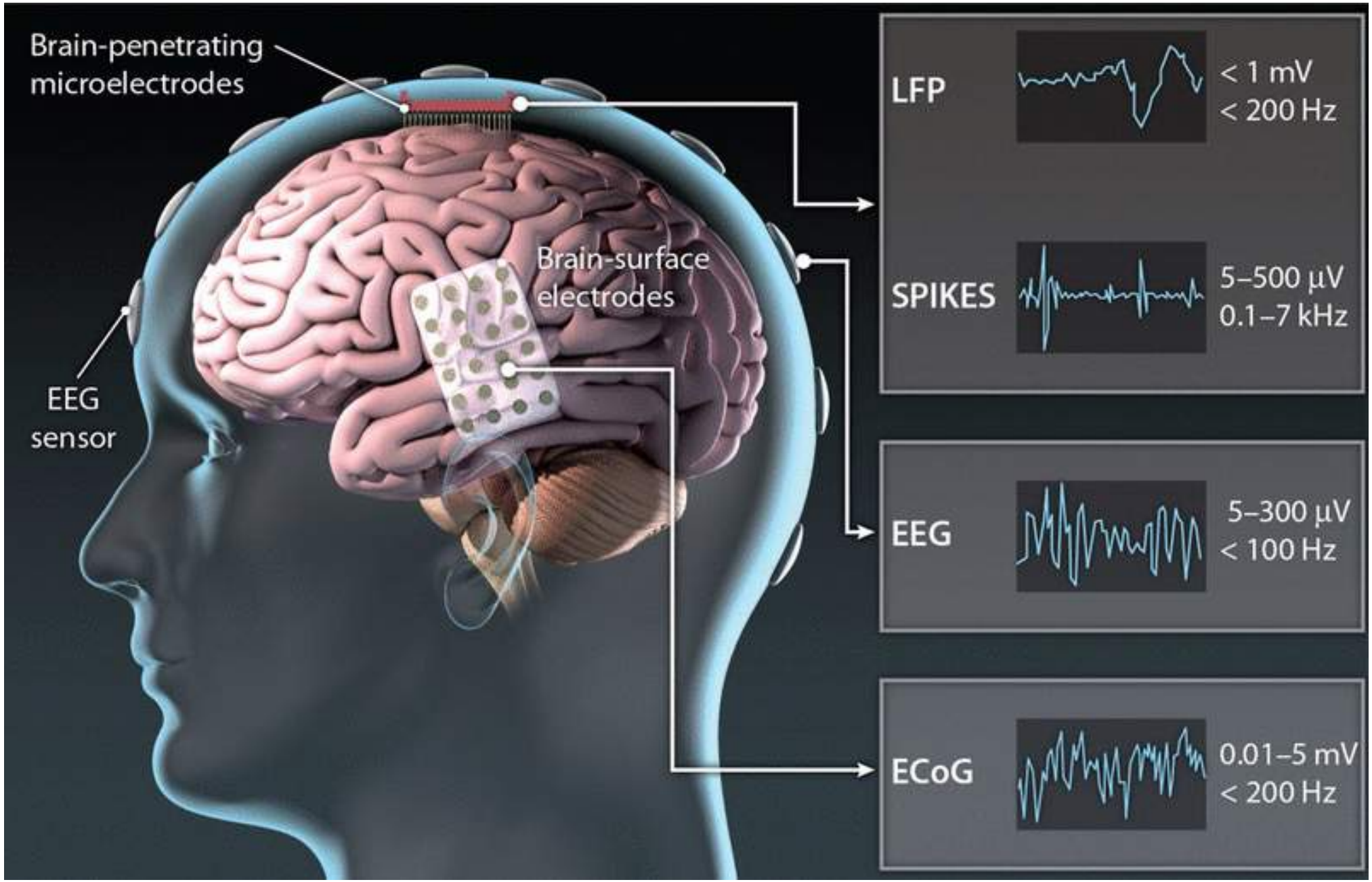
Brain Signal Acquisition

Background

- The brain communicates using spikes-- produced when the neuron receives enough input current from other neurons via synaptic connections.

- Recording brain activity are based on detecting changes in electrical potentials in neurons
  - invasive techniques based on implanting electrodes

- Or on detecting changes in large populations of neurons
  - noninvasive techniques such as electroencephalography or EEG


Brain-penetrating microelectrodes
Brain-surface electrodes
EEG sensor
LFP
< 1 mV
< 200 Hz
SPIKES
5–500 µV
0.1–7 kHz
EEG
5–300 µV
< 100 Hz
ECoG
0.01–5 mV
< 200 Hz

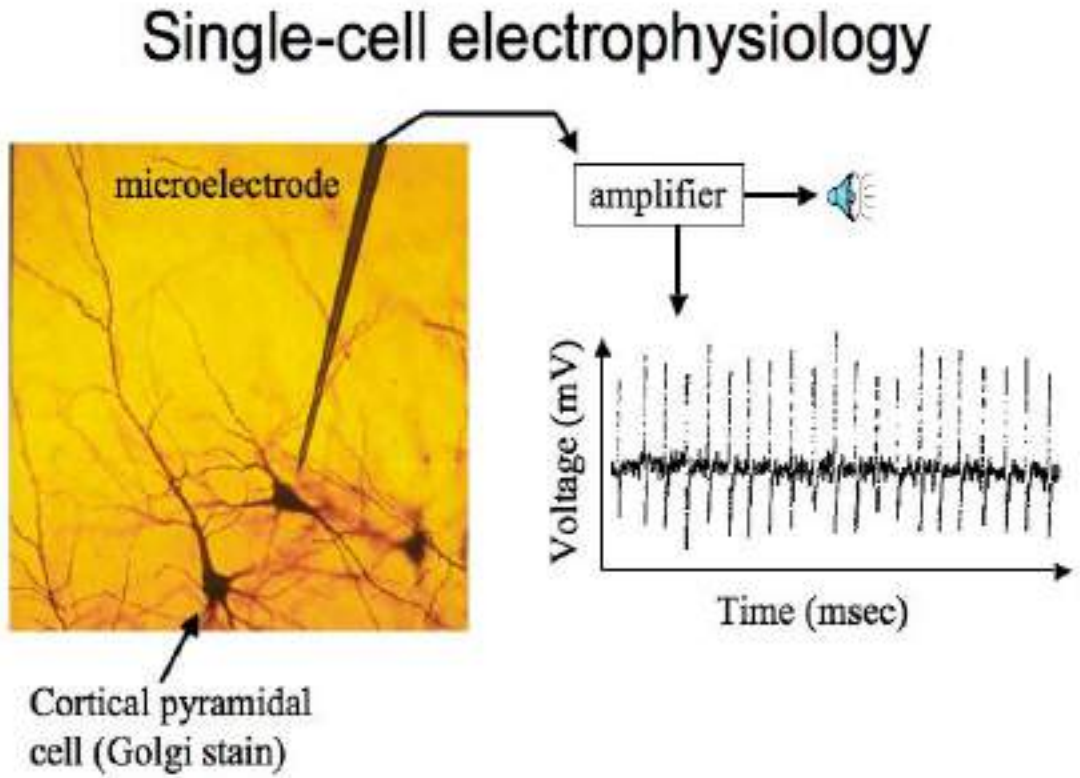
Recording signals from brain

Invasive Approaches:

- Techniques that allow recording from individual neurons in the brain are typically invasive.
- They involve some form of surgery,
  - A part of the skull is removed, an electrode or implant placed in the brain, and the removed part of the skull then replaced.
  - The recording itself is not painful because the brain has no internal pain receptors, but the surgery and recovery process can cause pain and involves risks such as infection.
- A major advantage of invasive recordings is that they allow recording of action potentials at the millisecond timescale.

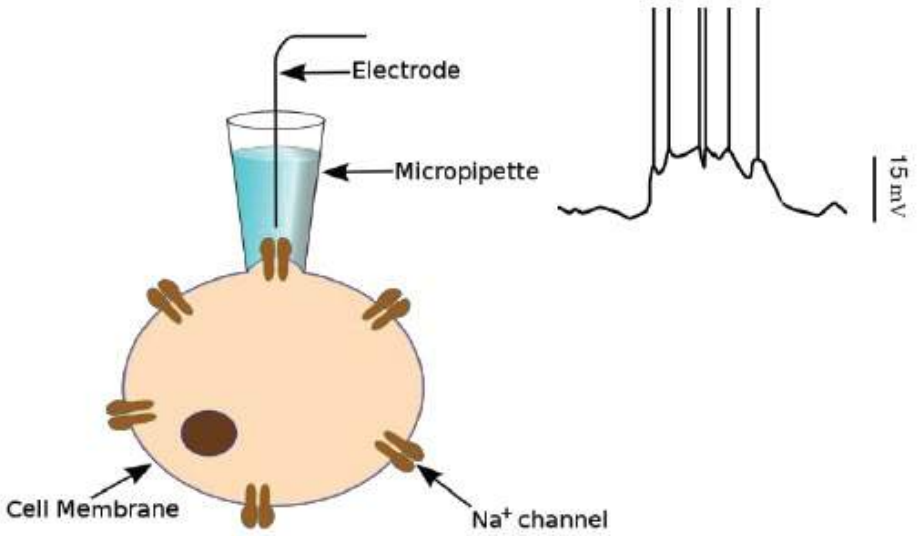
Invasive Approaches

- **<u>Microelectrode:</u>**
- A *microelectrode* is simply a very fine wire or other electrical conductor used to make contact with brain tissue.
- A typical electrode is made of tungsten or platinum- iridium alloy and is insulated except at the tip, which measures around 1µm in diameter (A neuron's cell body diameter is in the range of tens of µm).



Invasive Approaches

- **<u>Intracellular Recording:</u>**
- Measures the voltage or current across the membrane of the neuron.
- The most common technique, known as *patch clamp recording.*



- Very Delicate → Intracellular recordings are typically performed only on slices of brain tissue

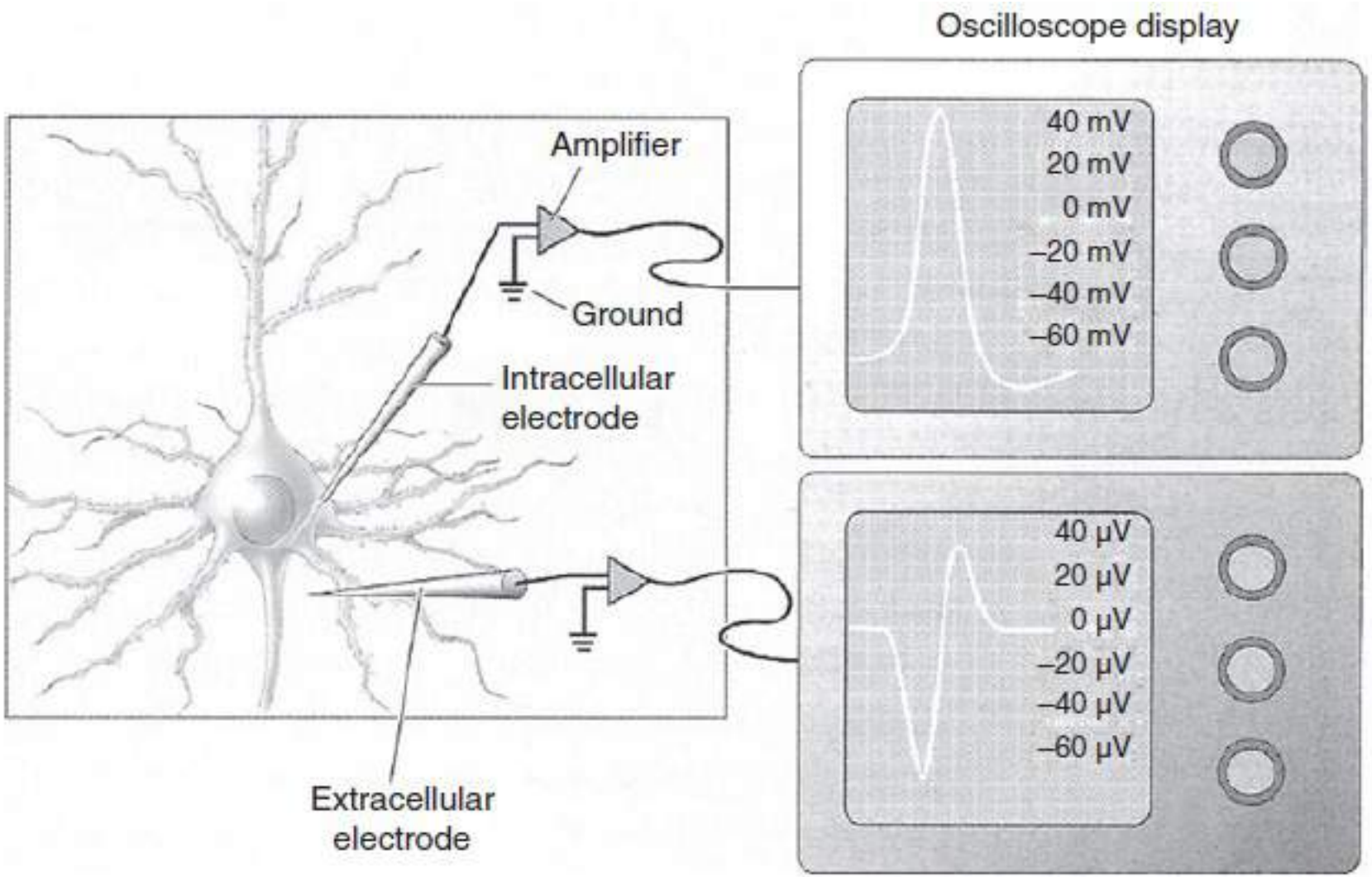
Invasive Approaches

- **Extracellular Recording:**
- Recording of a single neuron (or single“unit”): a tungsten or platinum-iridium microelectrode with a tip size of less than 10 microns is inserted into the target brain area.

- The magnitude of the recorded signal is usually less than a millivolt and thus requires the use of amplifiers to detect the signal.

- The signal from the amplifier is fed to a computer, which performs additional processing such as filtering noise and isolating the spikes (action potentials).

Invasive Approaches



(from Bear et al., 2007).

Invasive Approaches

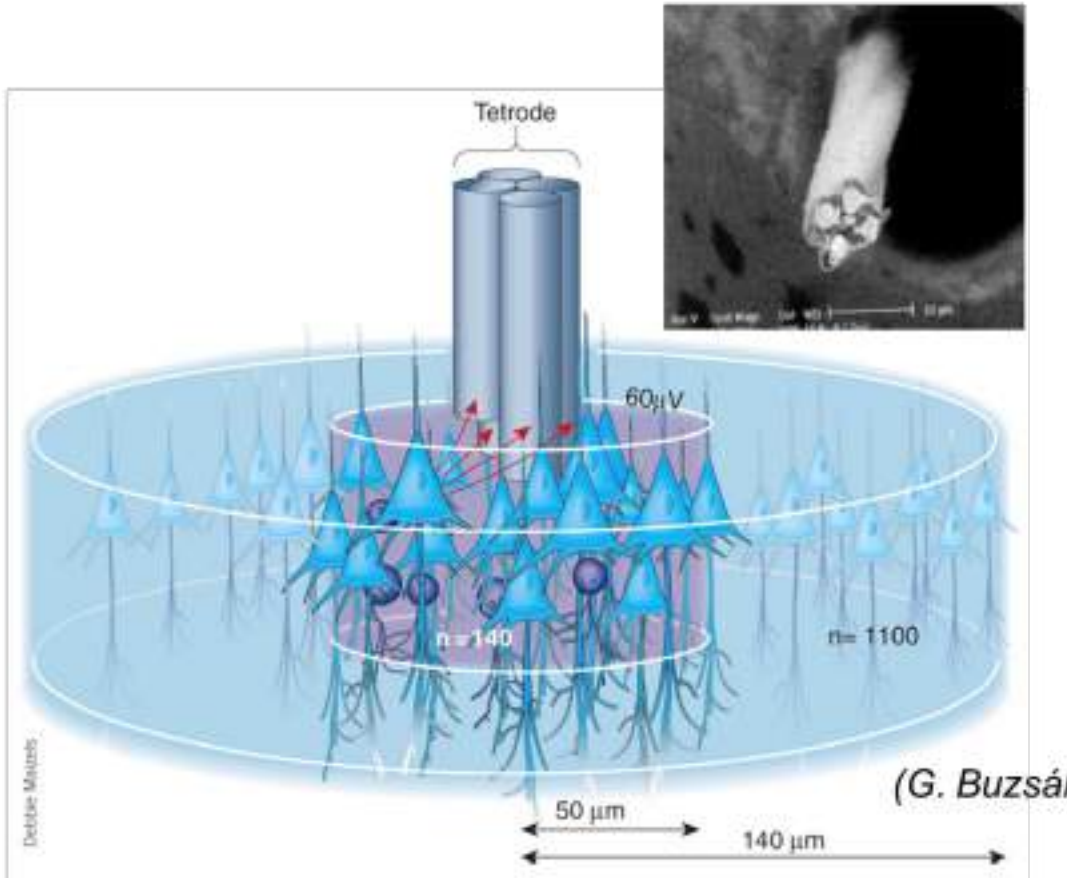
- When the neuron produces a spike, positive ions flow away from the extracellular electrode into the neuron, causing the initial negative deflection in the display. This is followed by a positive deflection as the action potential decreases and positive charges flow out of the neuron toward the extracellular electrode.

Invasive Approaches

Tetrodes and Multi-Unit Recording:

- To record from multiple neurons simultaneously by using more than one electrode.
- Four wires are tightly wound together in a bundle.

Tetrode recordings allow the monitoring of single neurons during behavior

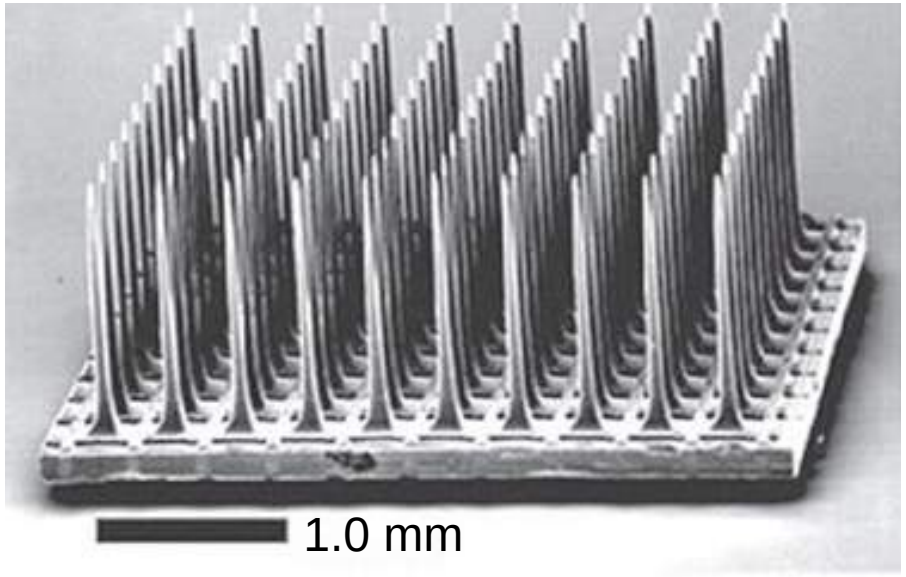


(G. Buzsáki, 2005)

Invasive Approaches

Multielectrode Arrays:

- To record from larger numbers of neurons, microelectrodes can be arranged in a grid-like structure to form a multielectrode array of $m \times n$ electrodes.



(adapted from Hochberg et al., 2006).

Invasive Approaches

- The most common types of implantable arrays are microwire, silicon-based, and flexible microelectrode arrays
- Increased spatial resolution
- The ability to record simultaneously from several dozens of neurons
- Opens the door to extracting complex types of information such as position or velocity signals that could be useful for controlling prosthetic devices.

Brain Signal Acquisition

Partially Invasive Approach

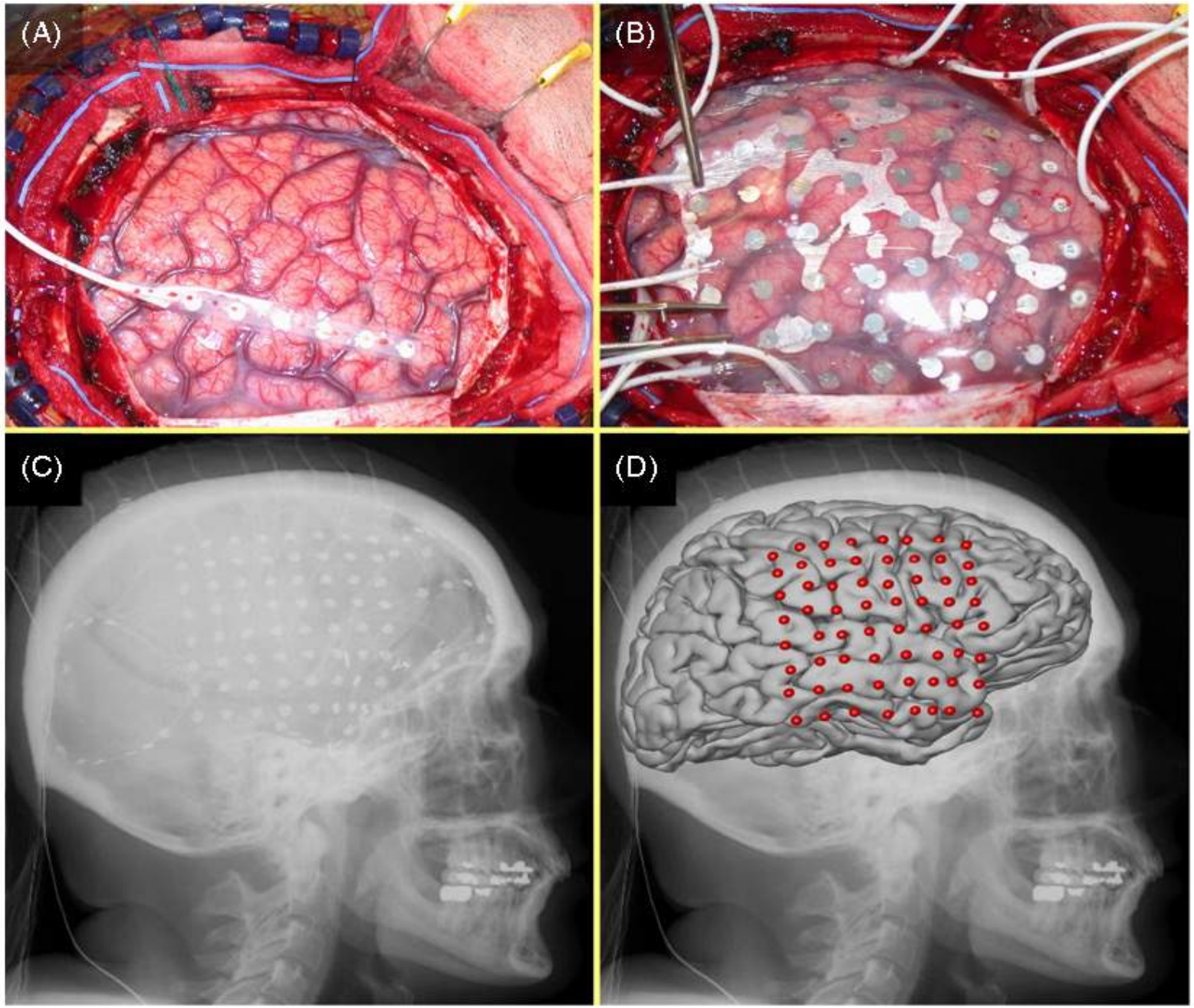
Electrocorticography (ECoG):

- *Electrocorticography* (ECoG) is a technique for recording brain signals that involves placing electrodes on the surface (cortex) of the brain.
- The procedure requires making a surgical incision into the skull to implant the electrodes on the brain surface

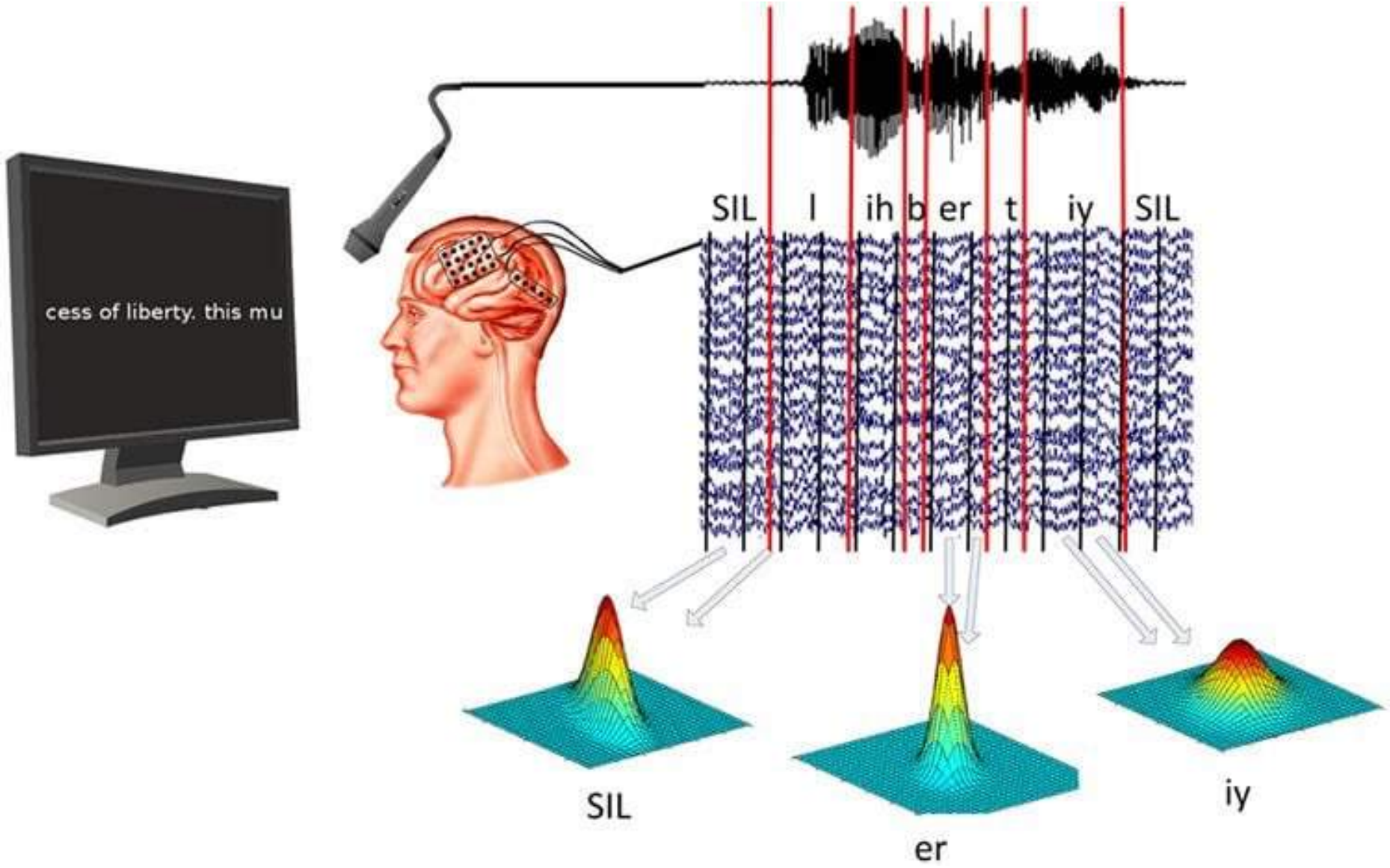
Partially Invasive Approach

- ECoG electrodes can record the electrical fluctuations caused by the coherent activity of large populations of neurons (several tens of thousands).
- Safer than arrays implanted inside the brain.
- ECoG electrodes may also be less likely to wear out compared to brain penetrating electrodes
- ECoG offers greater spatial resolution

https://www.youtube.com/watch?v=KEk3m4FwVt8



(from (Miller et al., 2007).



Automatic Speech Recognition from Neural Signals: A Focused Review" by Christian Herff and Tanja Schultz in *Frontiers in Neuroscience*. Published online September 27 2016 doi:10.3389/fnins.2016.00429

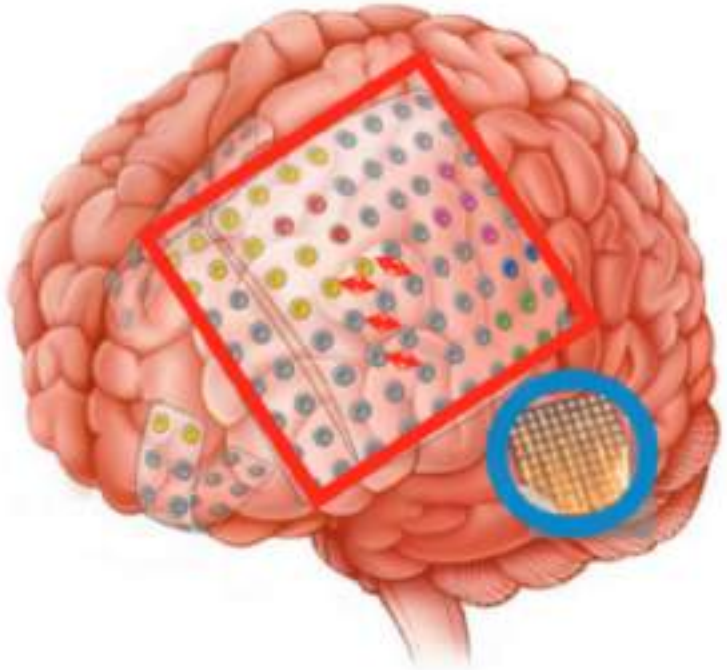
ECoG and audio data are recorded at the same time. Speech decoding software is then used to determine timing of vowels and consonants in acoustic data. ECoG models are then trained for each phone individually by calculating the mean and covariance of all segments associated with that particular phone.

Partially Invasive Approach

MicroECoG:

- One disadvantage of ECoG, is the relatively large size of ECoG electrodes
- These microelectrodes are only a fraction of a millimeter in diameter and spaced only 2–3 mm apart in a grid
- Allows detection of neural activity at a much finer resolution than traditional ECoG.
- Decoding fine movements, such as the movements of individual fingers, or even speech, without actually penetrating the brain.

