Measuring contrast processing in the visual system using the steady state visually evoked potential (SSVEP)

Alex R. Wade Daniel H. Baker

Department of Psychology and York Biomedical Research Institute, University of York, UK

Funding statement:

This work was supported by BBSRC grant BB/V007580/1 to DHB and ARW

Measuring contrast processing in the visual system using SSVEP

Abstract:

Contrast is the currency of the early visual system. Measuring the way that the computations underlying contrast processing depend on factors such as spatial and temporal frequency, age, clinical conditions, eccentricity, chromaticity and the presence of other stimuli has been a focus of vision science for over a century. One of the most productive experimental approaches in this field has been the use of the 'steady-state visually-evoked potential' (SSVEP): a technique where contrast modulating inputs are 'frequency tagged' (presented at well-defined frequencies and phases) and the electrical signals that they generate in the brain are analyzed in the temporal frequency domain. SSVEPs have several advantages over conventional measures of visually-evoked responses: they have relatively unambiguous ouput measures, a high signal to noise ratio (SNR), and they allow us to analyze interactions between stimulus components using a convenient mathematical framework. Here we describe how SSVEPs have been used to study visual contrast over the past 70 years (Dawson, 1954). Because our thinking about SSVEPs is well-described by simple mathematical models, we embed code that illustrates key steps in the modelling and analysis. This paper can therefore be used both as a review of the use of SSVEP in measuring human contrast processing, and as an interactive learning aid.

Keywords:

- EEG
- VEP
- SSVEP
- Vision
- Contrast

Introduction

Neurons in the visual areas of the brain are primarily responsive to changes in cone photoreceptor activations across time and space. This property, referred to as 'contrast', sets the fundamental limits of our visual abilities, which remain steady over a remarkably wide range of environmental light levels. The human response to contrast can be studied using many different techniques. Early work used psychophysical methods to measure contrast sensitivity Schade (1956), defined as the inverse of the lowest contrast that can be reliably detected. But neural responses can also be measured more directly using techniques such as functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG). Here we will describe how an EEG method known as the steady state visually evoked potential (SSVEP) technique has contributed to our understanding of human contrast processing in health, disease and throughout development.

The SSVEP is a continuous electrical response evoked in the brain by visual stimuli flickering at a constant frequency (Regan, 1966*a*). For contrast-defined stimuli, such as sine-wave gratings, it is strongest at the occipital pole, adjacent to the early visual areas that generate the signal, although careful analysis of individual visually-evoked potentials (VEPs) reveals multiple generators throughout visual cortex (Di Russo *et al.*, 2005, 2007). The flickering stimulus entrains neural population responses at multiples of the stimulus frequency, so continuous EEG data are typically analysed by taking the Fourier transform, and estimating the amplitude at these frequencies.

Two common stimulus variants involve sinusoidal on-off (or 'appearance/disappearance') flicker, where the stimulus alternates between a blank background and the peak contrast, and sinusoidal counterphase flicker, where the stimulus alternates in phase (i.e the black regions become white and the white regions become black). On-off flicker can excite independent populations of on- and off-cells in the visual system once per cycle. For spatial patterns, the contributions of individual cortical cells to these excitations are thought to sum at the scalp and generate an average of their onset responses on each cycle. For very low spatial frequencies (including zero) resposes at the scalp will be dominated by a single on- or off- cell type on each half cycle.

If the amplitudes and waveforms of responses to the stimulus appearance and disappearnce are not perfectly balanced, a response at the fundamental flicker frequency, known as 1F, and its integer harmonics can be produced. For stimuli of any spatial frequency (including zero), an imbalance can be caused by differences in either the onset- and offset- responses of either on- or off-cells. For example, if only on-sensitive neurons were present (firing to the onset, but not the offset of local increases in luminance) the response to the on/off (or 'appearance/disappearance') presentation of a grating of any spatial frequency below the resolution limit would consist entirely of odd harmonics.

