

Informatik/Mathematik/Komplexe Systeme, Neurobiologie

## Der kollektiven Signalverarbeitung von Nervenzellen auf der Spur

Bethge, Matthias;

Max-Planck-Institut für biologische Kybernetik, Tübingen  
Forschungsgruppe – Computational Vision and Neuroscience

Korrespondierender Autor

Bethge, Matthias

E-Mail: [matthias.bethge@tuebingen.mpg.de](mailto:matthias.bethge@tuebingen.mpg.de)

---

### Zusammenfassung

Untersuchungen zur neuronalen Signalverarbeitung beruhen auf Analysen, die die Aktivität von Nervenzellen mit Reizen korrelieren. Neue Fortschritte haben die simultane Aufnahme der Aktivitäten vieler Nervenzellen ermöglicht. Mit mathematischen Methoden werden diese dazu genutzt, die Antworten ganzer Netzwerke zu bestimmen. Dies ermöglicht Analysen, wie auf der Netzhaut eingehende Bilder von den Netzwerken der Sehbahn verarbeitet werden.

### *Abstract*

*Knowledge about neural signaling originates from experiments where the activity of a neuron is correlated with stimuli. Meanwhile, simultaneous recording from many neurons is possible. With new mathematical methods such experiments can be used to determine the response properties of neural networks rather than of individual cells. This allows us to analyze how retinal images are processed collectively by neural networks in the visual pathway.*

### **Erst die Verschaltung macht die Musik**

Vielleicht haben Sie auch einmal als Kind aus Neugier ein altes Radio auseinander genommen, um zu erfahren, wie so ein Gerät eigentlich funktioniert. Öffnet man das Gehäuse, so findet man viele verschiedene Einzelteile: Transistoren, Kondensatoren, Widerstände, Spulen und noch mehr. Doch die Funktionsweisen der einzelnen Bauteile alleine verraten noch lange nicht, wie ein Radio funktioniert. Denn es gilt: „Erst die Verschaltung macht die Musik“.

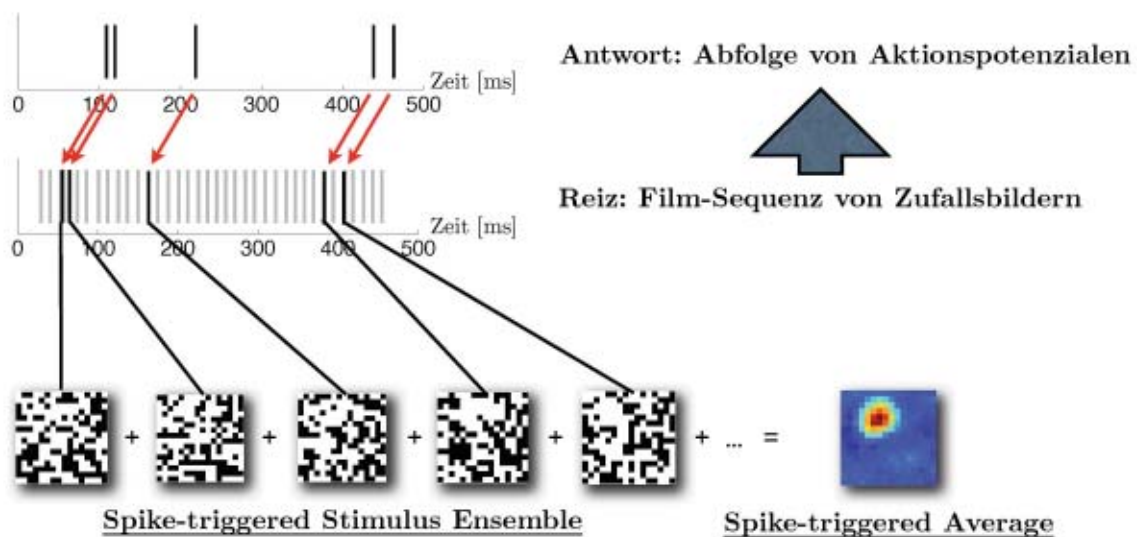
Die neurowissenschaftliche Forschung steht heutzutage vor einem ähnlichen Problem: Während die biophysikalischen Mechanismen von einzelnen Nervenzellen verhältnismäßig gut verstanden sind, ist die entscheidende Frage, wie die sensorischen Signale in den stark verschalteten Netzwerken von Nervenzellen verarbeitet werden, noch weitgehend ungeklärt. Der schnellste Informationsaustausch zwischen den Nervenzellen erfolgt durch kurzfristige Spannungspulse, deren Dauer ungefähr eine Millisekunde beträgt. Diese Pulse werden Aktionspotenziale oder auch einfach nur *Spikes* genannt. Die Ausbreitung der Aktionspotenziale von Neuron zu Neuron erfolgt über die verästelten Fasern der Nervenzellen, die so genannten Axone und Dendriten. Die Ausbreitungsgeschwindigkeit kann dabei bis zu 120 Meter pro Sekunde betragen, wenn es sich um myelinisierte Axone handelt.

