NEUROSCIENCE
INTERNATIONAL BRAIN IBRO RESEARCH ORGANIZATION

REVIEW

K. D. Himberger et al. / Neuroscience xxx (2018) xxx–xxx

Principles of Temporal Processing Across the Cortical Hierarchy

Kevin D. Himberger, Hsiang-Yun Chien and Christopher J. Honey*

Department of Psychological & Brain Sciences, Johns Hopkins University, Baltimore, MD, United States

Abstract—The world is richly structured on multiple spatiotemporal scales. In order to represent spatial structure, machine-learning models repeat a set of basic operations at each layer of a hierarchical architecture. These iterated spatial operations – including pooling, normalization and pattern completion – enable these systems to recognize and predict spatial structure, while robust to changes in the spatial scale, contrast and noisiness of the input signal. Because our brains also process temporal information that is rich and occurs across multiple time scales, might the brain employ an analogous set of operations for temporal information processing? Here we define a candidate set of temporal operations, and we review evidence that they are implemented in the mammalian cerebral cortex in a hierarchical manner. We conclude that multiple consecutive stages of cortical processing can be understood to perform temporal pooling, temporal normalization and temporal pattern completion.

This article is part of a Special Issue entitled: Sequence Processing. **© 2018 The Authors. Published by Elsevier Ltd on behalf of IBRO. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).**

Key words: hierarchy, predictive coding, timescales, sequence processing, temporal integration.

OVERVIEW

The mammalian cerebral cortex is organized as a functional hierarchy. The lowest levels of this hierarchy are located in primary sensory and motor cortices and higher levels are reached at increasing synaptic distance from the periphery (Fuster, 1997; Mesulam, 1998; Huntenburg et al., 2017). When measuring neural responses to spatially structured stimuli, recordings further up the cortical hierarchy reveal larger spatial receptive fields and increasing selectivity for feature configurations (Hubel and Wiesel, 1962; Grill-Spector and Malach, 2004; also Rauschecker and Tian, 2000). In response to time-varying stimuli, recordings further up the hierarchy reveal longer temporal receptive windows and increasing selectivity for coherent temporal structures (Xu et al., 2005; Hasson et al., 2008; Lerner et al., 2011; Honey et al., 2012; de Heer et al., 2017). Overall, circuits at low levels are affected by local spatiotemporal properties such as the pitch of sounds and the orientation of image patches, while higher order circuits are most affected by more complex and spatiotemporally distributed properties, such as those that determine semantics, animacy, and object-scene relationships.

The hierarchical organization of the cerebral cortex has prompted the hypothesis that a common set of input–output operations is applied at each stage of processing (Creutzfeldt, 1977; Bastos et al., 2012). This hypothesis motivates many computational models of hierarchical image processing, inspired by the structure of the primate ventral visual stream (Fukushima and Miyake, 1982; Serre et al., 2007; Yamins and DiCarlo, 2016). Three canonical operations used in hierarchical image recognition models include: spatial pooling, spatial normalization, and spatial pattern completion. Pooling between stages is thought to make higher order representations robust to scaling and translation of the input (Fukushima and Miyake, 1982; Serre et al., 2007). Normalization at each stage is thought to enhance feature selectivity and robustness to input gain (i.e. amplitude) (LeCun et al., 2010; Carandini and Heeger, 2012). Pattern completion is thought to reduce the effects of noise and missing information, and to produce low-dimensional representations that can accurately re-generate high-dimensional input (Hopfield, 1982; Dayan et al., 1995; Yuille and Kersten, 2006).

If these spatial operations are effective for analyzing and representing static images in a multi-stage processing stream, might similar operations work well for processing temporal structure? This manuscript considers the evidence for an analogous set of basic temporal operations: temporal pooling, temporal normalization, and temporal pattern completion. In Secti on "Motivation and scope" we motivate this hierarchical perspective and set the scope of our proposal. In Sectio n "Principles of temporal information processing" we define each of our proposed temporal operations and we consider the evidence that these operations are implemented in the brain. Finally, in Section "Challenges,

*Corresponding author.
E-mail address: chris.honey@jhu.edu (C. J. Honey).

Abbreviations: ECoG, electrocorticographic; EEG, electroencephalographic; ERPs, event-related potentials; fMRI, functional magnetic resonance imaging; MMN, mismatch negativity.

https://doi.org/10.1016/j.neuroscience.2018.04.030
0306-4522/© 2018 The Authors. Published by Elsevier Ltd on behalf of IBRO.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article in press as: Himberger KD et al. Principles of Temporal Processing Across the Cortical Hierarchy. Neuroscience (2018), https://doi.org/10.1016/j.neuroscience.2018.04.030

open questions, and predictions", we discuss the functional implications of this temporal processing perspective, as well as limitations, open questions and predictions.

