

Optimality in mono- and multisensory map formation

Moritz Bürck · Paul Friedel · Andreas B. Sichert ·
Christine Vossen · J. Leo van Hemmen

Received: 5 January 2010 / Accepted: 10 April 2010 / Published online: 26 May 2010
© Springer-Verlag 2010

Abstract In the struggle for survival in a complex and dynamic environment, nature has developed a multitude of sophisticated sensory systems. In order to exploit the information provided by these sensory systems, higher vertebrates reconstruct the spatio-temporal environment from each of the sensory systems they have at their disposal. That is, for each modality the animal computes a neuronal representation of the outside world, a monosensory neuronal map. Here we present a universal framework that allows to calculate the specific layout of the involved neuronal network by means of a general mathematical principle, viz., *stochastic optimality*. In order to illustrate the use of this theoretical framework, we provide a step-by-step tutorial of how to apply our model. In so doing, we present a spatial and a temporal example of optimal stimulus reconstruction which underline the advantages of our approach. That is, given a known physical signal transmission and rudimental knowledge of the detection process, our approach allows to estimate the possible performance and to predict neuronal properties of biological sensory systems. Finally, information from different sensory modalities has to be integrated so as to gain a unified perception of reality for further processing, e.g., for distinct motor commands. We briefly discuss concepts of multimodal interaction and how a multimodal space can evolve by alignment of monosensory maps.

Keywords Map formation · Sensory processing · Optimality

Moritz Bürck, Paul Friedel, Andreas B. Sichert and Christine Vossen have contributed equally to this work.

M. Bürck (✉) · P. Friedel · A. B. Sichert · C. Vossen · J. L. van Hemmen
Technical University of Munich, Munich, Germany
e-mail: mbuerck@ph.tum.de

1 Introduction

A mouse hears a rustling in the grass, sees some leaves moving and escapes from the predator. Thus, the perception of the outside world by sensory systems and the consequent translation of their response into a reliable neuronal representation that allows, for instance, directional motor commands is an essential concept for surviving. A neuronal representation of the external world is what we call a *map*. Depending on the map processing information from one or many sensory systems, the map is called uni- or multimodal. The advantageous concept of a neuronal map will be discussed in detail in the next section.

The processing of sensory information, from its generation to multimodal map formation, can in general be subdivided into three steps: physical mapping, optimal map formation, and multimodal integration; see Fig. 1. Before focusing on the aspects of optimality we analyze the “golden three” of sensory processing in turn.

1.1 Physical mapping

An object in the outside world reveals its presence by generating different signals that are transmitted along distinct physical pathways. A running animal may for instance generate sound and a changing visual image as it moves, as well as vibrations and an infrared profile. In concrete terms, given any signal that varies as a function of spatial position and time, it is possible to calculate the time-dependent response of the receptor neurons. That is, the physical mapping of the signal onto the neuronal detector response can be described by a set of *transfer functions*, indicated by the upper arrows in Fig. 1. The responses of the sensory systems then represent particular physical quantities such as sound, light intensity, volatile molecules, or heat originating from the object.

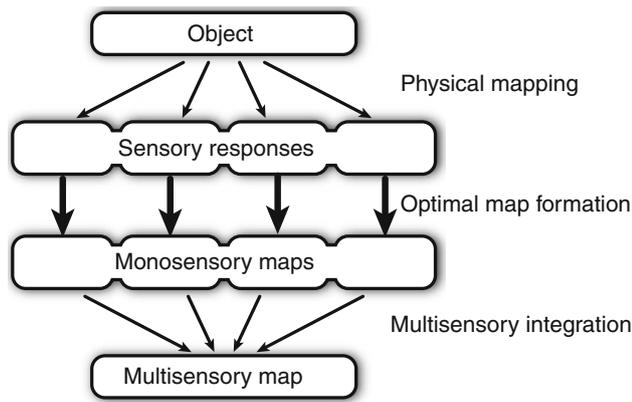


Fig. 1 The three steps leading to the formation of a unified multisensory map. An object in the outside world generates physical input signals, which can be detected by several different sensory systems. To form a map, the physical mapping must be “inverted” in some suitable way. After monosensory map formation, the distinct maps are combined into a unified multisensory map

1.2 Optimal map formation

From the sensory responses an observer needs to reconstruct a map that represents the spatio-temporal stimulus. For some sensory systems a pre-stage map may already exist inherently, e.g., on the retina in the visual system. For other systems, such as the auditory system (Carr and Konishi 1988, 1990; Kempter et al. 2001; Brand et al. 2002; Leibold and van Hemmen 2005), the lateral-line system (Franosch et al. 2003, 2005b; Goulet et al. 2008), or infrared vision (Sichert et al. 2006), spatial information is not readily available and a neuronal map must be constructed more explicitly. In either case, we want the map to represent the environment as accurately as possible. That is, we want an *optimal map*. So the task of the brain is now to obtain an optimal reconstruction of the signal (middle arrows in Fig. 1). The key to success is the choice of the right neuronal connections between the sensory systems and the corresponding maps. That is, the synaptic connections have to be chosen in such a way that the network “inverts” the physical mapping of the signal to the sensory and, thus, neuronal response (Oğuztöreli and Caelli 1985; Takeda and Goodman 1986; Zhou et al. 1988). The firing activity of the map neurons will then accurately represent the spatio-temporal signal.

1.3 Multimodal integration

In a final step of sensory processing, the monosensory maps must be merged into a single unambiguous multisensory map. Here two difficulties arise. First, proper map alignment is required for the successful fusion of unimodal maps. A collection of aligned maps is realized, for example, in the

superior colliculus (SC)¹ (Stein and Meredith 1993; King 1999; Calvert et al. 2004); see Sect. 2. Second, the monosensory maps should be combined optimally (Gu et al. 2008; Morgan et al. 2008) to increase the quality of the integrated map in comparison to that of the contributing maps. The integrated multimodal map then allows for a new concept in multimodal processing, the so-called “pooling” of information, an efficient way to identify and characterize objects.

The first processing step, physical mapping, is a purely physical description of the signal generation and detection process. In this prospect, we focus on optimal map formation and touch multimodal integration superficially. We begin with reviewing the concept of a neuronal map in Sect. 2. We then proceed in Sect. 3 by discussing a general framework that describes how a neuronal map can be built from a given sensory input in a *stochastically optimal* way (Mosegaard and Tarantola 2002). After the model has been derived we consider concrete examples of model predictions in Sect. 4. It turns out that the model can successfully describe several properties of monosensory map formation. Furthermore, we show that many known experimental findings and theoretical derivations can be explained within this unified framework. After having discussed carefully how the presented framework can describe monosensory maps, we move on to the integration of monosensory into multisensory maps in the final Sect. 5. We thereby review the current literature from the perspective of maps. To discuss how monosensory maps can form a single unambiguous representation, we present basic concepts such as “integration” and “pooling” of information. We conclude the final section with remarks on how a common sensory space can develop at all.

2 What’s in a map?

A major role in sensory processing is reserved for *maps*² (Knudsen et al. 1987; van Hemmen 2002). A neuronal map

¹ The SC is called *optic tectum* in non-mammals. In this review, we will simply use “SC” to refer to either the optic tectum or the superior colliculus, depending on the context.

² Although it has only relatively recently become possible to explicitly demonstrate the existence of neuronal maps, the idea that a map-like architecture underlies certain aspects of sensory processing is much older. Already in 1879, Helmholtz (1879) remarked “Dass durch das Entlangführen des tastenden Fingers an den Objecten die Reihenfolge kennen gelernt wird, in der sich ihre Eindrücke darbieten, dass diese Reihenfolge sich als unabhängig davon erweist, ob man mit diesem oder jenem Finger tastet, dass sie ferner nicht eine einläufig bestimmte Reihe ist, deren Elementen man immer wieder vor- oder rückwärts in derselben Ordnung durchlaufen muss, um von einem zum anderen zu kommen, also keine linienförmige Reihe, sondern ein flächenhaftes Nebeneinander, oder nach Riemann’s Terminologie, eine Mannigfaltigkeit zweiter Ordnung, das alles ist leicht einzusehen.” That is, Helmholtz already recognized that a neuronal representation of a two-dimensional surface constitutes a two-dimensional manifold in the brain.

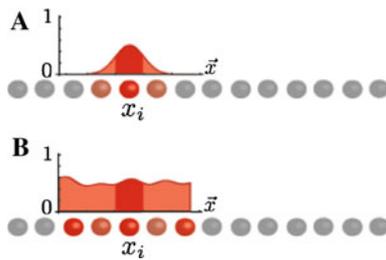


Fig. 2 The firing profile of a map encodes the likelihood of finding an object at a certain position in sensory space. Focusing on the firing rate of only a *single* neuron such as neuron x_i and ignoring its neighbors prevents a faithful perception of reality. A faithful perception can only be achieved if the activity of the whole neuronal map (just compare the ‘neighborhood’ of x_i with the rest) is taken into account. Consequently, even though the nominal value of the firing rate of neuron x_i is identical in case A and B, the represented physical reality differs significantly in both cases

is a neuronal representation of the external world quite often realized by an array of neurons in which neighboring neurons respond to similar sensory stimuli.

As an example, visual input in the mammalian brain is processed through multiple cortical layers that are organized according to the topography of the retinal input cells (“retinotopic organization”) (Wandell 1995; Kandel et al. 2000). In this way, neighboring neurons respond to visual input from neighboring points in space and thus form a spatial map. Spatial maps have been discovered in various sensory systems in many groups of vertebrates (Dräger and Hubel 1975; Hubel and Wiesel 1977; Hartline et al. 1978; Stein and Gaither 1981; Carr et al. 1982; Knudsen 1982; King and Palmer 1983; Middlebrooks and Knudsen 1984; Sullivan and Konishi 1986; King and Hutchings 1987; Zittlau et al. 1988; Bartels et al. 1990) and even in some non-vertebrates (Järvilehto 1985).

Neuronal maps are not limited to spatial representations as several examples show, for instance, in the auditory system. There, maps representing frequency, interaural time difference, interaural amplitude difference, and even amplitude modulation frequencies exist (Takahashi and Konishi 1986; Manley et al. 1988; Pickles 1988; Schreiner and Langner 1988; Olsen et al. 1989; Geisler 1990). Closely related to the concept of a neuronal map is that of receptive fields. A receptive field of a sensory neuron is defined as the region of space in which the presence of a stimulus alters the activity of the neuron. As we will see in Sects. 3.5, 4.2, and 4.3, our model can predict such receptive fields.

One might argue that neuronal maps exist simply because their neuronal architecture *follows* the sensory surface of their input modality. From this point of view, for instance the visual layers are retinotopically organized *because* they receive their input from the retina. Similarly, a frequency map just reflects the tonotopic organization of the cochlea.

Interestingly, this argument does not hold for every sensory map as it already becomes apparent in the existence of auditory maps for interaural time and amplitude differences. Regarding examples in the spatial domain, let us consider the sensory units of the frog lateral line system (Zittlau et al. 1988) or the snake infrared system (Hartline et al. 1978). These sensory units receive a complex superposition of input from several different spatial locations. It is certainly *not* straightforward to build a map from such a complex input (Franosch et al. 2003, 2005b; Goulet et al. 2008; Sichert et al. 2006).

A key question (van Hemmen 2002) is therefore: What is the *function* of a neuronal map? That is, why choose a map structure for neuronal processing? One argument is that, in contrast to *arbitrary* population coding, neuronal maps ensure a topographic neuronal organization. This organization then underlies the neuronal processing and allows the efficient representation of a continuously varying input signal. For instance, it allows the interpretation of a firing pattern on a spatial map as the *likelihood* to find a sensory object at a certain position (van Hemmen 2002; Pouget 2003; Jazayeri and Movshon 2006; Seung and Sompolinsky 1993; Denève et al. 2001). As illustrated by Fig. 2, in the map perspective one needs to consider the activity of the complete map in order to retrieve meaningful information from the firing rate of a single neuron.

There is an even more convincing argument supporting the importance of maps in the brain. The real computational power of neuronal maps can be appreciated only when the *interplay* of several maps is considered. A beautiful example of the connection between different maps is the well-studied SC (Stein and Meredith 1993; Calvert et al. 2004) to which all sensory systems project that provide information in a map-like form (Stein and Meredith 1993). Within the SC, multisensory as well as predominantly monosensory layers are found and are shown to process spatial information. All neuronal maps are mutually aligned to gain a unified multisensory representation of sensory space (Stein and Meredith 1993; King 1999; Calvert et al. 2004). The combined sensory information can then be used to generate directional motor responses (Krauzlis et al. 1997, 2004; Luksch 2008; Stein et al. 2004; van Opstal and Munoz 2004). This is only possible since, within the SC, there exist not only sensory but also motor maps, i.e., motor neurons organized in a map-like structure. Direct evidence for this hypothesis has recently been found in eye tracking experiments (Hafed et al. 2008).

Moreover, external *objects* can be *identified* by their position encoded through the firing pattern in a neuronal map. This does not mean that an object is adequately described if only the position is known but rather it implies that position serves as appropriate and—in contrast to higher computational levels (Eckhorn et al. 1988; Gray et al. 1989)—necessary information for defining a sensory object. When

combining different sensory systems, the spatial information is needed to bind information associated with the same position in the monosensory maps into one single multimodal percept for further processing.