Partially Invasive Approach



Current ECoGs
- Large area
- Low resolution

Current µECoGs
- Small area
- High resolution

Partially Invasive Approach

Optical Recording: Voltage-Sensitive Dyes and Two-Photon Calcium Imaging:

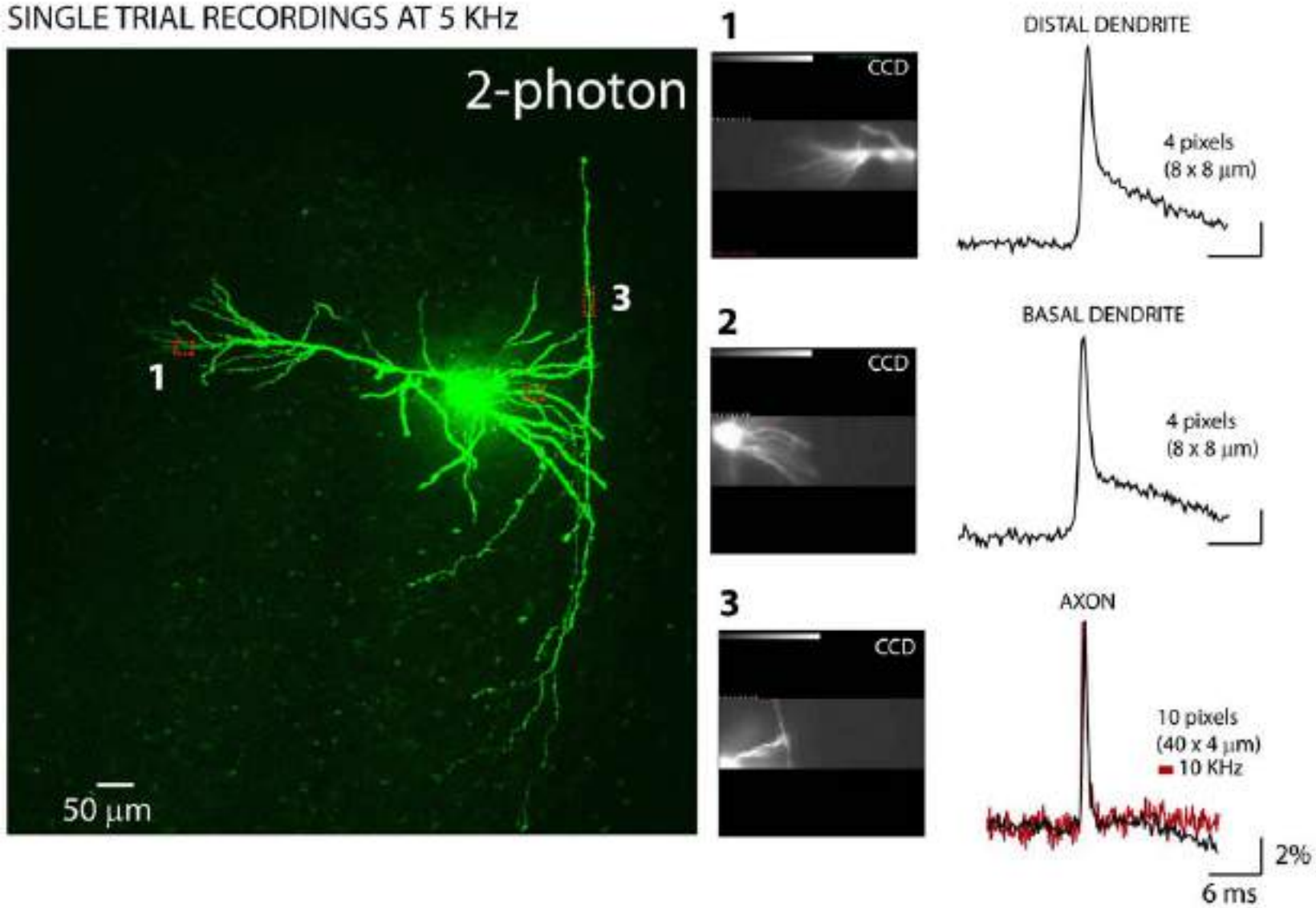
- Voltage-sensitive dyes
- Voltage-sensitive dye changes its absorbance and fluorescence intensity when membrane potential changes in a stained brain or heart tissue.
- By using voltage-sensitive dyes as chemical probes and capturing changes in light intensity with the use of a high-speed imaging device, it is possible to image in real time the activity of where, when, and how much excitation or inhibition occurred, in the brain and heart.

Partially Invasive Approach

Voltage-sensitive dyes
https://www.scimedia.com/applications/neuro/vsd/data/

- Neurons are stained with a voltage-sensitive dye
- Dye responds to changes in membrane potential by changing its absorption and/or fluorescence
- Recorded optical signals correspond to summed responses from several simultaneously active neurons.
- Useful for imaging macroscopic features of the brain.

VOLTAGE-SENSITIVE DYE IMAGING



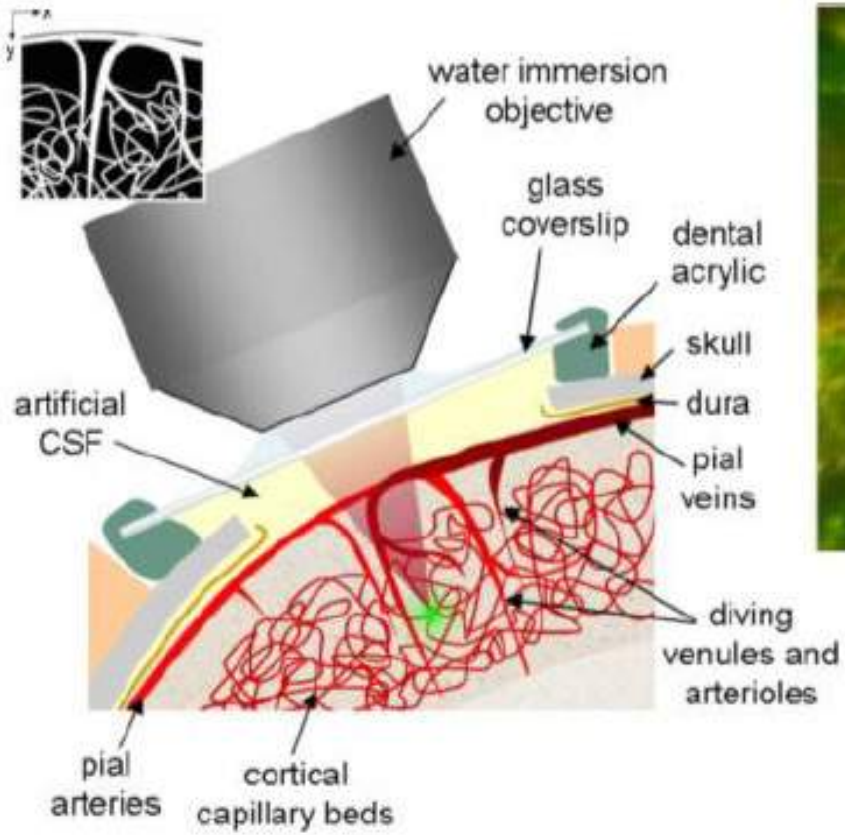
(image: Scholarpedia http://www.scholarpedia.org/article/Voltage-sensitive_dye).

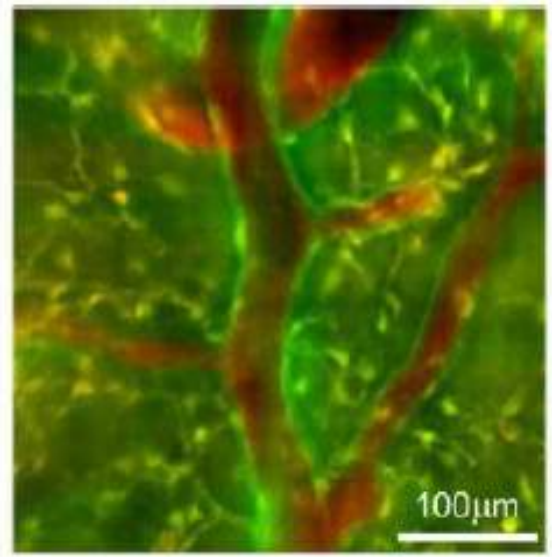
Partially Invasive Approach

- Two-photon calcium imaging
- https://elifesciences.org/articles/26839#content
- Based on the fact that electrical activity in neurons is typically associated with changes in calcium concentration.
- Two-photon calcium imaging allows us to observe the activity of multiple neurons up to ~500 µm below the cortical surface without cortical invasion.
- Photon calcium imaging involves:

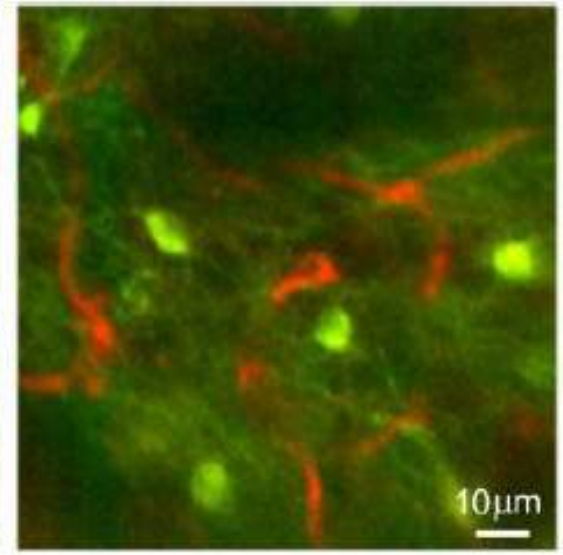
  (1) using pressure ejection to load neurons with fluorescent calcium-indicator dyes

  (2) monitoring changes in calcium fluorescence during neural activity using two-photon microscopy.





Oregon Green calcium sensitive dye stained neurons



Transgenic mouse expressing green fluorescent protein (GFP) in a subpopulation of neurons

Image from Kherlopian et al., 2008).

Brain Signal Acquisition

Non Invasive Approaches

Electroencephalography (EEG)

- EEG signals reflect the summation of postsynaptic potentials from many thousands of neurons that are oriented radially to the scalp.
- EEG predominantly captures electrical activity in the cerebral cortex, whose columnar arrangement of neurons and proximity to the skull favor recording by EEG.

Non Invasive Approaches

- The spatial resolution is typically poor (in the square centimeter range)
  - Due to lots of muscles between the source of signal and the electrodes placed on the scalp.
- The **temporal resolution is good** (in the milliseconds range)
- The measured signals are in the range of a few tens of microvolts, necessitating the use of powerful amplifiers and signal processing to amplify the signal and filter out noise.
- Artifacts in the EEG signal
  - eye movements, eye blinks, eyebrow movements, talking, chewing, and head movements

Non Invasive Approaches

- EEG recording involves the subject wearing a cap or a net into which the recording electrodes are placed
- A conductive gel or paste is injected into the holes of the cap before placing the electrodes.
- Over the years, new technology and innovations have introduced different types of electrodes. There are several types of electrodes: gel, water, and dry electrodes.

Gel electrodes:

- The gel electrodes are the most widely used type of electrodes. They are part of routine, clinical EEG recordings and have been the gold standard in EEG research for a long time.
- The electrode is commonly made of silver with a coating of silver chloride (Ag/AgCl).
- When a gel containing many chloride ions is applied between the skin and this electrode, conduction is improved and the skin-electrode interface impedance is reduced.
- Therefore, the gel between the skin and electrode allows for good-quality recording of biopotentials. These gel electrodes are disc-shaped and have a hole in the middle where gel can be applied with a syringe.
- The preparation time of a gel EEG cap requires time as the skin needs to be abrased and all electrodes need to be filled individually by a trained technician.

Gel electrodes:

Advantages of gel electrodes

- Allow for high-density EEG recordings.
- Very high signal quality.
- Less susceptible to mains interference and movement artifacts than dry and water electrodes.
- Stable recordings for a long time.
- It can be integrated with other research equipment (i.e. fNIRS)

Disadvantages of gel electrodes

- Skin needs to be prepared by lightly scratching the skin to reduce impedance.
- Inconvenient for researchers: Preparation time can be long, the head cap requires cleaning, and drying of the cap takes time.
- Inconvenient for participants: Hair needs to be cleaned and scratching the skin can feel uncomfortable.
- Requires a skilled technician.
- Conductive gel can dry out over time during recordings over 5 hours.

Gel electrodes:

Dry electrodes:

- Dry electrodes were first studied in the 90s, and are proposed as an alternative to overcome the common issues with wet electrodes.
- Dry EEG electrodes consist of an inert conductive material that mechanically couples with the skin for signal transduction, and eliminates the need for gel or skin preparation.
- Dry electrodes are composed using various materials and shapes, such as gold-plated electrodes, bristle-type electrodes, comb-like and multi-pin electrodes, silicone conductive rubber, or foam-based sensors.
- Since dry electrodes do not use a conductive gel or abrasive paste, there is a higher electrode impedance seen with dry electrodes than with wet electrodes. Also, this can lead to poor contact noise, increased signal instability, and more sensitivity to movement artifacts.

Dry electrodes:

Advantages of dry electrodes

- Quicker setup time than gel electrodes.
- Does not require skin preparation.
- Suitable for at-home testing.
- (Almost) no clean-up required.
- Possible without a trained technician in some situations.

Disadvantages of dry electrodes

- Difficulty keeping electrodes affixed onto the skin.
- Increased signal instability and higher impedances.
- More susceptible to mains interference and movement artifacts than gel electrodes.
- Limited actions are possible to improve the skin-electrode contact quality.
- Uncomfortable to the wearer.

Dry electrodes:

Water electrodes:

- Water electrodes are a novel type of electrodes, that like dry electrodes, have a really short preparation time and do not require the use of a conductive gel.
- These electrodes can also be called **semi-dry electrodes**. The key feature of water electrodes is that they use water or an electrolyte liquid. Some water electrodes use water sponges with tap water or saline water to increase the conduction between the skin surface and the electrode.
- Other water electrodes slowly and continuously release a tiny amount of electrolyte liquid to the scalp in a contained matter.
- Since these methods do not require the application of gel or abrasion of the skin, it also has a faster preparation time and clean-up than gel electrodes.

Water electrodes:

Advantages of water electrodes

- Quicker setup time than gel electrodes.
- Quicker clean-up than gel electrodes.
- Suitable for at-home testing.
- Possible without a trained technician in some situations.
- Overcomes problems with high impedance and signal instability seen in dry electrodes with water.

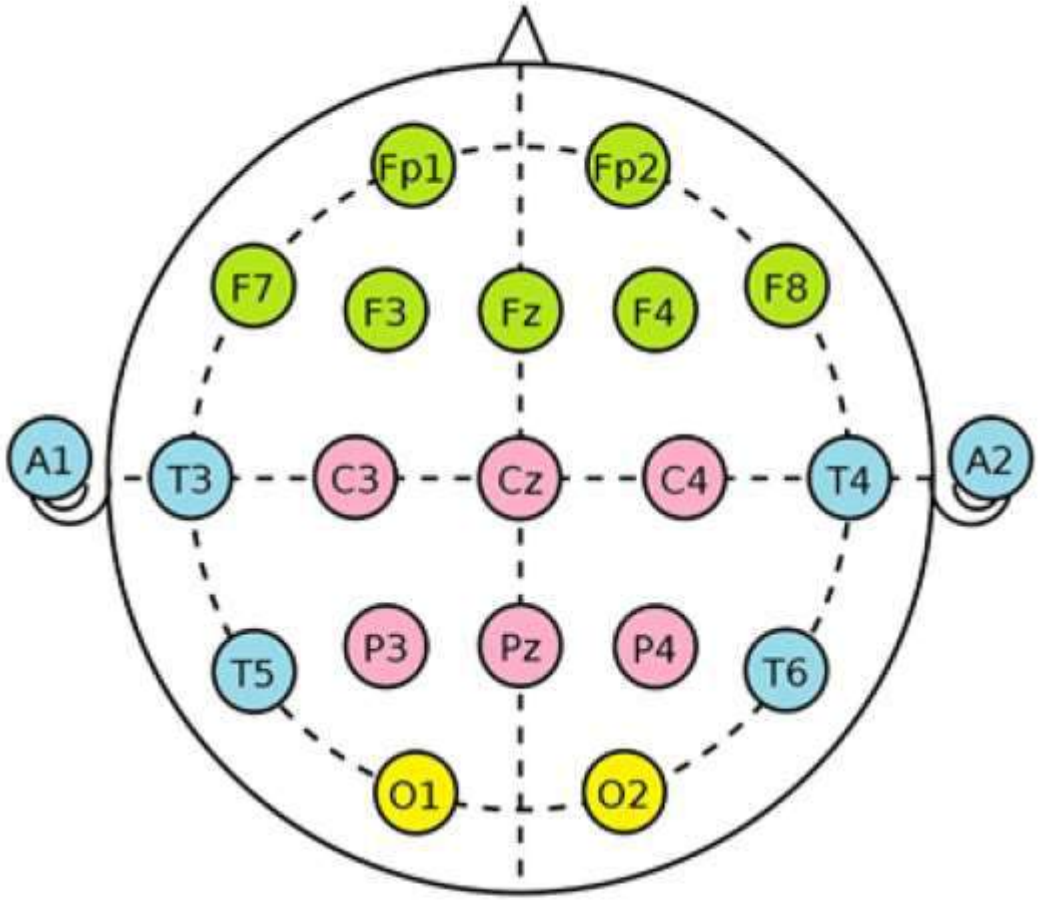
Disadvantages of water electrodes

- Dry out quicker compared to gel electrodes, so they need to be remoistened more often.
- More susceptible to mains interference and movement artifacts than gel electrodes.
- Limited actions are possible to improve the skin-electrode contact quality.

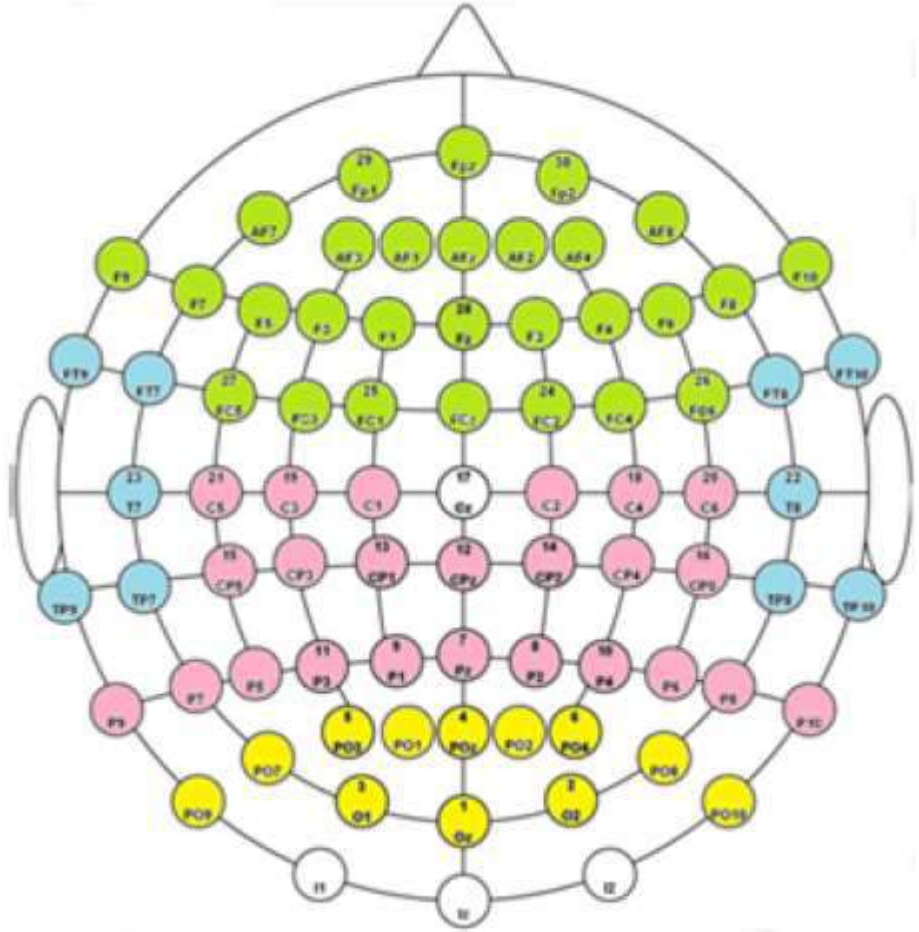
Water electrodes:

International 10–20 system

- The international 10–20 system is a convention used to specify standardized electrode locations on the scalp.
- C = central, P = parietal, T = temporal, F = frontal, Fp = frontal polar, O = occipital, A = mastoids
- Odd numbers on the left side and even numbers on the right side.



10-20 Electrode System



10-10 Electrode System

Frontal Lobe

Temporal Lobe

Parietal Lobe

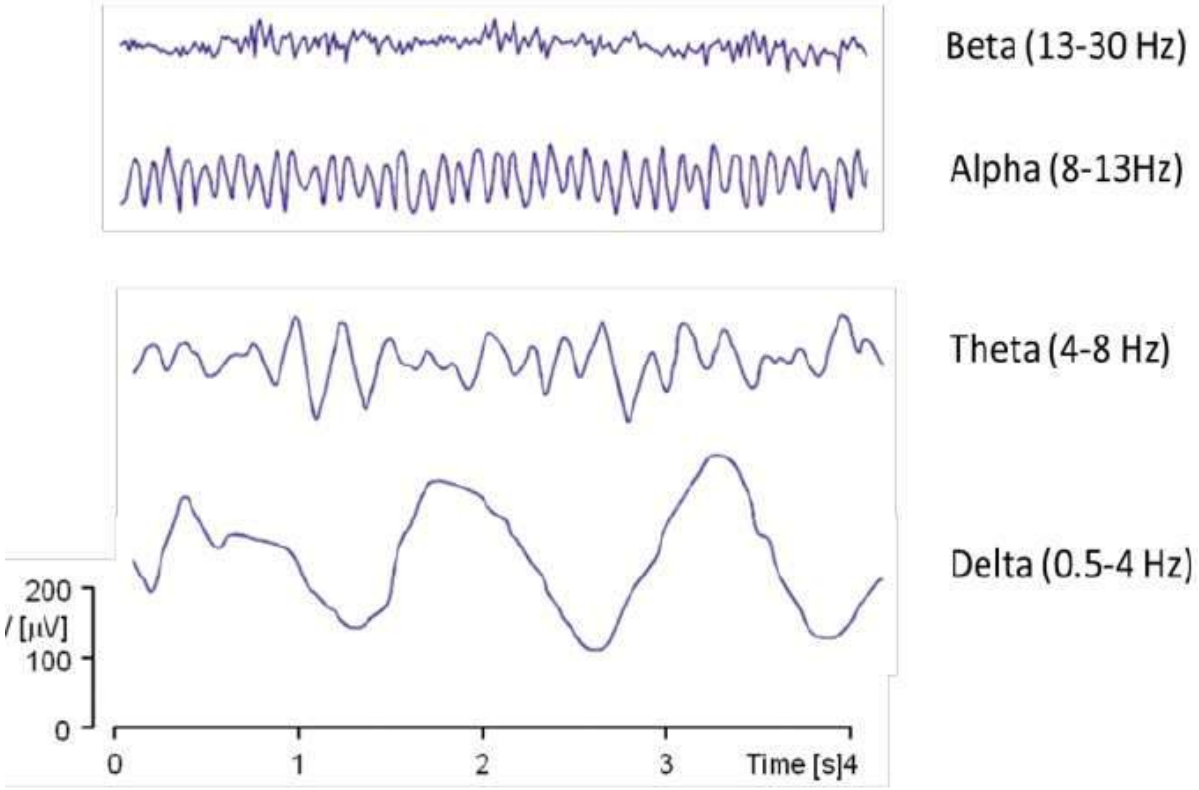
Occipital Lobe

International 10–20 system

- The mastoids reference electrode locations behind each ear (A1 and A2).
- Other reference electrode locations are nasion, at the top of the nose, level with the eyes; and inion, at the base of the skull on the midline at the back of the head.
- In a typical setup, each EEG electrode is connected to one input of a differential amplifier, and the other input is connected to a reference electrode

- The amplification of voltage between the active electrode and the reference is typically 1,000–100,000 times.
- The amplified signal is passed through a filter and then digitized via an A/D (analog to digital) converter.
- After digitization, the EEG signal may be additionally filtered by a 1–50 Hz bandpass filter.
  - Excludes noise and movement artifacts in the very low and very high frequency ranges.

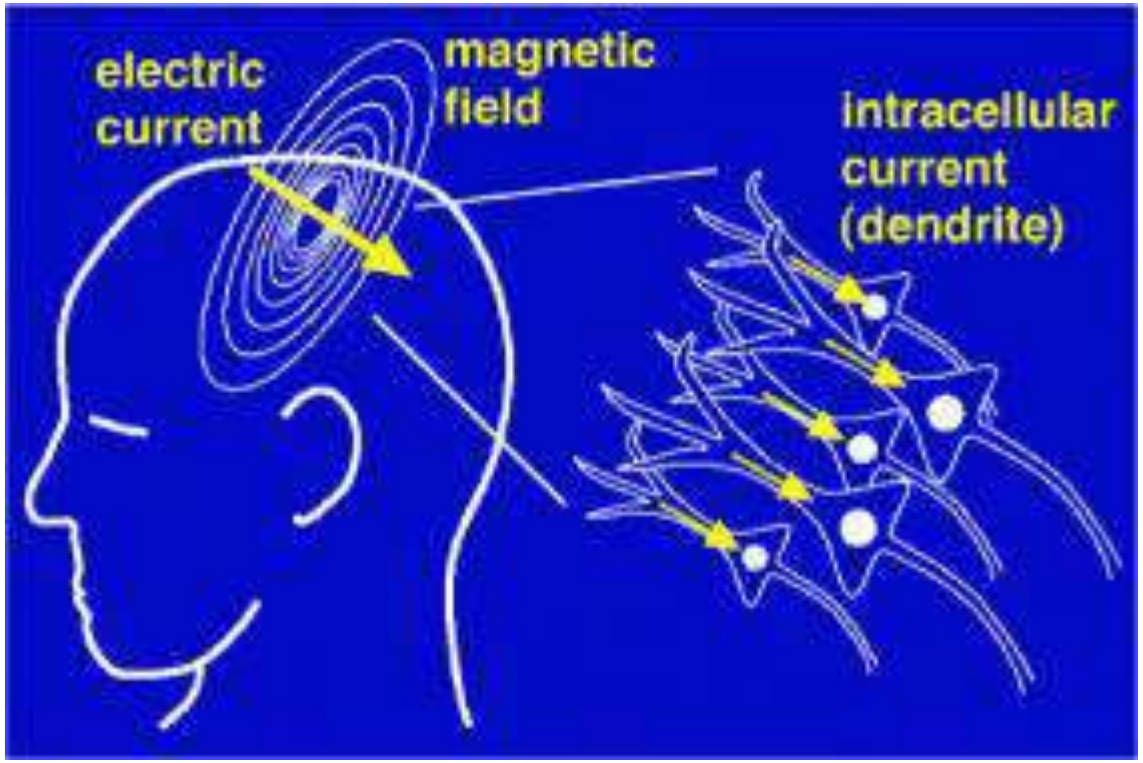
- EEG recordings are well-suited to capturing oscillatory brain activity or “brain waves” at a variety of frequencies
  - Alpha waves (8 to 13 Hz)
  - Beta waves (13 to 30 Hz)
  - Delta waves (0.5-4 Hz)
  - Theta waves (4-8 Hz)
  - Gamma waves (30-100 Hz or more)



Non Invasive Approaches

Magnetoencephalography (MEG):

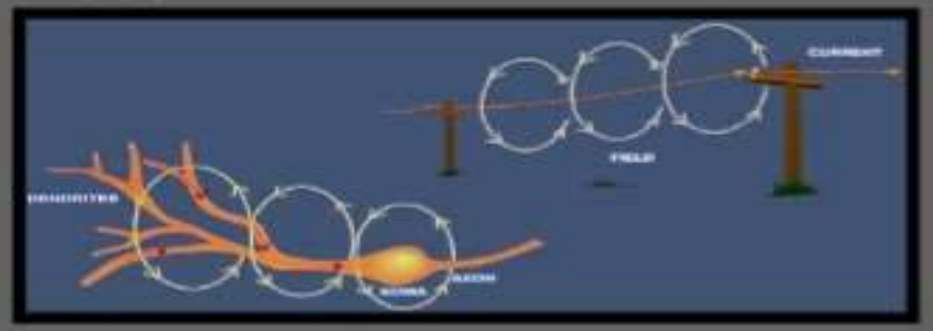
- Measures magnetic fields produced by activity of thousands of cortical neurons oriented perpendicular to the cortical surface
- Magnetic fields not distorted by skull and scalp
- Better spatial resolution than EEG
- Expensive and bulky
- Magnetically shielded rooms



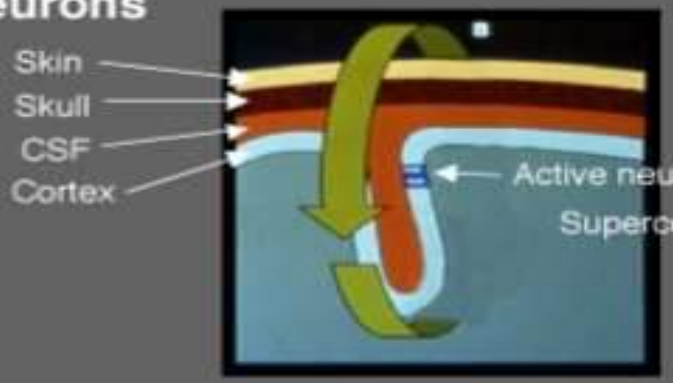
(image A: Wikimedia Commons;
image B: http://dateline.ucdavis.edu/photos_images/dateline_images/040309/DondersMEGOle_W2.jpg).