Eine weitere Gemeinsamkeit zwischen Radio und Gehirn besteht darin, dass Signale aus der Umwelt aufgenommen und entschlüsselt werden müssen. So wie das Radio elektromagnetische Wellen über die Antenne empfängt, so wird im Gehirn beim Sehen die Information aus der Umwelt mit Hilfe der Augen in neuronale Aktivitätsmuster umgewandelt. Von dort verzweigt sich die Information allerdings sehr schnell und beeinflusst die Aktivität von verschiedenen Nervenzellverbänden, die jeweils einen hohen Verschaltungsgrad aufweisen.

### Der *Spike-triggered Average*

Ein wichtiger Forschungsansatz, um die neuronale Signalverarbeitung beim Sehen zu verstehen, ist der Versuch, die Erregungsmuster in den ersten Verarbeitungsstufen der Sehbahn für beliebige sensorische Eingangssignale vorherzusagen. Der gleiche Ansatz funktioniert bei vielen elektronischen Schaltungen sehr gut: Durch geschickte Wahl verschiedener Testsignale und die Beobachtung, wie sich die Spannung an verschiedenen Stellen im Schaltkreis als Reaktion auf diese Testsignale ändert, lässt sich die Funktion einer Schaltung oft effizient bestimmen.

Solche Korrelationsanalysen werden auch dazu benutzt, um mehr über die visuelle Verarbeitung im Gehirn herauszufinden. In elektrophysiologischen Experimenten wird dazu die Abfolge von Aktionspotenzialen einer Nervenzelle aufgezeichnet, während dem Versuchstier ein Film gezeigt wird, der häufig aus einer schnellen Sequenz von Zufallsbildern besteht. Jedes Mal, wenn die Nervenzelle ein Aktionspotenzial generiert, wird die unmittelbar vorausgegangene Teilsequenz von Bildern gespeichert. Auf diese Weise erhält man das so genannte *Spike-triggered Stimulus Ensemble* (Abb. 1).



**Abb. 1:** Der *Spike-triggered Average*.

Der Film, der dem Versuchstier als visueller Reiz präsentiert wird, besteht aus einer Sequenz von Zufallsbildern, die durch senkrechte Striche oberhalb der Zeitachse angedeutet ist (mittlere Zeile). In dem Diagramm darüber ist die zeitliche Abfolge der Aktionspotenziale angedeutet, mit denen das Neuron auf den Reiz reagiert. Die Zufallsbilder, die den Spikes vorausgegangen sind, werden ausgewählt und gespeichert (mittlere und untere Zeile links). Auf diese Weise erhält man das *Spike-triggered Stimulus Ensemble*. Der Mittelwert dieses Ensembles, der *Spike-triggered Average*, spiegelt das Reizmuster wider, welches am stärksten mit der Spike-Antwort des Neurons korreliert ist (rechts unten).

Urheber: Max-Planck-Institut für biologische Kybernetik/Bethge

Wiederholt man das Experiment mit genau der gleichen Sequenz von Bildern, so erhält man in der Regel nicht noch mal die gleiche Abfolge von Aktionspotenzialen. Bereits die Ganglienzellen in der Netzhaut weisen ein erhebliches Maß an Variabilität in ihrem Antwortverhalten auf. Das heißt, die neuronalen Erregungsmuster sind durch die visuellen Reize nicht eindeutig festgelegt. Stattdessen lassen sich die Antworteigenschaften von Neuronen nur statistisch beschreiben. Am häufigsten wird dazu der Mittelwert des Spike-triggered Stimulus Ensembles, der so genannte *Spike-triggered Average* benutzt. Dieser stellt die zeitliche Bildsequenz dar, welche die größte Wahrscheinlichkeit aufweist, ein Aktionspotenzial hervorzurufen.