Bearing in mind these advantages of neuronal maps we now face the question of how to optimally construct a map from the responses of the different sensory systems.

3 Mathematical model

For the derivation of our mathematical model we recall our initial division of sensory processing into three major steps (see Fig. 1) and our focus on optimal map formation. We consequently need to decode the stimulus characteristics at the best from the sensory response as described in Sect. 1. Mathematically speaking we have to derive the *inverse transfer function* that can perform an optimal reconstruction of a particular stimulus from the sensory response. The inverse transfer function can then be translated into a neuronal connectivity pattern.

The derivation we provide below is based on two reasonable simplifications. First, we assume that all sensory maps are purely *monosensory*. Although it has been questioned whether maps without influence from other sensory systems exist (Stein and Stanford 2008), our assumption can be justified by the finding that many spatial maps are clearly dominated by a single sensory modality (Wallace and Stein 2007).

The second important assumption is a *linear* relation between the stimulus and the receptor response of the sensory system. That means that the detector responses change proportionally to the signal strength. For example, the moving speed (signal strength) of a submerged moving object translates linearly to the detectable water velocities at the lateral-line organs (Sichert et al. 2009; Franosch et al. 2005b). Nonlinear relations between stimulus and detector responses, e.g., a logarithmic response (Krueger 1989; Norwich and Wong 1997; Laming 1997; Johnson et al. 2002; Copelli et al. 2002), can in principle be treated with our model as well; see Appendix A for details.

3.1 Definition of the problem

An object generates a stimulus $s^{\mathbf{x}}(t)$ varying in time t and position \mathbf{x} in the external world. The corresponding signal may be, for instance, the time-dependent sound pressure at a particular location or may denote the presence of edges or movement at a particular position within the visual field.

The signal induces a response $r_i(t)$ in a set of N sensory detectors. Depending on the problem at hand a single detector i with $0 \leq i \leq N$ can be a complete sensory organ, such as the left ear, or a part of a detector array such as a specific interval of best frequencies in the cochlea. In principle, the

detector combines information from past signals within the whole sensory space. The response is therefore described by

$$r_i(t) = \int_{\text{all space}} d\mathbf{x} \int_{-\infty}^t d\tau s^{\mathbf{x}}(\tau) h_i^{\mathbf{x}}(t - \tau) \quad (1)$$

where the *transfer function* $h_i^{\mathbf{x}}(t)$ incorporates the physics of signal transmission and detection. The transfer function can be different for each detector i . Auditory transfer functions, for example, incorporate the position of sound source and ear with respect to the head midline and therefore differ between right and left ear. In general, we can safely assume that $h_i^{\mathbf{x}}(t) = 0$ for large values of $|\mathbf{x}|$ and t . This reflects our intuition that events occurring far away or long ago will not influence the state of a sensor. We will need this property later on. Moreover, since any detector can only react to temporal-causal, i.e., past signals we set $h_i^{\mathbf{x}}(t) = 0$ for $t < 0$. We can then rewrite the response function (1) with adapted integration limits as a convolution with respect to time,

$$\begin{aligned} r_i(t) &= \int d\mathbf{x} \int_{-\infty}^{\infty} d\tau s^{\mathbf{x}}(\tau) h_i^{\mathbf{x}}(t - \tau) \\ &=: \int d\mathbf{x} (s^{\mathbf{x}} \star h_i^{\mathbf{x}})(t). \end{aligned} \quad (2)$$

The above equation describes the response of an ideal system. In biological systems the quality of the detector response is limited by at least three factors.

First, information may get lost during the transfer from the outside object to the inside sensory system. Second, noise influences all steps in the detection and reconstruction process (Faisale et al. 2008). Finally, limitations of the neuronal hardware, for instance, the limited dynamic range of receptors, constrain possible solutions; see Sect. 3.5 for details.

Within our mathematical model we incorporate these three restrictive factors by introducing additional noise terms. Accordingly, a term describing background noise $\xi^{\mathbf{x}}(t)$ must be added to the signal. Furthermore, we assume that transfer function and sensory response are hampered by additional noise terms $\eta_i^{\mathbf{x}}(t)$ and $\chi_i(t)$, respectively. Consequently, (2) is modified so as to read

$$r_i(t) = \int d\mathbf{x} [(s^{\mathbf{x}} + \xi^{\mathbf{x}}) \star (h_i^{\mathbf{x}} + \eta_i^{\mathbf{x}})](t) + \chi_i(t). \quad (3)$$

To reconstruct the estimated signal from the detector responses $r_i(t)$, the above transformation must be “inverted” in some appropriate way. We therefore calculate the time-dependent inverse transfer functions $l_i^{\mathbf{x}}(t)$ between detector i and the map at position \mathbf{x} . When applying $l_i^{\mathbf{x}}(t)$ to the receptor responses at i we obtain the estimate

$$\hat{s}^{\mathbf{x}}(t) = \sum_i [r_i \star (l_i^{\mathbf{x}} + \lambda_i^{\mathbf{x}})](t) \quad (4)$$

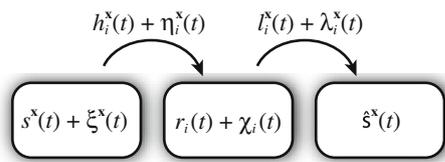


Fig. 3 *Physical mapping*: signal $s^x(t)$ with background noise $\xi^x(t)$ is mapped onto a noisy receptor response $r_i(t) + \chi_i(t)$ through the noisy transfer function $h_i^x(t) + \eta_i^x(t)$. *Optimal map formation*: the (possibly noisy) inverse transfer function $l_i^x(t) + \lambda_i^x(t)$ gives an estimate $\hat{s}^x(t)$ of the signal

of the original signal $s^x(t)$. Here the hat on $\hat{s}^x(t)$ denotes a reconstruction and the term $\lambda_i^x(t)$ represents the noise due to the concrete realization of the theoretical inverse transfer function. We note that in contrast to elsewhere (Oğuztöreli and Caelli 1985; Puetter et al. 2005) the present model is non-iterative. This will result in a purely feedforward network structure when it comes to a neuronal realization in Sect. 3.5.

Figure 3 illustrates the whole mathematical procedure of sensory information processing as described in the first two steps of Fig. 1. All the relevant terms are summarized in Table 1. In the next section, we will indicate how to calculate inverse transfer functions $l_i^x(t)$ that enable optimal signal reconstruction.

3.2 Optimal reconstruction

We want to tune our sensory system to *optimally* reconstruct not only one specific situation but the *typical* environment. In other words, biologically relevant signals belong to a class of signals that we denote as “typical”. Consequently, a specific sensory signal is a concrete realization of a class of typical, biologically relevant, signals. That is, it is a stochastic quantity. We therefore minimize the *expectation value* of the squared difference between signal and reconstruction.

This is possible because all quantities and functions (Figs. 1, 3) involved in both, the process of physical mapping (see Appendix B) and the neuronal process of optimal map formation (Sect. 3.5) are self-averaging. The mathematical definition of self-averaging allows for a description in terms of expectation values.

To derive the inverse transfer functions $l_i^x(t)$ that enable optimal signal reconstruction for a class of typical signals, we can next minimize the expectation value of the squared

error between estimated and real signal

$$E\{\mathbf{I}^x(t), t\} := \left\langle \int_{-T}^t dt' \int d\mathbf{x} [s^x(t') - \hat{s}^x(t')]^2 \right\rangle \tag{5}$$

$$= \int_{t-T}^t dt' \int d\mathbf{x} \langle [s^x(t') - \hat{s}^x(t')]^2 \rangle.$$

Here the brackets $\langle \cdot \rangle$ denote the expectation value with respect to the different types of noise and T is a typical processing time.

To be mathematically precise, an expectation value is an integral on a probability space with respect to a probability measure p . For arbitrary functions f and g , if $\langle |f - g|^2 \rangle = 0$ then $f = g$ with respect to p or, physically, looking at the world through p 's glasses: what p finds important pops up clearly whereas what p finds “irrelevant” has hardly any weight. The latter need not correspond to what we “think” ourselves (see van der Waerden 1957).

A quadratic form of the error term has been proven to be a reasonable and practical choice in many physical optimizing problems (see, e.g., Miller 1970). In case of independent Gaussian error terms, the formulation via a quadratic error is under certain conditions identical to results obtained by means of *maximum-likelihood* estimates (Johnson and Dudgeon 1993; Kay 1993); see Sect. 3.4.2.

Mathematically, the error (5) is a functional assigning to every set of inverse transfer functions one specific value. Minimization of functionals in the above integral form is a central and well-studied aspect of the calculus of variations (Clegg 1968; Gelfand and Fomin 1963; Jost and Li-Jost 1998; Van Brunt 2000). For the present situation the first variation with respect to every inverse transfer function $l_j(\mathbf{x}, t')$ is to vanish. That is,

$$\frac{\partial \langle [s^x(t') - \hat{s}^x(t')]^2 \rangle}{\partial l_j^x(t')} = 0 \quad \text{for every } j. \tag{6}$$

In order to solve (6), we have to substitute (4) for the estimate $\hat{s}^x(t)$ and replace $r_i(t)$ by its description (3). Expanding the square, we encounter expectation values of products consisting of varying combinations of noise and signal terms. Here we assume that all noise terms as well as the signal itself are stochastically independent of each other so that the expectation of a product of independent term factorizes; for

Table 1 Functions and error terms describing detection and processing of sensory information

Signal	$s^x(t) + \xi^x(t)$
Transfer function	$h_i^x(t) + \eta_i^x(t)$
Receptor response	$r_i(t) + \chi_i(t)$
Inverse transfer function	$l_i^x(t) + \lambda_i^x(t)$
Estimated signal	$\hat{s}^x(t)$

instance,

$$\langle s^{\mathbf{x}}(t)\eta_i^{\mathbf{x}'}(t') \rangle = \langle s^{\mathbf{x}}(t) \rangle \langle \eta_i^{\mathbf{x}'}(t') \rangle.$$

For a product consisting of the same kind of term we need to consider the definition of the autocorrelation of a quantity $f^{\mathbf{x}}(t)$ as given by

$$\langle f^{\mathbf{x}}(t)f^{\mathbf{x}'}(t') \rangle = \delta(\mathbf{x} - \mathbf{x}')\delta(t - t')(\mu_f^2 + \sigma_f^2) \tag{7}$$

with μ_f the mean and σ_f the variance of the quantity $f^{\mathbf{x}}(t)$. That is, we assume in a first step that the values for different spatio-temporal positions are completely uncorrelated, a kind of worst-case analysis.

Since the means of all noise terms μ_f vanish we get the following correlation terms

$$\langle \xi^{\mathbf{x}}(t)\xi^{\mathbf{x}'}(t') \rangle = \delta(\mathbf{x} - \mathbf{x}')\delta(t - t')\sigma_\xi^2, \tag{8a}$$

$$\langle \chi_i(t)\chi_j(t') \rangle = \delta_{ij}\delta(t - t')\sigma_\chi^2, \tag{8b}$$

$$\langle \eta_i^{\mathbf{x}}(t)\eta_j^{\mathbf{x}'}(t') \rangle = \delta_{ij}\delta(\mathbf{x} - \mathbf{x}')\delta(t - t')\sigma_\eta^2$$

with $|\mathbf{x}| < x^{\max}$ and $0 < t < t^{\max}$. (8c)

Through the final equation we take into account that the noise $\eta_i^{\mathbf{x}}(t)$ vanishes for large values of t and $|\mathbf{x}|$, in the same way as for the transfer function $h_i^{\mathbf{x}}(t)$.

The autocorrelation (7) of the signal $s^{\mathbf{x}}(t)$ itself depends on the problem at hand. Either the detectors of the sensory system measure absolute signal strengths (μ_s), e.g., vision, or modulations of a mean value of the signal (deviation σ_s), e.g., audition. In any case, one has to choose the corresponding biologically relevant term and put the others equal to zero. In the following, we choose the expectation value μ_s^2 of the signal as the appropriate quantity and therefore take σ_s^2 zero,

$$\langle s^{\mathbf{x}}(t)s^{\mathbf{x}'}(t') \rangle = \delta(\mathbf{x} - \mathbf{x}')\delta(t - t')\mu_s^2. \tag{9}$$

While (8) incorporates reasonable assumptions for all noise terms, the correlation (9) for the signal is a strong hypothesis. Signals are namely characterized by spatio-temporal continuity (e.g., objects and their corresponding signals usually do not disappear from one point in time to the next). A Gaussian correlation term

$$\langle s^{\mathbf{x}}(t)s^{\mathbf{x}'}(t') \rangle = A \exp\left(-|\mathbf{x} - \mathbf{x}'|^2/(2\sigma_x^2)\right) \exp\left(-|t - t'|^2/(2\sigma_t^2)\right), \tag{10}$$

for instance, can take into account correlations between neighboring points in space and time. Here σ_x and σ_t are typical spatial and temporal correlation scales. The application of such a Gaussian correlation, however, does not greatly alter the further derivation (see Appendix D for details) but

only smoothens the final estimated signal. For reasons of clarity, we will therefore stick to the relation (9).