Basic Principles of MEG

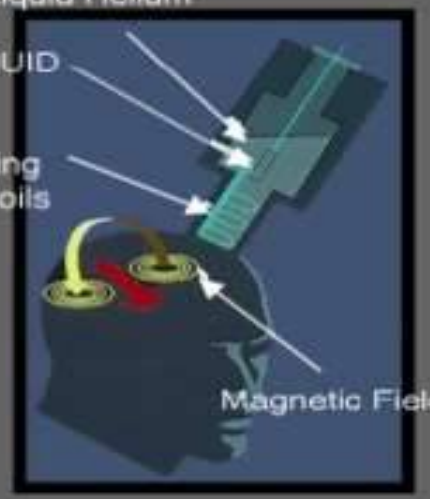
Sources of Magnetic Fields



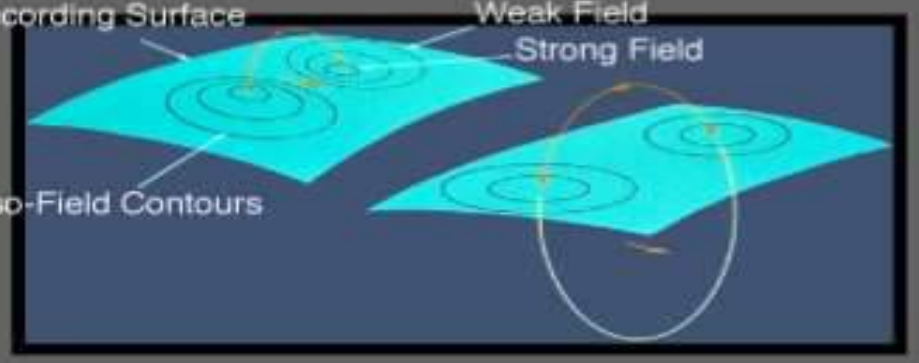
Orientation of Neurons



Detection Device

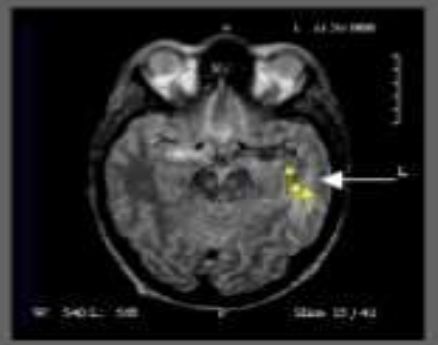


Magnetic Field Pattern

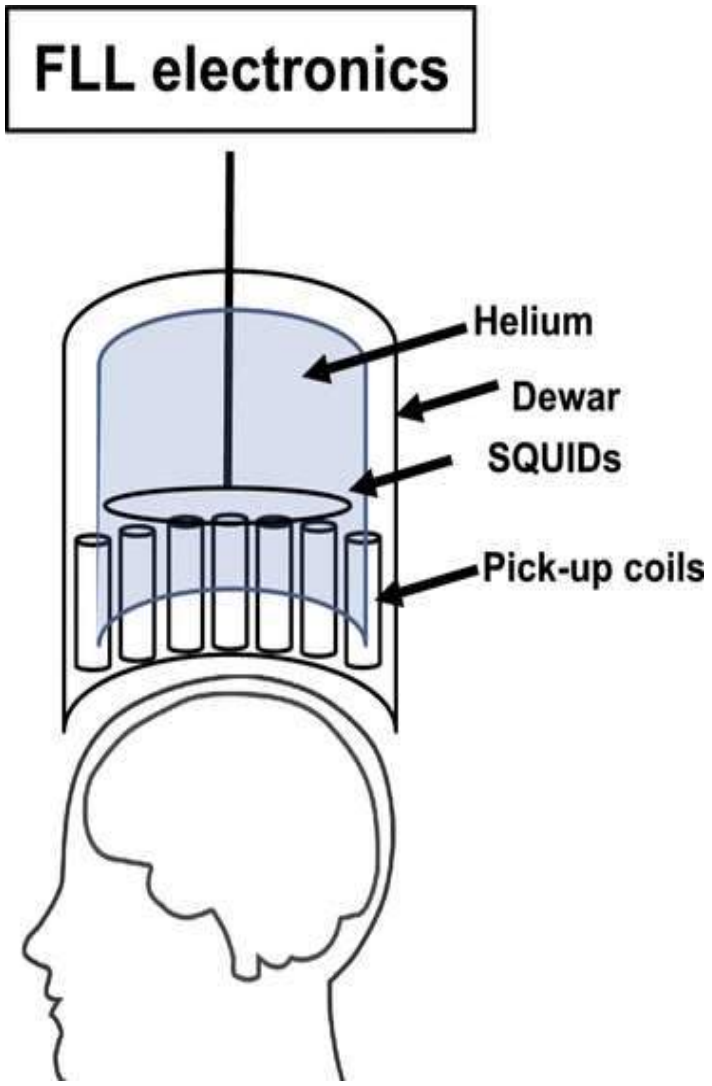


Model

Result



Sources of epileptic spikes



- MEG has very exquisitely sensitive sensors, filtering, and a method to shield the sensors from recording outside noise.
- conventional MEG systems use sensitive magnetic field sensors called superconducting quantum interference devices or SQUIDs.
- SQUIDs are made from materials that become superconductors at extremely low temperatures, meaning that the material can conduct electricity without resistance.
- Many of the MEG systems today use niobium for the SQUIDs because it can reliably reach a superconductive state at low temperatures and return back to room temperature.
- Pick-up coils are linked to the SQUID by an input coil and are kept in a superconducting state by liquid helium at −269°C.
- The SQUID then produces a small voltage current that can be detected through what is called a flux-locked loop (FLL) electronics system.
- The electrical output can then be transformed into a digital signal through optical cables. This output is usually displayed on a computer system connected to the MEG.

Source: https://journals.physiology.org/doi/full/10.1152/jn.00530.2020

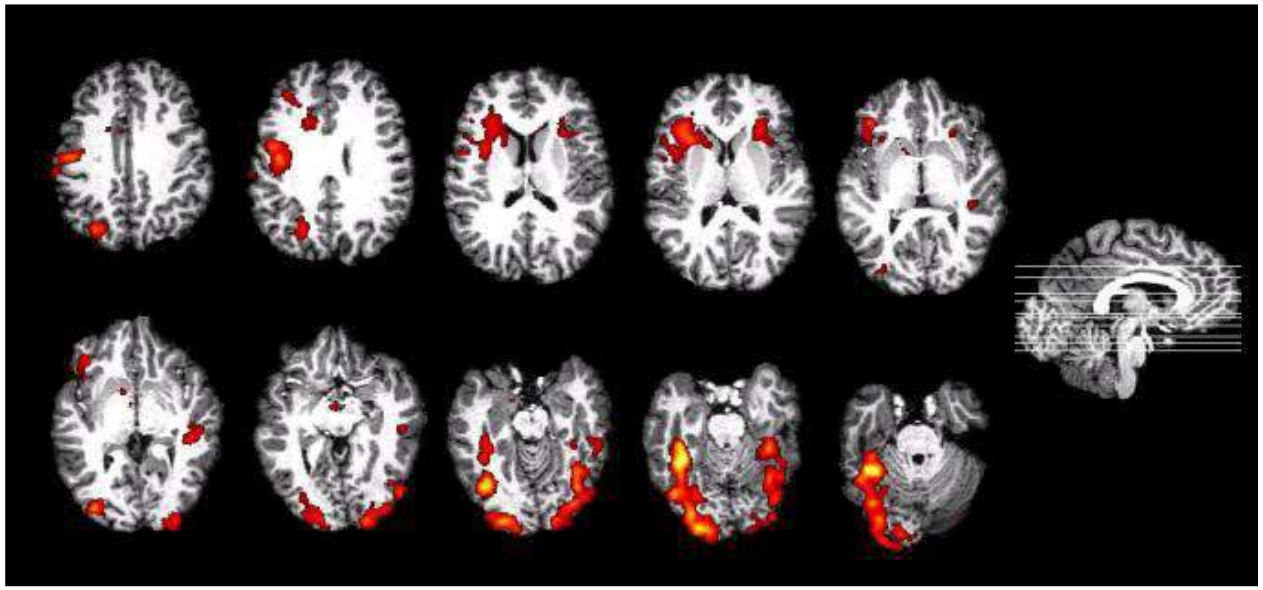
Non Invasive Approaches

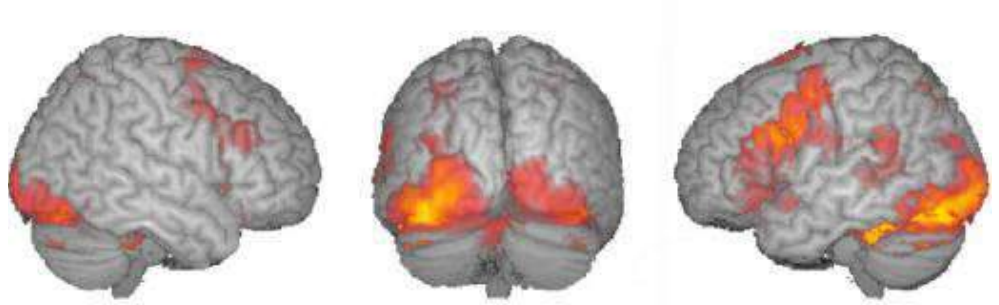
Functional Magnetic Resonance Imaging (fMRI) :

- Measures changes in blood flow due to increased activation of neurons in an area
- Relies on paramagnetic properties of oxygenated and deoxygenated hemoglobin in the blood
- Produces images showing blood-oxygenation-level-dependent signal changes (BOLD)

Non Invasive Approaches

- In the 1890s it has been known that changes in blood flow and blood oxygenation in the brain (collectively known as **hemodynamics**) are closely linked to neural activity.
- When neurons become active, local blood flow to those brain regions increases, and oxygen-rich (oxygenated) blood displaces oxygen-depleted (deoxygenated) blood. Oxygen is carried by the hemoglobin molecule in red blood cells.
- Deoxygenated hemoglobin is more magnetic (paramagnetic) than oxygenated hemoglobin (Hb), which is virtually resistant to magnetism (diamagnetic).

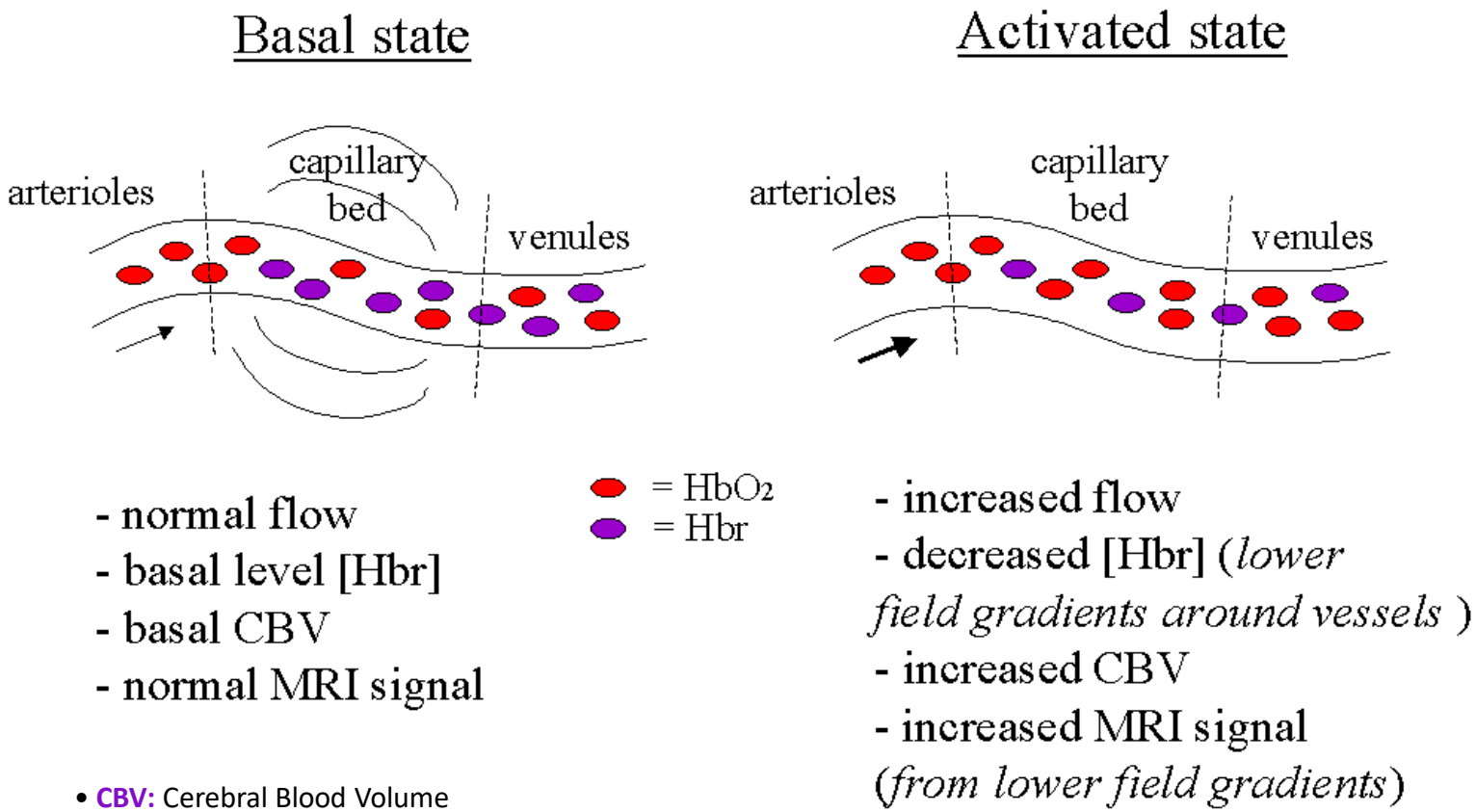




Example fMRI Images (word reading task)

Blood Oxygen Level Dependent (BOLD) Signal

↑neural activity ➔ ↑ blood flow ➔ 🡻 deoxyhemoglobin ➔ ↑ MR signal



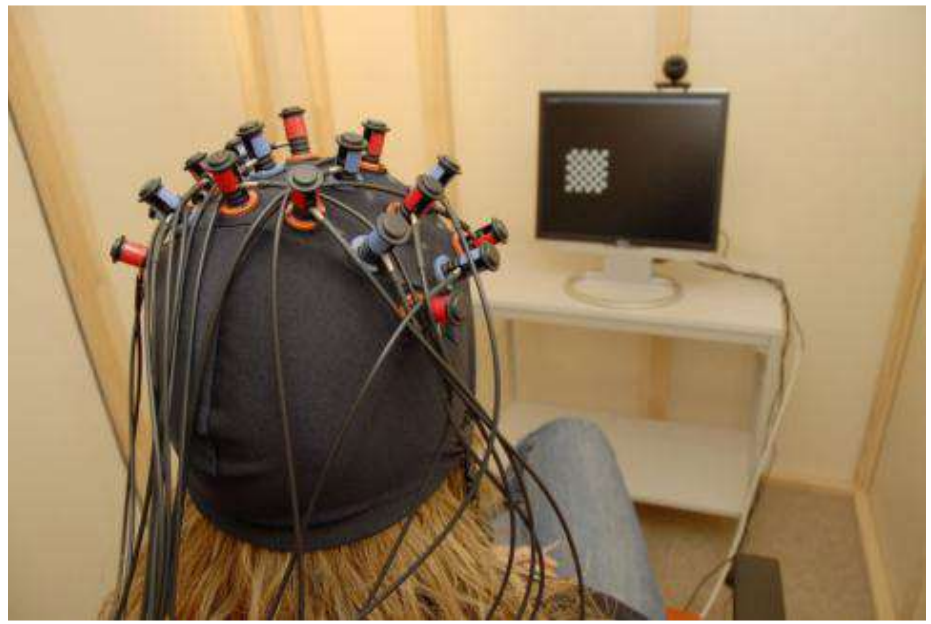
- **CBV:** Cerebral Blood Volume
- **CBF:** Cerebral Blood Flow
- **HBr:** Deoxy- Hemoglobin

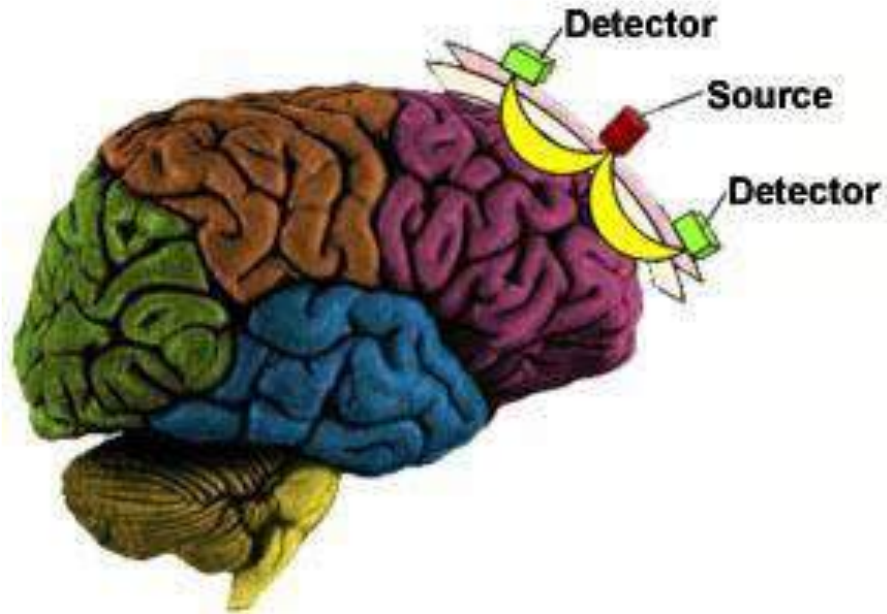
Source: Jorge Jovicich

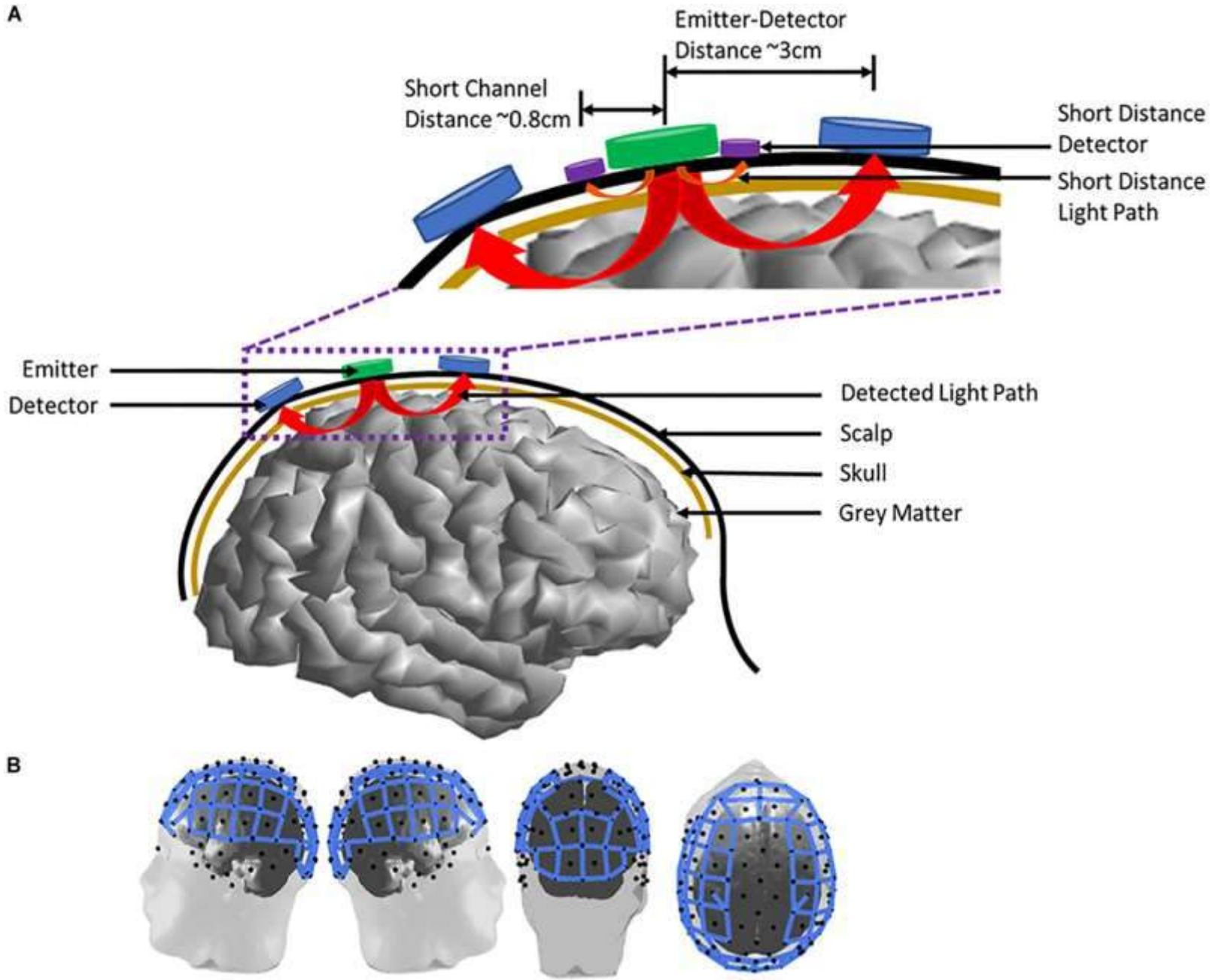
Non Invasive Approaches

Functional Near-Infrared Spectroscopy (fNIRS)

- Measures change in blood oxygenation level caused by increased neural activity in the brain.
- Based on detecting near-infrared light absorbance of hemoglobin in the blood with and without oxygen.
- Maps neural activity using "*optodes*" (emitters and detectors)




Detector
Source
Detector



Source: https://www.frontiersin.org/articles/10.3389/fnins.2020.00724/full

Non-Invasive Approaches

- **Positron Emission Tomography (PET):**
- Measures emissions from radioactively labeled, metabolically active chemicals that have been injected into the bloodstream for transportation to the brain.
  - The labeled compound is called a *radiotracer*.

- Sensors in the PET scanner detect the radioactive compound
  - As a result of metabolic activity caused by brain activity.
- Generate two-or three-dimensional images indicating the amount of brain activity.

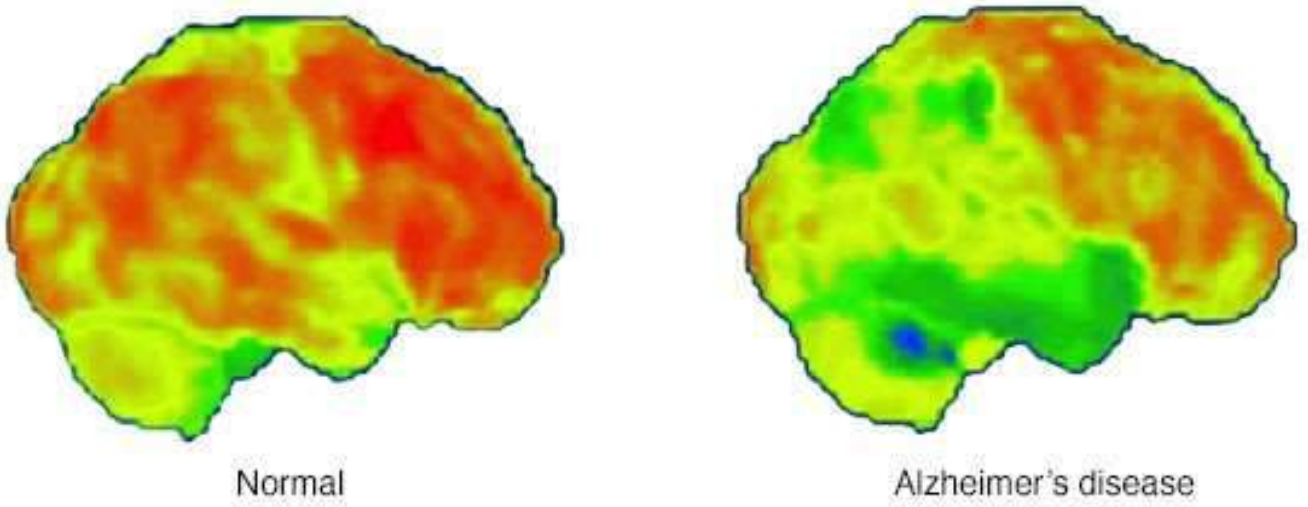
Non-Invasive Approaches

- **Positron Emission Tomography (PET):**

What does a PET scan check for?

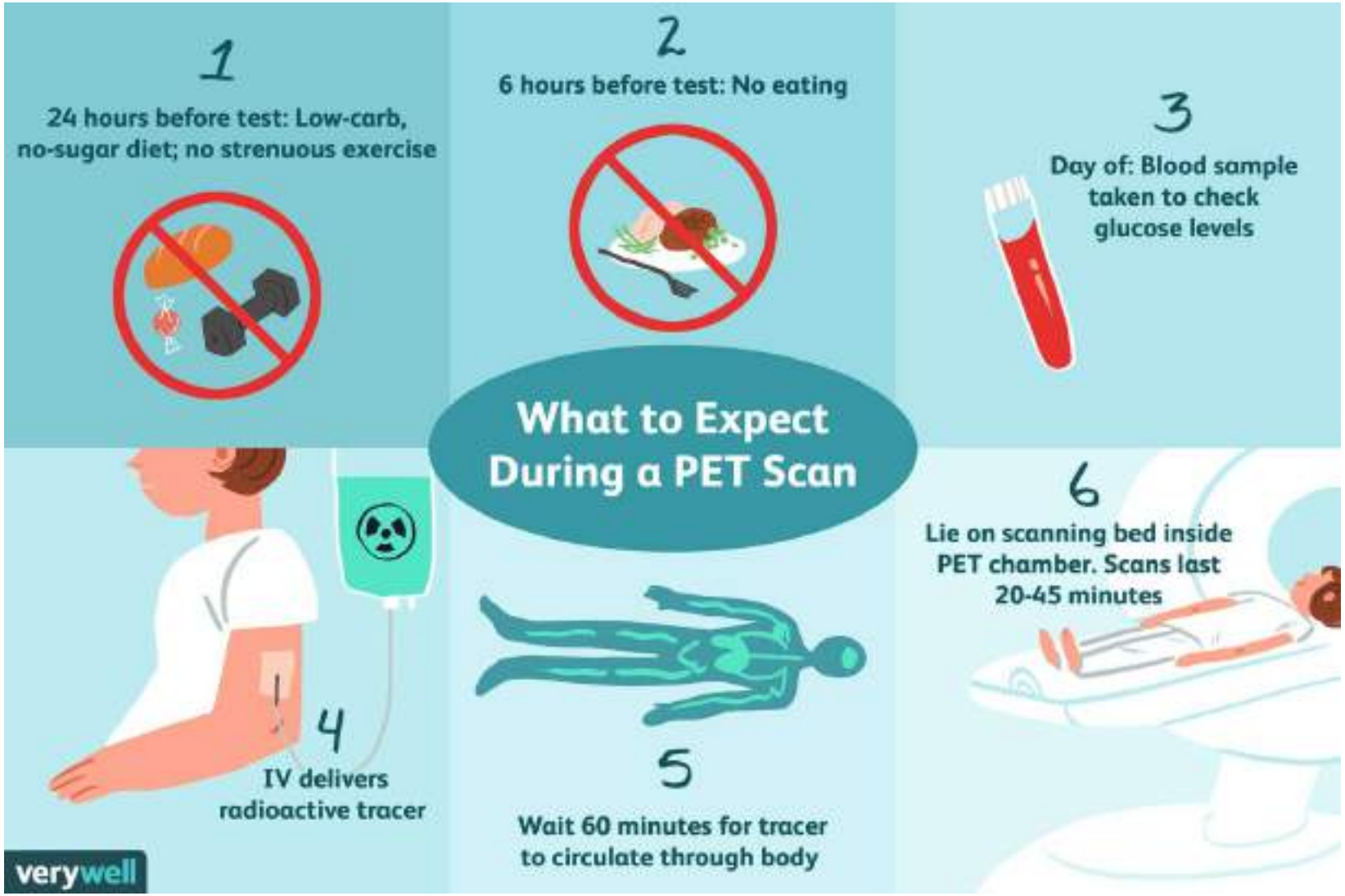
- Cancer, including breast cancer, lung cancer and thyroid cancer.
- Coronary artery disease, heart attack or other heart problems.
- Brain disorders, such as brain tumors, epilepsy, dementia and Alzheimer's disease.

Positron Emission Tomography (PET):



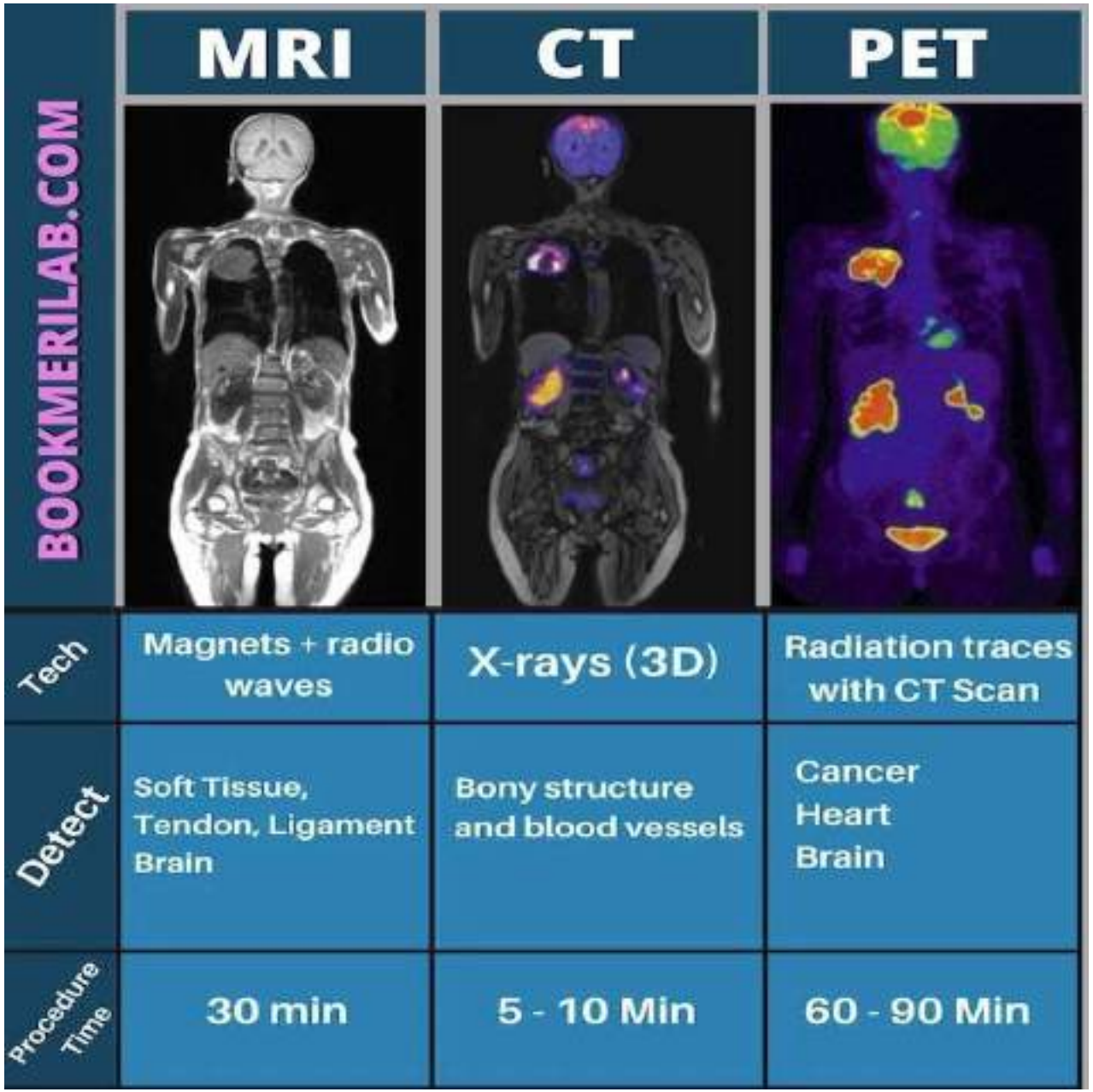
A PET scan can compare a normal brain (left) with one affected by Alzheimer's disease (right). An increase in blue and green colors shows decreased brain metabolic activity due to Alzheimer's disease.