## Das Standardmodell der sensorisch-visuellen Verarbeitung

Vieles von dem, was wir heute über das visuelle System wissen, basiert auf dem *Spike-triggered Average*. Die gängige Lehrbuch-Darstellung der Signalverarbeitung in der Sehbahn bis zu der Stelle, an der die retinale Bildinformation den primären visuellen Kortex erreicht, ist die einer linearen Filterbank, bei der die Nervenzellen unabhängig voneinander verschiedene Filter implementieren. Jeder Filter kann durch einen bevorzugten Reiz, das so genannte „Rezeptive Feld“, charakterisiert werden. Je mehr der aktuelle Reiz mit dem Rezeptiven Feld übereinstimmt, desto größer ist die Wahrscheinlichkeit, dass das Neuron Aktionspotenziale generiert. Zusätzlich zu der bereits von Hubel und Wiesel gefundenen Orts- und Orientierungsselektivität der Rezeptiven Felder gibt es auch noch Unterschiede in der räumlichen Durchlassfrequenz, die die Filtereigenschaften der Neuronen im primären visuellen Kortex charakterisieren.

Zwei wesentliche Eigenschaften des Standardmodells, die Linearität und die Unabhängigkeit der Antworten, stellen weniger ein gemessenes Ergebnis dar, sondern spiegeln viel mehr die wesentlichen Annahmen wider, die dem *Spike-triggered Average* zugrunde liegen. Obwohl es schon immer erhebliche Zweifel an diesen Eigenschaften des Standardmodells der frühen visuellen Verarbeitung gegeben hat, so ist es bisher nicht gelungen, die vielen Einzelbeobachtungen, die dem Standardmodell widersprechen, in ein verständliches Gesamtbild zu integrieren.

In der Tat ist die Beschränkung auf unabhängige lineare Filter zur Modellierung von Nervenzellantworten nicht so abwegig, wie es auf den ersten Blick vielleicht erscheinen mag. Im Gegensatz zu vielen anderen ad hoc Untersuchungen ermöglicht die lineare Systemanalyse eine eindeutige Charakterisierung der neuronalen Antworteigenschaften, welche die Spike-Aktivität auch für beliebige, vorher nicht getestete Reize mit messbarem Erfolg vorhersagen kann.

Mit anderen Worten, um die so genannten nicht-klassischen Eigenschaften von Rezeptiven Feldern in ein schlüssiges Gesamtbild zu bringen, bedarf es mehr als nur der bloßen Einsicht, dass die Neuronen sich gegenseitig beeinflussen und Nichtlinearitäten aufweisen. Ein Gesamtmodell ist nötig, welches den Raum der denkbaren Nichtlinearitäten und Wechselwirkungen zwischen den Nervenzellen aufspannt und zu einer messbaren Verbesserung in der Vorhersage der neuronalen Antworten führt. Zum Teil sind solche Modelle bereits in den späten achtziger Jahren vorgeschlagen worden, die jedoch zunächst keine besonders große Beachtung gefunden haben. Das mangelnde Interesse zu der Zeit kann man damit begründen, dass damals noch nicht die Messmethoden zur Verfügung standen, mit denen man die Aktivität von vielen Zellen gleichzeitig hätte aufnehmen können.