By comparison, counterphase flickering patterns generate two essentially identical transients per cycle and therefore does not produce a response at 1F, only even harmonics: 2F, 4F, 6F

and so on. Because square-waves are spectrally broad-band, square wave flicker (either on/off or contrast reversig) tends to produce additional spectral harmonics compared to sine-wave flicker.

The higher harmonics of the steady-state signal are generally thought to reflect nonlinear processing in the visual system (Regan & Regan, 1988). Because they can arise from different neuronal computations, different populations and even different stages of the visual system, different harmonics can exhibit different input-output functions. For example, 1F and 2F responses can have different thresholds (Bobak *et al.*, 1984) and scalp topographies (Regan, 1973) while Kaester et al (Kaestner *et al.*, 2024) show both different response slopes and thresholds for 1F and 2F components generated by dynamic noise.

SSVEP signals can also be elicited by periodic changes of stimulus properties other than achromatic and chromatic contrast, such as motion, stereo depth, and facial identity or expression (see Norcia *et al.*, 2015, for an overview); however our focus here is on the contrast response.

Why measure responses to contrast?

Contrast is one of the most fundamental pieces of information that the eye transmits to the brain. It can be defined as the change in cone photoreceptor activity over space ('spatial contrast') or time ('temporal contrast'). Cone photoreceptors - which drive precortical opponent pathways - contribute to both chromatic and achromatic contrast, and although most of the research we describe here focuses on achromatic contrast, SSVEPs have proven to be an excellent measure of early chromatic processing as well McKeefry *et al.* (1996) (see also (Baseler & Sutter, 1997)).

Contrast is typically specified as the percentage deviation of a uniform stimulus from the background. So, for example, a disk of 100 units of cone activation (I_{stim}) surrounded by a 'background' of 50 units of activation ($I_{\text{background}}$) has a contrast of $\frac{I_{\text{stim}}-I_{\text{background}}}{I_{\text{background}}} = 100\%$. Where patterns are more complex (for example, the sine-wave gratings or Gabor patches common in vision science), the Michelson (1927) definition of contrast is specified by the maximum and minimum excursions from the mean:

$$\frac{I_{\text{stimmax}} - I_{\text{stimmin}}}{I_{\text{stimmax}} + I_{\text{stimmin}}}. \tag{1}$$

These contrast definitions are appropriate both to photometric measures of stimulus contrast (for example, luminance; Lennie *et al.* (1993)) and also to definitions based on cone excitations (MacLeod & Boynton, 1979; Derrington *et al.*, 1984) which are more common in work on chromatic processing.

Although its mathematical definition is straightforward, the computations that underlie contrast processing in the brain have been the subject of intense research for many decades. The

neural code for contrast, even in the earliest parts of visual cortex, is not simply a linear transform of the contrast at the retina - instead, contrast signals undergo a cascade of nonlinear processing stages that, broadly, attempt to normalise the output relative to the spatiotemporal environment. This normalization, achieved through a computation called 'contrast gain control' (Heeger, 1992; Foley, 1994; Carandini & Heeger, 2011) maximises the sensitivity of the visual system by making optimal use of neuronal bandwidth. As an example, a grating placed at the centre of a low-contrast background typically appears more intense than the same grating when superimposed on a high contrast background (see Figure **??**; note that the code used to produce all figures in this review is available in python and R at: https://github.com/wadelab/contrastReviewPaperVNS).