Returning to the (6) we have to solve it and in so doing apply the correlations (8) and (9) so as to arrive at

$$L_j^{\mathbf{x}}(t) \left[\sigma_\chi^2 + (\mu_s^2 + \sigma_\xi^2) \int_{\substack{|\mathbf{y}| < y^{\max} \\ 0 < \tau < t^{\max}}} d\mathbf{y}d\tau \sigma_\eta^2 \right] + (\mu_s^2 + \sigma_\xi^2) \sum_i \int d\mathbf{y} \left[(h_i^{\mathbf{y}} \star L_i^{\mathbf{x}}) \circ h_j^{\mathbf{y}} \right](-t) = \mu_s^2 h_j^{\mathbf{x}}(-t); \tag{11}$$

for details see Appendix C. The open circle \circ denotes the autocorrelation integral

$$(a \circ b)(t) := \int_{-\infty}^{\infty} d\tau a(\tau)b(t + \tau). \tag{12}$$

In order to simplify (11), we define two new noise measures,

$$\tau^2 := \frac{\sigma_\xi^2}{\mu_s^2} \tag{13}$$

and

$$\sigma^2 := \frac{\sigma_\chi^2}{\mu_s^2} + \int_{\substack{|\mathbf{y}| < y^{\max} \\ 0 < \tau < t^{\max}}} d\mathbf{y}d\tau \frac{\sigma_\eta^2(\mu_s^2 + \sigma_\xi^2)}{\mu_s^2}. \tag{14}$$

The parameter τ represents an inverse signal-to-noise ratio. It is therefore often reasonable to assume a small value of τ . The parameter σ , on the other hand, describes the overall measurement noise by relating *detection* and *transmission* noise, σ_χ and σ_η , to the signal mean amplitude μ_s . A priori, its value cannot be assumed to be small and has to be adjusted according to the situation at hand.

In order to further simplify Eq. 11 we switch to Fourier space, where convolution (2) and correlation (12) become ordinary multiplications combined with complex conjugations. Denoting Fourier transforms by capital letters and the complex conjugation by an overline, (11) simplifies to

$$\sum_i L_i^{\mathbf{x}} \left[\sigma^2 \delta_{ij} + (1 + \tau^2) \int d\mathbf{y} H_i^{\mathbf{y}} \overline{H_j^{\mathbf{y}}} \right] = \overline{H_j^{\mathbf{x}}} \tag{15}$$

where we have used (13) and (14).

Equation 15 is the main result of our derivation. In principle, it allows us to calculate the inverse transfer functions $L_i^{\mathbf{x}}$ for optimal signal reconstruction. A calculation of the second variation—see (51) in the Appendix for details—then confirms that the inverse transformation we have found indeed minimizes the error. For convenience we will introduce an alternative notation in the next section.

3.3 Matrix notation

To rewrite (15) in a more practical notation we introduce “matrices” \mathcal{H} and \mathcal{L} by putting

$$\mathcal{H}_{[ix]} = H_i^x, \quad \mathcal{L}_{[xi]} = L_i^x. \tag{16}$$

The notations illustrate that transfer functions and inverse transfer functions are linear transformations from a continuous space (the outside world) into a discrete space (the neuronal map) and vice versa. \mathcal{H} and \mathcal{L} are therefore only formally matrices with a spatial coordinate \mathbf{x} varying in \mathbb{R} . The matrix multiplication involving the spatial coordinate must consequently be understood as an integration. A discretization of space, as is usual in numerics, would lead to a true matrix formulation.

In addition, we introduce the covariance matrix $\mathcal{C}(\mathbf{R})$ of the receptor response \mathbf{R} as described, e.g., in Johnson and Dudgeon (1993); Kay (1993). In our case we find

$$\mathcal{C}(\mathbf{R}) := \left\langle (\mathbf{R} - \langle \mathbf{R} \rangle)(\mathbf{R} - \langle \mathbf{R} \rangle)^T \right\rangle \tag{17}$$

$$= \mu_s^2 \left(\sigma^2 \mathbb{1} + \tau^2 \overline{\mathcal{H}} \cdot \mathcal{H}^T \right) \tag{18}$$

where the superscript T denotes the matrix transpose and $\mathbb{1}$ the identity matrix. Equation 15 now simplifies to

$$\mathcal{M} \cdot \mathcal{L}^T = \overline{\mathcal{H}} \quad \text{with } \mathcal{M} := \mu_s^{-2} \mathcal{C} + \overline{\mathcal{H}} \cdot \mathcal{H}^T. \tag{19}$$

Given \mathcal{M} as an invertible matrix, denoted as the ‘model matrix’, the solution for \mathcal{L} turns out to be

$$\mathcal{L} = \left(\mathcal{M}^{-1} \overline{\mathcal{H}} \right)^T = \overline{\mathcal{H}}^T \left(\mu_s^{-2} \mathcal{C} + \mathcal{H} \cdot \overline{\mathcal{H}}^T \right)^{-1}. \tag{20}$$

This equation gives a unique solution for the optimal reconstruction for any given set of transfer functions and noise constants (σ , τ). Using (4) in matrix form we find

$$\hat{\mathbf{S}} = \mathcal{L} \cdot \mathbf{R} \tag{21}$$

as estimated signal from the measured response vector \mathbf{R} .

3.4 Relation to common methods

The challenge of signal reconstruction has a long tradition, and, accordingly, one may ask how the above formalism relates to methods that have been established in this field. In the following, we will discuss the relation of our model to methods based on the pseudo-inverse and to the maximum-likelihood approach.

3.4.1 Pseudo-inverse

If the noise terms can be neglected, the interpretation of Eq. 19 is straightforward. In this case the covariance matrix \mathcal{C} vanishes and the resulting equation leads to

$$\mathcal{L} = \overline{\mathcal{H}}^T \cdot \left(\mathcal{H} \cdot \overline{\mathcal{H}}^T \right)^{-1}. \tag{22}$$

The inverse transfer functions \mathcal{L} that we have just found fulfill the properties of the Moore–Penrose *pseudo-inverse*³ of \mathcal{H} (Ben-Israel and Greville 2003). With hindsight, this makes sense since the pseudo-inverse generates an approximate inverse matrix that minimizes the quadratic error (see Appendix E). An exact inversion may not be possible for a matrix \mathcal{H} that is, e.g., rectangular instead of square or of incomplete rank.

But even in the more general situation of non-vanishing noise terms, we can observe strong relations between our framework and methods based on the pseudo-inverse. The point is that, for the calculation of the pseudo-inverse, a regularization has to be introduced to suppress noise terms, typically high-frequency variations. The so-called *Tikhonov–Miller regularization* adds a positive term $\alpha \mathbb{1}$ to make it more stable (Miller 1970; Tikhonov et al. 1977, 1995; Press et al. 2007). The regularized equation then reads

$$\left(\alpha \mathbb{1} + \overline{\mathcal{H}} \cdot \mathcal{H}^T \right) \cdot \mathcal{L}^T = \overline{\mathcal{H}}. \tag{23}$$

Comparing this equation with (19) we see that α corresponds exactly to our term σ^2 if $\tau = 0$. Hence, in this special case our general approach is identical to methods using the Tikhonov–Miller regularization.

3.4.2 Maximum-likelihood approach

The *maximum-likelihood* analysis (Johnson and Dudgeon 1993; Kay 1993) is a common tool in the interpretation of measurement data. Within the maximum-likelihood scheme one computes the stimulus that is the most likely one given a set of detector responses \mathbf{R} . Experiments have shown that optimal or near-optimal stimulus combinations can indeed describe several phenomena of sensory processing (Ernst and Banks 2002; Hürlimann and Kiper 2002; Körding and Wolpert 2004; Alais and Burr 2004; Helbig and Ernst 2007; Morgan et al. 2008). A method of optimal stimulus combination like the maximum-likelihood approach is therefore highly relevant to neuronal information processing and ought to be included into our model.

The maximum-likelihood approach tries to find the most probable input signal \mathbf{S} given the detector responses \mathbf{R} , a known transfer function \mathcal{H} , and *no* apriori knowledge about the signal ($\sigma_s = \infty$). We now assume a linear relation

$$\mathbf{R} = \mathcal{H} \mathbf{S} + \chi \tag{24}$$

³ The explicit expression for \mathcal{L} in (22) only holds if $(\mathcal{H} \cdot \overline{\mathcal{H}}^T)$ is invertible.

with χ representing the noise. We assume the noise to follow a Gaussian distribution with zero mean and the standard deviation σ_χ .

The method minimizes the noise χ . That is, based on the fundamental definitions of Bayesian statistics, it maximizes the conditional probability density function

$$p(\mathbf{R}|\mathbf{S}) \propto \exp \left[-\frac{1}{2\sigma_\chi^2} (\mathbf{R} - \mathcal{H}\mathbf{S})^T (\mathbf{R} - \mathcal{H}\mathbf{S}) \right] \quad (25)$$

with respect to the signal \mathbf{S} . This leads to a linear system of equations

$$\mathbf{S} = \underbrace{(\overline{\mathcal{H}}^T \mathcal{H})^{-1}}_{=: \mathcal{L}_{ML}} \overline{\mathcal{H}}^T \mathbf{R}. \quad (26)$$

Using the above assumptions for our model, viz., $\sigma_s = \infty$, $\eta = 0$, and $\xi = 0$, (19) reduces to

$$(\overline{\mathcal{H}} \mathcal{H}^T) \mathcal{L}^T = \overline{\mathcal{H}}. \quad (27)$$

To test whether the two filters are equal, we insert \mathcal{L}_{ML} into (27). Application of the transposition rules shows that with the assumptions we used $\mathcal{L} = \mathcal{L}_{ML}$ and therefore the two strategies are identical; for details we refer to elsewhere (Johnson and Dudgeon 1993; Kay 1993; Sarkar et al. 1981; Rosenfeld 2002).

3.5 Neuronal realization of the model

In this subsection, we translate the general mathematical algorithm of optimal stimulus reconstruction into a concrete neuronal context. We therefore have to verify first whether the assumptions we have made in Sect. 3.2 are fulfilled in neuronal processing. That is, we need to check whether the neuronal quantities and functions of optimal map formation are *self-averaging*. To this end we note on the one hand that firing of neurons is correlated with neuronal input and that neuronal noise can be described by a stochastic process, e.g., a Gaussian one; we will see in a minute why. Our framework can cope with any distribution of neuronal noise as long as the mean is zero. On the other hand the optimal inverse transfer functions $l_i^x(t)$ are *learned* synaptic connections between the maps associated with different modalities and hence reflect properties of the underlying learning process. Effective learning is slow because it needs many independent repetitions. Accordingly time scales for learning and individual realizations of an external signal can be separated. In other words, learning is a self-averaging process where only *averaged* quantities enter by the very nature of the process; (see [Kempler et al. 1999](#)). As mentioned before, quantities and functions within the physical mapping process are self-averaging as well; please see also Appendix B. In

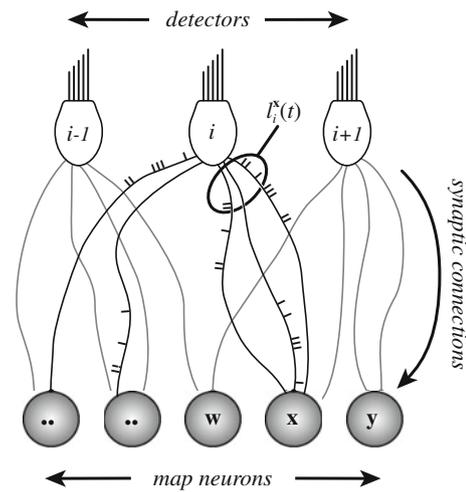


Fig. 4 Neuronal realization of unimodal map formation. Each sensor (here hair cells labeled by i) connects to several map neurons. The map neurons (encoding the location \mathbf{x}) may receive (multiple) connections from each sensor. Each connection has a well-defined strength and temporal delay t . In this way, the transformation $l_i^x(t)$ can be reliably represented in a neuronal network ([Franosch et al. 2003](#))

conclusion, the conditions needed to exploit the mathematical framework as derived in Sect. 3 are fulfilled.

Consequently, we can now translate the inverse transfer functions $l_i^x(t)$ into neuronal hardware. In such an architecture, the actual processing is performed by the synaptic connections between neurons and detectors. Spatial processing is governed by the topographic structure of the network; that is, which detector is connected to which neuron. Temporal processing on the other hand is determined by the distribution of delays within the set of connections. Figure 4 shows an example of such a neuronal setup.

In the above derivation we have already taken into account the discrete character of detectors and the ensuing map through a discrete number of inverse transfer functions. Furthermore, the discrete, “spiky” character of response and reconstruction by the neuronal realization is already taken care of by the noise terms χ_i and λ_i^x . That is, we are left with the temporal discretization of the inverse transfer functions $l_i^x(t)$. This discretization is realized by a sampling procedure where a number of dendrites with appropriate delays is chosen to represent the complete $l_i^x(t)$. It has indeed been shown that a limited number of synaptic connections suffices to sample the time course of $l_i^x(t)$ ([Franosch et al. 2003](#)). Even more so, the map-neuron response is robust with respect to the sampling method of the temporal delays ([Lingenheil 2004](#)) as well.

