimensionality reduction methods.

Code from the authors' of LLE and Isomap was downloaded and run over the data structures used. Code for PCA, MDS, clustering, and modeling was written by Neda.

## *Broad types*

We used the dimensions computed by the Isomap algorithm to classify interactions into functional types. We formalized classification using an iterative algorithm that at each step merged the most functionally similar interactions into the same cluster and averaged their corresponding transmission filters together. By examining the similarity of the clusters that are merged at each step of this algorithm, we can assess the significance of the merger: when two clusters are very similar, their merger score will be close to zero, and when two clusters are very different, their score will be greater than one.

Figure 6A shows the results of this clustering algorithm when applied to 133 ganglion cells recorded from eleven retinas in the salamander, whose amacrine transmission profiles are shown in figure 6B. There was a clear break in the similarity of cell clusters, shown by a dashed line. At this point, there were seven distinct clusters with multiple members—monophasic ON, biphasic ON, biphasic OFF, monophasic OFF, biphasic weak-ON, fast biphasic, and slow biphasic (shown in colors)—as well as eight cells that belonged to their own cluster (shown in gray).

The distinct clusters had members recorded from at least half of the eleven retinas used in this study and were routinely observed in other experiments, so we treated them as broad types. Unique cells were observed in a single retinal patch and were not commonly seen in other experiments, so we treated them as unclassified. The classification of cells into seven broad types is a robust property of the temporal dynamics of the transmission filters.