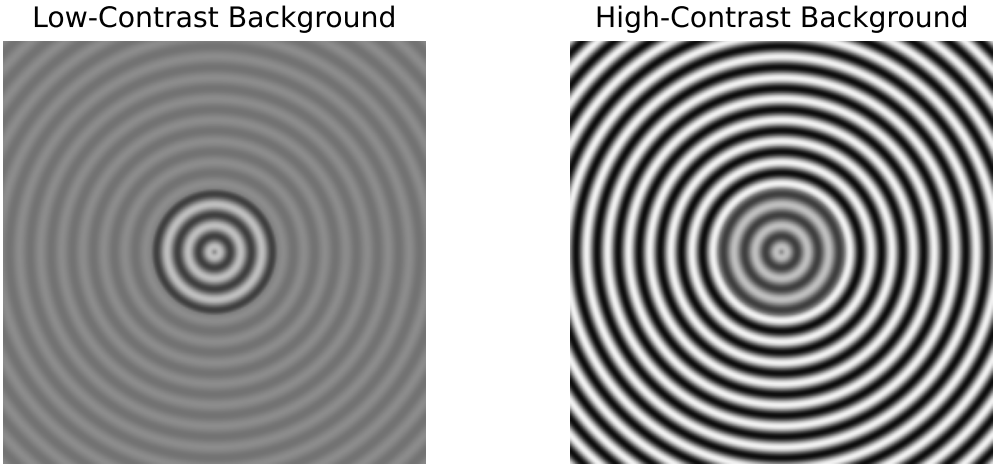

Figure 1: The perceived contrast of a stimulus depends on its context. A high contrast surround reduces the apparent contrast of the central 'probe' region.

A significant body of research into contrast processing is concerned with how these normalization mechanisms depend on colour (Chen *et al.*, 2000), orientation (Foley, 1994), eye of origin (Legge, 1979; Baker *et al.*, 2007), spatial and temporal frequency (Meese & Baker, 2009), location (Polat & Sagi, 1993; Tadin *et al.*, 2003; Petrov *et al.*, 2005), age (Betts *et al.*, 2005), and the presence of neurological disorders (Porciatti *et al.*, 2000; Tsai *et al.*, 2011). The SSVEP has proven to be invaluable in this research because it provides an objective readout of contrast representation at different stages of the visual system, and allows us to 'tag' the probe and background at separate frequencies.

Because it provides a direct read-out of neural population activity, the SSVEP signal can reveal key features of neural signal transduction. For example, by varying the peak stimulus contrast parametrically, a 'contrast vs response' function (CRF) can be measured - where the 'response' is typically defined as the amplitude of the SSVEP frequency component at the stimulus frequency, or a low multiple thereof. This corresponds closely to similar functions reported by studies measuring single unit activity or local field potentials in the cortex (Shapley

& Victor, 1980; Morrone *et al.*, 1982). However the SSVEP has the advantage that it is non-invasive, and so can be measured in awake, behaving human participants.

Although the SSVEP does seem to reflect the activity of relatively well-tuned neuronal populations , changing some aspect of the stimulus may change the nature of those populations. For example, Campbell and Maffei (Campbell & Maffei, 1970) noted that densely-sampled measurements of contrast responses can reveal the presence of two, qualitatively different types of neurons that exhibit different log-linear contrast functions (see also Souza *et al.* (2007)). Modeling such two-limbed contrast response functions using a single sigmoidal function is therefore an approximation.

To understand the utility of the contrast SSVEP, it is helpful to identify the cascade of processing stages in the early visual system that give rise to it. In the following section we illustrate how a typical SSVEP signal measured over early visual cortex might contain information about a large number of early visual computations.

Contrast processing - linear and nonlinear

Neurons have a limited dynamic range, yet they can transmit information about visual stimuli that span many orders of magnitude. In the domain of contrast, to some extent this is accomplished at a population level - individual neurons typically implement a non-linear, sigmoidal CRF transducer (Tolhurst *et al.*, 1981; Albrecht & Hamilton, 1982) and different neurons exhibit peak sensitivity (defined as the maximum slope of the function) at different contrast levels (Carandini & Heeger, 1994; Carandini *et al.*, 1998; Busse *et al.*, 2009). A neuronal population will therefore span a sensitivity range greater than any individual member.

Individual neurons at multiple stages of the visual hierarchy also change their sensitivity depending on the average spatiotemporal contrast energy of their environment. This "normalisation" process is dynamic and nonlinear and is well-modeled by a hyperbolic ratio function in which the response of each neuron is modulated by a local 'gain pool' composed of the summed responses of the local neuronal population (Heeger, 1992; Busse *et al.*, 2009; Carandini & Heeger, 2011; Baker & Wade, 2017).