Consequently, as illustrated by Fig. 4, our unified framework can be implemented by means of a simple feedforward network of excitatory and inhibitory connections in order to form a unimodal map from arbitrary input ([Franosch et al. 2003](#); [Sichert et al. 2006](#); [Lingenheil 2004](#)). It does not,

however, explain how such a connectivity pattern is established in a real biological system. Here the correct synaptic connections have to be *learned*. It has been shown (Franosch et al. 2005a; Friedel and van Hemmen 2008) that a teacher such as the visual system can generate correct synaptic strengths so that a map can indeed develop in other modalities by means of (supervised) STDP; for details see Sect. 5.2. Thanks to the present method we can compare the learned connectivity pattern with the optimal one as given by Eqs. (15) and (20).

A meaningful comparison of the mathematically optimal network architecture with an actual biological setup, though, may not be straightforward. In real biological systems, error reduction as in Eq. 5 to its minimum—that is, realizing the optimal connectivity—may not be possible because of neuronal limitations. The limited neuronal accuracy that results can be included into our framework by reducing the error only below a certain error threshold, which may even vary in space. For instance, the sampling arrays of animal eyes are non-uniform, with different parts of the visual field being sampled with different spatial and spectral resolution (Hughes 1977; Stavenga 2002; Zeil and Hemmi 2006). Such a focus on specific spatio-temporal domains can mathematically be realized by introducing a positive weighting function into the integral of Eq. 5. Accordingly, when reducing the global error below a certain threshold, the areas within the focus of the weight function have to reach a higher level of optimization, i.e., of resolution, than the rest.

As indicated in Sect. 2, the concept of receptive fields is included in our formalism. It may be well to remember, though, that there are two mapping functions (\mathcal{LH}) and \mathcal{L} projecting directly onto the map. Since the rows of these mapping functions contain the information from which areas a specific map neuron receives input, the rows describe the receptive fields.

Taken together, the formalism of optimal map formation is capable to deliver an optimal neuronal connectivity pattern, just as illustrated in Fig. 4, and hereby directly gives a forecast of how the receptive fields are shaped.

4 Exploring the model

In the previous section, we have shown that an optimal connectivity pattern between sensory system and map can be calculated (Fig. 3 and Eq. 15) and that it can be realized neurally (Fig. 4). We now focus on concrete applications of our framework. To this end, we provide a *simple* “recipe” that summarizes the mathematical concepts discussed above. Following this recipe step by step we then demonstrate through two examples how to arrive at an optimal map in both the spatial and the temporal domain.

4.1 A recipe of making maps

To bring to life the mathematical framework of Sect. 3, we present an easy step-by-step “recipe” to find the optimal connectivity in a realistic biological setup:

- First, we derive the transfer function $h_i^{\mathbf{x}}(t)$ that determines the response of the detector i to a stimulus pulse that occurred t time units ago at position \mathbf{x} .
- Next, we calculate the Fourier transform $H_i^{\mathbf{x}}$ of the transfer function $h_i^{\mathbf{x}}(t)$.
- We choose suitable values of τ and σ . In general the noise-to-signal ratio τ can be assumed to be much smaller than 1 for any measurable signal. In contrast, σ needs to be estimated in dependence upon the situation at hand (Franosch et al. 2003, 2005b; Sichert et al. 2006).
- We then calculate the matrix entries M_{ij} as given by Eq. 19 and invert the model matrix \mathcal{M} .
- We multiply the inverted matrix \mathcal{M}^{-1} by the vector $\overline{H_i^{\mathbf{x}}}$ so as to find the input connection strengths $L_i^{\mathbf{x}}$.
- Finally, we calculate the inverse Fourier transform of $L_i^{\mathbf{x}}$ so as to find the connection strengths $l_i^{\mathbf{x}}(t)$.

In the following, we will demonstrate the above recipe through two examples, starting with the derivation of optimal map formation in the spatial domain.

4.2 Spatial example: visual processing

Within the visual system each sensory neuron is basically tuned to a particular spatial position. In mathematical terms, every retinal neuron i receives input from a spatial position \mathbf{x}_i , its preferred position, and neighboring positions within a region determined by resolution ρ . The transfer function corresponding to such a sensory system is

$$h_i^{\mathbf{x}}(t) = \exp\left(-\frac{|\mathbf{x} - \mathbf{x}_i|^2}{2\rho^2}\right)\delta(t), \tag{28}$$

and its Fourier transform reads

$$H_i^{\mathbf{x}} = \exp\left(-\frac{|\mathbf{x} - \mathbf{x}_i|^2}{2\rho^2}\right). \tag{29}$$

Within our exemplary setup we assume that the signal position $\mathbf{x} = (u, v)$ encodes positions $u, v \in [-1/2, 1/2]$. As a reminder, we have rescaled positions so as to make them dimensionless and fit in the square $[-1/2, 1/2]^2$. From the above ansatz (28) and (15) we calculate the matrix components

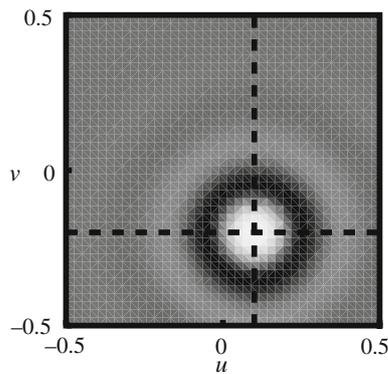


Fig. 5 Spatial receptive field. Connection strengths to a map neuron encoding the position $(u, v) = (0.1, -0.2)$. The sensory neurons are distributed on a 40×40 grid with preferred positions $u, v \in [-1/2, 1/2]$ and a tuning curve width $\rho = 0.9$. We chose $\sigma = 1$ and $\tau = 0$. A clear center-surround receptive field emerges. Receptor neurons that have a preferred position matching that of the map neuron have excitatory connections (white spot). Receptor neurons having a slightly off-set position inhibit the map neuron (dark circle). Neurons with preferred positions far away from the map neuron have connection strength zero (gray)

$$M_{ij} = \sigma^2 \delta_{ij} + (1 + \tau^2) \exp\left(-\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{4\rho^2}\right) \times \left[\operatorname{erf}\left(\frac{u_i + u_j - 1}{2\rho}\right) - \operatorname{erf}\left(\frac{u_i + u_j + 1}{2\rho}\right) \right] \times \left[\operatorname{erf}\left(\frac{v_i + v_j - 1}{2\rho}\right) - \operatorname{erf}\left(\frac{v_i + v_j + 1}{2\rho}\right) \right] \quad (30)$$

where $\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x \exp(-y^2) dy$ is the error function. To find the connection strengths I_i^x , we numerically calculate the model matrix \mathcal{M} for a discretized space and parameters $\sigma = 1$ and $\tau = 0$. With the matrix \mathcal{M} we then determine the connection strengths \mathbf{L} . By an inverse Fourier transformation we can numerically obtain I_i^x for each map position \mathbf{x} as shown in Fig. 5. Here the connections from all receptors to a map neuron i , i.e., its receptive field, are plotted for an arbitrary preferred position $\mathbf{x}_i = (0.1, -0.2)$. Clearly, the receptors encoding the preferred position have strong projections to the map neuron (bright spot in Fig. 5) but, interestingly, the receptors that encode slightly differing locations contribute negatively (dark circle in Fig. 5).

Such a center-surround profile is called “Mexican hat” and is, e.g., realized by lateral inhibition, a well-known phenomenon first described by Mach (1866) in the visual system in 1866. Up to now this mechanism, studied in the mammalian visual system (Wandell 1995; Kandel et al. 2000), has been discovered as well in, for instance, insect vision (Järvilehto 1985), snake infrared vision (Stanford and Hartline 1980; Sichert et al. 2006), electric field detection in electric fish (Shumway 1989), and surface wave detection in the back swimmer (Murphy 1973).

In contrast to many models such as pop-out (Knierim and van Essen 1992) or saliency detection (Itti and Koch 2001; Li 2002) a center-surround receptive field is a natural consequence of our model and thus explains lateral inhibition as optimal for map-formation purposes.

4.3 Temporal example: auditory processing

The framework of map formation is not restricted to spatial dimensions. In the following we demonstrate this by applying our recipe of making maps to the processing of temporal information in the auditory system.

The response of the auditory system to an acoustic signal is usually corrupted by noise and reflections, i.e., echoes. In a first-order approximation an echo is a delayed, weakened repetition of the signal itself. In addition, the signal may well be smeared out. Assuming one single auditory sensor, we choose a simple transfer function incorporating the above characteristics of an echo is

$$h^x(t) = \left[\exp\left(-\frac{t^2}{2\rho^2}\right) + \alpha \exp\left(-\frac{(t-\Delta t)^2}{2\rho^2}\right) \right] \delta(\mathbf{x}). \quad (31)$$

Here we ignore any spatial information, Δt is the delay between signal and echo, ρ the width of the temporal smearing, and α the strength of the echo relative to the signal. The Fourier transform of the transfer function (31) is

$$H(\omega) = \sqrt{2\pi}\rho \left\{ \exp\left(-\frac{\rho^2\omega^2}{2}\right) + \alpha \exp\left[-\frac{1}{2}\omega(2i\Delta t + \rho^2\omega)\right] \right\}. \quad (32)$$

Without the spatial dimension, the model matrix \mathcal{M} now consists of only one entry

$$M = \sigma^2 + 2\pi\rho^2(1 + \tau^2) \left[1 + \alpha^2 + 2\alpha \cos(\omega\Delta t) \right] \times e^{-i\omega\Delta t} e^{-\rho^2\omega^2} \quad (33)$$

and the input connection strengths are given by L . To arrive at the connection strengths $I(t)$ we need to numerically perform an inverse Fourier transform. The result of a typical parameter set is shown in Fig. 6 where the connection strengths are plotted as a function of the delay. This temporal receptive field can be segregated into two functional subunits. First, similarly to the center-surround profile in the last section, it consists of half a Mexican hat that transmits and sharpens the original signal. Second, it features an inverse Mexican hat that, just as expected, suppresses the echo.

Altogether, the application of our recipe to echo suppression can indeed make valuable predictions for biological systems dealing with echoes. In a more detailed study to be published elsewhere we show how the mathematical concept of optimality explains the actual physiology of neuronal echo suppression.

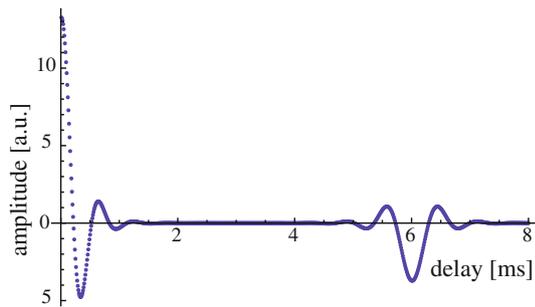


Fig. 6 Temporal receptive field. Echo-canceling inverse transfer function against delay. In a way similar to the spatial case, we find Mexican hat profiles in time. The half Mexican hat at 0 ms transmits and sharpens the signal, the upside-down Mexican hat at 6 ms suppresses the echo. Parameters are $\alpha = 0.5$, $\rho = 0.2$ ms, $\Delta t = 6$ ms, $\sigma = 0.01$, and $\tau = 0$

5 Multimodality

In deriving our framework in Sect. 3, we have assumed all sensory maps to be purely monosensory. In spite of this assumption, animals and humans perceive their environment through *several*, multimodal, sensory systems. To fully access the complete information of all monosensory maps, their information therefore has to be combined. Physiological and behavioral experiments indeed show that the monosensory perceptions are not independent but mutually interact with each other (Shams et al. 2000; Ernst et al. 2000; Bresciani et al. 2005; Alais and Burr 2004). This interaction can then lead to the formation of *multisensory* maps, i.e., maps receiving input from more than one sensory system (Morgan et al. 2008; Kaas and Collins 2004).

A number of brain areas, such as the midbrain in mammals, more precisely the SC (Stein and Meredith 1993), or even higher brain areas, such as the Anterior Ectosylvian Sulcus (AES) (Carriere et al. 2007), contain clearly distinct monosensory, as well as multisensory neurons. Since the perceptual and behavioral role of higher brain areas nevertheless remains vague we use the well-studied SC as an example for multimodal interaction. The SC features a layered organization of spatial maps from all sensory systems that dispose of topographic, map-like information (such as vision, but not olfaction) (Stein and Meredith 1993). All these maps, uni- or multisensory, are mutually aligned (Stein and Meredith 1993; King 1999; Calvert et al. 2004) and thus, provide a *common* reference system of sensory space.

Bearing in mind the above we now turn to the functional possibilities that arise from multimodal interaction.

5.1 Multimodal interaction

In general, we can distinguish two categories of multimodal interaction: *integration* and *pooling* of information.