## Rezeptive Felder für neuronale Populationen

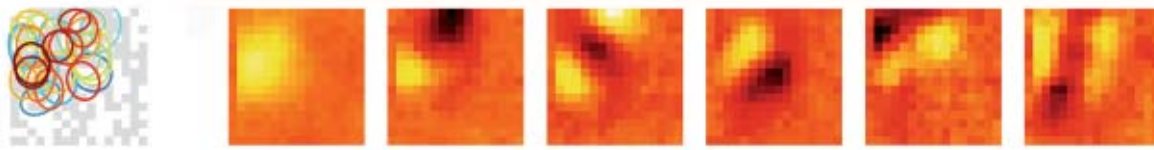
Dank großer experimenteller Fortschritte in den letzten Jahren gibt es heute aber immer mehr Laboratorien, die in der Lage sind, solche simultanen Vielzellenableitungen durchzuführen. Dadurch kommt der Entwicklung von neuen, nichtlinearen Modellen, die auch die Netzwerkeigenschaften der Nervenzellen berücksichtigen, jetzt eine große Bedeutung zu. In der Tat stellt die Analyse der Daten aus solchen Experimenten eine beachtliche Herausforderung dar, weshalb das eigentliche Potenzial dieser neu erschlossenen experimentellen Möglichkeiten bisher noch kaum genutzt werden konnte, um ein neues Standardmodell von der frühen visuellen Verarbeitung zu entwickeln.

Eine relativ einfache Methode, die viel dazu beitragen kann, dass die Signalverarbeitung von Nervenzellgruppen besser verstanden und auch intuitiv zugänglicher wird, ist die Erweiterung des Konzepts des Rezeptiven Feldes von der einzelnen Zelle auf die ganze Population [1]. Während der *Spike-triggered Average* das Reizmuster angibt, dessen Amplitude aus den Spikes eines einzelnen Neurons rekonstruiert werden kann, möchte man gern analog bei vielen Neuronen wissen, welche Komponenten der Bildreize am besten aus der Populationsaktivität rekonstruiert werden können.

Diese Frage lässt sich jedoch nicht mehr durch *Spike-triggering* beantworten, da es nun bei  $n$  Neuronen, anstelle der binären Unterscheidung zwischen „Spike“ und „kein Spike“, zu jedem Zeitpunkt tatsächlich  $2^n$  verschiedene räumliche Spikemuster gibt, die beobachtet werden könnten. Für jedes mögliche Muster ein Rezeptives Feld zu ermitteln wird bei wachsender Anzahl von Neuronen sehr schnell unmöglich. Stattdessen kann man aber fragen, welche Stimulusparameter besonders gut von einem nachgeschalteten Neuron aus der Populationsaktivität herausgelesen werden könnten.

In einer sinnvollen Verallgemeinerung des *Spike-triggered Average* muss man daher ein doppeltes Optimierungsproblem lösen: Im Prinzip muss man für jede mögliche Reizkomponente den dazugehörigen Filter bestimmen, mit dem diese Reizkomponente optimal aus der Populationsaktivität dekodiert werden kann. Anschließend kann man dann beurteilen, welche der verschiedenen Reizkomponenten am besten von der Populationsaktivität repräsentiert wird.

Zur Lösung dieses doppelten Optimierungsproblems bietet sich die so genannte „Kanonische Korrelationsanalyse“ an, ein grundlegender Algorithmus, der seinen Ursprung in der Statistik hat. Sie findet eine eindeutige, optimale Lösung: Man erhält eine geordnete Menge von unabhängigen Reizkomponenten, aus deren Gesamtheit man durch Kombination jeden beliebigen Reiz erhalten kann. Die Anordnung richtet sich aber danach, wie gut sich die jeweiligen Komponenten aus der neuronalen Aktivität der Population rekonstruieren lassen. Entsprechend dieser Anordnung wird deshalb zwischen dem ersten, zweiten, dritten, ... Rezeptiven Feld der Population unterschieden. In **Abbildung 2** wird die Methode für ein Beispiel einer Population von 32 retinalen Ganglienzellen demonstriert. Tatsächlich bieten die „Populationsbezogenen Rezeptiven Felder“ eine viel bessere Zusammenfassung der Signalübertragungseigenschaften der Gesamtheit der Nervenzellen als dies auf der Basis der „Einzelzellen-Rezeptiven Felder“ möglich wäre.



**Abb. 2:** Populationsbezogene Rezeptive Felder.

Die 32 klassischen Rezeptiven Felder der einzelnen retinalen Ganglienzellen (links) können nur sehr groben Aufschluss darüber geben, welche Bildmerkmale von dieser Population kodiert werden. Zum Vergleich sind auf der rechten Seite die ersten sechs „populationsbezogenen Rezeptiven Felder“ dargestellt, angeordnet nach ihrer Wichtigkeit von links nach rechts. Diese vermeiden nicht nur die Unübersichtlichkeiten, die durch die Überlappungen der „Einzelzell-Rezeptiven Felder“ zustande kommen, sondern berücksichtigen auch Unterschiede in der Zuverlässigkeit sowie Korrelationen zwischen den verschiedenen Neuronen. Tatsächlich entspricht die Anordnung genau dem Signal-Rausch-Verhältnis, mit dem die verschiedenen Komponenten übertragen werden, so dass die Filtereigenschaften der Population durch die Populationsbezogenen Rezeptiven Felder exakt charakterisiert sind.