Positron Emission Tomography (PET):



Source: Internet

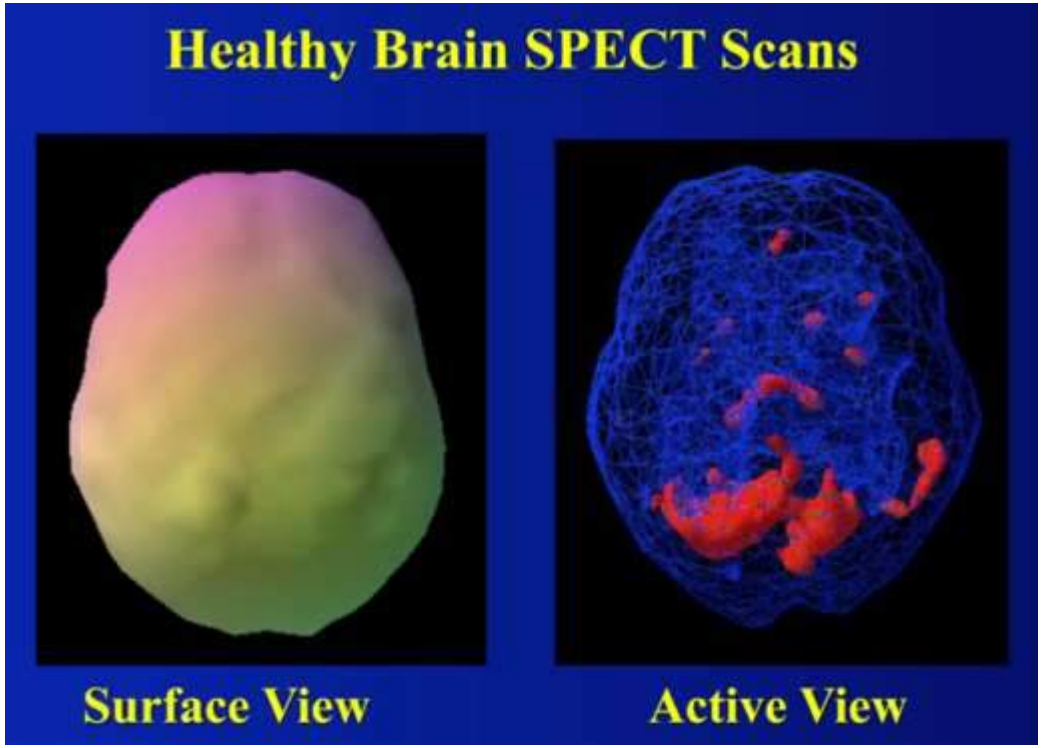
Comparison between MRI, CT and PET Scan



Source: Internet

Non Invasive Approaches

- **Single-photon emission-computed tomography (SPECT):**
- SPECT is a nuclear medicine technique that uses gamma rays to study the brain.
- A radioactive substance is injected into the patient's body and is scanned using a SPECT machine.
- Allows doctors to see how blood flows into tissues and organs.
  - Active, inactive, or overactive.
- Averages the brain activity over a few minutes and generates an image.



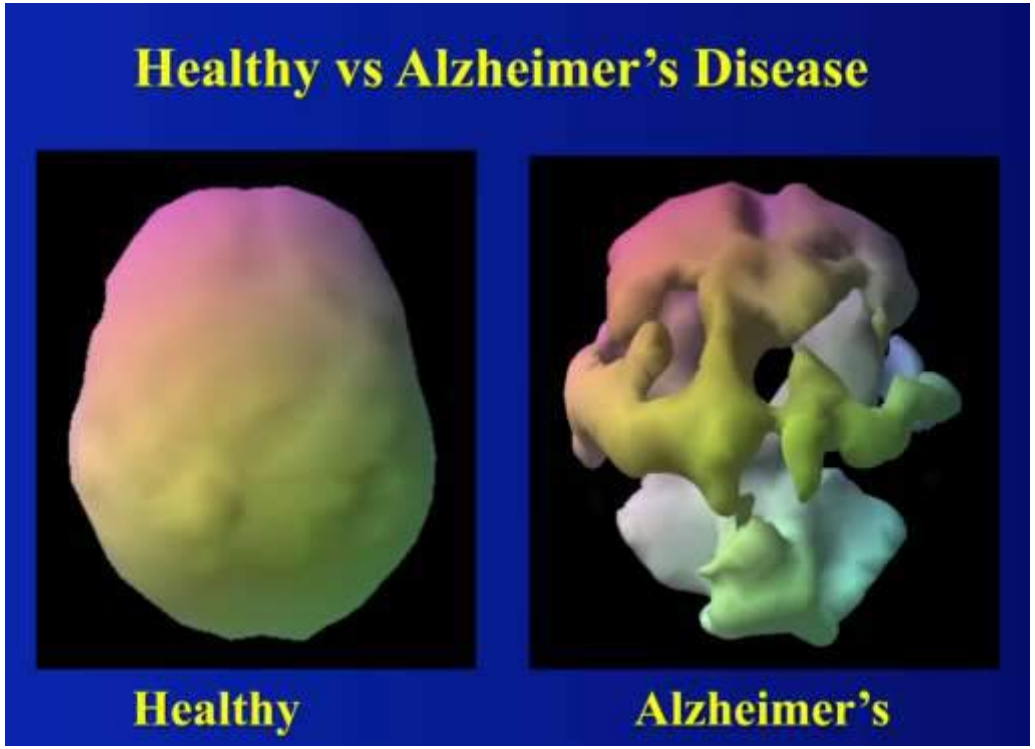


Image credit: https://hamnaqasim.medium.com/brain-scans-a-lesson-to-learn-9dfe67abd841

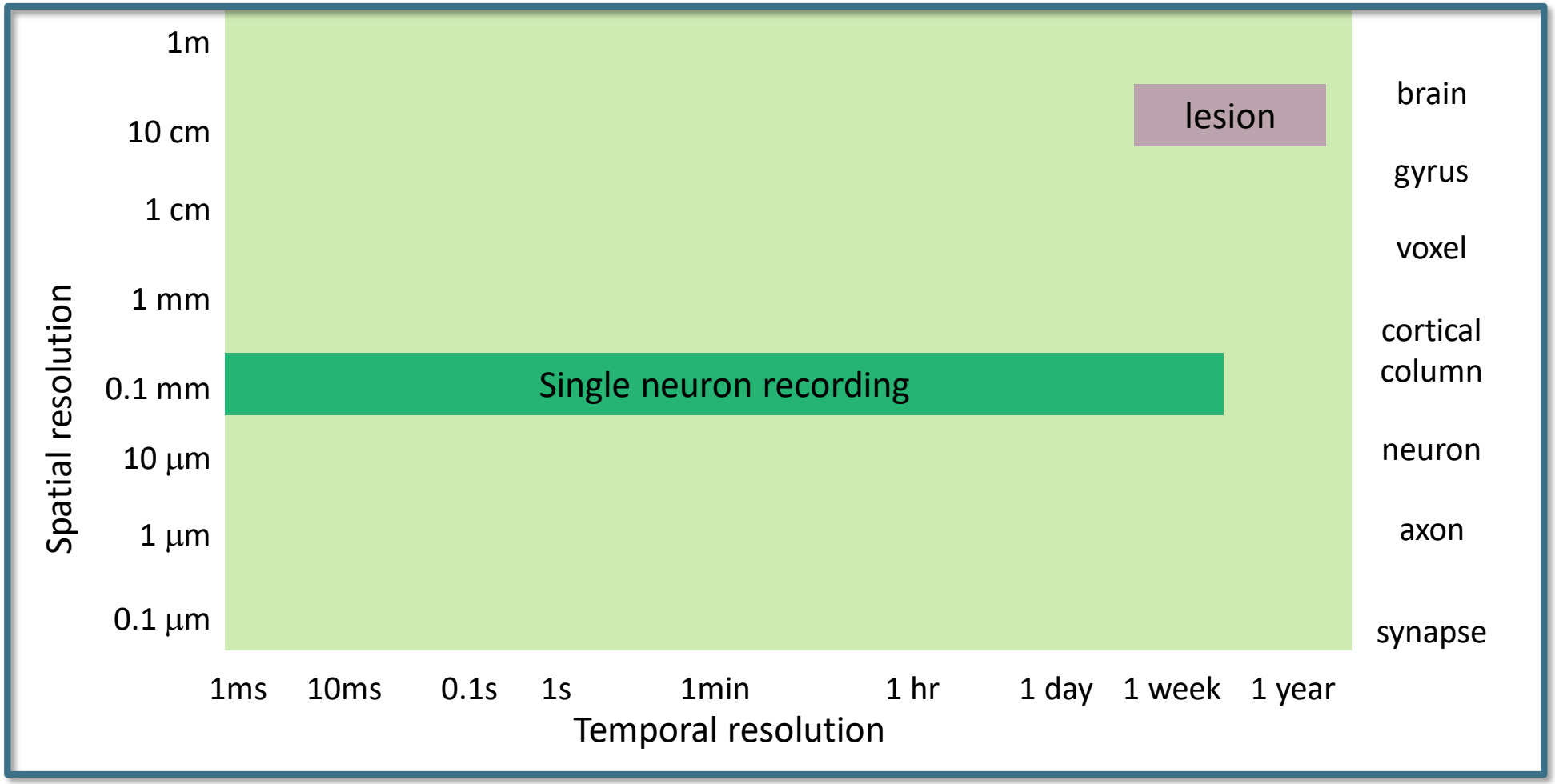
What spatial and temporal scale is relevant for studying systems neuroscience questions?

Adapted from: PS Churchland and TJ Sejnowski, Perspectives on cognitive neuroscience, 1988 Science, Vol 242, Issue 4879, 741-745

What spatial and temporal scale is relevant for studying systems neuroscience questions?

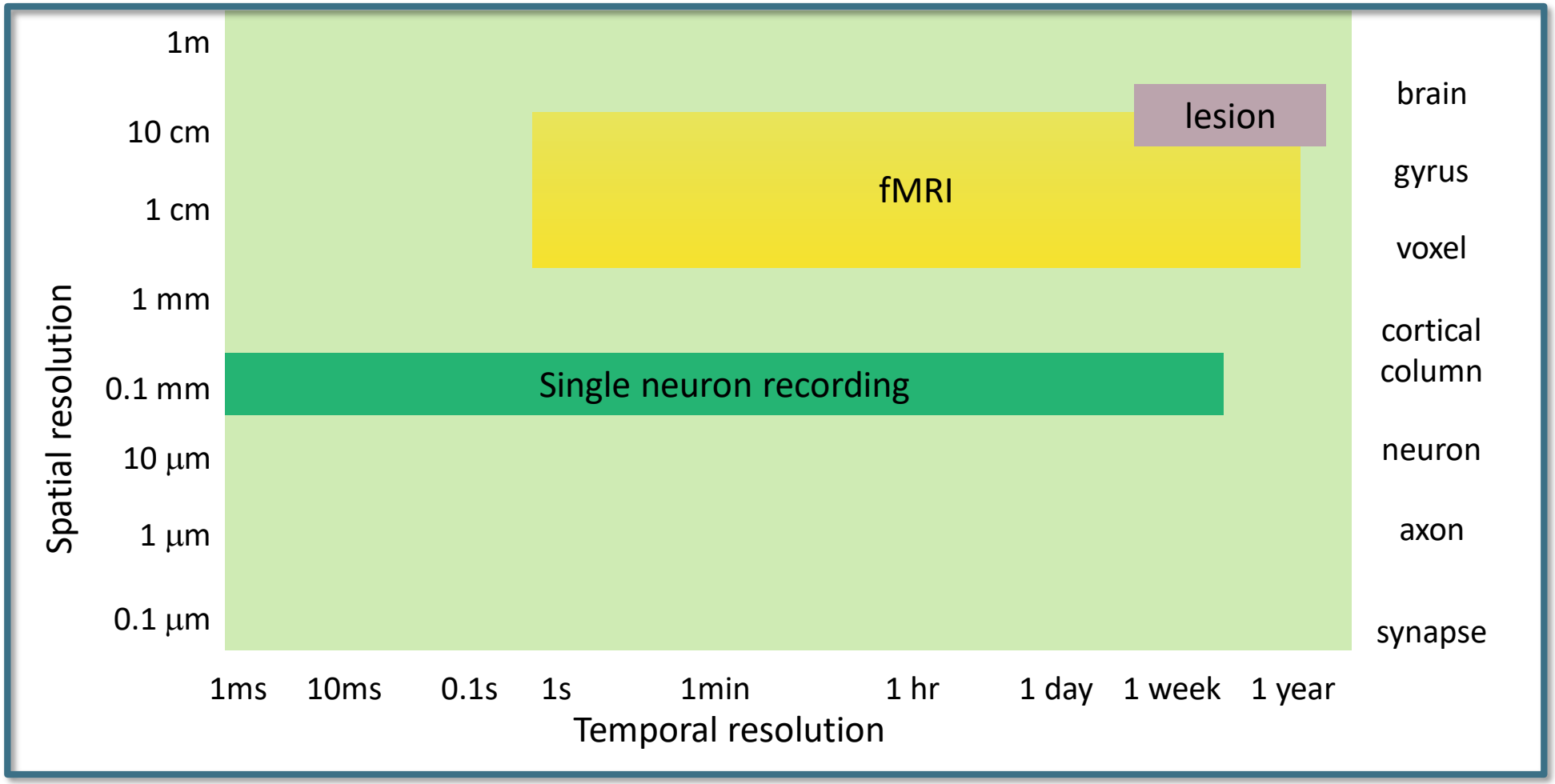
Adapted from: PS Churchland and TJ Sejnowski, Perspectives on cognitive neuroscience, 1988 Science, Vol 242, Issue 4879, 741-745

What spatial and temporal scale is relevant for studying systems neuroscience questions?



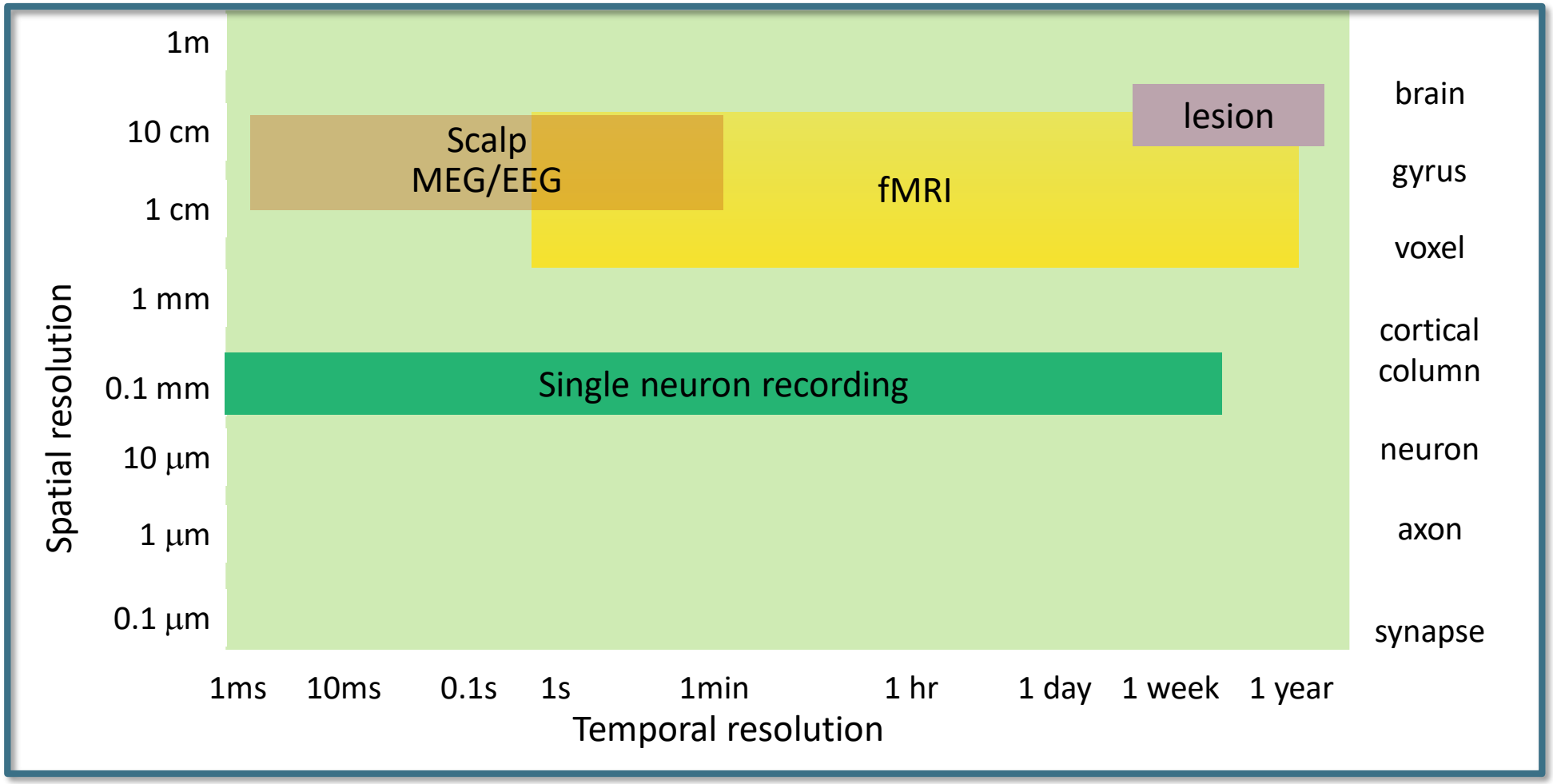
Adapted from: PS Churchland and TJ Sejnowski, Perspectives on cognitive neuroscience, 1988 Science, Vol 242, Issue 4879, 741-745

What spatial and temporal scale is relevant for studying systems neuroscience questions?



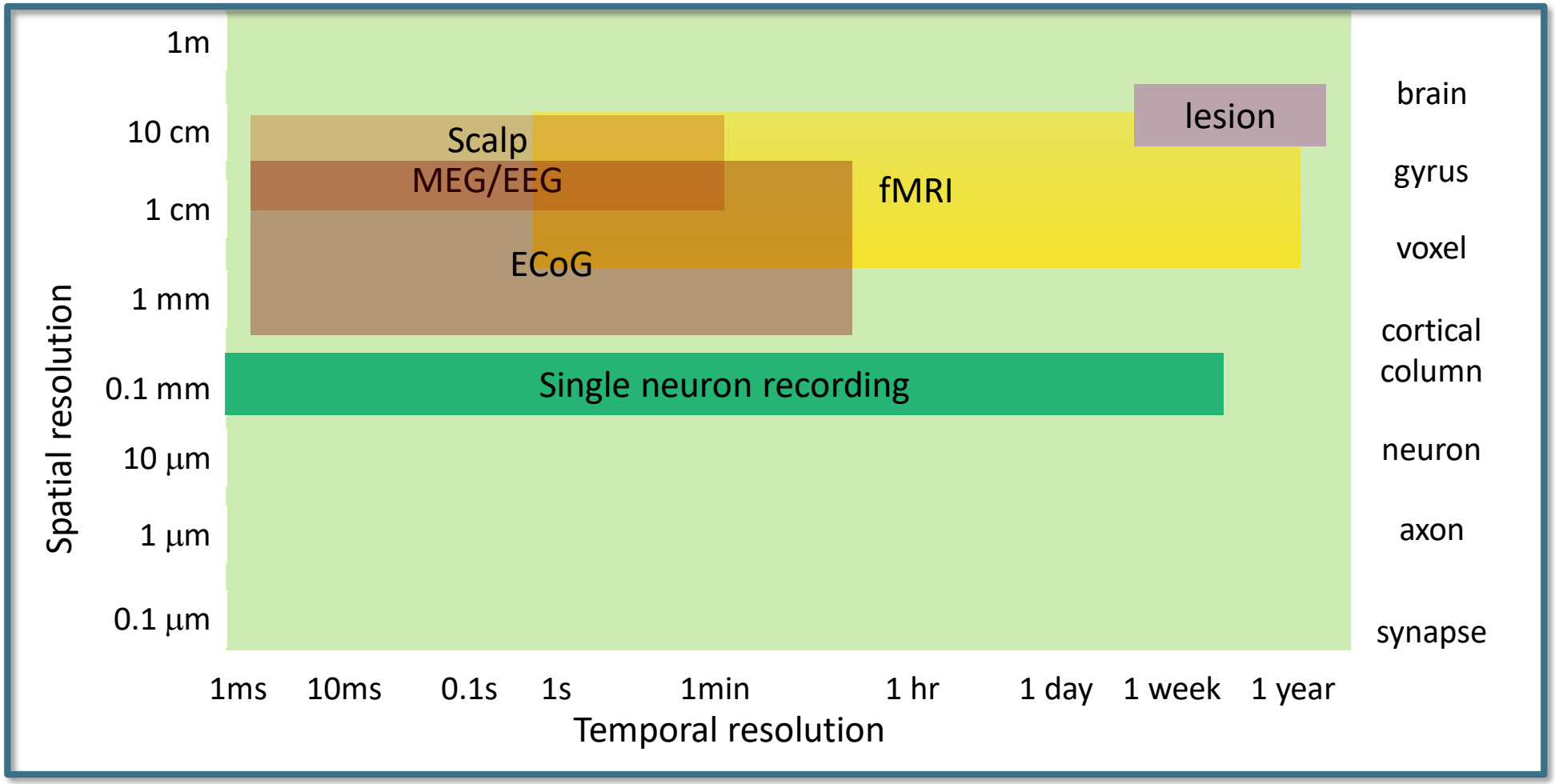
Adapted from: PS Churchland and TJ Sejnowski, Perspectives on cognitive neuroscience, 1988 Science, Vol 242, Issue 4879, 741-745

What spatial and temporal scale is relevant for studying systems neuroscience questions?



Adapted from: PS Churchland and TJ Sejnowski, Perspectives on cognitive neuroscience, 1988 Science, Vol 242, Issue 4879, 741-745

What spatial and temporal scale is relevant for studying systems neuroscience questions?



Adapted from: PS Churchland and TJ Sejnowski, Perspectives on cognitive neuroscience, 1988 Science, Vol 242, Issue 4879, 741-745

Brain Computer Interaction

EEG Artifacts

Course Instructors

Dr. Sreeja S R

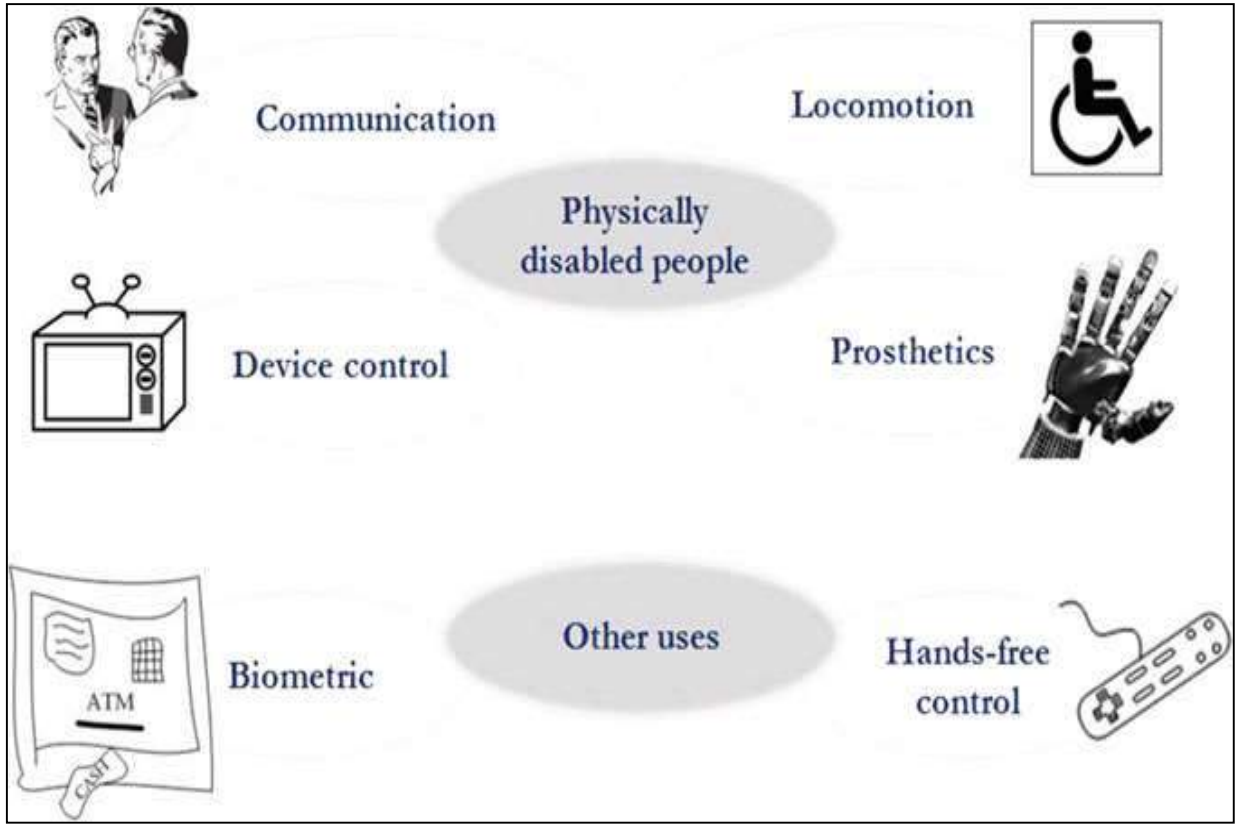
Assistant Professor

Indian Institute of Information Technology

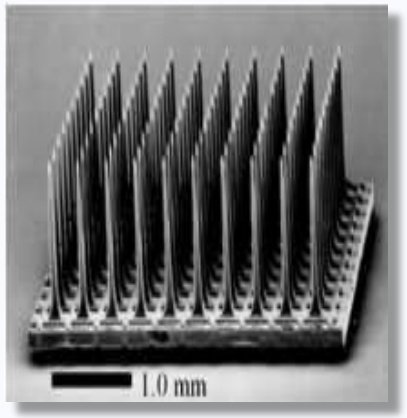
IIIT Sri City

What is BCI?

A Brain Computing Interface (BCI) is a “communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles” [Wolpaw, 2002].