5.1.1 Integration

Congruent spatial information from different sensory systems can be integrated into a single merged and, hence, multimodal map. Such an integrated map, as compared to unimodal information processing, features increased information reliability and saliency as well as an improved sensitivity in both space and time (Krauzlis et al. 1997, 2004; Hafed et al. 2008; Rowland et al. 2007a). For example, if visual and auditory sensory system both register a signal, e.g., “brown ahead” and “barking ahead”, it is very probable that the signal corresponds to an actual object rather than to a sensory artefact. At the same time, the integrated signal will be stronger and allows for faster reactions (e.g., “escape!”). In some cases an integrated signal is even *optimal* (Gu et al. 2008; Morgan et al. 2008).

More general neuronal models describing multimodal integration and based on statistical methods have been presented elsewhere (Denève et al. 1999, 2001; Denève and Pouget 2004). Concrete theoretical models of multimodal integration within the SC have been developed as well (Anastasio et al. 2000; Patton et al. 2002; Rowland et al. 2007b,c; Magosso et al. 2008; Ursino et al. 2009).

5.1.2 Pooling

Not only can the monosensory maps be merged into a more reliable multisensory map, but the diverse information, thus, signal characteristics within the monosensory maps can be accessed simultaneously as well. This simultaneous accessing is only possible since all monosensory maps are aligned and consequently space-time can serve to link the different modalities. Consequently, an object at one specific position can be identified and characterized in order to select motor responses in a complex environment. For example, a rattle snake may detect spatial coherent activity in its visual and/or infrared map. Only if the encoded object is visible *and* warm will it be identified as a living prey object. If it is visible and *not* warm the snake will discard the information. Experimental evidence for such a pooling of information is provided by neuronal AND and OR processing steps for the combination of visual and infrared map (Newman and Hartline 1981, 1982). These prominent examples of pooling in the SC could enable target selection and thus ensure appropriate motor commands in a complex environment.

Despite increased reliability of an integrated map, its individual input streams cannot be distinguished anymore. That is, the information of which monosensory map has determined the position is lost. Within the above example the multimodal map may indicate a multimodal event ahead, but the triggering modality, that is, visual, auditory, or yet another modality remains unresolved.

In summary, integration of information allows for a reliable spatial determination of an object, the key task of *object formation*. On the other hand, pooling of information assures an access to the details of an object necessary for *object identification*. Switching between integration and pooling corresponds to a switch between parallel and serial data processing to best fit different tasks.

To enable efficient multimodal interaction such as integration and pooling, alignment of the different mono- and multisensory maps is of crucial importance. Only then can a multimodal stimulus at a specific spatial location be identified. An alignment of sensory maps, however, is not present at birth and must be *learned* (van Hemmen 2002; Stein and Stanford 2008; Knudsen 2002), as discussed in the next section.

5.2 Development of multisensory space

We now discuss the question of how sensory maps can be aligned, i.e., how a common multimodal space can evolve. An obvious solution to such an alignment process would be the existence of one dominant modality as reference for all other modalities (Knudsen and Brainard 1991; Knudsen 2002). This reference map would then automatically lead to modifications of all other maps.

And indeed, experimental and physiological studies have shown that, in many animals, destruction or disturbance of the visual pathway leads to disorganized and abnormal sensory maps in non-visual modalities. These findings have been obtained in hamster (Mooney et al. 1987), cat (Wallace and Stein 2007; Wallace et al. 2004), clawed frog (Claas 1994), ferret (King et al. 1988), barn owl (Knudsen and Knudsen 1985), and in snakes (Grace et al. 2001). Psychophysical experiments with congenitally blind and normally sighted humans have shown that visual input early in life is necessary for multimodal interaction to occur (Hötting et al. 2004; Putzar et al. 2007; Röder et al. 2004). Consequently, vision seems to serve as “teacher” for non-visual modalities.

A plausible argument supporting the idea of vision as a teacher input is the intrinsic topographic order of the retina. It is known that layers of neurons can self-organize into topographic maps, provided that initially a small set of correctly organized neurons exists (Willshaw and Malsburg 1976). For a review the reader is referred to Udin and Fawcett (1988). This together with the subsequent development of layers in the visual cortex (for mice, see Jiang et al. 2007) may allow the intrinsic topography of the retina to step-by-step dictate the organization and alignment of higher visual and, potentially, also multimodal maps.

The general mechanism facilitating such an alignment of maps is spike-timing-dependent plasticity (STDP) (Bi and Poo 1998, 2001; Dan and Poo 2004; Gerstner et al. 1996; van Hemmen 2001; Kempter et al. 1999; Markram et al. 1997; Song et al. 2000; Zhang et al. 1998). An example

where the alignment has been studied in detail, both experimentally and theoretically, is audio-visual integration within the SC of the barn owl. Here experiments (Hyde and Knudsen 2001; Knudsen and Brainard 1991) have shown that the auditory map follows systematic changes within the visual input. Although the precise nature of this teaching signal has not been clarified experimentally, *selective neuronal disinhibition*, or gating, seems to play a key role (Gutfreund et al. 2002; Winkowski and Knudsen 2006). Theoretical studies have confirmed that excitatory and inhibitory teaching input can account for proper map alignment and thus development of multimodal space (Friedel and van Hemmen 2008; Davison and Frégnac 2006). It is, however, only by *inhibitory* teaching input that an already aligned map can be re-aligned later on (Friedel and van Hemmen 2008).

In summary, the above studies support the idea of vision as teacher modality to align other monosensory maps, but there are contradicting findings as well. We can summarize these findings into two major points. First, vision is not needed at all as teacher input for the learning process of sensory maps. Second, vision shows plasticity as it is influenced by other modalities and as it improves during development. Concerning the first point, it has been shown both theoretically and experimentally that, although imprecise, a map of *azimuthal* sound location can be learned *without* any visual input (Kempter et al. 2001; Knudsen and Brainard 1991) though admittedly on a genetically determined substrate. In addition, non-visual modalities can influence each other as well, e.g., audition can influence haptics (Bresciani et al. 2005). Moreover, somatosensory receptive fields already shrink in a postnatal phase when only auditory, but no visual neurons are present (Wallace and Stein 1997; Wallace et al. 2006). For the second aspect, behavioral and psychophysical studies show that visual perception can even be influenced by other modalities such as haptics (Ernst et al. 2000) or audition (Shams et al. 2000; Frens and van Opstal 1998; Steenken et al. 2008). More importantly, vision itself can improve, respectively, sharpen as found in the visual system of young cats (Wallace and Stein 1997; Wallace et al. 2006).

Altogether the experimental and theoretical findings we have presented above put into question the current picture of vision-guided map alignment (King 2009). Wallace and Stein (1997) have pointed out that the development of different modalities starts in parallel and in temporal coincidence with the appearance of multimodal integration. They hereby suggest a common mechanism driving both map development and multimodal integration.

6 Discussion

In summary, we have started this review by defining the ‘golden three’ of sensory processing: physical mapping,

optimal map formation, and multimodal integration, cf. Fig. 1. Based on these concepts we have formulated a mathematical framework that is able to quantify how a signal stimulates a detector and how the detector response is processed so as to lead to a “reconstruction” of the original signal. In the context of neuronal information processing our framework unifies long-standing concepts of stochastically optimal signal processing by seamless transitions to maximum likelihood, Wiener filter, and the Tikhonov-Miller regularization. Moreover, the present model extends these established techniques by linking its parameters to easily accessible experimental quantities. By the use of the mathematical principle of *stochastic optimality* we have arrived at a discrete and optimal representation of the outside world—a map.

Most importantly, we have then translated our mathematical setup into neuronal architecture. That is, by a discretization in space-time of our mathematical model we have derived synaptic connection patterns between detector and map neurons. To illustrate the relation to real biological settings, we have provided a step-by-step recipe that offers the possibility of solving concrete problems of map formation. Two such problems, a spatial and a temporal one, have been treated as illustrative examples and have been shown to reproduce experimental findings such as receptive fields. The generality of the method of optimal map formation can now be tested to model and analyze experimental results. In particular, the measurement of internal connections, for instance in the SC, as well as receptive fields for specific sensory systems would provide a possibility to experimentally access the inverse transfer functions as defined in Sect. 3.1.

On the basis of our understanding of monosensory map formation, we have proceeded to multimodal interaction and the development of multisensory space. Here the concept of neuronal maps as compared to single neuron effects can deliver new perspectives on *multimodal* interaction, viz., *integration* and *pooling* of information. While integration of information allows for a reliable spatial localization of an object, pooling of information assures access to the details of an object. The latter observation would answer the question as to why we find multiple maps in the SC instead of a single multimodal map.

Finally, we have emphasized the importance of proper map alignment for multimodal interaction. We have shown that STDP learning algorithms with an inhibitory teacher signal can account for both, initial alignment and even subsequent re-alignment of maps. Further experimental studies on inhibitory teacher input, e.g., within the SC, are nevertheless needed to clarify the precise role of inhibition in the alignment process. In addition, such experiments could answer the crucial question of which sensory systems determine the formation of multimodal space.

In other words, we need more experimental evidence in order to understand how multimodal interaction is realized

and established on an anatomical level. For example, through which mechanism could a collection of aligned maps allow the pooling of information? Does such a mechanism also include feature selection in a common sensory space? How does multimodal interaction of maps contribute to the formation of such a common sensory space? And to what extent does such a finding contradict the current picture of vision guiding the map alignment? Furthermore, it is important to test the generality of the mathematical framework that we have so far substantiated only by findings in the SC through additional applications to other areas of the brain, such as the AES, a well-defined multisensory cortical area in cat.

In other words, we are able to explain fundamental ingredients of the ‘golden three’ of sensory processing. Through the present work we provide not only a synopsis of the current state-of-knowledge but also helpful tools for verifiable predictions of sensory processing by means of maps.

Acknowledgments The authors thank Jochen Zeil and André Longtin for helpful comments on earlier versions of the manuscript and the BCCN-Munich for funding.

Appendix

A Nonlinearities in information processing

The presented model assumes a linear relation between stimulus and detector response. For a number of sensory systems, however, we find non-linearities in the mapping process. First, the transfer function h can be a non-linear function \hat{h} . Second, the neuronal detector response can be nonlinear, typically logarithmic (Krueger 1989; Norwich and Wong 1997; Laming 1997; Johnson et al. 2002; Copelli et al. 2002). In case of a logarithmic response and a nonlinear transfer function the biological detector response \tilde{r} has to be rewritten from (1) as

$$\tilde{r}_i(t) = \log \left[\int_{\text{all space}} d\mathbf{x} \int_{-\infty}^t d\tau s^{\mathbf{x}}(\tau) \tilde{h}_i^{\mathbf{x}}(t - \tau) \right]. \quad (34)$$

To apply our model we first have to incorporate an additional computational step canceling the logarithm. In a biological system this can be realized, e.g., by neurons with exponential firing behavior. Assuming such a neuronal step $\tilde{r}_i(t)$ as in (34) reduces to (1) with a nonlinear \hat{h} . We can linearize a nonlinear transfer function by a redefinition of the signal $s \rightarrow \tilde{s}$. That is, we identify appropriate characteristics of the stimulus that are linearly related to r . For example, instead of looking at the heat distribution $T(\mathbf{x}, t)$ we can consider the intensity distribution of the corresponding radiation $\sim T^4(\mathbf{x}, t)$ due to the Stefan-Boltzmann Law. In this way, a *reasonable* redefinition of detector response and signal can allow for an optimal linear stimulus reconstruction.

B Self-averaging and laws of large numbers

Why can physical but noisy input quantities be expected to be self-averaging? To understand this valuable property we assume a detector receives input signals f_i with $1 \leq i \leq N$ as a sum $\eta_N \sum_{i=1}^N f_i$ where η_N is a scaling factor. For the sake of convenience we take $\eta_N = 1/N$ to get a decent scaling behavior as $N \rightarrow \infty$ but for N large but finite there is ample choice. Moreover, we assume the f_i to be stochastic random variables with mean a_i and finite variance. Finally, if the $1 \leq i \leq N$ represent for example different positions in space, in biological reality the stochastic correlation between positions that are far apart is small. Hence we write $f_i = a_i + \phi_i$ where we start by taking the noise terms ϕ_i , cf. Fig. 3, to be independent random variables that by construction all have zero mean.

For inputs of the sum form

$$\sum_{i=1}^N f_i = \left(\sum_{i=1}^N a_i \right) + \left(\sum_{i=1}^N \phi_i \right) =: A_N + \Phi_N \tag{35}$$

we see that A_N is some deterministic number and regarding Φ_N we can immediately invoke the strong law of large numbers (see below), so as to conclude that, as $N \rightarrow \infty$, we find $\eta_N \Phi_N \rightarrow 0$ whatever specific realization of the $\{\phi_i\}$ we meet. The latter circumstance is exactly what we need in practical work since we never know the realization until it is all over. The strong law of large numbers guarantees that Φ_N vanishes as N becomes large.