MOTIVATION AND SCOPE

Many theoretical frameworks assume that a common set of input–output mappings is applied across multiple stages of cortical processing (Hawkins and Blakeslee, 2007; Kiebel et al., 2008; Bastos et al., 2012; Hasson et al., 2015; Heeger, 2017). One motivation for this approach is from anatomy: if similar circuitry is observed across diverse cortical areas (Douglas and Martin, 2018), then perhaps similar functional operations are being implemented as well (Creutzfeldt, 1977). A second and more recent motivation is the success of neural networks for image recognition and analysis. These models from machine learning have bolstered the argument that hierarchically repeated computations are an effective way to recognize the properties of spatial images (e.g. Jarrett et al., 2009) and to form multi-level spatial representations (Hinton et al., 2006). More recently, time-varying linguistic stimuli have also been successfully analyzed and represented using multi-scale architectures (Chung et al., 2016; Bradbury et al., 2016; Gehring et al., 2017).

Hierarchical processing architectures may be especially effective for processing stimuli that possess a hierarchical structure. In the spatial domain, this can be seen in the way that edge-like features compose contours, which compose objects (Hubel and Wiesel, 1959, 1962; Kourtzi and Connor, 2011), with roughly corresponding representations in consecutive levels of the models (Zeiler and Fergus, 2014). Since information is also organized hierarchically in time (e.g. phonemes inside of words inside of prosodic curves; motions inside of actions inside of events; Gibson, 1982; Poeppel, 2003) then perhaps a hierarchical architecture is also effective in the time domain, and may be used in the brain.

In addition to the theoretical motivations, there are strong empirical reasons to seek a set of canonical temporal operations. Empirically, we observe that temporal properties affect information processing ubiquitously across the cortical hierarchy. In the auditory pathway, neuronal responses depend on stimulus history up to many seconds earlier, and this effect is already present in primary auditory cortex (Ulanovsky et al., 2004; Rosburg et al., 2008) and its inputs (Parras et al., 2017). In the visual cortex, even the earliest stages of processing depend on prior stimulus properties from hundreds of milliseconds (Gavornik and Bear, 2014) or seconds earlier (Homann et al., 2017). More generally, history-dependent responses – manifesting as sequential "mismatch effects" – have been demonstrated in diverse cortical systems in electroencephalographic (EEG) and electrocorticographic (ECoG) recordings in visual and auditory paradigms (Pazo-Alvarez et al., 2003; Näätänen et al., 2007; Bekinschtein et al., 2009; Kremláček et al., 2016).

There is further empirical evidence that temporal processing has a hierarchical character. Early sensory regions tend to be affected by more recent properties of the input stream, while higher order regions are affected by more complex features and by longer windows of prior stimulus context (Xu et al., 2005; Hasson et al., 2008; Bekinschtein et al. 2009; Wacongne et al., 2011; Lerner et al., 2011; Honey et al. 2012; Gauthier et al., 2012). Thus, the classic computational notion of sensory pathways as a hierarchy of growing receptive fields might be more accurately conceived as a hierarchy of spatiotemporal fields. A critical remaining question is then, what kinds of temporal operations are performed at each level of the hierarchy?

We define "temporal operations" broadly as any input–output relationship where temporal properties of the input can affect the output. Thus, we do not restrict our focus to aspects of behavior and perception related to "timing" per se (e.g. duration estimation (Grondin, 2010) or temporal sensitivity (Watson, 1986)) or to stimulus parameters that constrain motion processing (Exner, 1875; Wertheimer, 1912; Sekuler, 1996). Our goal is not to describe minimal delays or fusion thresholds, but instead to specify which classes of input–output operations might be ubiquitously and generically applied in the cerebral cortex. Because almost all information in the world has a temporal structure, we focus on temporal operations that are observable across many different sensory pathways, and across many stages of processing.

Because we seek operations that may be applied to temporal structure in many areas, we contrast and combine data across modalities (vision and audition) and across aspects of cognition (language, perception, and semantics). For example, the concept of temporal pooling can be understood to apply both in visual psychophysics experiments (Barlow, 1958) and in studies of temporally reversed acoustics (Saberi and Perrott, 1999). Thus, although many parameters of auditory and visual processing may differ (Kubovy and Van Valkenburg, 2001), we believe that there are fundamental anatomical and functional similarities that motivate a search for common operations (Von Melchner et al., 2000). Therefore, we focus on experiments for which (i) basic temporal properties are manipulated, and (ii) neural responses to these manipulations have been measured for more than one area or type of input. Such studies have the best chance to reveal information processing motifs that are repeated across levels and neural pathways. Often, electrophysiological and behavioral data provide some of the best information about precise timing and about mechanism, while methods with a broad spatial view, such as functional magnetic resonance imaging (fMRI) enable us to explore variations in response across regions (e.g. Lerner et al., 2011; Gauthier et al., 2012; Zhou et al., 2017).

The empirical and theoretical considerations sketched above indicate that history-dependent processes unfold ubiquitously in the neocortex. In the text that follows, we marshal evidence for three candidate operations – temporal pooling, temporal normalization, and temporal pattern completion – that may be applied at each stage

UNCORRECTED PROOF

Download English Version:

https://daneshyari.com/en/article/11017862

Download Persian Version:

https://daneshyari.com/article/11017862

Daneshyari.com