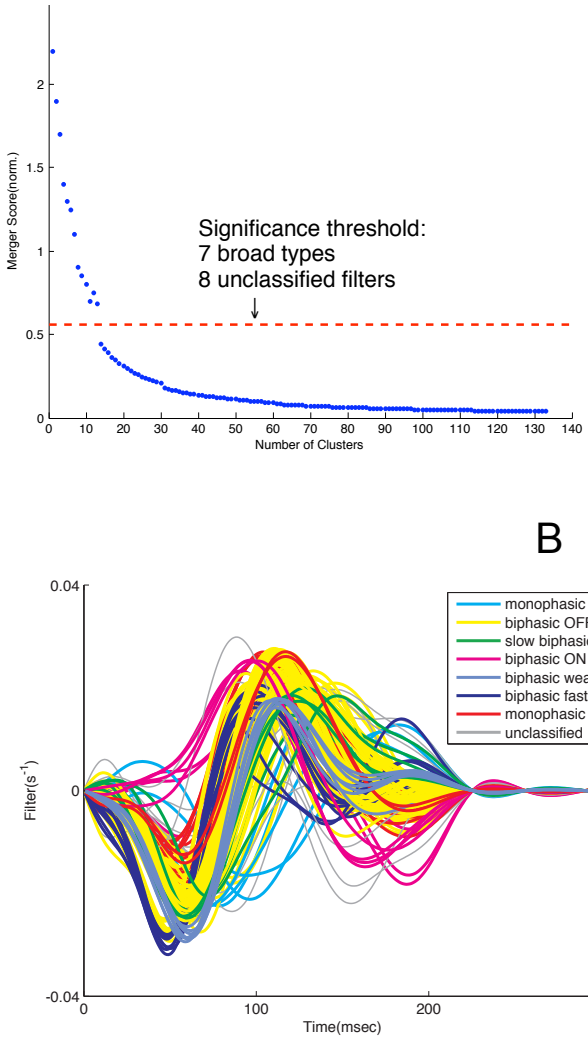


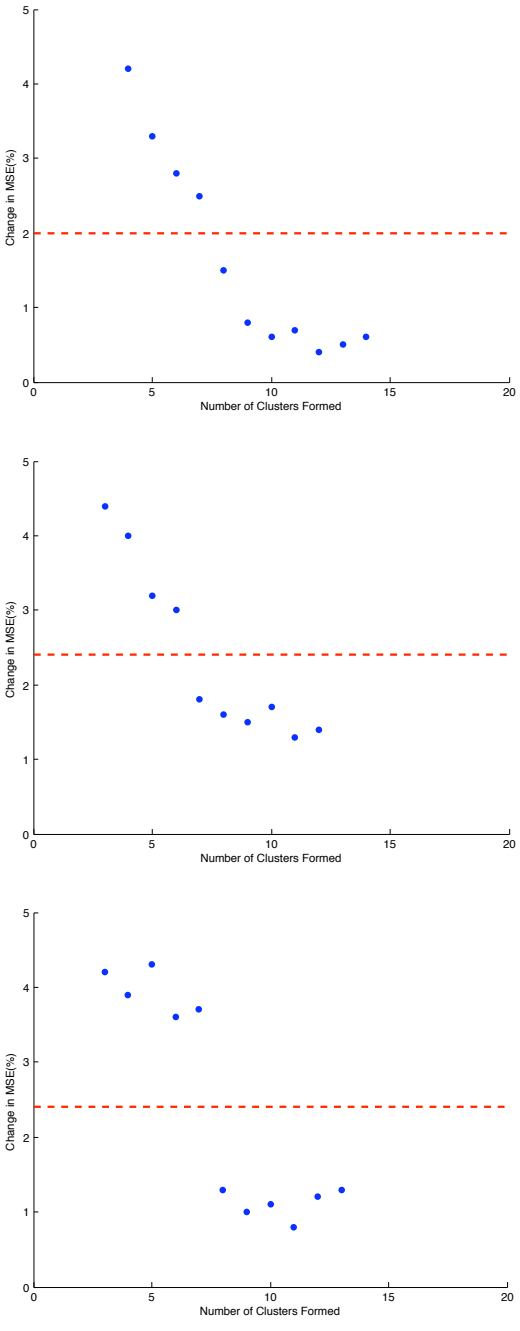

Figure 6. *A*: merger score as function of number of clusters. Significance threshold (red dashed line) identifies 7 broad types and 8 unique cells. *B*: amacrine transmission filters, of 133 ganglion cells measured from 11 retinas with broad type indicated by color

## *Fine types*

Because agglomerative clustering algorithms lump the transmission population into no more than seven broad functional types, we wanted to explore other clustering schemes that might resolve more types. Our approach was motivated by the observation that for data recorded from a single patch of the retina, where ganglion cells presumably shared inputs from some of the same amacrine interneurons and also both amacrine and ganglion cells see the same visual stimulus, we often found several transmissions with exceptionally similar functional properties. Building on this observation, we used a fine classification scheme, where we considered only cells from a single retinal patch. To divide the population into

the maximal number of cell types allowed by the data, we used K-means clustering to define cluster boundaries.

Figure 7 shows examples of fine types formed from ganglion cells recorded in three different retinal patches. As more clusters were formed, the total difference between the transmission profile of individual ganglion cells and cluster averages decreased (Figure 7A). For the first retinal patch (*top row*), this decrease $\Delta^{(K)}$ was large when the number of clusters was 7 or less, and dropped significantly when >7 clusters were formed. As a result, we divided this group of transmissions into 7 fine types (Figure 7B). For the second and third retinal patches, a transition in the clustering score $\Delta^{(K)}$ was found after six clusters. Consistent results were found for other retinal patches, with a total number of fine types ranging ≤7, depending in part on how many cells were recorded in a single patch. Our fine classification scheme was consistent with the broad scheme: fine functional types were either the same as a broad type or they were subtypes within a single broad type; the fine types never combined cells from different broad types.