The only, minor, drawback of all this is twofold. First, in reality the ϕ_i are never perfectly independent. Nevertheless, as long as correlations fall off fast enough as the distance $|i - j|$ becomes large, the strong law of large numbers still holds; the adjective “strong” indicates that $\Phi_N \rightarrow 0$ as $N \rightarrow \infty$ for (almost) all realizations of the $\{\phi_i\}$; here “almost” is an epithet nearly everybody can forget about. Second, in practical work N is and remains finite. Then the central limit theorem in the paragraph below tells us that for independent ϕ_i , whatever their distribution provided the second moment $\langle \phi_i^2 \rangle$ is finite, and N large

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \phi_i \tag{36}$$

has a Gaussian distribution with mean zero; its standard deviation gives information about the width of the Gaussian. The same holds true for weakly dependent ϕ_i . The weak dependence being a consequence of biophysical reality, we are done.

To finish the argument, we need to discuss briefly the various assumptions that we have used above to clarify the issue of self-averaging. The discussion below is practically identical with Appendix A of van Hemmen (2001). The textbook by Durrett (2004) is a general, though advanced, background

for various formulations of the laws of large numbers listed below. To begin with, let us suppose that the f_i are independent, identically distributed random variables with mean zero. If the mean $\langle f \rangle$ is nonzero, we subtract it and consider $f_i := f_i - \langle f \rangle$ instead. There is no harm in taking the f_i to be real variables. Furthermore, we require the second moment $\langle f^2 \rangle$ to be finite. By Cauchy-Schwarz, $\langle |f| \rangle \leq \langle f^2 \rangle^{1/2} < \infty$, and the variance $\sigma^2 := \langle (f - \langle f \rangle)^2 \rangle$ is finite too. Let

$$S_n = \sum_{i=1}^n f_i \tag{37}$$

be the sum of the random variables f_i . Then the following theorems hold:

- Strong law of large numbers: $\lim_{n \rightarrow \infty} n^{-1} S_n = 0$ with probability 1. Since the f_i are sampled from a probability distribution, this means that, as $n \rightarrow \infty$, the configurations where the above equality does not hold have probability zero. In plain English, they do not occur. One also says that the above equality holds ‘almost surely’ (a.s.). All that is needed is $\langle |f| \rangle < \infty$.
- Central limit theorem: As $n \rightarrow \infty$, $n^{-1/2} S_n$ has a Gaussian distribution with mean zero and variance σ^2 .

Etemadi (1981) has given an “elementary” proof of the strong law of large numbers for pairwise independent, identically distributed random variables under the minimal condition $\langle |f| \rangle < \infty$. Slick proofs (occasionally with some extra conditions, say, finite fourth moment) have been given by Lamperti (1966). Breiman (1968) treats the two theorems in their full generality.

The theorems also hold for independent, not necessarily identically distributed random variables (Lamperti 1966; Breiman 1968; Gnedenko and Kolmogorov 1968). They even allow a weak dependence. For example, let $R_{ij} := \langle f_i f_j \rangle - \langle f_i \rangle \langle f_j \rangle$, and suppose the f_i do not have too wide a distribution, e.g., $\sup_i |R_{ii}| < \infty$. Then the strong law of large numbers holds (Gnedenko and Kolmogorov 1968, p. 265; Halmos 1956), provided $R_{ij} \rightarrow 0$ as $|i - j| \rightarrow \infty$; that is to say, the correlations between f_i and f_j should not have too long a range. For the central limit theorem to hold, trickier conditions are required, e.g., stationarity of the sequence f_1, f_2, \dots , and some kind of mixing (Durrett 2004, Ch. 7.7c) so that, say, $\sum_j |R_{ij}| < \infty$. Then the variance of the Gaussian limit distribution is given by

$$\sigma^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{ij} \langle f_i f_k \rangle = \langle f_1^2 \rangle + 2 \sum_{k=2}^{\infty} \langle f_1 f_k \rangle. \tag{38}$$

Dropping stationarity, the reader may consult Scott (1973) for an advanced account.

C Remaining derivation steps leading to (15)

In the following we elaborate some steps skipped in the derivation of (15) in the main text. In doing so we take advantage of ideas due to the calculus of variations (Jost and Li-Jost 1998). We therefore start from (6) as a condition to minimize the expectation value of the quadratic error with respect to the optimal reverse transfer functions, the connection strengths $l_j^x(t)$. This leads to

$$\frac{\partial \left\langle [s^x(t') - \hat{s}^x(t')]^2 \right\rangle}{\partial l_j^x(t')} = 0 \quad \text{for every } j.$$

$$\Leftrightarrow \left\langle [s^x(t') - \hat{s}^x(t')] \frac{\partial \hat{s}^x(t')}{\partial l_j^x(t')} \right\rangle = 0. \tag{39}$$

In order to solve (39), we expand the estimate $\hat{s}^x(t)$ using Eqs. (3) and (4) from the main text giving

$$\hat{s}^x = \sum_i \{ \chi_i \star l_i^x + \chi_i \star \lambda_i^x + \int dy [s^y \star h_i^y \star l_i^x + s^y \star h_i^y \star \lambda_i^x + s^y \star \eta_i^y \star l_i^x + s^y \star \eta_i^y \star \lambda_i^x + \xi^y \star h_i^y \star l_i^x + \xi^y \star h_i^y \star \lambda_i^x + \xi^y \star \eta_i^y \star l_i^x + \xi^y \star \eta_i^y \star \lambda_i^x] \}. \tag{40}$$

The variation of \hat{s} in (40) leads to

$$\frac{\partial \hat{s}^x(t')}{\partial l_j^x(t')} = \left[\chi_j + \int dy \left(s^y \star h_j^y + s^y \star \eta_j^y + \xi^y \star h_j^y + \xi^y \star \eta_j^y \right) \right] (0). \tag{41}$$

As before, we assume on the one hand that all noise terms as well as the expectation of the input are stochastically independent of each other. On the other hand we use the fact that all noise terms have zero mean and the signal is self-averaging. With these two assumptions, the expectation values $\langle s \partial \hat{s} / \partial l \rangle$ and $\langle \hat{s} \partial \hat{s} / \partial l \rangle$ from (39) can be written

$$\left\langle s^x(t') \frac{\partial \hat{s}^x(t')}{\partial l_j^x(t')} \right\rangle = \int dy \left\langle s^x(t') (s^y \star h_j^y)(0) \right\rangle \tag{42}$$

and

$$\left\langle \hat{s}^x(t') \frac{\partial \hat{s}^x(t')}{\partial l_j^x(t')} \right\rangle = \sum_i \{ \langle (\chi_i \star l_i^x)(t') \chi_j(0) \rangle + \int dy dy' \left[\langle (s^y \star h_i^y \star l_i^x)(t') (s^{y'} \star h_j^{y'})(0) \rangle + \langle (s^y \star \eta_i^y \star l_i^x)(t') (s^{y'} \star \eta_j^{y'})(0) \rangle \right] \}$$

$$+ \left\langle (\xi^y \star h_i^y \star l_i^x)(t') (\xi^{y'} \star h_j^{y'})(0) \right\rangle + \left\langle (\xi^y \star \eta_i^y \star l_i^x)(t') (\xi^{y'} \star \eta_j^{y'})(0) \right\rangle \}. \tag{43}$$

In the next step we substitute the correlation terms as given in (9) of the main text. To illustrate the calculations, which simplify (42) and (43), we analyze two isolated terms from (43) as an example. The other terms are treated in a similar way. We first simplify

$$\sum_i \int dy dy' \left\langle (s^y \star h_i^y \star l_i^x)(t') (s^{y'} \star h_j^{y'})(0) \right\rangle = \sum_i \int dy dy' d\tau d\tau' d\tau'' \left\langle s^y(t' - \tau - \tau') h_i^y(\tau) l_i^x(\tau) s^{y'}(-\tau'') h_j^{y'}(\tau'') \right\rangle. \tag{44}$$

Exploiting the correlation assumptions this expression becomes

$$\mu_s^2 \sum_i \int dy dy' d\tau d\tau' d\tau'' \delta(y - y') \delta(t' - \tau - \tau' + \tau'') h_i^y(\tau') l_i^x(\tau) h_j^{y'}(\tau'') = \mu_s^2 \sum_i \int dy d\tau d\tau'' h_i^y(t' - \tau + \tau'') l_i^x(\tau) h_j^{y'}(\tau'') = \mu_s^2 \sum_i \int dy ((h_i^y \star l_i^x) \circ h_j^{y'})(t') \tag{45}$$

with the open circle \circ denoting the autocorrelation integral defined in (12). We focus on the second term in the right-hand side of (43). That is,

$$\sum_i \int dy dy' \left\langle (s^y \star \eta_i^y \star l_i^x)(t') (s^{y'} \star \eta_j^{y'})(0) \right\rangle = \sum_i \int dy dy' d\tau d\tau' d\tau'' \left\langle s^y(t' - \tau - \tau') \eta_i^y(\tau) l_i^x(\tau) s^{y'}(-\tau'') \eta_j^{y'}(\tau'') \right\rangle, \tag{46}$$

which simplifies to

$$\mu_s^2 \sigma_\eta^2 \sum_i \int_{\substack{|y| < y^{\max} \\ 0 < \tau' < t^{\max}}} dy dy' d\tau d\tau' d\tau'' \delta(y - y') \delta(t' - \tau - \tau' + \tau'') \delta_{ij} \delta(y - y') \delta(t' - \tau'') l_i^x(\tau) = \mu_s^2 \sigma_\eta^2 \int_{\substack{|y| < y^{\max} \\ 0 < \tau' < t^{\max}}} dy d\tau d\tau' \delta(-\tau + t') l_j^x(\tau) = \mu_s^2 \sigma_\eta^2 \int_{\substack{|y| < y^{\max} \\ 0 < \tau' < t^{\max}}} dy d\tau' l_j^x(t'). \tag{47}$$

Altogether the final expressions for the expectation values become

$$\left\langle s^{\mathbf{x}}(t') \frac{\partial \hat{s}^{\mathbf{x}}(t')}{\partial l_j^{\mathbf{x}}(t')} \right\rangle = \mu_s^2 h_j^{\mathbf{x}}(-t') \tag{48}$$

and

$$\begin{aligned} \left\langle \hat{s}^{\mathbf{x}}(t') \frac{\partial \hat{s}^{\mathbf{x}}(t')}{\partial l_j^{\mathbf{x}}(t')} \right\rangle &= \sigma_\chi^2 l_j^{\mathbf{x}}(t') + \sigma_\eta^2 (\mu_s^2 + \sigma_\xi^2) \\ &\int_{\substack{|\mathbf{y}| < y^{\max} \\ 0 < \tau < t^{\max}}} d\mathbf{y} d\tau l_j^{\mathbf{x}}(t') + (\mu_s^2 + \sigma_\xi^2) \\ &\times \sum_i \int d\mathbf{y} \left[(h_i^{\mathbf{y}} \star l_i^{\mathbf{x}}) \circ h_j^{\mathbf{y}} \right](t'). \end{aligned} \tag{49}$$

Equation 39 therefore transforms into

$$\begin{aligned} l_j^{\mathbf{x}}(t) &\left[\sigma_\chi^2 + (\mu_s^2 + \sigma_\xi^2) \int_{\substack{|\mathbf{y}| < y^{\max} \\ 0 < \tau < t^{\max}}} d\mathbf{y} d\tau \sigma_\eta^2 \right] + (\mu_s^2 + \sigma_\xi^2) \\ &\sum_i \int d\mathbf{y} \left[(h_i^{\mathbf{y}} \star l_i^{\mathbf{x}}) \circ h_j^{\mathbf{y}} \right](-t) = \mu_s^2 h_j^{\mathbf{x}}(-t). \end{aligned} \tag{50}$$

Inserting the parameter σ and τ and applying a Fourier transformation finally leads to (15).

In order to test whether the extremum is indeed a minimum, we have to calculate the second variation, which reads

$$\begin{aligned} &\frac{\partial^2 \left[[s^{\mathbf{x}}(t') - \hat{s}^{\mathbf{x}}(t')]^2 \right]}{\left(\partial l_j^{\mathbf{x}}(t') \right)^2} \\ &= 2 \frac{\partial}{\partial l_j^{\mathbf{x}}(t')} \left[\left\langle s^{\mathbf{x}}(t') \frac{\partial \hat{s}^{\mathbf{x}}(t')}{\partial l_j^{\mathbf{x}}(t')} \right\rangle - \left\langle \hat{s}^{\mathbf{x}}(t') \frac{\partial \hat{s}^{\mathbf{x}}(t')}{\partial l_j^{\mathbf{x}}(t')} \right\rangle \right] \\ &= 0 + 2 \left[\sigma_\chi^2 + \sigma_\eta^2 (\mu_s^2 + \sigma_\xi^2) \int_{\substack{|\mathbf{y}| < y^{\max} \\ 0 < \tau < t^{\max}}} d\mathbf{y} d\tau \right. \\ &\quad \left. + (\mu_s^2 + \sigma_\xi^2) \int d\mathbf{y} \int d\tau \left[h_j^{\mathbf{y}}(\tau) \right]^2 \right]. \end{aligned} \tag{51}$$

Since the squares are positive, so is the second derivative and thus the extremum is a minimum.