BCI Devices



Sub Electrode Array

ECoG

Functional MRI

Positron Emersion Tomography

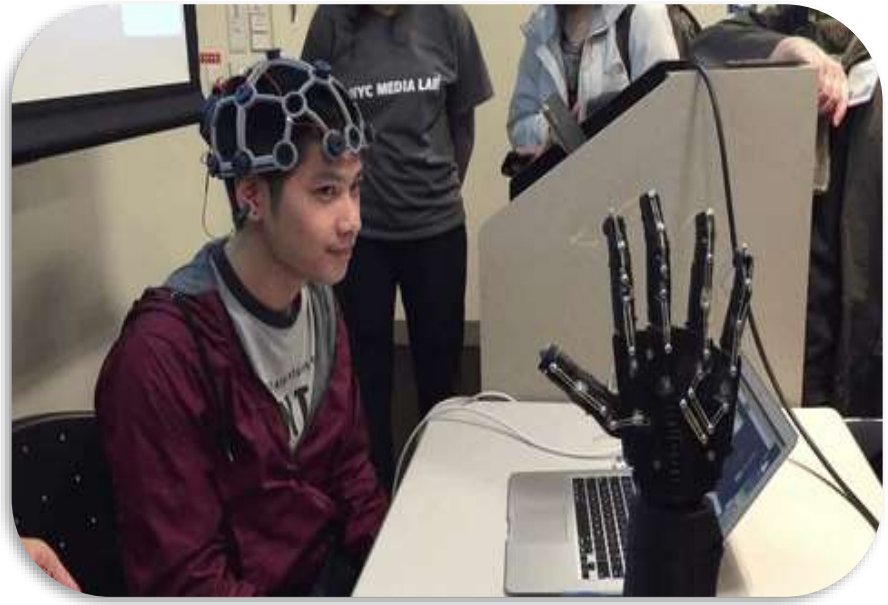
MEG

Functional NIRS

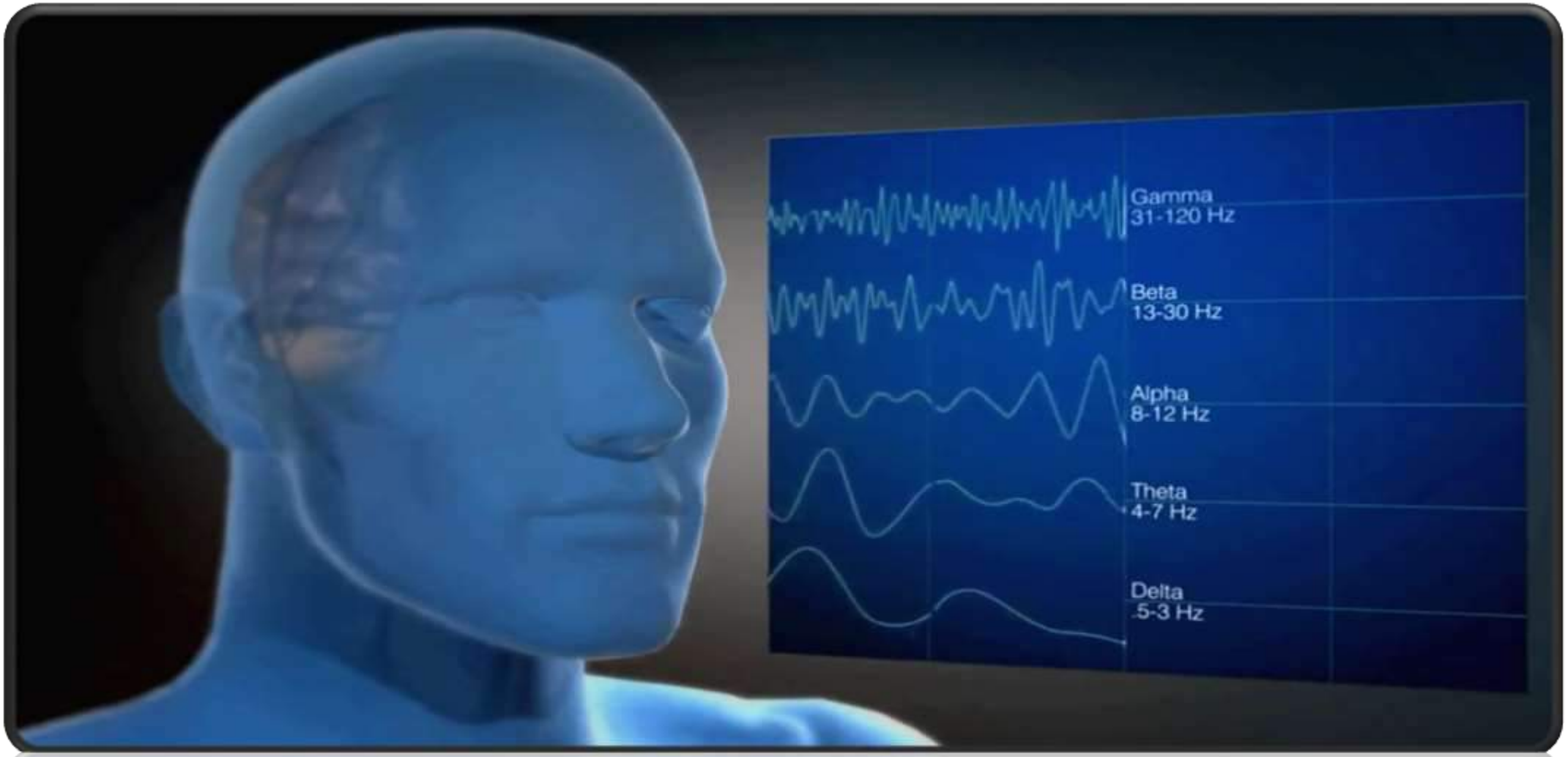
EEG

Why EEG?


Less expensive
Simple way to capture the brain signal accurately
Wireless
Non-invasive
User-friendly
EEG devices are



Signal bands present in the EEG signal



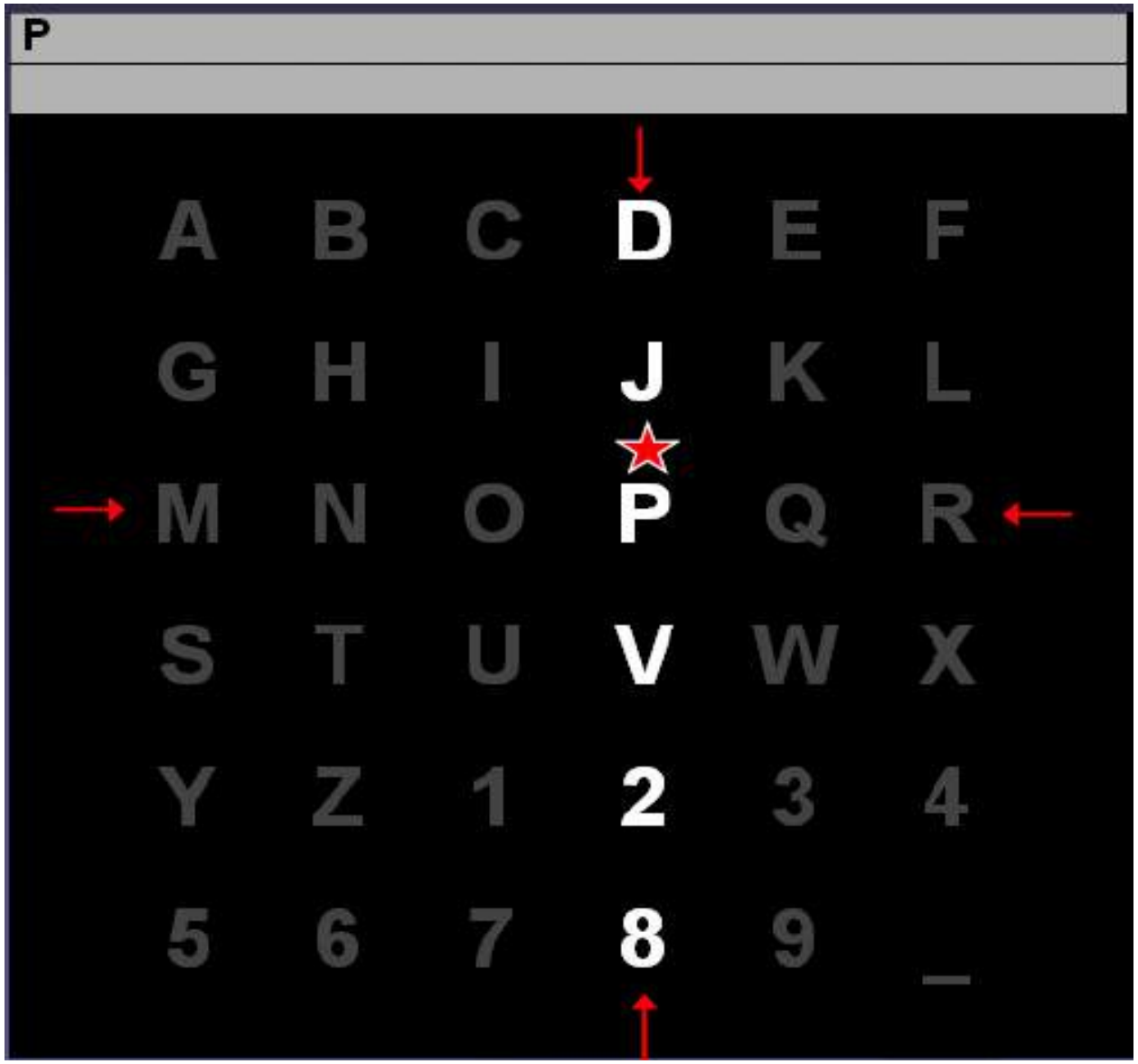
EEG Paradigms

01 P300

02 SCPs

03 SSVEPs

04 SMRs



Useful Links

https://www.frontiersin.org/articles/10.3389/fnhum.2020.583358/full

https://www.youtube.com/watch?v=9SIBtW1QqT8

https://www.youtube.com/watch?v=XIr2cRKFolY

https://www.youtube.com/watch?v=z9_R2YM48Jk

https://www.youtube.com/watch?v=8bgvbYVJplM

https://www.youtube.com/watch?v=U_WxaDHNw6I

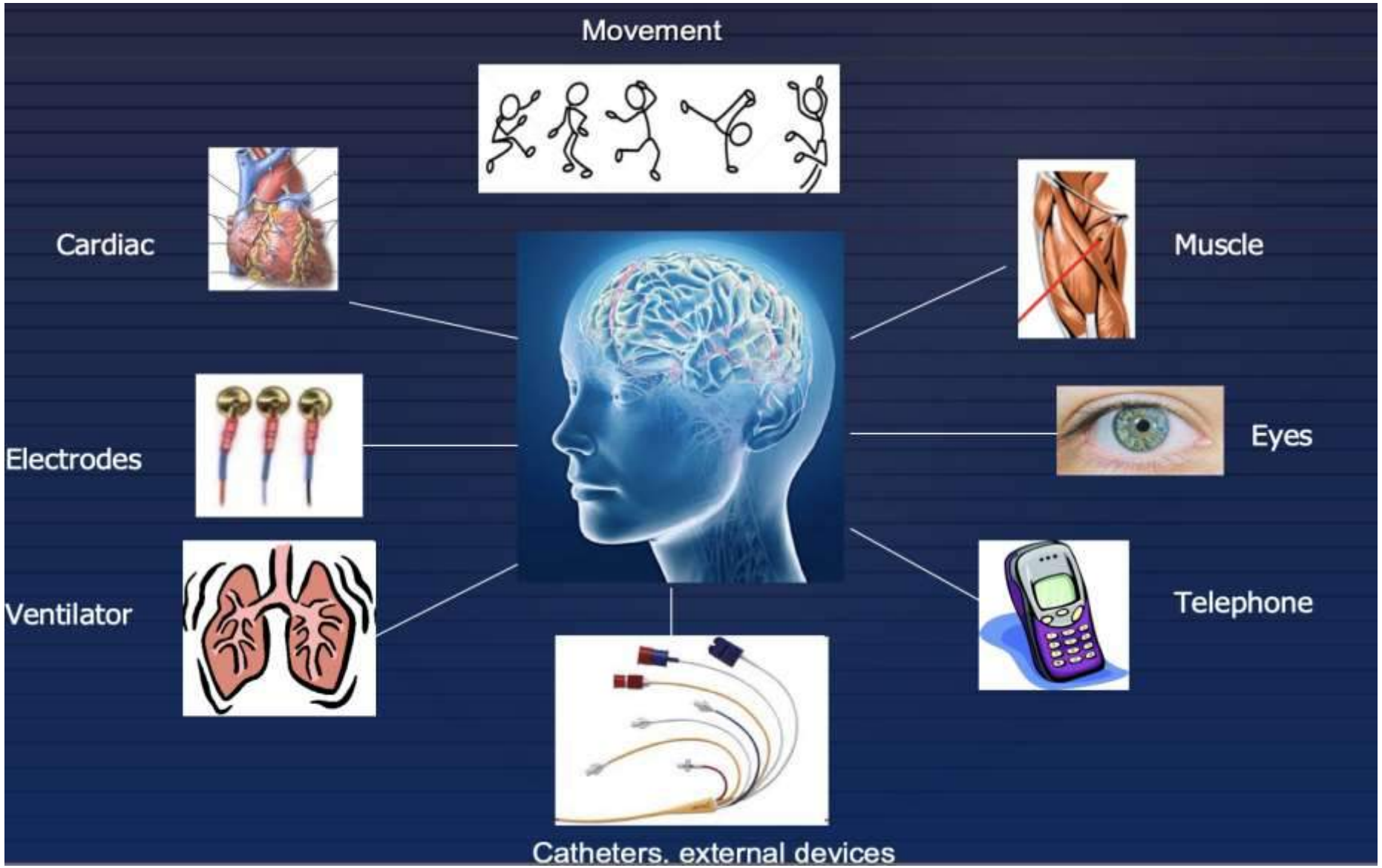
EEG Artifacts

- The electrical artifacts that is not of cerebral origin.
- Anything that is NOT of cerebral origin is termed as **ARTIFACT**
- Physiological and Electrophysiological artifacts.
- Physiological – source (generated other than brain ie. Body)
- Electrophysiological – arise outside the body - equipment and environment
- Some readily distinguished, others closely resemble cerebral activity.

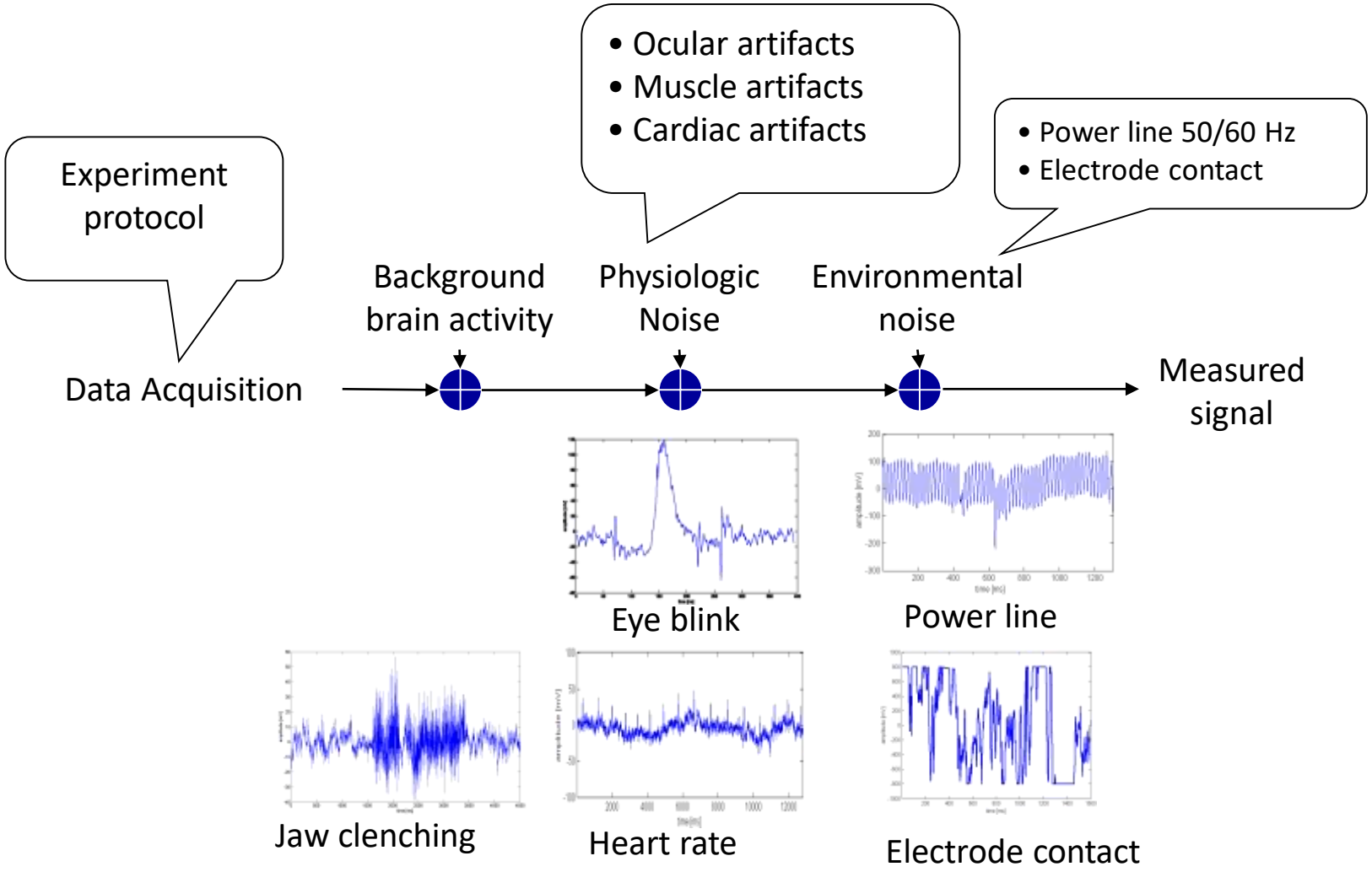
KEY TO AN ARTIFACT FREE RECORDING

- Good, clean preparation
- Good hook-up, neatly bundled electrodes
- Place jack-box close to patients head
- Keep the subject cool, not cold
- Unplug all electrical items close to patient, i.e. bed, radio, fan, etc.

EEG Artifacts

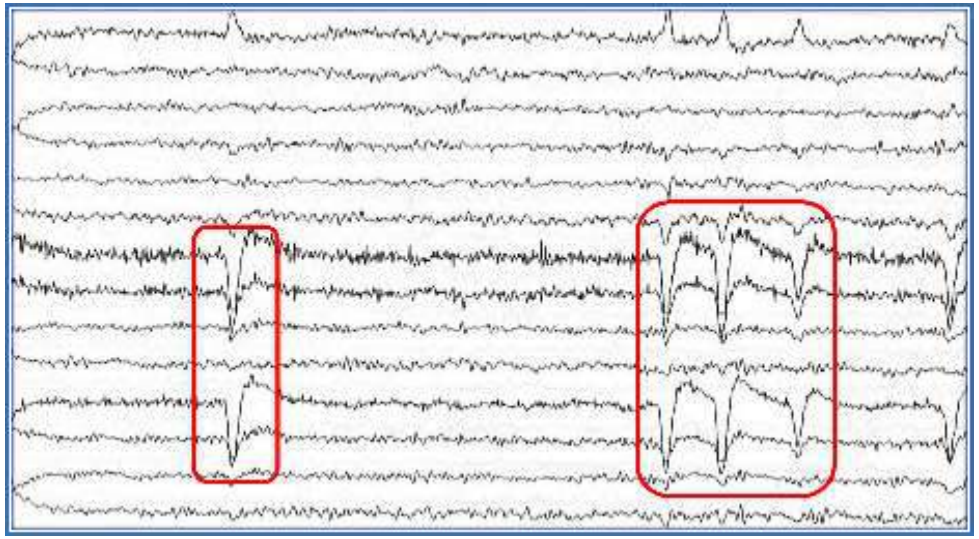


Captured EEG Signal

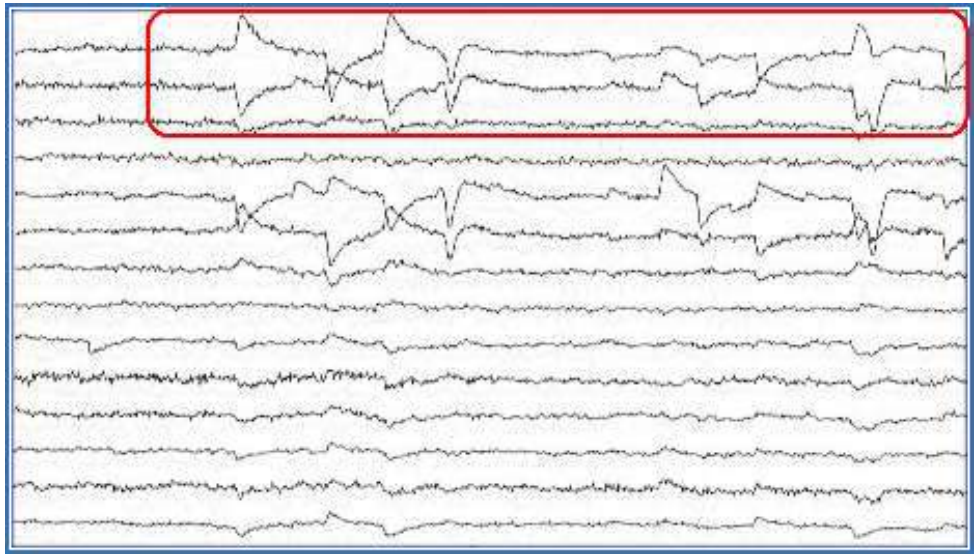


Ocular Artifacts

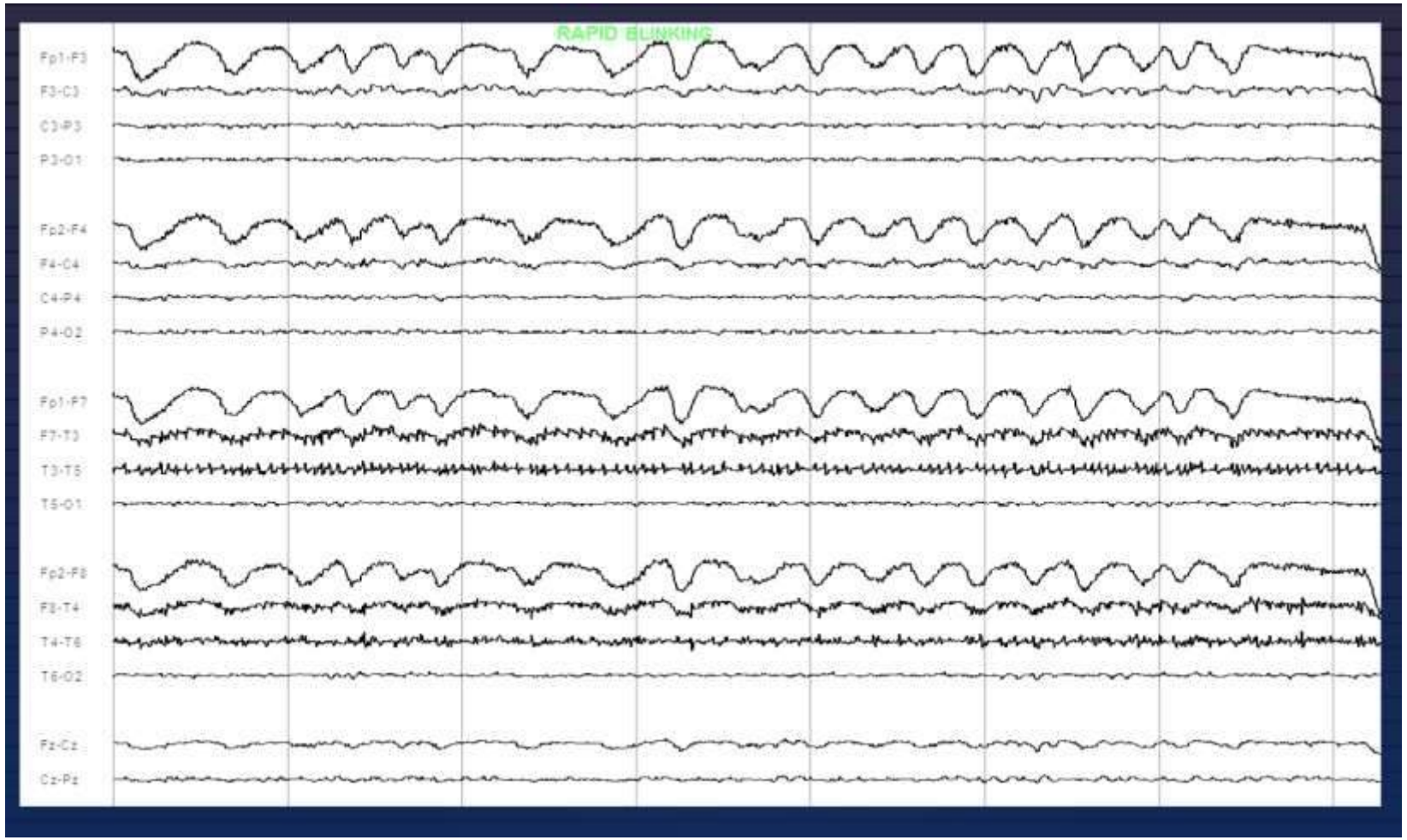
- Blinks
- Eye flutter
- Lateral gaze
- Slow/roving eye movement
- Rapid eye movement
- Electroretinogram (ERG) - that measures the electrical activity of the retina in response to a light stimulus. The ERG arises from currents generated directly by retinal neurons.



Blink



Slow Eye Movement


RAPID BLINKING
Fp1-F3
F3-C3
C3-P3
P3-O1
Fp2-F4
F4-C4
C4-P4
P4-O2
Fp1-F7
F7-T3
T3-T5
T5-O1
Fp2-F8
F8-T4
T4-T6
T6-O2
Fz-Cz
Cz-Pz

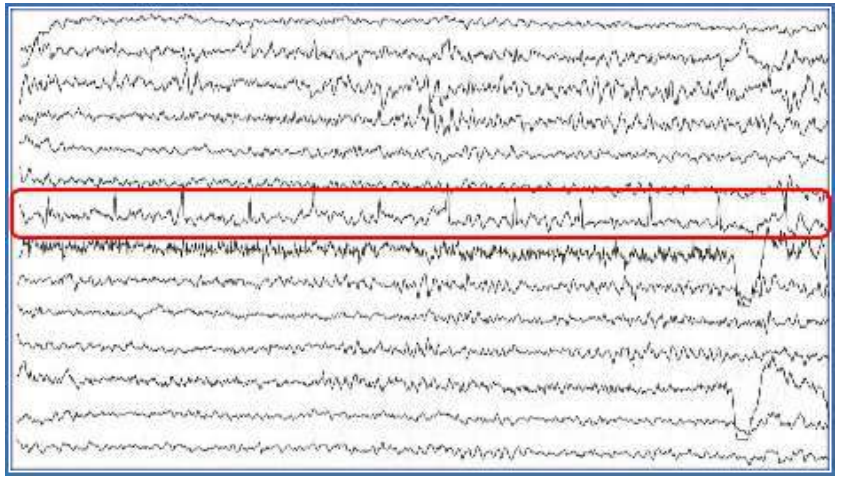
Eye Flutter

Cardiac Artifacts

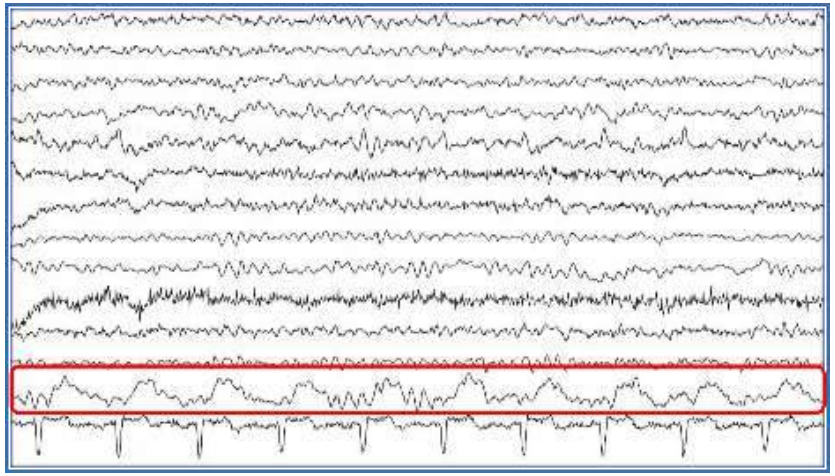
- Mechanical and Electrical
- ECG, Pacemaker - Electrical
- Pulse, Ballistocardiographic – Mechanical

✓ Mostly these are high in amplitude and prominent in babies, obese and short neck persons.

✓ Referential montages picks up cardiac artifacts.



Cardiac (Electrical)



Cardiac (Mechanical)

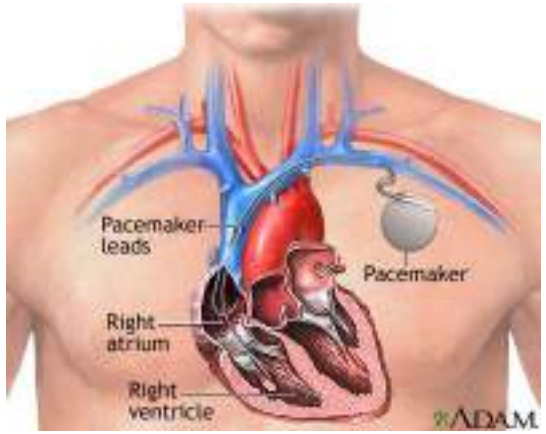
Cardiac Artifacts

ECG:

- The artifact created by the electrical signals
- generated by the heart
- Looks like a QRS complex
- More prominent in pts with short, thick necks
- More prominent in montages that use ear as reference

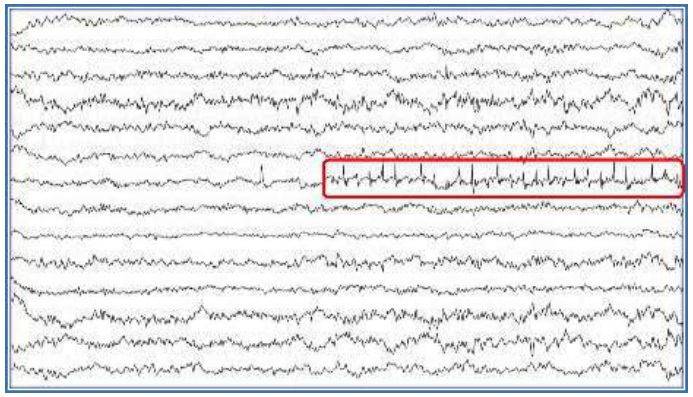
Pulse:

- The artifact created by placing an electrode over a pulsating artery.
- Most common in central, temporal electrodes.
- Irregular heart rate will create arrhythmic delta.

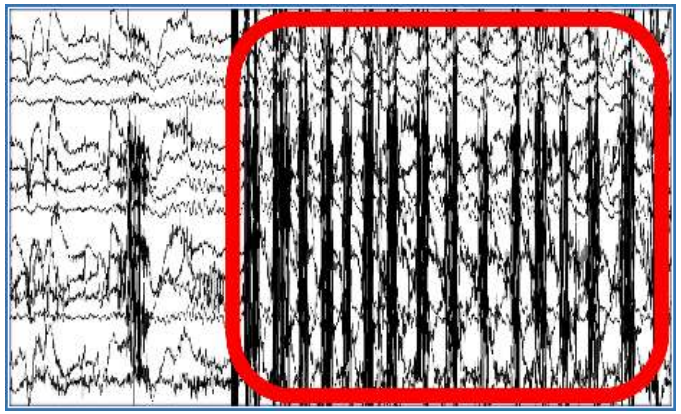


Muscle Artifacts

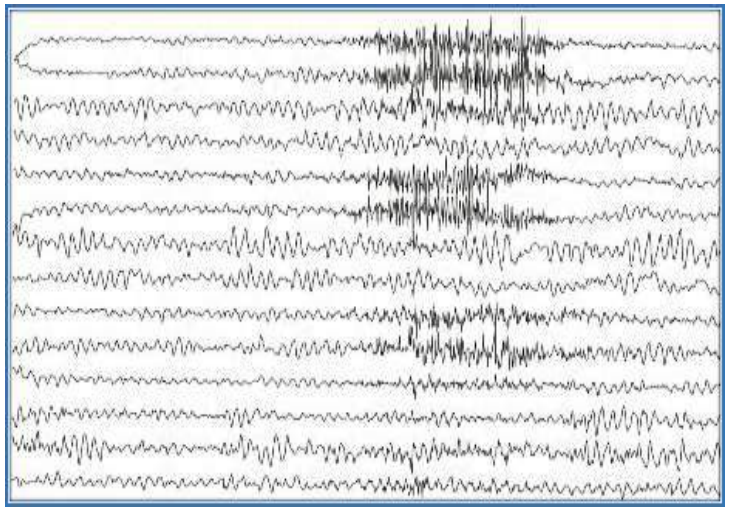
- ✓ Glossokenetic (related to tongue movements, Chew and swallow)
- ✓ Photomyogenic/ Photo-myoclonic (When flash of light falls over the face, the activity occurs due to myoclonus of the facial muscles).
- ✓ Surface EMG (Electromyography) – used to measure electrical activity during muscle contractions and relaxation cycles.