An additional complexity is introduced by the fact that the EEG is an average population response and that the visual system contains many different types of neurons. As a stimulus changes in, for example, contrast it may selectively activate qualitatively different neuronal populations. This type

To understand these processes better, we will show how sinusoidal input signals might be processed by the visual system to produce SSVEPs. Figure **??** illustrates how sine waves of different contrasts are processed in a linear system. The first panel shows the input sine wave, which would be used to modulate stimulus amplitude over time. Notice that there are five peaks in the waveform during the one second sample, so the stimulation frequency is 5Hz (F1). The second panel shows Fourier transform of the waveform, which contains a substantial peak

at this frequency. If we change the stimulus contrast (i.e. the amplitude of the waveform), the amplitude of the F1 component increases linearly with contrast (right panel).

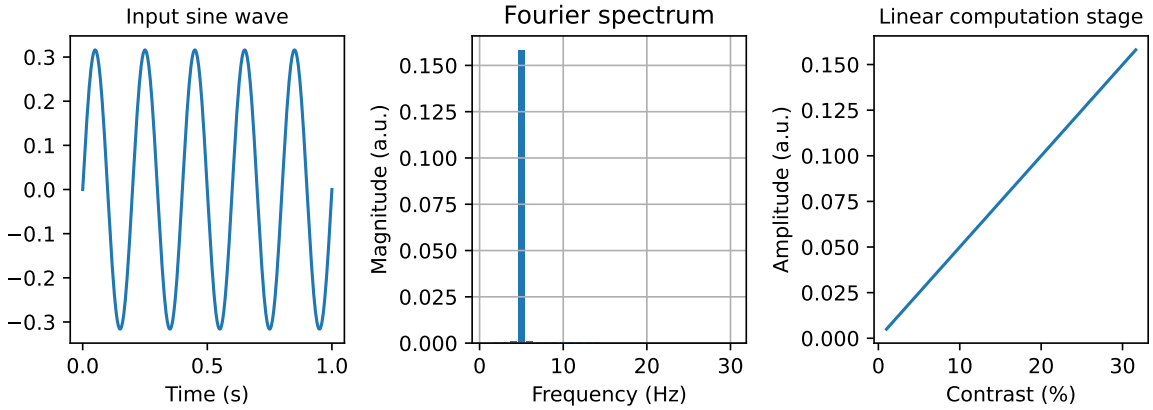

Figure 2: Illustration of a sinusoidal input signal (left), its Fourier spectrum (middle), and how the amplitude of the first harmonic component increases with contrast (right). These calculations assume an entirely linear system.

Next we can consider the impact of two types of nonlinearity in processing on the responses.

First, we consider the fact that neurons in the early visual system code either positive or negative contrast. The responses of each different type of neuron is an increase in firing rate to the increasing levels of the prefered contrast polatiry and these neurons typically exhibit zero (or baseline) response to the non-prefered polarity. We can model the combined responses of these two cell populations to a time varying contrast signal by full-wave-rectification. The effect of this computation is, effectively, to double the input freuquency in the population measurement and to introduce additional higher (even) harmonics due to the discontinuity at the contrast reversals. The effect of full-wave rectification on the SSVEP signal is shown in Figure **??**

One of the simplest nonlinearities is the function that describes a cell's response to different levels of contrast. In Figure **??** this is modeled by a hyperbolic ratio function resulting in a saturating non-linearity:

$$R_{\max} = \frac{C_{\text{in}}^n}{C_{50}^n + C_{\text{in}}^n}, \tag{2}$$

where $R_{\max}$ describes the maximum response level, C_{in} is the input contrast (or the time-varying waveform), C_{50} is the 'semi-saturation constant' (the point at which the response is at half-maximum) and n controls the steepness of the curve (with a typical value around $n = 2$.