D Gaussian blurred signal

In this subsection, we present an equation equivalent to (15) of the main text so as to derive an expression for a Gaussian

blurred signal. As in (10), a realistic signal would fulfill some kind of Gaussian relation for the expectation value

$$\left\langle s^{\mathbf{x}}(t) s^{\mathbf{x}'}(t') \right\rangle = A \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2\sigma_x^2}\right) \exp\left(-\frac{|t - t'|^2}{2\sigma_t^2}\right). \tag{52}$$

For this case we can, analogously to Appendix C, derive an equation like (15). Since for the signal the Gaussian correlations, however, replace the delta functions, e.g., in (47), integrals over space and time cannot be evaluated directly. Instead they can only be restricted to the region where the Gaussian is non-negligible. Denoting these temporal and spatial limits by ϵ_t and ϵ we can derive the analogue to (15), viz.,

$$\begin{aligned} &\int d\epsilon d\epsilon_t A \exp\left(-\frac{|\epsilon|^2}{2\sigma_x^2}\right) \exp\left(-\frac{|\epsilon_t|^2}{2\sigma_t^2}\right) \\ &\quad h_j^{\mathbf{x}+\epsilon}(t'' + \epsilon_t) \\ &= \sigma_\chi^2 l_j^{\mathbf{x}}(t'') + \sigma_\eta^2 \sigma_\xi^2 \int_{\substack{|\mathbf{y}| < y^{\max} \\ 0 < \tau < t^{\max}}} d\mathbf{y} d\tau l_j^{\mathbf{x}}(t'') \\ &\quad + \sigma_\xi^2 \sum_i \int d\mathbf{y} \left[(h_i^{\mathbf{y}} \star l_i^{\mathbf{x}}) \circ h_j^{\mathbf{y}} \right](t'') \\ &\quad + \sigma_\eta^2 \int_{\substack{|\mathbf{y}| < y^{\max} \\ 0 < \tau < t^{\max}}} d\mathbf{y} d\tau A \exp\left(-\frac{\epsilon_t^2}{2\sigma_t^2}\right) \\ &\quad l_j^{\mathbf{x}}(t'' + \epsilon_t) \\ &\quad + \sum_i \int d\mathbf{y} d\epsilon d\epsilon_t A \exp\left(-\frac{|\epsilon|^2}{2\sigma_x^2}\right) \exp\left(-\frac{\epsilon_t^2}{2\sigma_t^2}\right) \\ &\quad \left((h_i^{\mathbf{y}} \star l_i^{\mathbf{x}}) \circ h_j^{\mathbf{y}+\epsilon} \right)(t'' + \epsilon_t). \end{aligned} \tag{53}$$

The effect of the additional remaining spatio-temporal integrals as compared to (15) is a smoothening of the final reconstruction. Not only is the value at a specific point in space and time (\mathbf{y}, t'') taken into account but neighboring points in a nearby area surrounding it are included as well.

E Pseudoinverse

A pseudoinverse matrix \mathcal{B} of an arbitrary $m \times n$ matrix \mathcal{A} with its elements $\in \mathbb{R}, \mathbb{C}$ is a generalization of the inverse matrix. Following Moore (1920); Penrose (1955) the pseudoinverse \mathcal{B} is defined by four equations,

$$\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}, \tag{54a}$$

$$\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}, \tag{54b}$$

$$\overline{(\mathcal{A}\mathcal{B})}^T = \mathcal{A}\mathcal{B}, \tag{54c}$$

$$\overline{(\mathcal{B}\mathcal{A})}^T = \mathcal{B}\mathcal{A}, \tag{54d}$$

where the overbar denotes the complex conjugation. The pseudoinverse is used for a matrix \mathcal{A} that is of incomplete rank and therefore cannot be inverted directly but only due to auxiliary constructions (54 a, b). The last two equations (54 c, d) tell us that the product of matrix \mathcal{A} with its pseudoinverse \mathcal{B} is Hermitian.

One of the most famous applications of the pseudoinverse is to calculate the least-square solution of a system of linear equations.

$$\mathcal{A}\mathbf{x} = \mathbf{b} \rightarrow \hat{\mathbf{x}} = \mathcal{B}\mathbf{b} \quad (55)$$

Given a matrix \mathcal{A} and a vector \mathbf{b} , the above solution $\hat{\mathbf{x}} = \mathcal{B}\mathbf{b}$ minimizes the Euclidean norm $\|\mathcal{A}\hat{\mathbf{x}} - \mathbf{b}\|$. If the inverse matrix exists, the pseudoinverse reduces to the normal inverse matrix. For a general review and applications the reader is referred to, e.g., (Albert 1972).

References

- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14:257–262
- Albert A (1972) Regression and the Moore-Penrose pseudoinverse. Academic, New York
- Anastasio TJ, Patton PE, Belkacem-Boussaid K (2000) Using bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Comput* 12:1165–1187
- Bartels M, Münz H, Claas B (1990) Representation of the lateral line and electrosensory systems in the midbrain of the axolotl, *Ambystoma mexicanum*. *J Comp Physiol A* 167:347–356
- Ben-Israel A, Greville TNE (2003) Generalized inverses: theory and applications, 2nd edn. Springer, New York, NY
- Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J Neurosci* 18:10,464–10,472
- Bi GQ, Poo MM (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu Rev Neurosci* 24:139–166
- Brand A, Behrend O, Marquardt T, McAlpine D, Grothe B (2002) Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417:543–547
- Breiman L (1968) Probability. Addison-Wesley, Reading, MA
- Bresciani JP, Ernst MO, Drewing K, Bouyer G, Maury V, Kheddar A (2005) Feeling what you hear: auditory signals can modulate tactile taps perception. *Exp Brain Res* 162:172–180
- Calvert G, Spence C, Stein BE (2004) The handbook of multisensory processes. MIT, Cambridge, MA
- Carr CE, Konishi M (1988) Axonal delay lines for time measurement in the owl's brainstem. *Proc Natl Acad Sci USA* 85:8311–8315
- Carr CE, Konishi M (1990) A circuit for detection of interaural time differences in the brain stem of the barn owl. *J Neurosci* 10:3227–3246
- Carr CE, Maler L, Sas E (1982) Peripheral organization and central projections of the electrosensory nerves in gymnotiform fish. *J Comp Neurol* 211:139–153
- Carriere BN, Royal DW, Perrault TJ, Morrison SP, Vaughan JW, Stein BE, Wallace MT (2007) Visual deprivation alters the development of cortical multisensory integration. *J Neurophysiol* 98:2858–2867
- Claas B (1994) Removal of eyes in early larval stages alters the response of the clawed toad, *Xenopus laevis*, to surface waves. *Physiol Behav* 56:423–428
- Clegg JC (1968) Calculus of variations. Oliver and Boyd, Edinburgh
- Copelli M, Roque AC, Oliveira RF, Kinouchi O (2002) Physics of psychophysics: Stevens and Weber-Fechner laws are transfer functions of excitable media. *Phys Rev E* 65:60,901
- Dan Y, Poo MM (2004) Spike timing-dependent plasticity of neural circuits. *Neuron* 44:23–30
- Davison AP, Frégnac Y (2006) Learning cross-modal spatial transformations through spike timing-dependent plasticity. *J Neurosci* 26:5604–5615
- Denève S, Pouget A (2004) Bayesian multisensory integration and cross-modal spatial links. *J Physiol (Paris)* 98:249–258
- Denève S, Latham PE, Pouget A (1999) Reading population codes: a neural implementation of ideal observers. *Nat Neurosci* 2(8):740–745
- Denève S, Latham PE, Pouget A (2001) Efficient computation and cue integration with noisy population codes. *Nat Neurosci* 4(8):826–831
- Dräger UC, Hubel DH (1975) Responses to visual stimulation and relationship between visual, auditory, and somatosensory inputs in mouse superior colliculus. *J Neurophysiol* 38:690–713
- Durrett R (2004) Probability: theory and examples, 3rd edn. Duxbury Press, Belmont, MA
- Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitboeck HJ (1988) Coherent oscillations: a mechanism of feature linking in the visual cortex. *Biol Cybern* 60:121–130
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433
- Ernst MO, Banks MS, Bühlhoff H (2000) Touch can change visual slant perception. *Nat Neurosci* 3:69–73
- Etemadi N (1981) An elementary proof of the strong law of large numbers. *Z Wahrscheinlichkeitstheorie verw Geb* 55:119–122
- Faisal AA, Selen LPJ, Wolpert DM (2008) Noise in the nervous system. *Nat Rev Neurosci* 9:292–303
- Franosch JMP, Sobotka MC, Elepfandt A, van Hemmen JL (2003) Minimal model of prey localization through the lateral-line system. *Phys Rev Lett* 91:158101
- Franosch JMP, Lingenheil M, van Hemmen JL (2005a) How a frog can learn what is where in the dark. *Phys Rev Lett* 95:78,106
- Franosch JMP, Sichert AB, Suttner MD, van Hemmen JL (2005b) Estimating position and velocity of a submerged moving object by the clawed frog *Xenopus* and by fish—a cybernetic approach. *Biol Cybern* 93:231–238
- Frens MA, van Opstal AJ (1998) Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Res* 46:211–224
- Friedel P, van Hemmen JL (2008) Inhibition, not excitation, is the key to multimodal sensory integration. *Biol Cybern* 98:597–618
- Geisler CD (1990) From sound to synapse: physiology of the mammalian ear. Oxford University, Oxford
- Gelfand IM, Fomin SV (1963) Calculus of variations. Prentice-Hall, Englewood Cliffs, NY
- Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383:76–78
- Gnedenko B, Kolmogorov A (1968) Limit distributions for sums of independent random variables. Addison-Wesley, Reading, MA
- Goulet J, Engelmann J, Chagnaud BP, Franosch JMP, Suttner MD, van Hemmen JL (2008) Object localization through the lateral line system of fish: theory and experiment. *J Comp Physiol A* 194:1–17
- Grace MS, Woodward OM, Church DR, Calisch G (2001) Prey targeting by the infrared-imaging snake *Python*: effects of experimental and congenital visual deprivation. *Behav Brain Res* 119:23–31
- Gray CM, König P, Engel AK, Singer W (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338:334–337