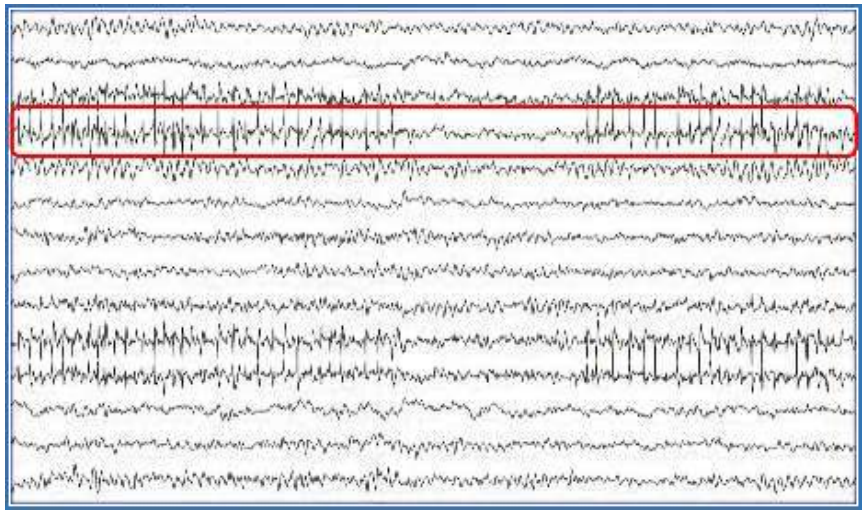
Photomyogenic



Chewing



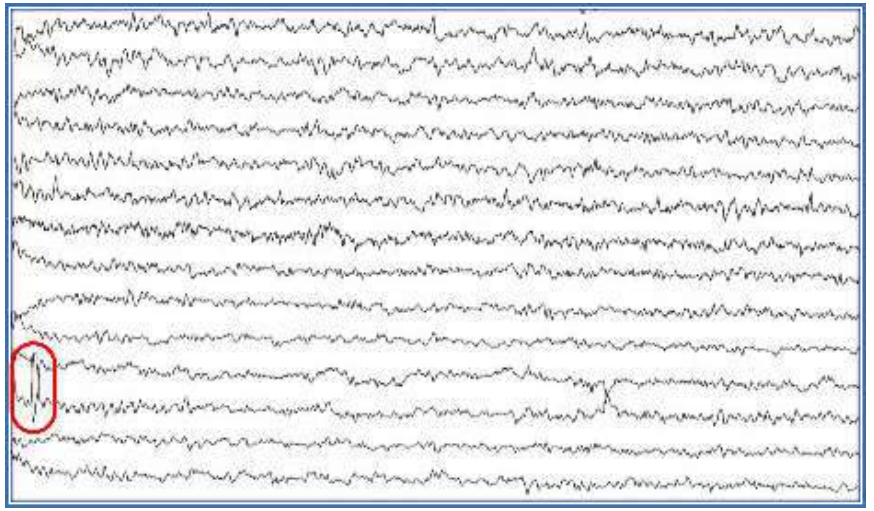
Electromyography (Scalp)



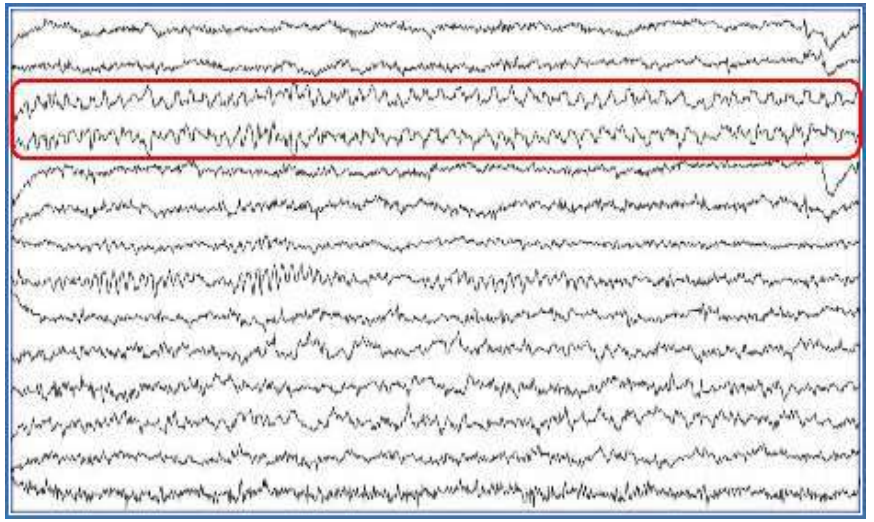
Electromyography (Facial)

Electrode and Equipment Artifacts

- ✓ Electrode pop, electrode contact, electrode movement
- ✓ Perspiration – the process of sweating
- ✓ salt bridge – differs from perspiration by low amplitude.
- ✓ Movement artifacts - Movement of head, body and limbs produce irregular high voltage potentials
- ✓ 50/60 Hz ambient electrical noise.
- ✓ Ventilators, circulatory pumps.
- ✓ Telephone, mobile.



Electrode Pop



Electrode Movement

Electrode and Equipment Artifacts

- Seen due to smearing of the electrode paste between electrodes or presence of perspiration across the scalp
- Forms an unwanted electrical connection between the electrodes forming a channel

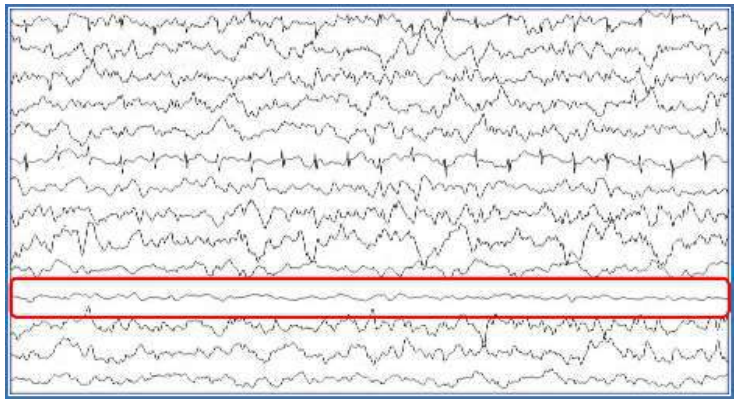
✓Perspiration artifact

- manifests as low amplitude
- undulating (smooth) waves
- duration is typically greater than 2 sec

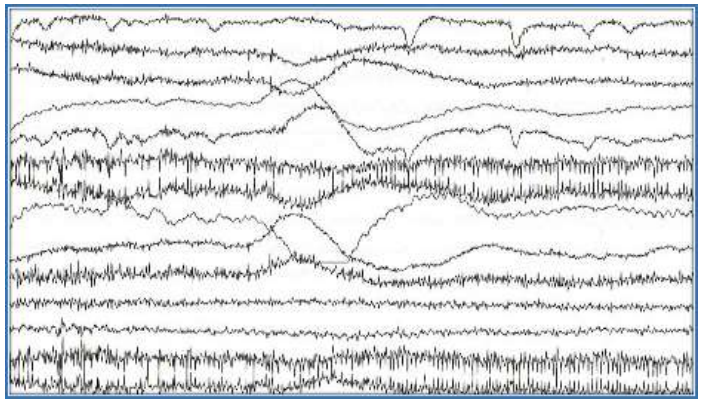
✓Slat bridge artifact

- lower in amplitude
- not wavering with low frequency oscillation - typically include only one channel
- It may appear flat and close to isoelectric

Electrode and Equipment Artifacts

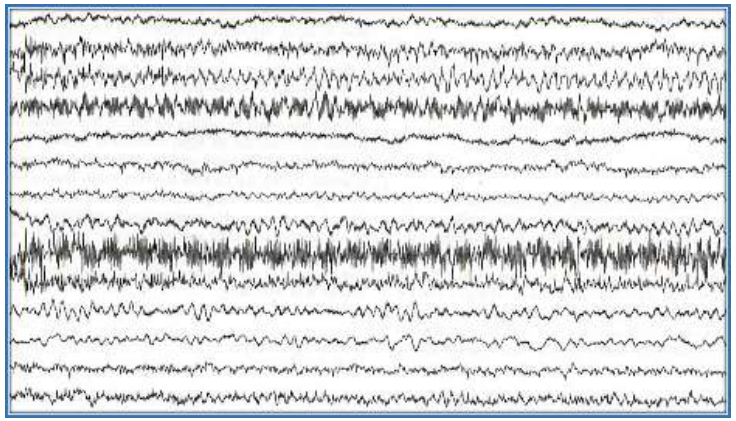


Salt Bridge

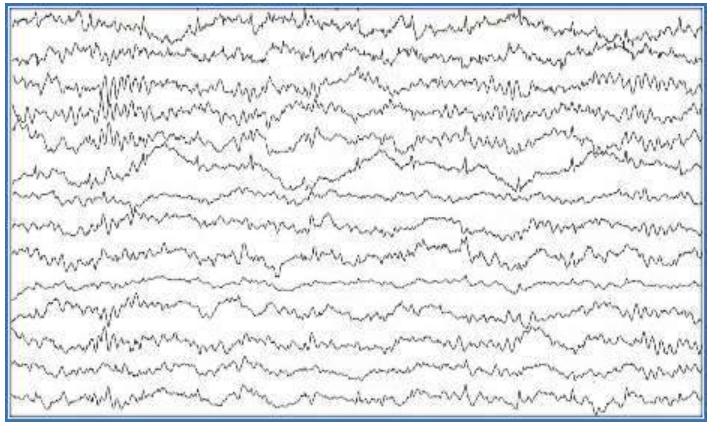


Electrode Lead Movement

Electrical Motor



60 Hz



Perspiration

Phone

Thank You!

Brain Computer Interaction

EEG Variables

Course Instructors

Dr. Sreeja S R

Assistant Professor

Indian Institute of Information Technology

IIIT Sri City

EEG

- Electroencephalogram (EEG) signals are useful for diagnosing various mental conditions such as epilepsy, memory impairments and sleep disorders.
- EEGs can indicate the general conscious state of a person, e.g., asleep, awake, anaesthetized, since each state is correlated with particular EEG patterns.
- A flat EEG (no electrical activity) is clinical evidence of death.

Why EEG?

- Hardware costs are significantly lower than those of most other techniques.
- EEG sensors can be used in more places than fMRI, SPECT, PET, MRS, or MEG, as these techniques require bulky and immobile equipment.
- EEG has very high temporal resolution, on the order of milliseconds rather than seconds, commonly recorded at sampling rates between 250 and 2000 Hz thus a valuable tool for research and diagnosis.
- EEG is relatively tolerant of subject movement, unlike most other neuro imaging techniques. There even exist methods for minimizing, and even eliminating movement artifacts in EEG data.
- EEG is silent, which allows for better study of the responses to auditory stimuli.
- EEG does not involve exposure to high-intensity (>1 Tesla) magnetic fields, as in some of the other techniques, especially MRI and MRS. These can cause a variety of undesirable issues with the data, and also prohibit use of these techniques with participants that have metal implants in their body, such as metal-containing pacemaker.
- EEG can be used in subjects who are incapable of making a motor response.
- EEG is a powerful tool for tracking brain changes during different phases of life. EEG sleep analysis can indicate significant aspects of the timing of brain development, including evaluating adolescent brain maturation.

EEG Disadvantages

- Low spatial resolution on the scalp. fMRI, for example, can directly display areas of the brain that are active, while EEG requires intense interpretation just to hypothesize what areas are activated by a particular response.
- EEG poorly determines neural activity that occurs below the upper layers of the brain (the cortex).
- Unlike PET and MRI, cannot identify specific locations in the brain at which various neurotransmitters, drugs, etc. can be found.
- Often takes a long time to connect a subject to EEG, as it requires precise placement of dozens of electrodes around the head and the use of various gels, saline solutions, and/or pastes to keep them in place. Where as a general rule it takes considerably less time to prepare a subject for MEG, fMRI, MRS, and PET.

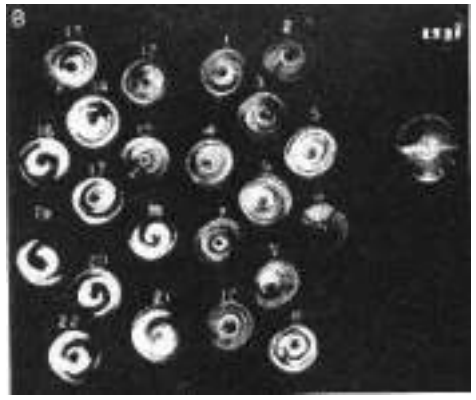
EEG Pioneers

In 1929, Hans Berger

- Recorded brain activity from the closed skull
- Reported brain activity changes according to the functional state of the brain
  - Sleep
  - Hypnothesis
  - Pathological states (epilepsy)

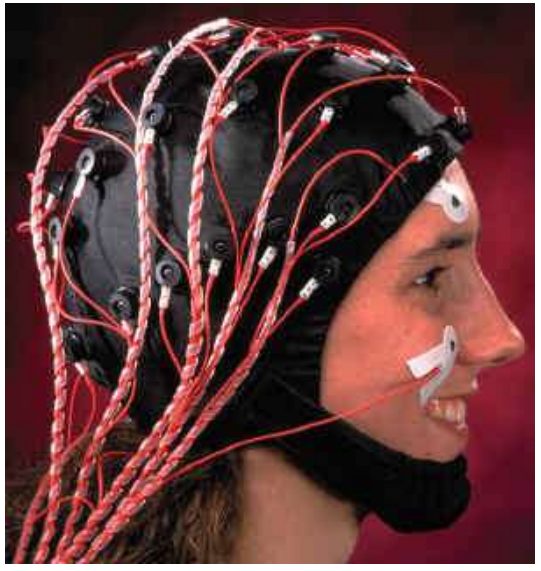
In 1957, Gray Walter

- Makes recordings with large numbers of electrodes
- Visualizes brain activity with the toposcope
- Shows that brain rhythms change according to the mental task demanded



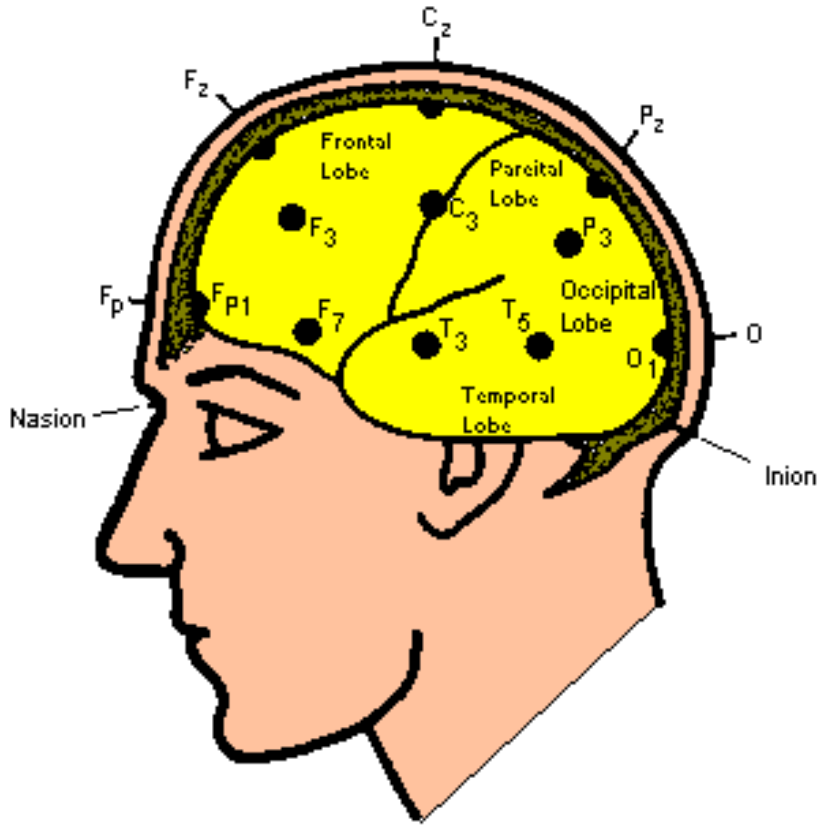
EEG CAPS

- EEGs require electrodes attached to the scalp with sticky gel
- Require physical connection to the machine



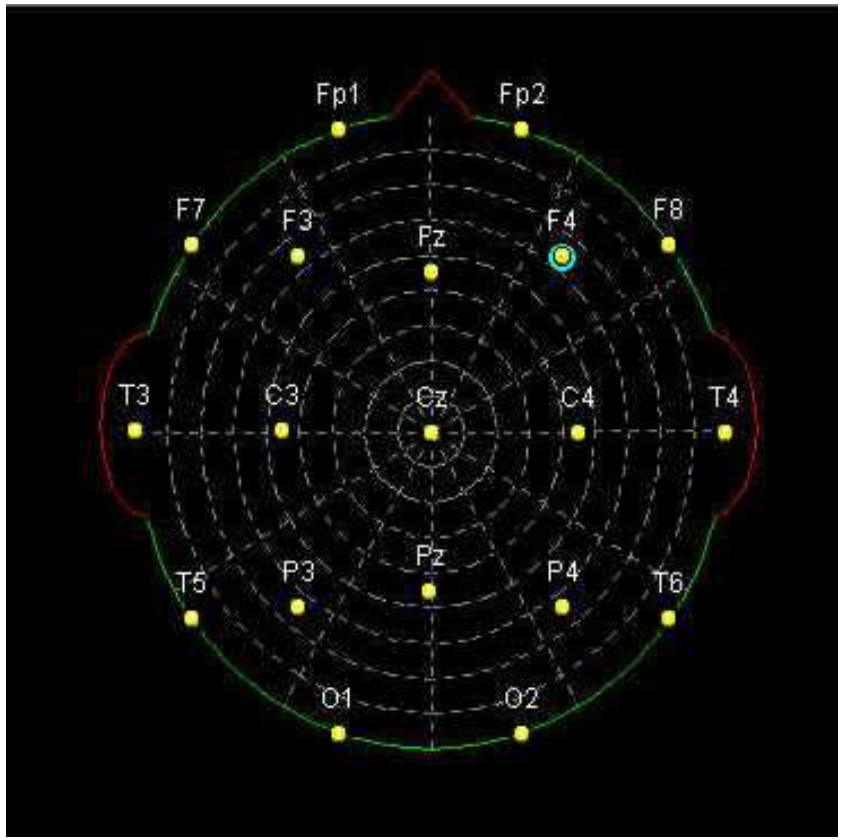
Electrode Placement

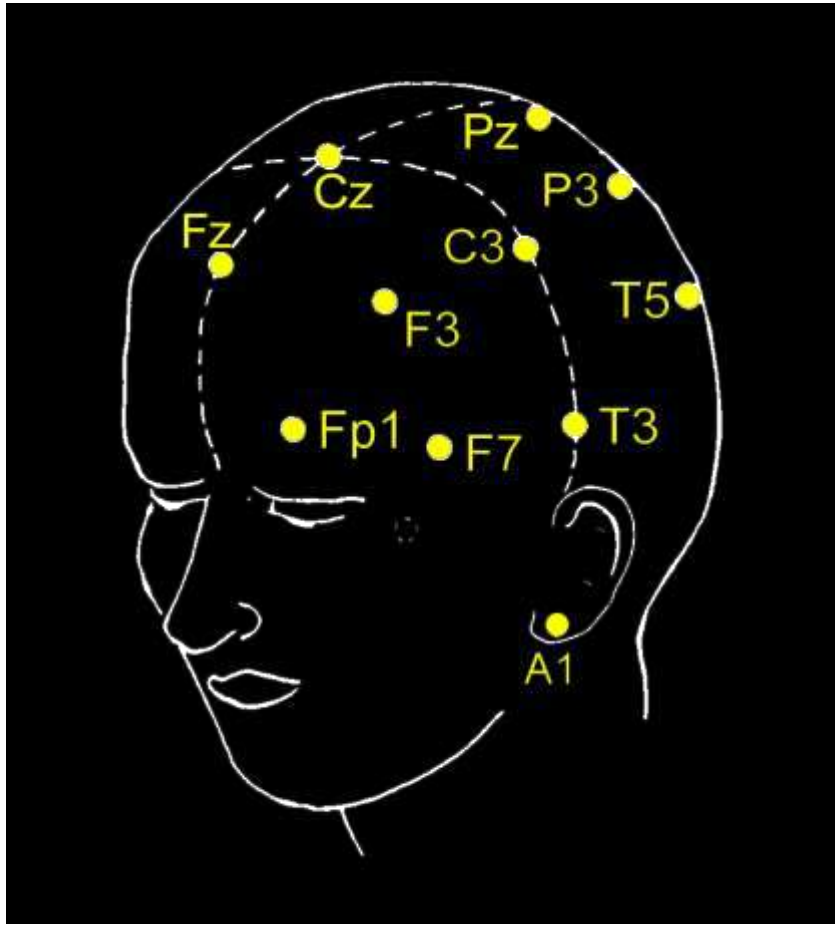
- Standard “10-20 System”
- Spaced apart 10-20%
- Nasion – point between the forehead and the skull
- Inion – Bump at the back of the skull
- Letter for region
  - F - Frontal Lobe
  - T - Temporal Lobe
  - C - Center
  - O - Occipital Lobe
- Number for exact position
  - Odd numbers - left
  - Even numbers - right



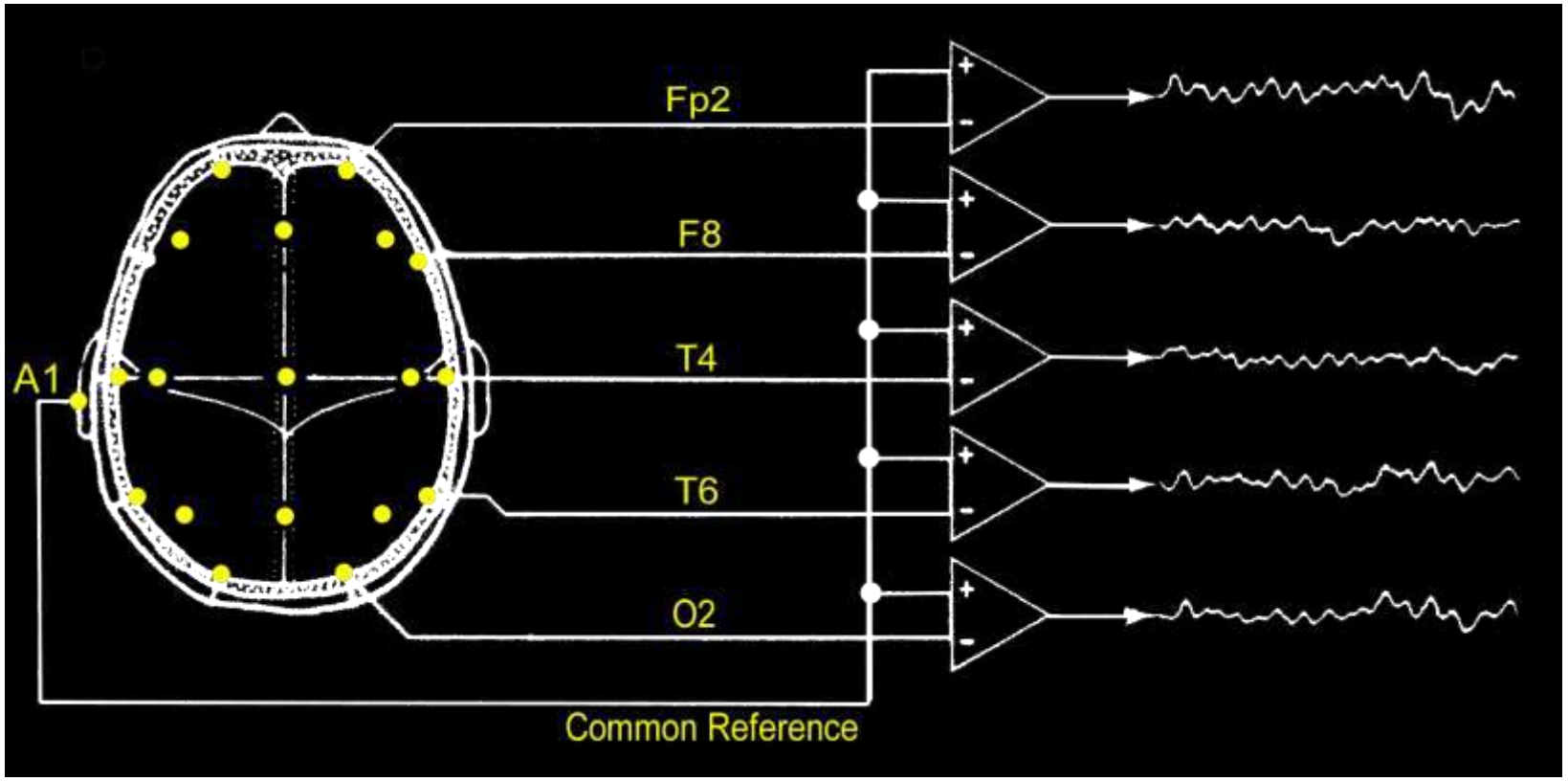
Electrode Placement

- International 10/20 system





EEG Channels



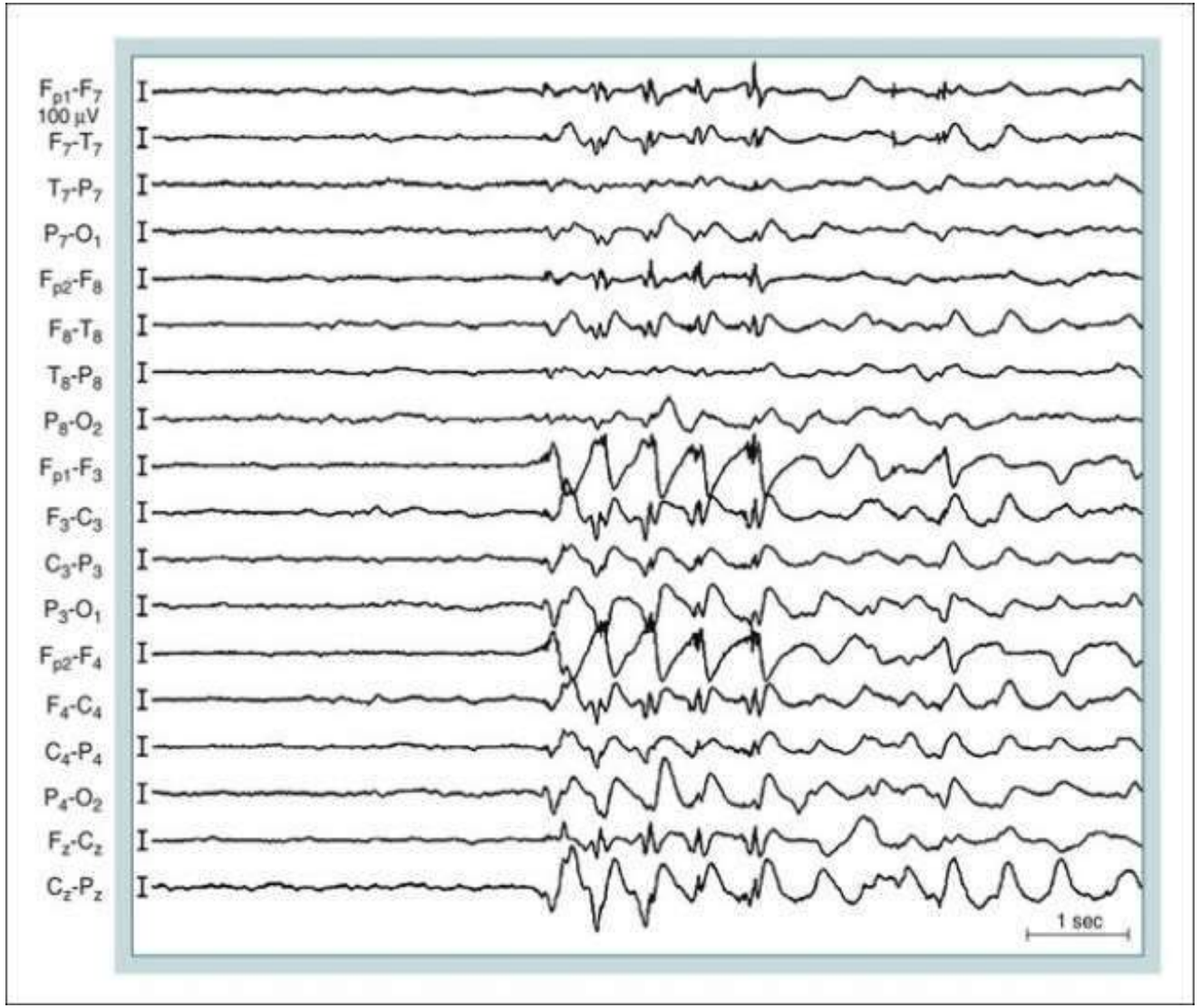
Channel: Recording from a pair of electrodes (here with a common reference: A1 – left ear)

Multichannel EEG recording: up to 64 channels recorded in parallel

Representation of EEG channels:

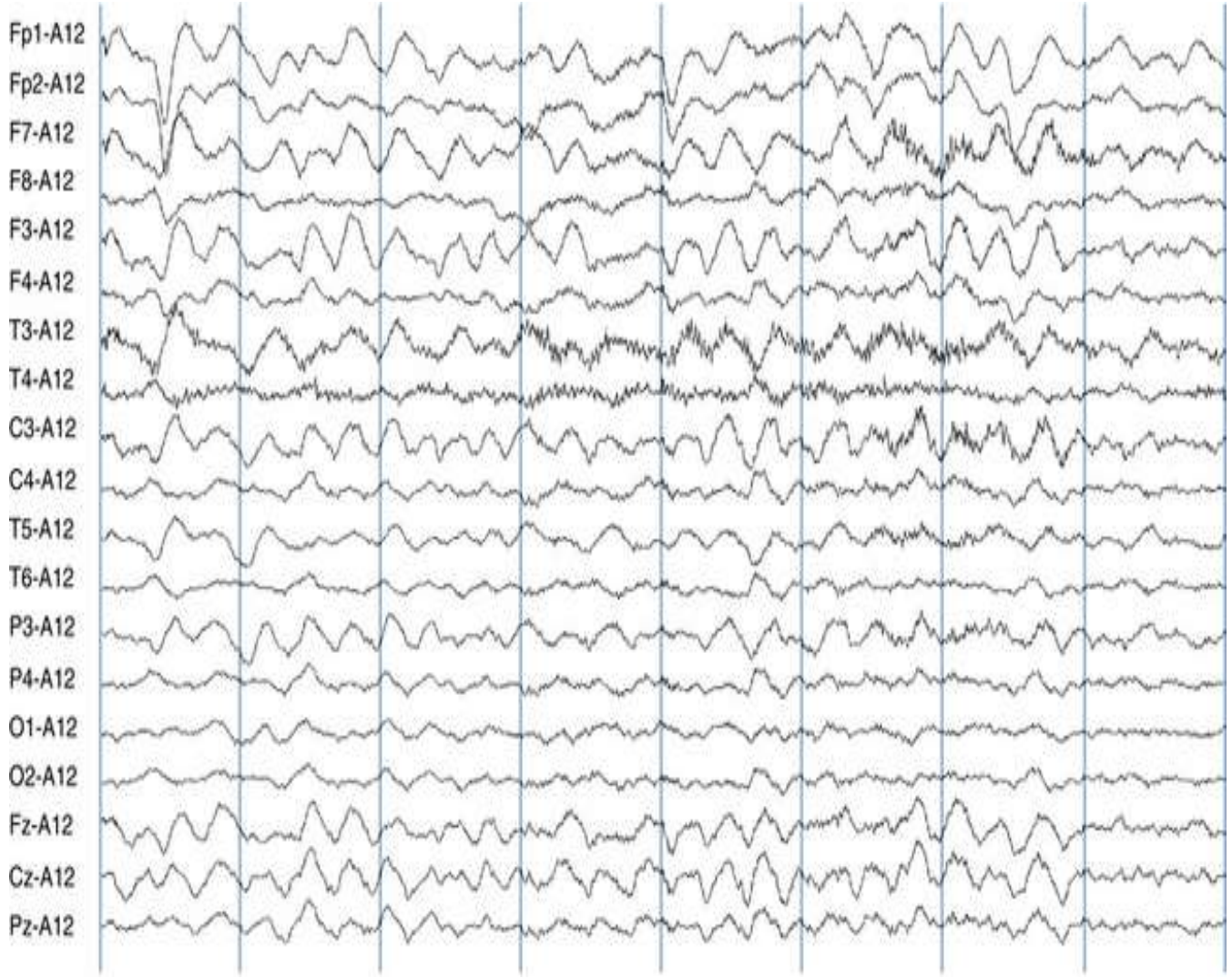
The representation of the EEG channels (i.e., waveform) is referred to as a **montage**.

