crcns.org hc2 data description
version 1.2 (Jan 18, 2012)

Animals and surgery

Three male Long-Evans rats (rat ID; ec13, ec14, ec16, 250-400 g) were implanted with a 4-shank or 8-shank silicon probe in layer CA1 of the right dorsal hippocampus. The individual silicon probes were attached to respective micromanipulators and moved independently. Each shank had 8 recording sites (160 μm^2 each site; 1-3 MΩ impedance). These recordings sites were staggered to provide a two-dimensional arrangement (20 µm vertical separation; Fujisawa et al., 2008). The shanks were aligned parallel to the septo-temporal axis of the hippocampus (45 degrees parasagittal), positioned centrally at anteroposterior=3.5 mm from bregma and mediolateral=2.5 mm from midline. Two stainless steel screws inserted above the cerebellum were used as indifferent and ground electrodes during recordings. All protocols were approved by the Institutional Animal Care and Use Committee of Rutgers University.

Behavioral testing

After recovery from surgery (~1 week), physiological signals were recorded during the open field task in which the rats chased randomly dispersed drops of water or pieces of Froot Loops (~25 mg) on an elevated square platform (180 cm x 180 cm, or 120 cm x 120 cm). The animals were trained on the platform at least 3 days before the recordings.

Data collection and analysis

During the recording sessions, neurophysiological signals were amplified (1,000X), bandpass-filtered (1 Hz – 5kHz) and acquired continuously at 20 kHz on a 128-channel DataMax system (16-bit resolution; RC Electronics, .dat file). After recording, local field potential (LFP) was down-sampled to 1250 Hz for additional analysis (.eeg file).

For offline spike sorting, the wideband signals were digitally high-pass filtered (0.8-5 kHz) and the waveforms were re-sampled (Csicsvari et al., 1999). Neurophysiological and behavioral data were explored using NeuroScope (http://neuroscope.sourceforge.net; Hazan et al., 2006). Spike sorting was performed automatically, using KlustaKwik (http://klustawik.sourceforge.net), followed by manual adjustment of the clusters (using "Klusters" software package; http://klusters.sourceforge.net).

Each experiment dataset consists of:

1. Binary raw data file(s) (FileBase.dat) and binary LFP signal (FileBase.eeg) the binary file format is as follows: each sample is stored as type short integer (2 bytes) in the order Channel_1:Sample_1, Channel_2:Sample_1, ..., Channel_N:Sample_1, Channel_1:Sample_2, Channel_2:Sample_2, ... etc. The number of channels is not always a multiple of 8 (eight recording sites per shank) because bad channels (due to very high impedance or broken shank) were removed from the data.

2. For each shank/electrode N there are several files containing the information about spikes extracted from this electrode:
    a. Text file of spike time stamps (FileBase.res.N, where N – shank index). The time stamps are in samples at 20kHz sampling rate.
    b. Text file of cluster identities (FileBase.clu.N). The first number – total number of clusters, where cluster with index 0 represents artifacts, 1 – noise/nonclusterable units, 2 and above – isolated units.
    c. Spikes recorded in multiple sessions on the same day were sorted together and the same cell ID numbers were assigned in the sessions.
    Example: In sessions ec013.527, ec013.528 and ec013.529, a given cell has the same cell ID in the .clu files in all three sessions, as they were clustered together.
    d. Binary file of spike waveshape (FileBase.spk.N). The waveshape of each spike across channels are written consecutively in the multiplexed form as in .dat file. The width of each spike is 32 samples, with sample #15 being the trough of spike.
    e. Text file of waveshape features for each spike (FileBase.fet.N). First line – total number of dimensions in the feature space. Typically, for each channel there are 3 PCAs, and on top of this there are 3 features coding peak-to-trough, peak-to-baseline and trough- to-baseline of the spike on the channel of largest amplitude, 1 feature coding for spike width and the last one – time of the spike (identical to FileBase.res.N)
    f. Several auxiliary files used by Klusters software. (".m1m2." and ".mm." in name).