- Gu Y, Angelaki DE, DeAngelis GC (2008) Neural correlates of multisensory cue integration in macaque MSTd. *Nat Neurosci* 11: 1201–1210
- Gutfreund Y, Zheng W, Knudsen EI (2002) Gated visual input to the central auditory system. *Science* 297:1556–1559
- Hafed ZM, Goffart L, Krauzlis RJ (2008) Superior colliculus inactivation causes stable offsets in eye position during tracking. *J Neurosci* 28:8124–8137
- Halmos PR (1956) *Ergodic theory*. Chelsea, New York
- Hartline PH, Kass L, Loop MS (1978) Merging of modalities in the optic tectum: infrared and visual integration in rattlesnakes. *Science* 199:1225–1229
- Helbig HB, Ernst MO (2007) Optimal integration of shape information from vision and touch. *Exp Brain Res* 179:595–606
- Helmholtz H (1879) *Die Thatsachen in der Wahrnehmung*. August Hirschwald, Berlin
- Hötting K, Rösler F, Röder B (2004) Altered auditory-tactile interactions in congenitally blind humans: an event-related potential study. *Exp Brain Res* 159:370–381
- Hubel DH, Wiesel TN (1977) Functional architecture of macaque monkey visual cortex. *Proc Roy Soc Lond B* 198:1–59
- Hughes A (1977) The topography of vision in mammals of contrasting life style: comparative optics and retinal organization. In: Crescitelli F (ed) *Handbook of sensory physiology*, vol VII/5, chap. 8. Springer, Berlin, Heidelberg, New York, pp 613–756
- Hürlimann F, Kiper DC, Carandini M (2002) Testing the bayesian model of perceived speed. *Vis Res* 42:2253–2257
- Hyde PS, Knudsen EI (2001) A topographic instructive signal guides the adjustment of the auditory space map in the optic tectum. *J Neurosci* 21:8586–8593
- Itti L, Koch C (2001) Computational modeling of visual attention. *Nat Rev Neurosci* 2(3):194–203
- Järvilehto M (1985) The eye: vision and perception. In: Kerkut GA, Gilbert LI (eds) *Nervous system: sensory, comprehensive insect physiology, biochemistry, and pharmacology*, vol 6. Pergamon, Oxford pp 355–429
- Jazayeri M, Movshon (2006) Optimal representation of sensory information by neural populations. *Nat Neurosci* 9:690–696
- Jiang B, Treviño M, Kirkwood A (2007) Sequential development of long-term potentiation and depression in different layers of the mouse visual cortex. *J Neurosci* 27:9648–9652
- Johnson DH, Dudgeon DE (1993) *Array signal processing: concepts and techniques*. Prentice-Hall, Upper Saddle River, NJ
- Johnson KO, Hsiao SS, Yoshioka T (2002) Neural coding and the basic law of psychophysics. *Neuroscientist* 8:111–121
- Jost J, Li-Jost X (1998) *Calculus of variations*. Cambridge University, Cambridge
- Kaas JH, Collins CE (2004) The resurrection of multisensory cortex in primates: connection patterns that integrate modalities. In: Calvert G, Spence C, Stein BE (eds) *The handbook of multisensory processes*, chap 17. MIT, Cambridge, MA, pp 285–293
- Kandel ER, Schwartz JH, Jessell TM (2000) *Principles of neural science*, 4th (international) edn. McGraw-Hill, New York
- Kay SM (1993) *Fundamentals of statistical signal processing*. Prentice Hall, Upper Saddle River, NJ
- Kempler R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. *Phys Rev E* 59:4498–4514
- Kempler R, Leibold C, Wagner H, van Hemmen JL (2001) Formation of temporal-feature maps by axonal propagation of synaptic learning. *Proc Natl Acad Sci USA* 98:4166–4171
- King AJ (1999) Sensory experience and the formation of a computational map of auditory space in the brain. *Bioessays* 21:900–911
- King AJ (2009) Visual influences on auditory spatial learning. *Phil Trans R Soc B* 364:331–339
- King AJ, Hutchings ME (1987) Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret: a map of auditory space. *J Neurophysiol* 57:596–624
- King AJ, Palmer AR (1983) Cells responsive to free-field auditory stimuli in guinea-pig superior colliculus: distribution and response properties. *J Physiol* 342:361–381
- King AJ, Hutchings ME, Moore DR, Blakemore C (1988) Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature* 332:73–76
- Knierim JJ, van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol* 67:961–980
- Knudsen EI (1982) Auditory and visual maps of space in the optic tectum of the owl. *J Neurosci* 2:1177–1194
- Knudsen EI (2002) Instructed learning in the auditory localization pathway of the barn owl. *Nature* 417:322–328
- Knudsen EI, Brainard MS (1991) Visual instruction of the neural map of auditory space in the developing optic tectum. *Science* 253: 85–87
- Knudsen EI, Knudsen PF (1985) Vision guides the adjustment of auditory localization in young barn owls. *Science* 230:545–548
- Knudsen EI, Lac Sd, Esterly SD (1987) Computational maps in the brain. *Annu Rev Neurosci* 10:41–65
- Körding K, Wolpert DM (2004) Bayesian integration in sensorimotor learning. *Nature* 427:244–247
- Krauzlis RJ, Basso MA, Wurtz RH (1997) Shared motor error for multiple eye movements. *Science* 276:1693–1695
- Krauzlis RJ, Liston D, Carello CD (2004) Target selection and the superior colliculus: goals, choices and hypotheses. *Vision Res* 44: 1445–1451
- Krueger LE (1989) Reconciling Fechner and Stevens: toward a unified psychophysical law. *Behav Brain Sci* 12:251–320
- Laming D (1997) *The measurement of sensation*. Oxford University, Oxford
- Lamperti J (1966) *Probability*. Benjamin, New York
- Leibold C, van Hemmen JL (2005) Spiking neurons learning phase delays: how mammals may develop auditory time-difference sensitivity. *Phys Rev Lett* 94:168102
- Li Z (2002) A saliency map in primary visual cortex. *Trends Cogn Neurosci* 6(1):9–16
- Lingenheil M (2004) *Theorie der Beuteortung beim Krallenfrosch*. Master's thesis, Technische Universität München
- Luksch H (2008) Sensorimotor integration: optic tectum. In: Binder MD, Hirokawa N, Windhorst U, Hirsch MC (eds) *Encyclopedia of neuroscience*. Springer, New York
- Mach E (1866) Über die physiologische Wirkung räumlich vertheilter Lichtreize. *Sitzungsber Akad Wiss Wien II* 54:393–408
- Magosso E, Cuppinia C, Serino A, Di Pellegrino G, Ursino M (2008) A theoretical study of multisensory integration in the superior colliculus by a neural network model. *Neural Netw* 21:817–829
- Manley GA, Köppl C, Konishi M (1988) A neural map of interaural intensity differences in the brainstem of the barn owl. *J Neurosci* 8:2665–2676
- Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275:213–215
- Middlebrooks JC, Knudsen EI (1984) A neural code for auditory space in the cat's superior colliculus. *J Neurosci* 4:2621–2634
- Miller K (1970) Least squares methods for ill-posed problems with a prescribed bound. *SIAM J Math Anal* 1:52–74
- Mooney RD, Klein BG, Rhoades RW (1987) Effects of altered visual input upon the development of the visual system and somatosensory representations in the hamster's superior colliculus. *Neurosci* 20:537–555

- Moore EH (1920) On the reciprocal of the general algebraic matrix. *Bull Am Math Soc* 26:394–395
- Morgan ML, DeAngelis GC, Angelaki DE (2008) Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* 59:662–673
- Mosegaard K, Tarantola A (2002) Probabilistic approach to inverse problems. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) *International handbook of earthquake and engineering seismology*, chap 16. Academic, San Diego, CA, pp 237–265
- Murphey RK (1973) Mutual inhibition and the organization of a non-visual orientation in *Notonecta*. *J Comp Physiol A* 84:31–40
- Newman EA, Hartline PH (1981) Integration of visual and infrared information in bimodal neurons of the rattlesnake optic tectum. *Science* 213:789–791
- Newman EA, Hartline PH (1982) The infrared “vision” of snakes. *Sci Am* 246(3):98–107
- Norwich KH, Wong W (1997) Unification of psychological phenomena: the complete form of Fechner’s law. *Percept Psychophys* 59:929–940
- Oğuztöreli MN, Caelli TM (1985) An inverse problem in neural processing. *Biol Cybern* 53:239–245
- Olsen JF, Knudsen EI, Esterly SD (1989) Neural maps of interaural time and intensity differences in the optic tectum of the barn owl. *J Neurosci* 9:2591–2605
- Patton PE, Belkacem-Boussaid K, Anastasio TJ (2002) Multimodality in the superior colliculus: an information theoretic analysis. *Cogn Brain Res* 14:10–19
- Penrose R (1955) A generalized inverse for matrices. *Camb Philos Soc* 51:406–413
- Pickles JO (1988) *An introduction to the physiology of hearing*, 2nd ed. Academic Press, London
- Pouget A, Dayan P, Zemel RS (2003) Inference and computation with population codes. *Annu Rev Neurosci* 26:381–410
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) *Numerical recipes: the art of scientific computing*, 3rd ed. Cambridge University, Cambridge
- Puetter RC, Gosnell TR, Yahil A (2005) Digital image reconstruction: deblurring and denoising. *Annu Rev Astron Astrophys* 43:139–194
- Putzar L, Goerendt I, Lange K, Rösler F, Röder B (2007) Early visual deprivation impairs multisensory interactions in humans. *Nat Neurosci* 10:1243–1245
- Röder B, Rösler F, Spence C (2004) Early vision impairs tactile perception in the blind. *Curr Biol* 14:121–124
- Rosenfeld D (2002) New approach to gridding using regularization. *Magn Reson Med* 48:193–202
- Rowland BA, Quessy S, Stanford TR, Stein BE (2007a) Multisensory integration shortens physiological response latencies. *J Neurosci* 27:5879–5884
- Rowland BA, Stanford TR, Stein BE (2007b) A bayesian model unifies multisensory spatial localization with the physiological properties of the superior colliculus. *Exp Brain Res* 180:153–161
- Rowland BA, Stanford TR, Stein BE (2007c) A model of the neural mechanisms underlying multisensory integration in the superior colliculus. *Perception* 36(10):1431–1443
- Sarkar TK, Weiner DD, Jain VK (1981) Some mathematical considerations in dealing with the inverse problem. *IEEE Trans Antennas Propag* 29:373–379
- Schreiner CE, Langner G (1988) Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol* 60:1823–1840
- Scott DJ (1973) Central limit theorems for martingales and for processes with stationary independent increments using Skorohod representation approach. *Adv App Probab* 5:119–137
- Seung HS, Sompolinsky H (1993) Simple models for reading neuronal population codes. *Proc Natl Acad Sci USA* 90:10,749–10,753
- Shams L, Kamitani Y, Shimojo S (2000) What you see is what you hear. *Nature* 408:788
- Shumway CA (1989) Multiple electrosensory maps in the medulla of weakly electric gymnotiform fish. I. Physiological differences. *J Neurosci* 9:4388–4399
- Sichert AB, Friedel P, van Hemmen JL (2006) Snake’s perspective on heat: reconstruction of input using an imperfect detection system. *Phys Rev Lett* 97:68,105
- Sichert AB, Bamler R, van Hemmen JL (2009) Hydrodynamic object recognition: when multipoles count. *Phys Rev Lett* 102:058,104
- Song S, Miller KD, Abbott LF (2000) Competitive hebbian learning through spike-timing-dependent plasticity. *Nat Neurosci* 3:919–926
- Stanford LR, Hartline PH (1980) Spatial sharpening by second-order trigeminal neurons in crotaline infrared system. *Brain Res* 185:115–123
- Stavenga DG (2002) Reflections on colourful ommatidia of butterfly eyes. *J Exp Biol* 205:1077–1085
- Steenken R, Colonius H, Diederich A, Rach S (2008) Visual and auditory interaction in saccadic reaction time: effects of auditory masker level. *Brain Res* 1220:150–156
- Stein BE, Gaither NS (1981) Sensory representations in reptilian optic tectum: some comparisons with mammals. *J Comp Neurol* 202:69–87
- Stein BE, Meredith MA (1993) *The merging of the senses*. MIT Press, Cambridge, MA
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9:255–266
- Stein BE, Jiang W, Stanford TR (2004) Multisensory integration in single neurons of the midbrain. In: Calvert G, Spence C, Stein BE (eds) *The handbook of multisensory processes*, chap 15. MIT, Cambridge, MA, pp 243–264
- Sullivan WE, Konishi M (1986) Neural map of interaural phase difference in the owl’s brainstem. *Proc Natl Acad Sci USA* 83:8400–8404
- Takahashi T, Konishi M (1986) Selectivity for interaural time difference in the owl’s midbrain. *J Neurosci* 6:3413–3422
- Takeda M, Goodman JW (1986) Neural networks for computation: number representations and programming complexity. *Appl Optics* 25:3033–3046
- Tikhonov AN, Arsenin VY, John F (1977) *Solution of ill-posed problems*. V. H. Winston & Sons, Washington, DC. English translation of the original Russian text
- Tikhonov AN, Goncharsky AV, Stepanov VV, Yagola AG (1995) *Numerical methods for the solution of ill-posed problems*. Kluwer Academic Publishers, Dordrecht. English translation of the original Russian text
- Udin S, Fawcett J (1988) Formation of topographic maps. *Annu Rev Neurosci* 11:289–327
- Ursino M, Cuppinia C, Magosso E, Serino A, Di Pellegrino G (2009) Multisensory integration in the superior colliculus: a neural network model. *J Comput Neurosci* 26(1):55–73
- Van Brunt B (2000) *The calculus of variations*. Springer, Heidelberg
- van der Waerden BL (1957) *Mathematische Statistik*. Springer, Berlin
- van Hemmen JL (2001) Theory of synaptic plasticity. In: Moss F, Gielen S (eds) *Neuro-informatics, neural modelling, handbook of biological physics*, vol 4. Elsevier, Amsterdam pp 771–823
- van Hemmen JL (2002) The map in your head: how does the brain represent the outside world? *Chem Phys Chem* 3:291–298
- van Opstal AJ, Munoz DP (2004) Auditory-visual interactions subserving primate gaze orienting. In: Calvert G, Spence C, Stein BE (eds) *The handbook of multisensory processes*, chap 23. MIT, Cambridge, MA, pp 373–393

- Wallace MT, Stein BE (1997) Development of multisensory neurons and multisensory integration in cat superior colliculus. *J Neurosci* 17:2429–2444
- Wallace MT, Stein BE (2007) Early experience determines how the senses will interact. *J Neurophysiol* 97:921–926
- Wallace MT, Perrault TJ Jr, Hairston WD, Stein BE (2004) Visual experience is necessary for the development of multisensory integration. *J Neurosci* 24:9580–9584
- Wallace MT, Carriere BN, Perrault TJ Jr, Vaughan JW, Stein BE (2006) The development of cortical multisensory integration. *J Neurosci* 26(46):11,844–11,849
- Wandell BA (1995) *Foundations of vision*. Sinauer Associates, Sunderland, MA
- Willshaw DJ, Malsburg C (1976) How patterned neural connections can be set up by self-organization. *Proc Roy Soc Lond B* 194:431–445
- Winkowski DE, Knudsen EI (2006) Top-down gain control of the auditory space map by gaze control circuitry in the barn owl. *Nature* 439:336–339
- Zeil J, Hemmi JM (2006) The visual ecology of fiddler crabs. *J Comp Physiol A* 192:1–25
- Zhang LL, Huizong WT, Holt CE, Harris WA, Poo MM (1998) A critical window for cooperation and competition among developing retinotectal synapses. *Nature* 395:37–44
- Zhou YT, Chellappa R, Vaid A, Jenkins BK (1988) Image restoration using a neural network. *IEEE Trans Acoust Speech Sign Process* 36:1141–1151
- Zittlau KE, Claas B, Münz H (1988) Horseradish peroxidase study of tectal afferents in *Xenopus laevis* with special emphasis on their relationship to the lateral-line system. *Brain Behav Evol* 32: 208–219