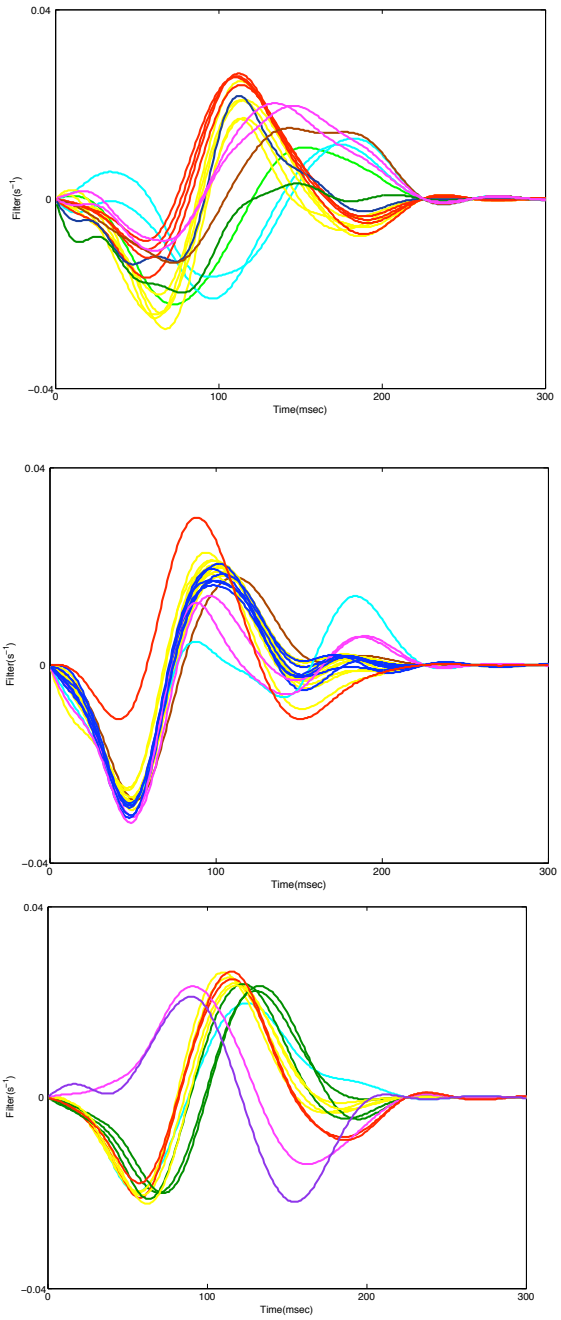



Figure 7. *A*: change in the mean-squared difference between cluster centers and individual cells $\Delta^{(K)}$ plotted as a function of the number of clusters *K* for three retinal patches. Clusters were defined using K-means clustering. The decrease in mean-squared difference dropped sharply after (7/6/6) clusters were defined (top/middle/bottom) and reached a similar small value, indicating that these subsequent distinctions were not significant (dashed red line). *B*: transmission filters of cells simultaneously recorded from three retinal patches. Cells are divided into fine types, shown by their color.

## 4. Conclusion

We relied on quantitative methods of functional classification, involving several choices of clustering algorithm as well as several choices of response characteristic (using dimensionality reduction). One issue that should be addressed is how we can compare the performance of different dimensionality reduction methods, e.g. LLE versus Isomap. Moreover, because any method of functional classification requires some arbitrary choices, we should supplement this approach by some other measures of functional similarity such as analyzing the shared information between ganglion cells that can be calculated during stimulation with natural scenes and makes minimal assumptions about how ganglion cell spike trains encode the visual world.

## 5. Acknowledgements

I thank professor Stephen Baccus for valuable and critical discussions; and Mihai Manu for help with data collection.

## References

[1] Olveczky BP, Baccus SA, Meister M. 2003. Segregation of object and background motion in the retina. Nature 423:401-8

[2] Baccus SA. 2007. Timing and computation in inner retinal circuitry. Annu. Rev. Physiol 69:271-90

[3] Chichilnisky EJ. 2001. A simple white noise analysis of neuronal light responses. Network. 12:199-213

area MT. J. Neurophysiol. 88: 3469–3476.

[4] Cox TF and Cox MA. Multidimensional Scaling. Second edition. 2000

[5] Tenenbaum JB, de Silva V, Langford JC. 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290(5500):2319-2323 Proc. IEEE Conf. Comput. Vision Pattern Recog.: 321-326.

[6] Roweis S and Saul L. 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290(5500) 2323-2326

[7] Ridder DD and Duin RPW. 2002. Locally linear embedding for classification. IEEE Trans. on Pattern Analysis and Machine Intelligence, Number PH-2002-01

[8] Segev R, Puchalla J, and Berry MJ. 2006. Functional organization of ganglion cells in the salamander retina. J Neurophysiol 95: 2277-2292

[9] Duda RO, Hart PE, and Stork DG. Pattern Classification. New York: Wiley-Interscience, 2000