- **Sequential montage**: Each channel represents the difference between two adjacent electrodes. The entire montage consists of a series of these channels. For example, the channel "Fp1-F3" represents the difference in voltage between the Fp1 electrode and the F3 electrode. The next channel in the montage, "F3-C3," represents the voltage difference between F3 and C3, and so on through the entire array of electrodes.



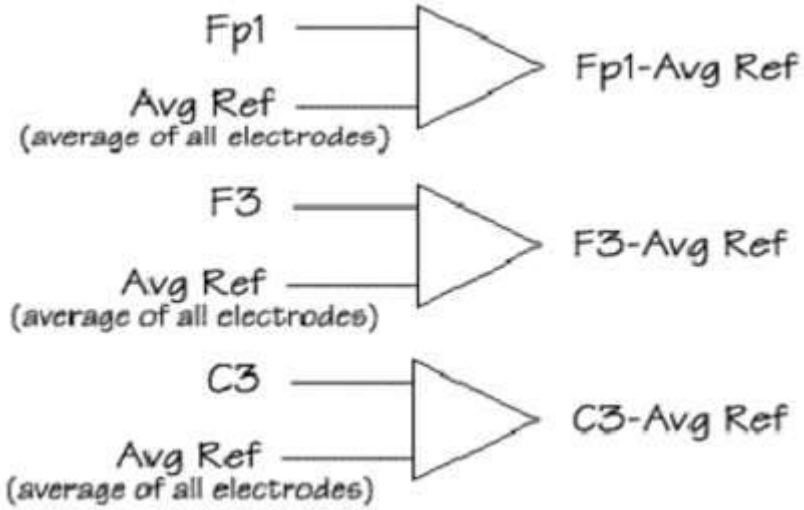
Representation of EEG channels:

- **Referential montage**: Each channel represents the difference between a certain electrode and a designated reference electrode. There is no standard position for this reference; it is, however, at a different position than the "recording" electrodes. Midline positions are often used because they do not amplify the signal in one hemisphere vs. the other. Another popular reference is "linked ears," which is a physical or mathematical average of electrodes attached to both earlobes or mastoids.

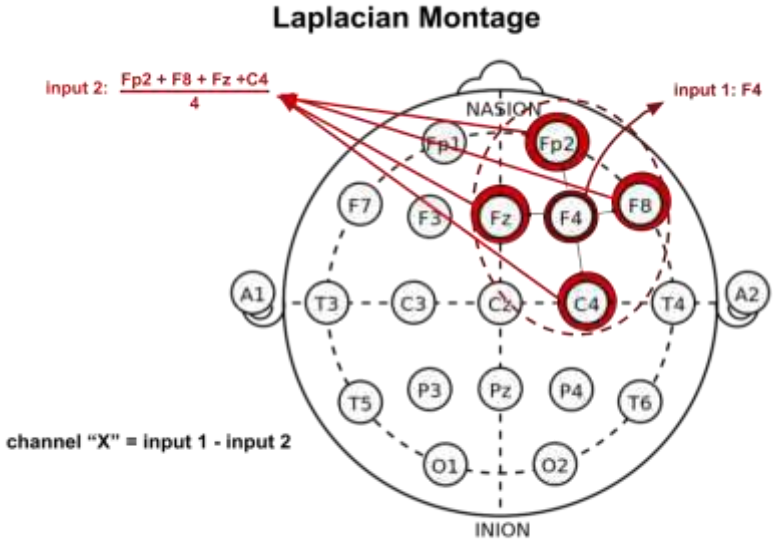


Representation of EEG channels:

- **Average reference montage:** The outputs of all of the amplifiers are summed and averaged, and this averaged signal is used as the common reference for each channel.



- **Laplacian montage:** Each channel represents the difference between an electrode and a weighted average of the surrounding electrodes.



EEG Rhythms

- Generally grouped by frequency: (amplitudes are about 100µV max)

Type	Frequency	Location	Use
Delta	<4 Hz	everywhere	occur during sleep, coma
Theta	4-7 Hz	temporal and parietal	correlated with emotional stress (frustration & disappointment)
Alpha	8-15 Hz	occipital and parietal	reduce amplitude with sensory stimulation or **mental imagery**
Beta	16-30 Hz	parietal and frontal	can increase amplitude during intense **mental activity**
Gamma	>30 Hz	Somatosensory cortex	A decrease in gamma-band activity is associated with **cognitive decline**
Mu	8-12 Hz	frontal (motor cortex)	diminishes with movement or **intention of movement**
Lambda	sharp, jagged	occipital	correlated with visual attention
Vertex			higher incidence in patients with epilepsy or encephalopathy

Alpha Rhythm

Frequency: 8 – 15 Hz

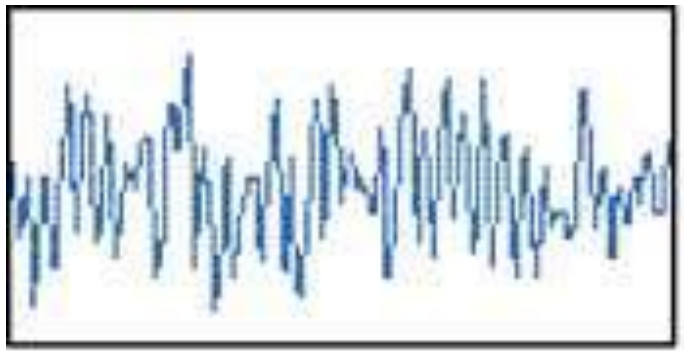
Amplitude: 5 – 100 microVolt

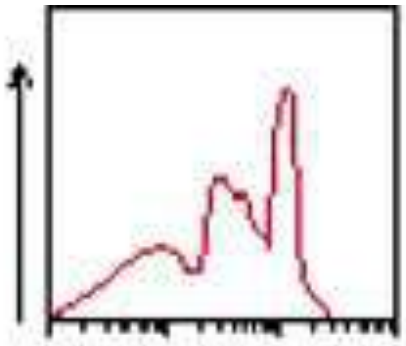
Location: Occipital, Parietal

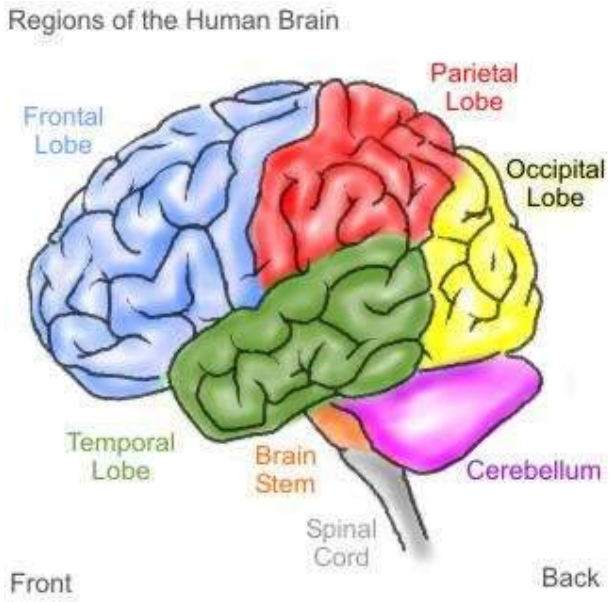
State of Mind: Alert Restfulness

Source: Oscillating thalamic pacemaker neurons

Alpha blockade occurs when new stimulus is processed







Beta Rhythm

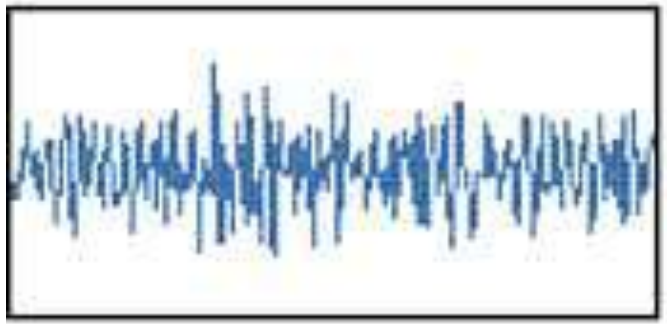
Frequency: 16 – 30 Hz

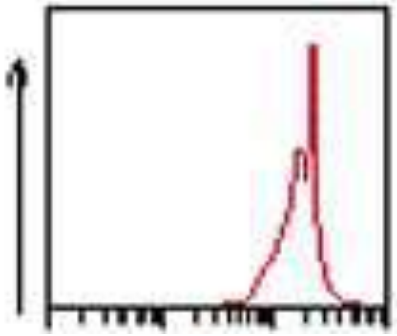
Amplitude: 2 – 20 microVolt

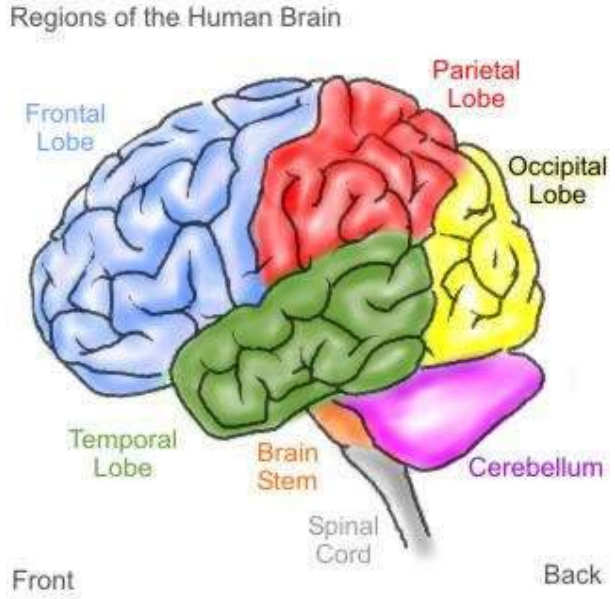
Location: Frontal

State of Mind: Mental Activity

Reflects specific information processing between cortex and thalamus







Delta Rhythm

Frequency: 1 – 4 Hz

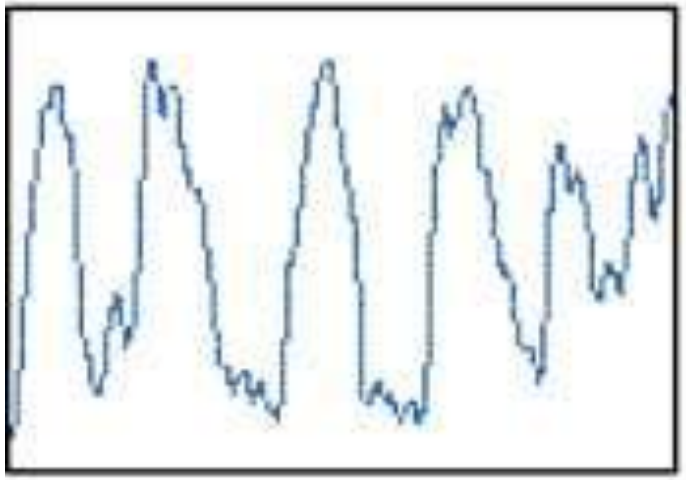
Amplitude: 20 – 200 microVolt

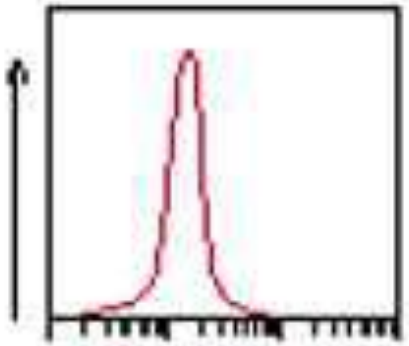
Location: Variable

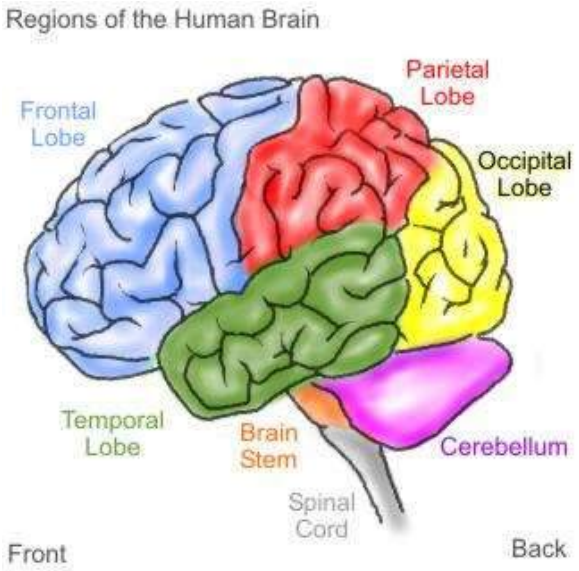
State of Mind: Deep sleep

Oscillations in Thalamus and deep cortical layers

Usually inibited by ARAS (Ascending Reticular Activation System)







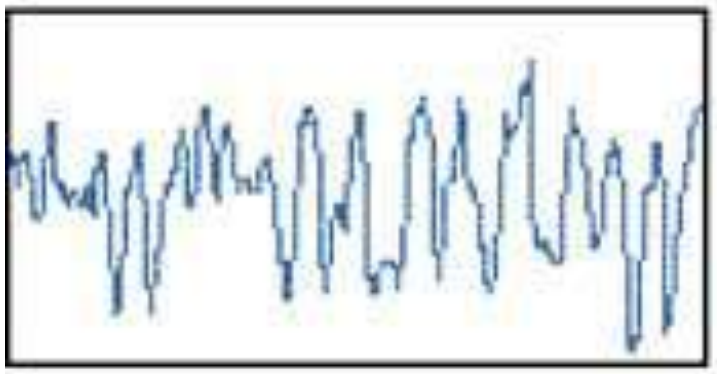
Theta Rhythm

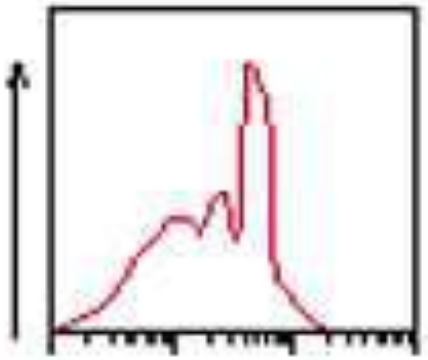
Frequency: 4 – 7 Hz

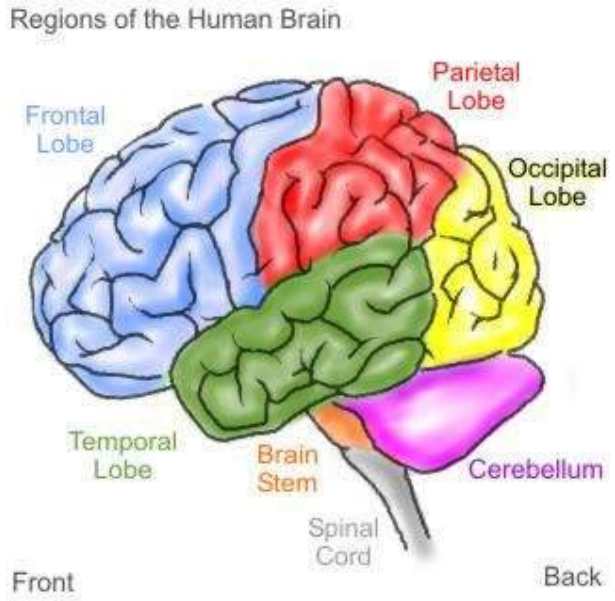
Amplitude: 5 – 100 microVolt

Location: Frontal, Temporal

State of Mind: Sleepiness







Mu Waves

- Studied since 1930s
- Found in Motor Cortex
- Amplitude suppressed by Physical Movements, or *intent to* move physically
- (Wolpaw, et al 1991) trained subjects to control the mu rhythm by visualizing motor tasks to move a cursor up and down (1D)
- (Wolpaw and McFarland 2004) used a linear combination of Mu and Beta waves to control a 2D cursor.
- Weights were learned from the users in real time.
- Cursor moved every 50ms (20 Hz)
- 92% “hit rate” in average 1.9 sec

Alpha and Beta Waves

- Studied since 1920s
- Found in Parietal and Frontal Cortex
- Relaxed - Alpha has high amplitude
- Excited - Beta has high amplitude
- So, Relaxed -> Excited

  means Alpha -> Beta

Variables used in EEG measurement:

Frequency:

- Frequency refers to rhythmic repetitive activity (in Hz). The frequency of EEG activity can have different properties including:
- **Rhythmic**. EEG activity consisting in waves of approximately constant frequency.
- **Arrhythmic**. EEG activity in which no stable rhythms are present.
- **Dysrhythmic**. Rhythms and/or patterns of EEG activity that characteristically appear in patient groups or rarely or seen in healthy subjects.

Variables used in EEG measurement:

Voltage: Voltage refers to the average voltage or peak voltage of EEG activity.

- **Attenuation** (synonyms: suppression, depression). Reduction of amplitude of EEG activity resulting from decreased voltage.
- **Hypersynchrony**. Seen as an increase in voltage and regularity of rhythmic activity, or within the alpha, beta, or theta range. The term implies an increase in the number of neural elements contributing to the rhythm.
- **Paroxysmal**. Activity that reaching (usually) quite high voltage and ending with an abrupt return to lower voltage activity. Though the term does not directly imply abnormality, much abnormal activity is paroxysmal.

Variables used in EEG measurement:

Morphology: Morphology refers to the shape of the waveform. The shape of a wave or an EEG pattern is determined by the frequencies that combine to make up the waveform and by their phase and voltage relationships. Wave patterns can be described as being:

- **Monomorphic**. Distinct EEG activity appearing to be composed of one dominant activity
- **Polymorphic**. Distinct EEG activity composed of multiple frequencies that combine to form a complex waveform.
- **Sinusoidal**. Waves resembling sine waves. Monomorphic activity usually is sinusoidal.
- **Transient**. An isolated wave or pattern that is distinctly different from background activity.
  - Spike: a transient with a pointed peak and duration from 20 to less than 70 msec.
- Sharp wave: a transient with a pointed peak and duration of 70-200 msec.

Variables used in EEG measurement:

Synchrony:

- Synchrony refers to the simultaneous appearance of rhythmic or morphologically distinct patterns over different regions of the head, either on the same side (unilateral) or both sides (bilateral).

Periodicity:

- Periodicity refers to the distribution of patterns or elements in time (e.g., the appearance of a particular EEG activity at more or less regular intervals). The activity may be generalized, focal or lateralized.

Thank You!

Brain Computer Interaction

EEG Signal Pre-processing - Epoching

Course Instructors

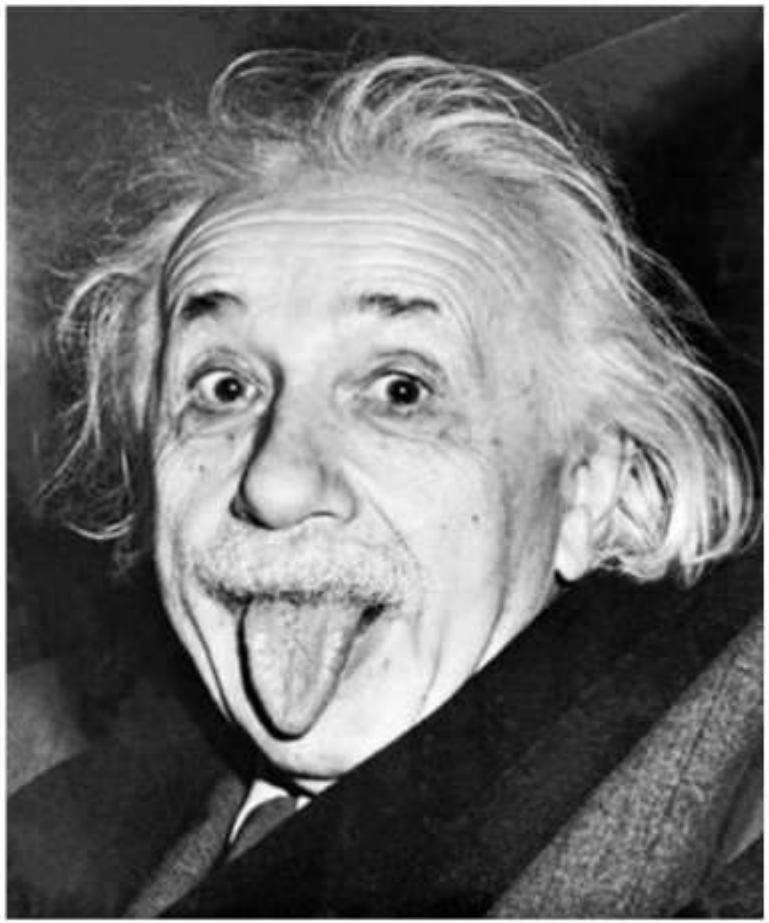
Dr. Sreeja S R

Assistant Professor

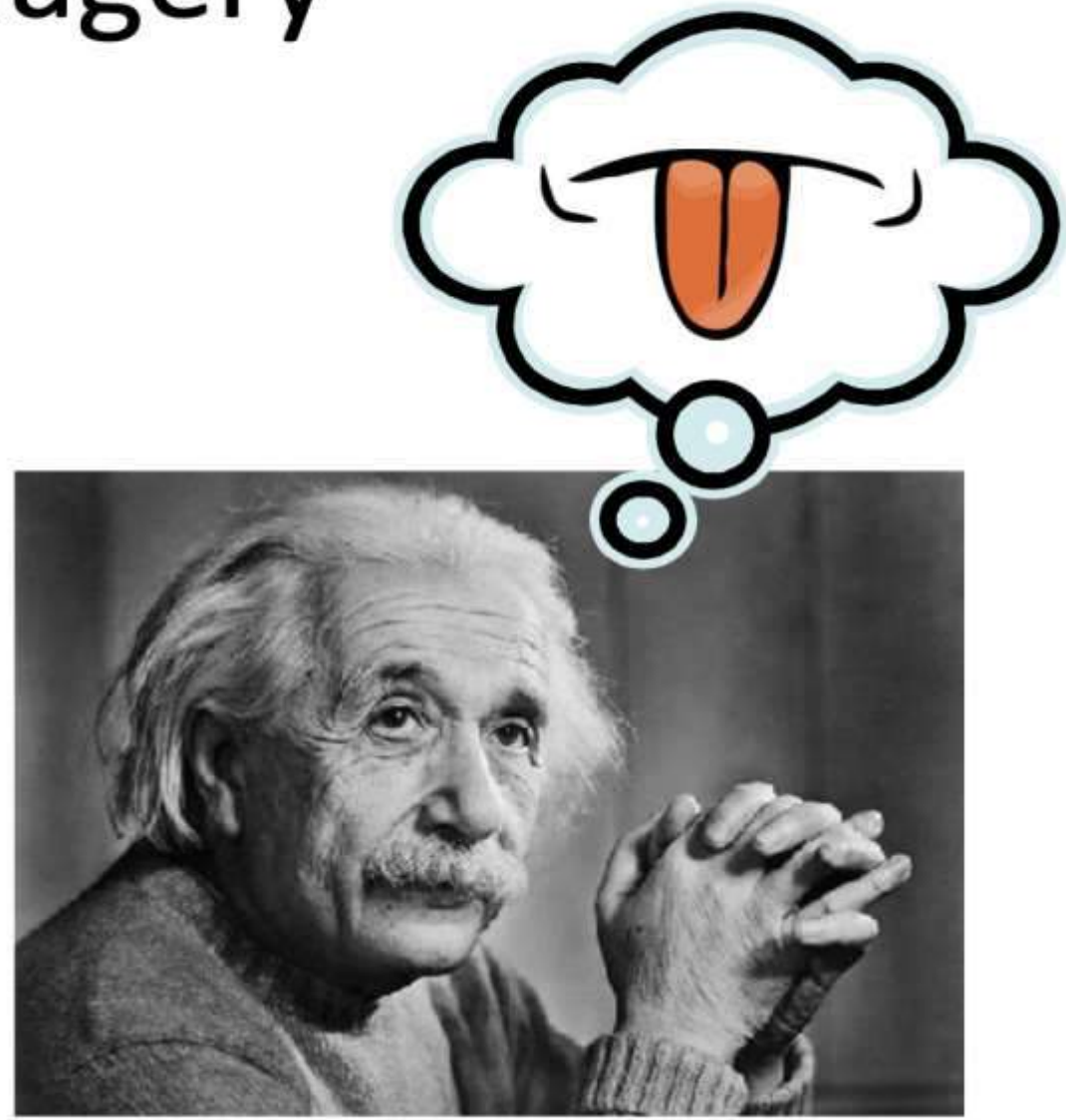
Indian Institute of Information Technology