Urheber: Max-Planck-Institut für biologische Kybernetik/Bethge

### Methodische Erweiterungen für nichtlineare Effekte

Die bisher vorgestellte Methode der „Populationsbezogenen Rezeptiven Felder“ ist nach wie vor auf lineare Signalverarbeitungsprozesse beschränkt. Es gibt jedoch diverse Erweiterungen, die auch nichtlineare Effekte auf der Ebene von Nervenzellverbänden mit einbeziehen: Zunächst lässt sich die Kanonische Korrelationsanalyse durch Kombination mit so genannten Kernmethoden [2, 3] so erweitern, dass sich auch verschiedenste nichtlineare Effekte berücksichtigen lassen, sowohl hinsichtlich der neuronalen Verarbeitung des Reizes als auch hinsichtlich der neuronalen Kodierung, die bei der Dekodierung der Reizkomponenten unterstellt wird [1]. Des Weiteren kann man den binären Charakter von Aktionspotenzialssequenzen durch eine entsprechende Erweiterung von korrelationsbasierten Methoden berücksichtigen. Konzeptionell bieten sich hierzu insbesondere die aus der statistischen

Physik zur Modellierung von Spingläsern bekannten Ising-Modelle an, die zur Klasse der so genannten Maximum-Entropie-Modelle gehören. Die zum Teil erheblichen Schwierigkeiten in der Praxis, solche Modelle tatsächlich an gegebene Daten anzupassen, konnten teilweise durch ein effizientes Annäherungsverfahren überwunden werden [4].

Zum Abschluss sei noch die Klasse der „Generalisierten Linearen Modelle“ als eine dritte Erweiterungsmöglichkeit erwähnt, die in vieler Hinsicht einen guten Kompromiss zwischen der Flexibilität in der Modellierung nichtlinearer Effekte und der Schätzbarkeit der Modellparameter darstellt. Darüber hinaus kann die Schätzung der Parameter „bayesianisch“ durchgeführt werden, so dass man auch eine Aussage darüber erhält, wie groß die Unsicherheit in den verschiedenen Parametern ist [5].

Alle diese Arbeiten haben das Ziel, simultane Ableitungen von vielen Zellen für eine systematische Erforschung der Frage nutzbar zu machen, wie große Populationen von Neuronen kollektiv die auf der Netzhaut eingehenden Bilder verarbeiten. Vielleicht gelingt es dadurch eines Tages zu verstehen, welche Berechnungen in der Sehbahn durchgeführt werden – schließlich wäre es uns ohne diese Berechnungen nicht möglich, die Welt zu sehen und einen Text wie diesen gelesen zu haben.

**Literaturhinweise**

- [1] J. H. Macke, G. Zeck, M. Bethge:  
Receptive fields without spike-triggering.  
Advances in Neural Information Processing Systems **21**, ...–... (2007).
- [2] B. Schölkopf, A. Smola:  
Learning with Kernels.  
MIT Press, Cambridge, MA (2002).
- [3] K. Fukumizu, F. R. Bach, A. Gretton:  
Statistical consistency of kernel canonical correlation analysis.  
Journal of Machine Learning Research **8**, 361–383 (2007).
- [4] M. Bethge, P. Berens:  
Near-maximum entropy models for binary neural representations of natural images.  
Advances in Neural Information Processing Systems **21**, ...–... (2007).
- [5] S. Gerwinn, J. Macke, M. Seeger, M. Bethge:  
Bayesian inference for spiking neuron models with a sparsity prior.  
Advances in Neural Information Processing Systems **21**, ...–... (2007).

**Weiterführende Links**

- [1] Forschergruppe Matthias Bethge  
Internetseiten der Forschergruppe  
<http://www.kyb.mpg.de/bethgegroup/index.html>

**Drittmittelfinanzierung**

BMBF