3. Configuration file(s) associated with each session file (extension .xml) This file is created by the neuroscope software, and it's content can be accessed via the GUI interface of this software. You can also read the file in any editor will give you an idea of various fields there. Most important is the section on the acquision system, which contains number of channels, sampling rate, voltage range (not always verified), and amplification. Also important is the channelGroups field, which contains the list of channels that represent one shank of the probe. The number of channels per electrode group (shank) is not always 8 because bad channels were removed from the data. For more information about the content of the .xml files see the manual of the neuroscope software at http://neuroscope.sourceforge.net/.

4. Animal tracking data.
    a. MPEG-2 video file (FileBase.mpg). 2 LEDs on animal head are used to identify it's position. One more stationary LED is blinking and is used to synchronize video with data acquisition. The synchronization pulse is recorded in the FileBase.led file (at 20kHz sampling rate). Format is one short integer (2 bytes) per sample.
    b. Text file of the position of the animal (FileBase.whl) extracted from the video file. Each row – a frame in the movie. Sampling rate – 39.06 Hz. The samples are

synchronized with the electrophysiology. First two columns code for (x,y) coordinates, respectively, of the first LED, last two – of the second. -1 means adequate tracking was not possible for these frames.

Reference

Csicsvari,J., Hirase,H., Czurko,A., Mamiya,A., and Buzsáki,G. (1999). Oscillatory coupling of hippocampal pyramidal cells and interneurons in the behaving rat. J. Neurosci. *19*, 274-287.

Hazan,L., Zugaro,M., and Buzsáki,G. (2006). Klusters, NeuroScope, NDManager: a free software suite for neurophysiological data processing and visualization. J. Neurosci. Methods *155*, 207-216.

Specific information for each session:

rat ID	date	session	maze	shanks	channels	reward	duration (minutes)	size (GB)
ec13	2-Feb-06	ec013.527	180 cm	4	31	food	17.7	1.3
ec13	2-Feb-06	ec013.528	180 cm	4	31	food	26.7	2.0
ec13	2-Feb-06	ec013.529	180 cm	4	31	water	30.8	2.3
ec13	10-Feb-06	ec013.713	180 cm	4	31	water	37.8	2.7
ec13	10-Feb-06	ec013.714	180 cm	4	31	water	38.0	2.7
ec13	13-Feb-06	ec013.754	180 cm	4	31	water	53.9	3.8
ec13	13-Feb-06	ec013.755	180 cm	4	31	food	32.8	2.4
ec13	13-Feb-06	ec013.756	180 cm	4	31	water	40.4	2.8
ec13	13-Feb-06	ec013.757	180 cm	4	31	water	40.3	2.8
ec13	16-Feb-06	ec013.808	180 cm	4	31	water	36.8	2.6
ec13	18-Feb-06	ec013.844	180 cm	4	31	water	30.6	2.2
ec14	22-Feb-07	ec014.277	180 cm	8	64	water	91.9	15
ec14	25-Feb-07	ec014.333	180 cm	8	64	water	93.5	13
ec14	15-Mar-07	ec014.793	120 cm	8	64	water	47.4	6.9
ec14	15-Mar-07	ec014.800	120 cm	8	64	water	49.3	7.2
ec14	19-Mar-07	ec015.041	120 cm	8	64	water	47.1	6.8
ec14	19-Mar-07	ec015.047	120 cm	8	64	water	49.4	7.4
ec16	19-Sep-07	ec016.397	180 cm	8	55	water	90.8	12
ec16	21-Sep-07	ec016.430	180 cm	8	55	water	106.8	13
ec16	22-Sep-07	ec016.448	180 cm	8	55	water	90.8	12
ec16	1-Oct-07	ec016.582	180 cm	8	56	water	90.9	11

Legend for table:

date – recording date

maze – size of square platform. Either 180 cm x180 cm or 120 cm x 120 cm.

channels – number of channels recorded. Is not always a multiple of 8 (number of sites per shank) because bad channels were removed from the data.

duration – duration of experiment in minutes.

size – size of downloadable data (compressed) in gigabytes, not including .mpg file.