IIIT Sri City

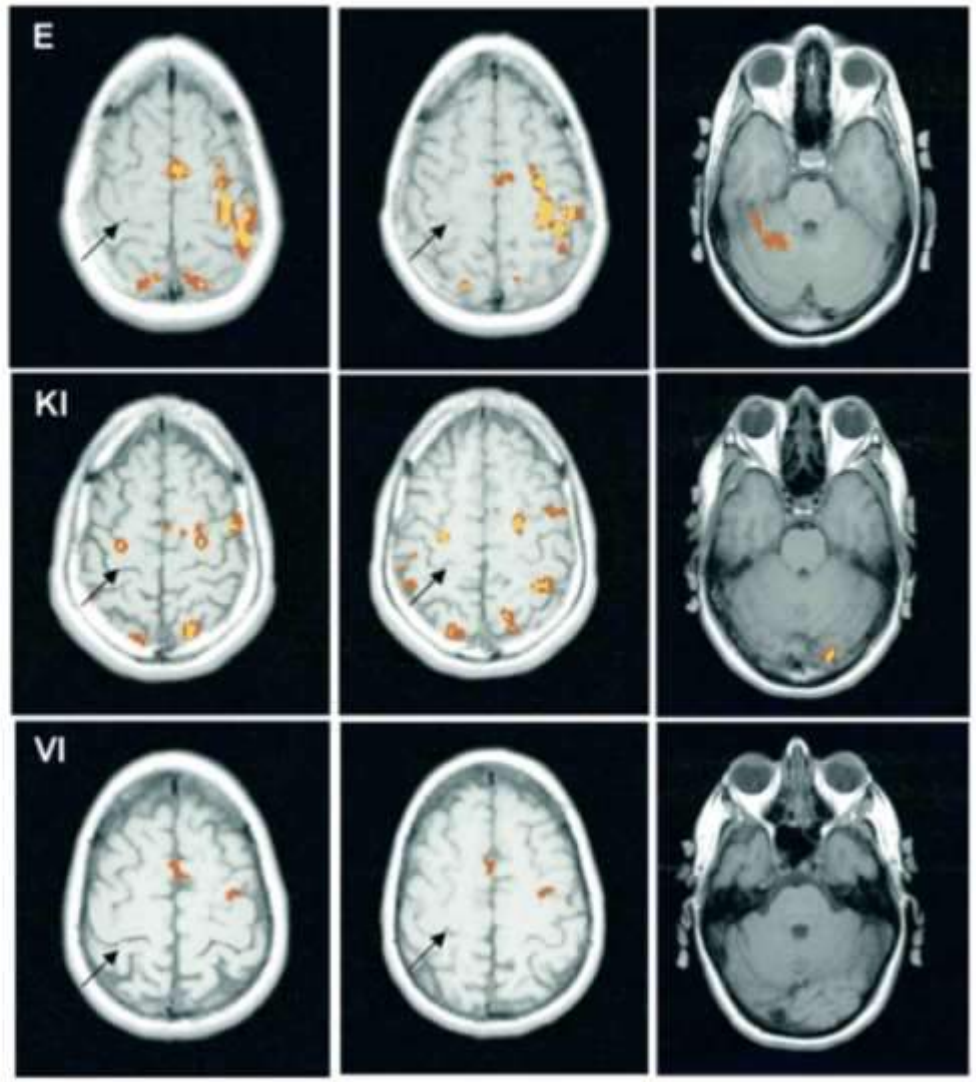
Motor Imagery



VS

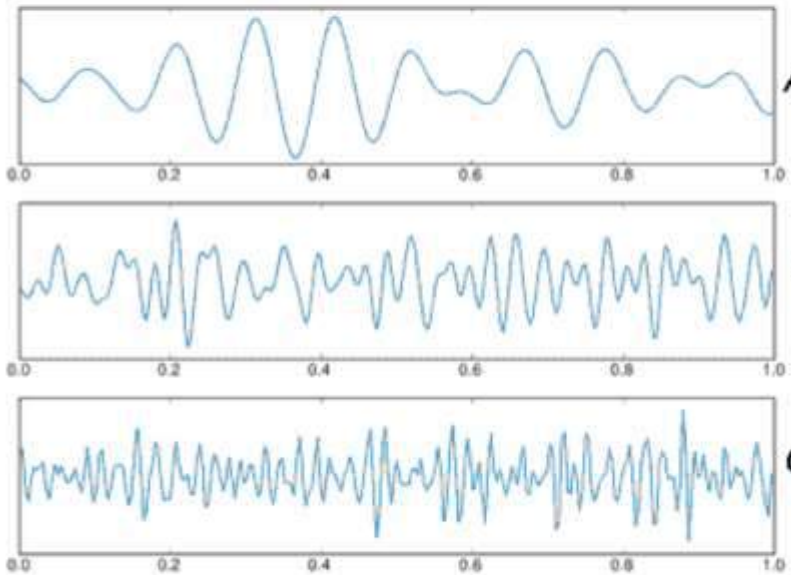


Brain activity during motor action vs imagery



Solodkin et al, 2004

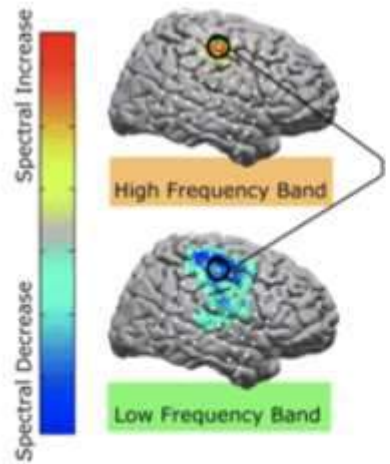
Motor rhythms in cortex

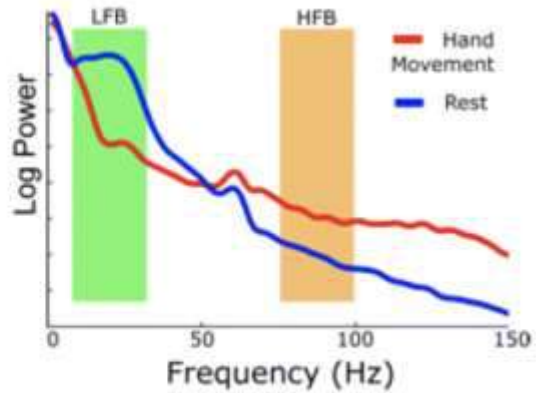


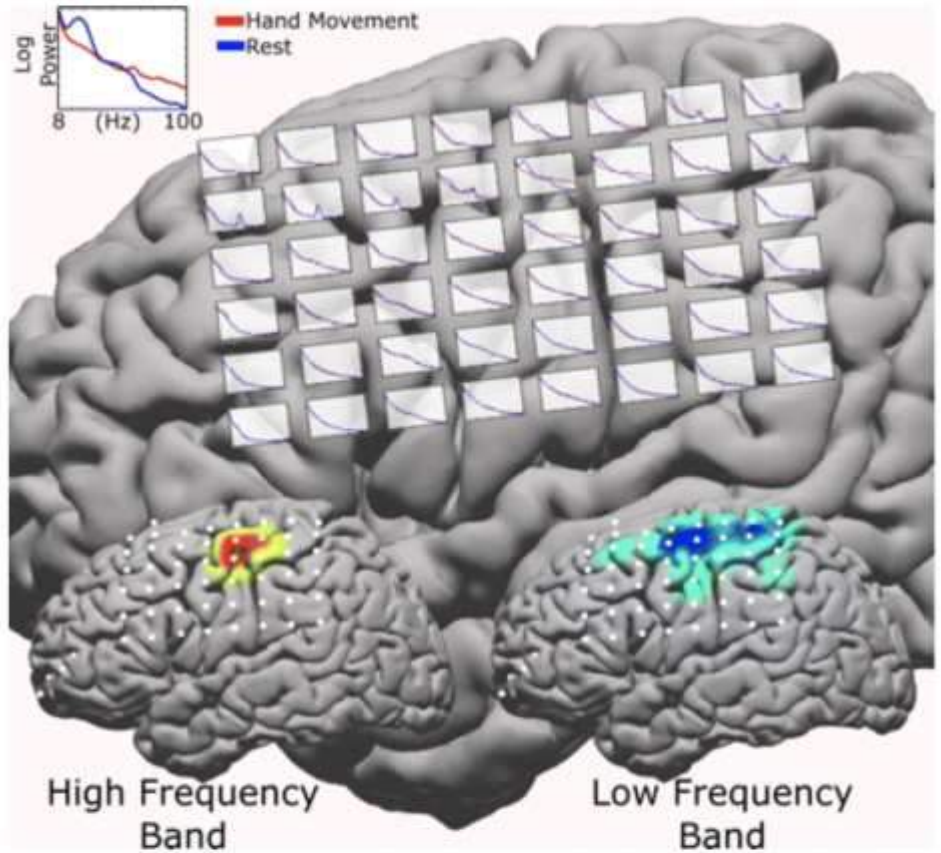
Apiha waves
8-13 Hz

Beta waves
14-25 Hz

Gama waves
25-150 Hz







Miller et al, 2007

Goal of the work

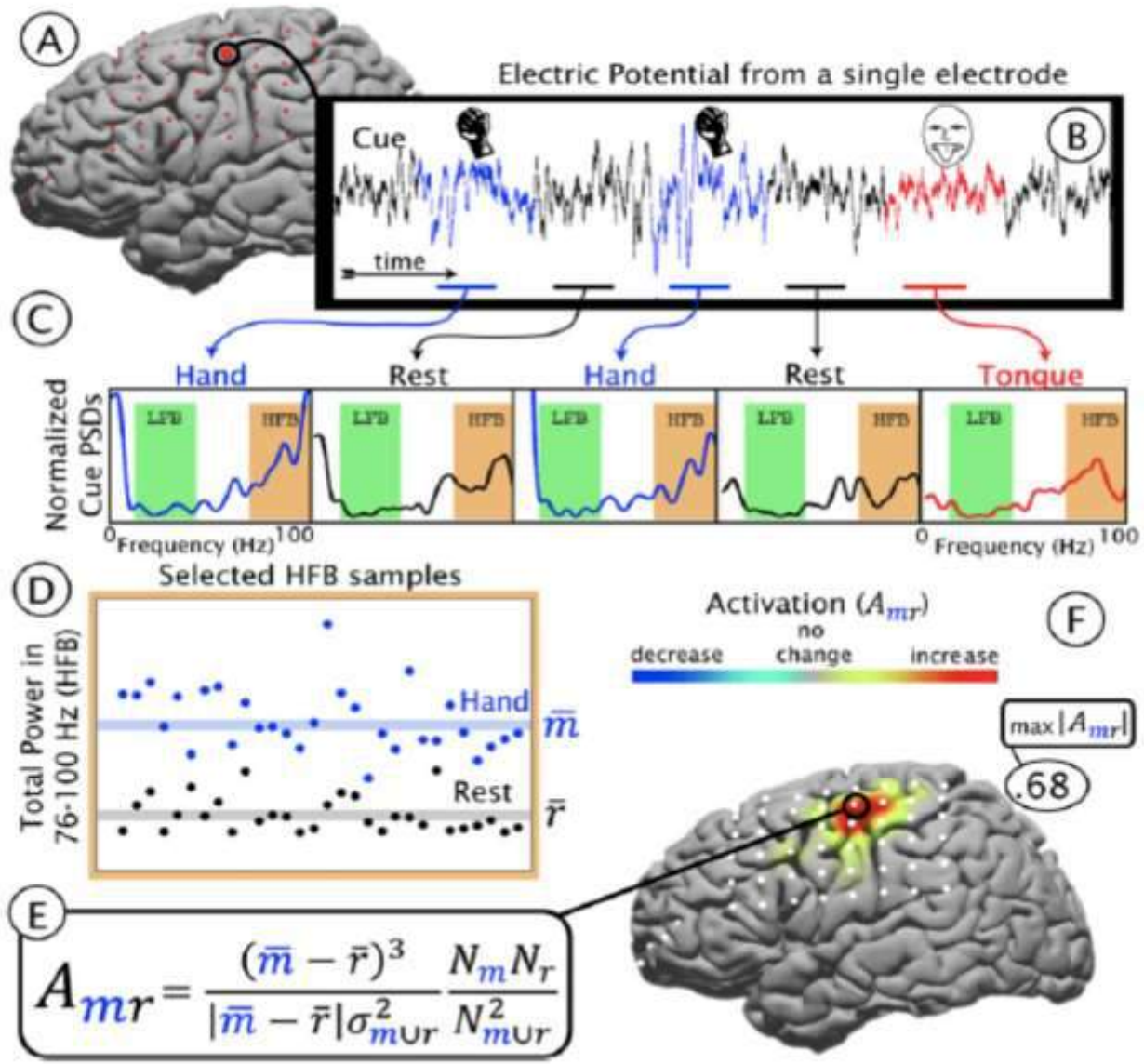
To check whether real movements can be compared with imagery movements.

Methods Used:

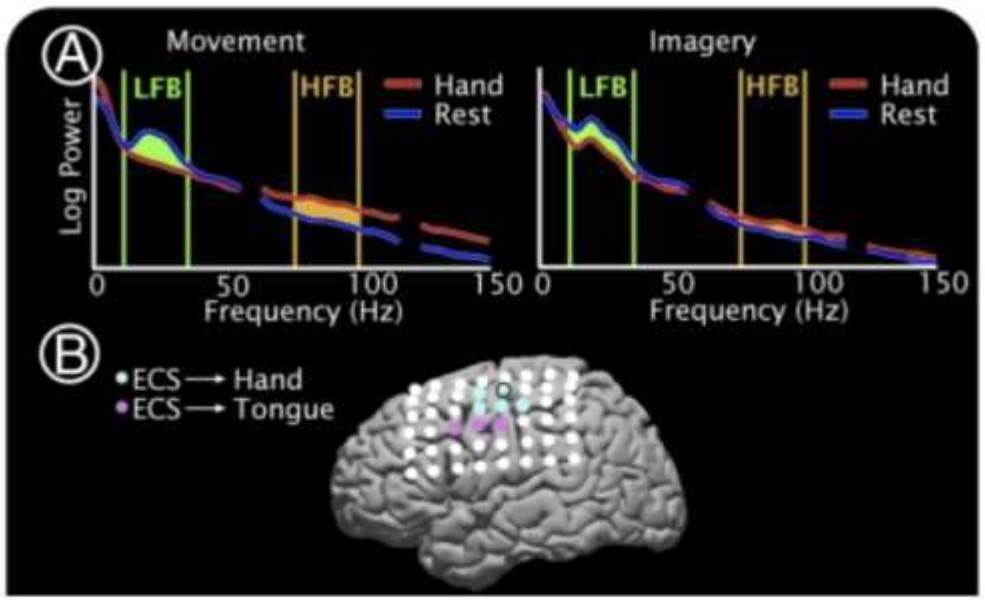
- 8 patients implanted with 4x8 or 8x8 grid

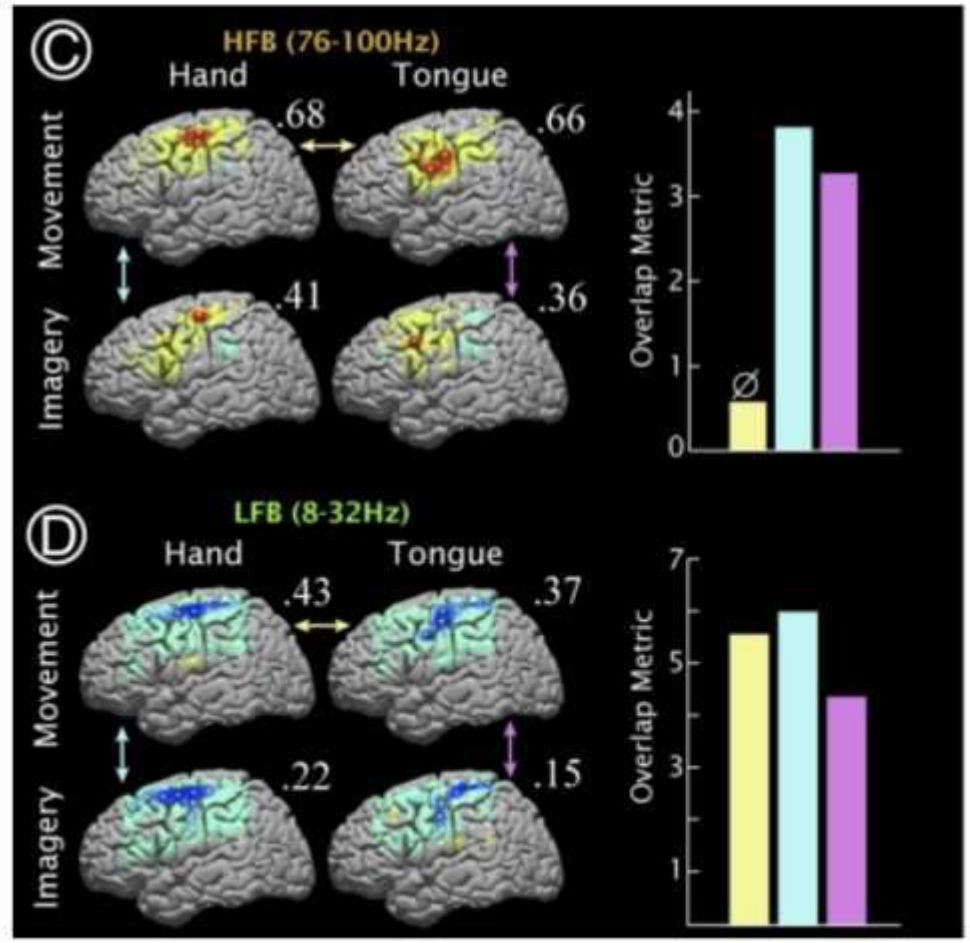
1) Active movements –clench – release of hand, stick out tongue, shoulder shrug, say the word “move”
2) Image movements – same as before

Quantification of brain activity



Cortical activity during real and imagined movements





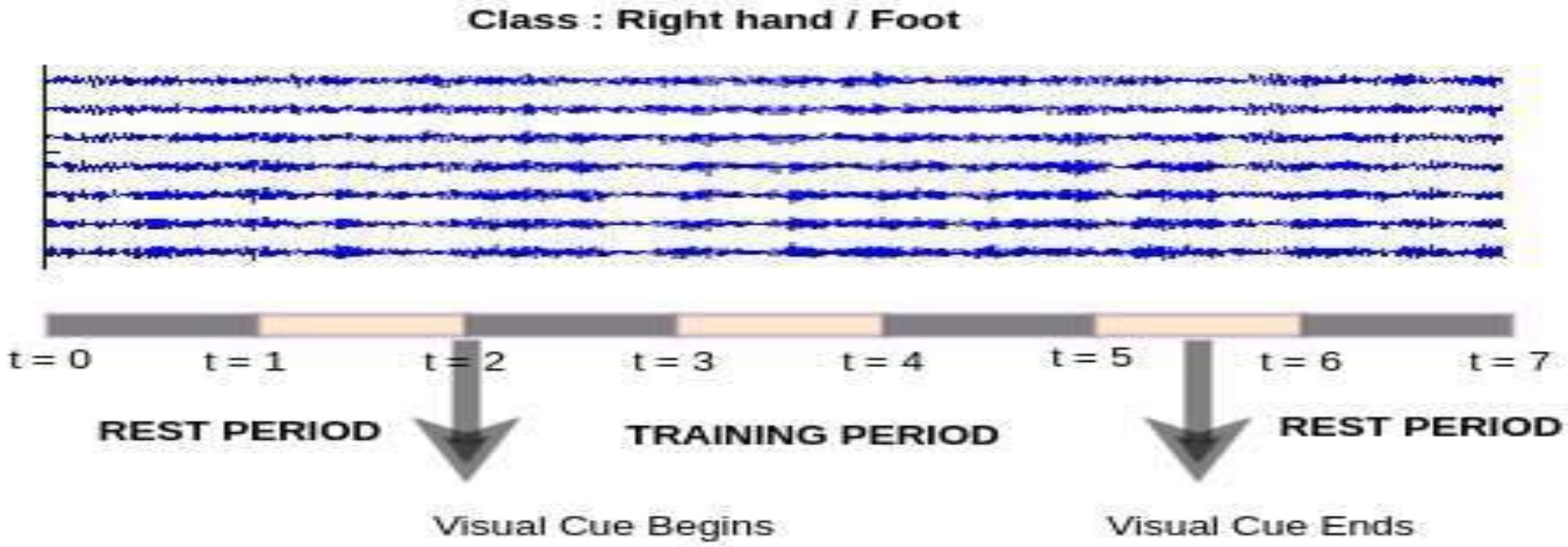
Data Epoching

Selecting data epochs

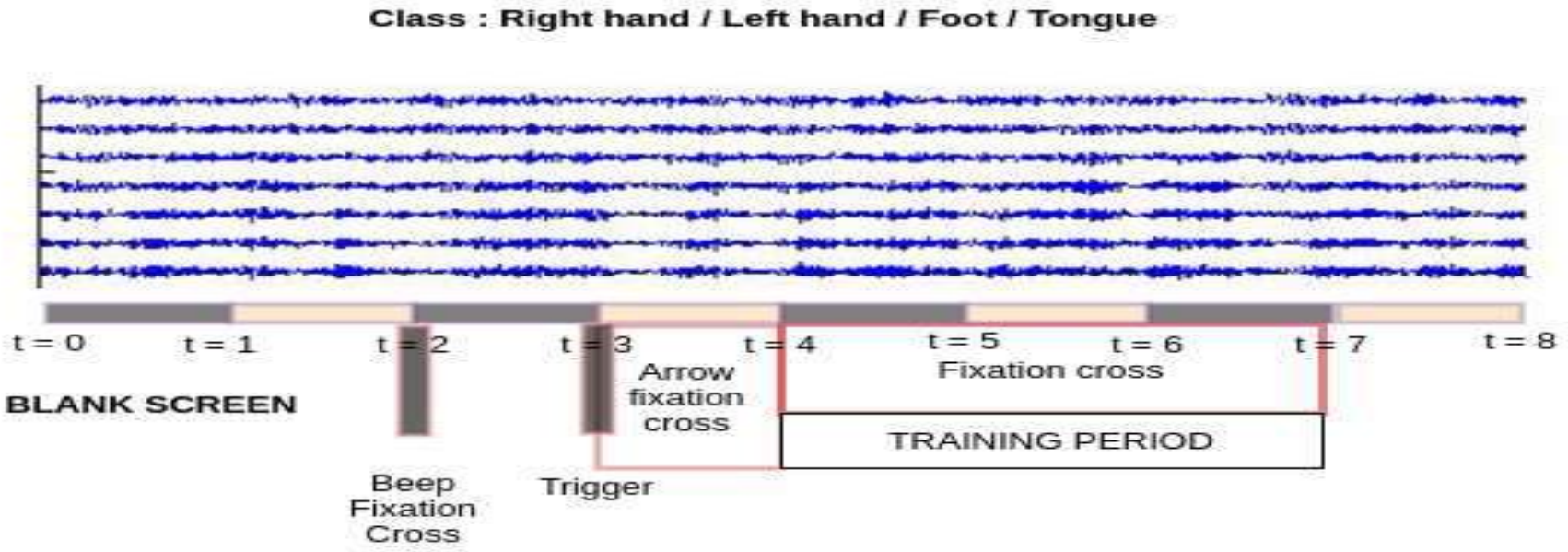
There is no real good reason to select subsets of data epochs. When comparing conditions – performed by creating contrast at the STUDY level (the group analysis interface which may also be used for single-subject analysis) – one may ignore specific data epochs.

- Non-overlapping segments
- Overlapping segments
    - Fully overlapping segments
    - Partially overlapping segments

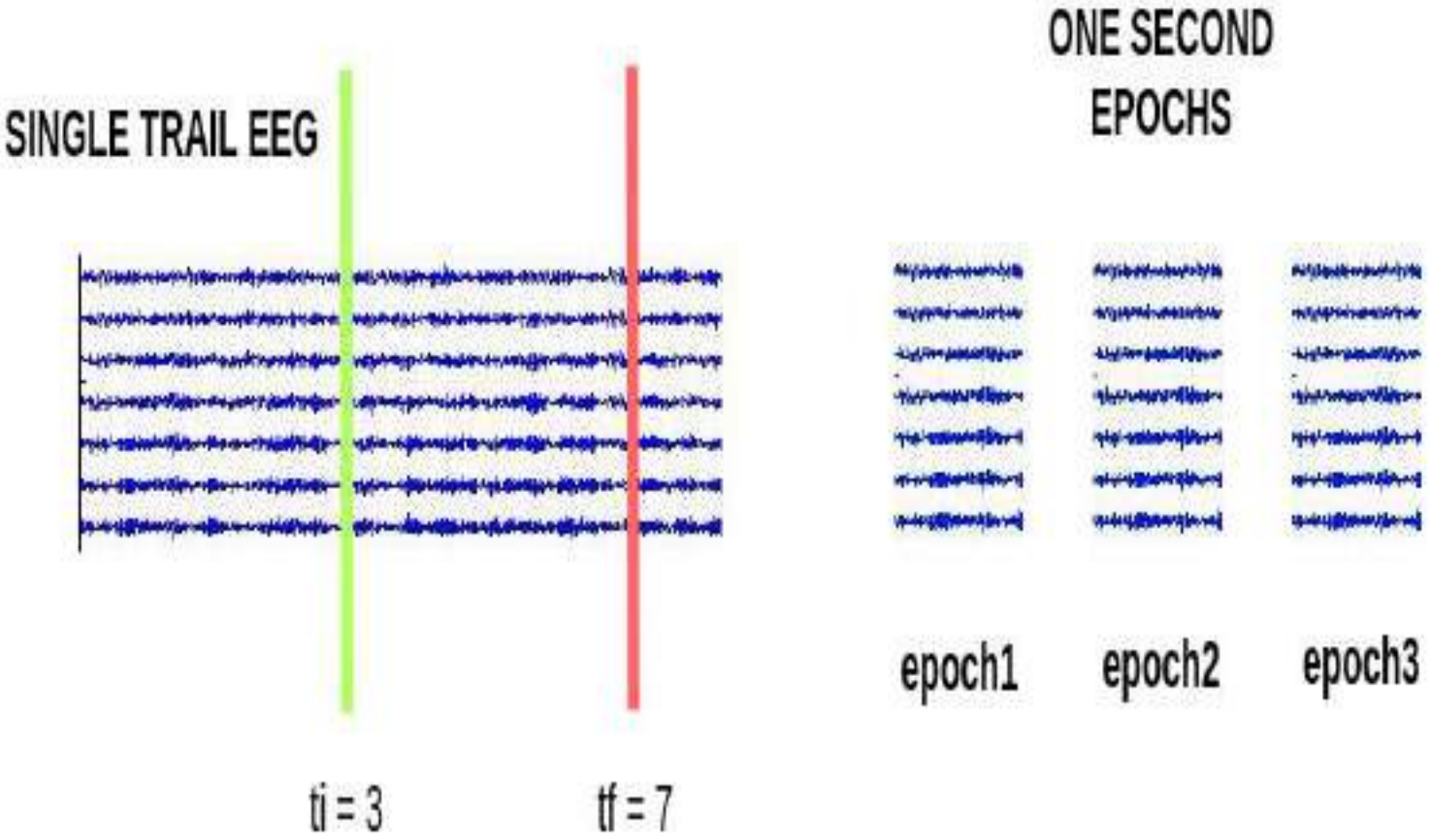
DATASET 1 SINGLE TRIAL

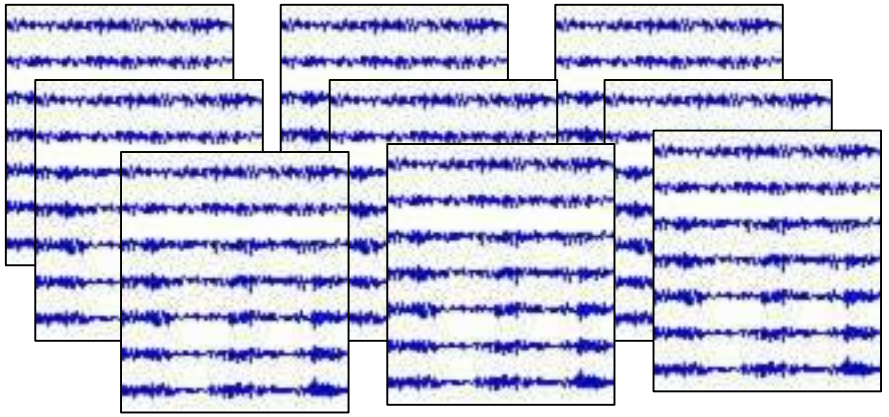


DATASET 2 SINGLE TRIAL

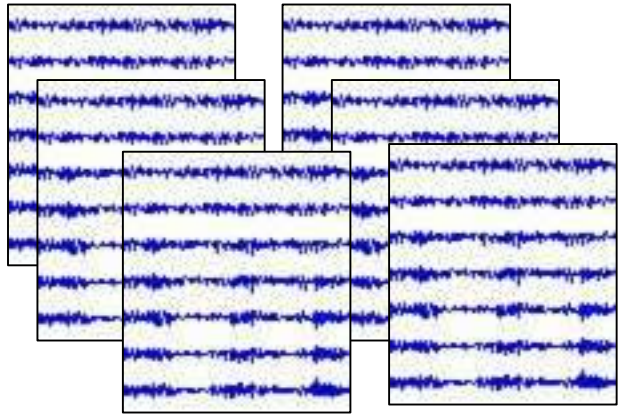


Data Epoching





TRAINING SET(80%)

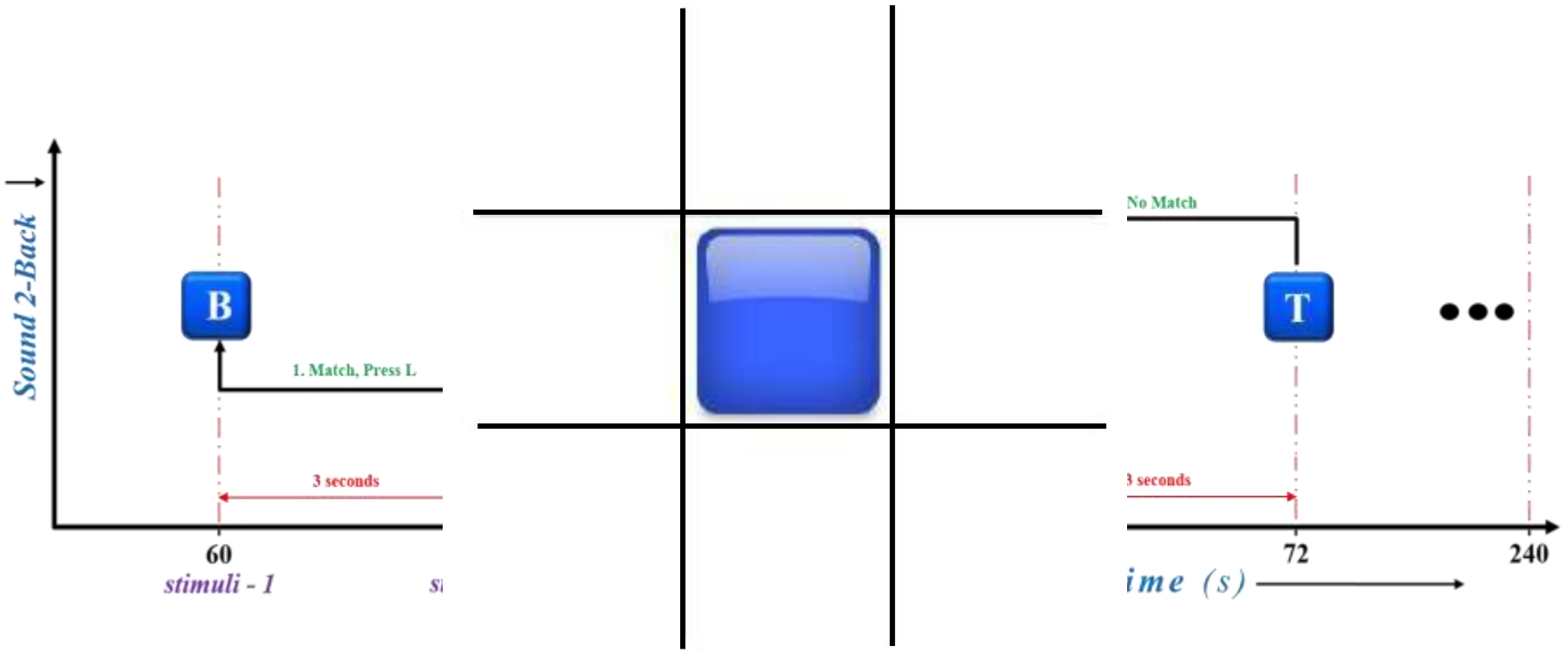


TEST SET(20%)

Cognitive Workload

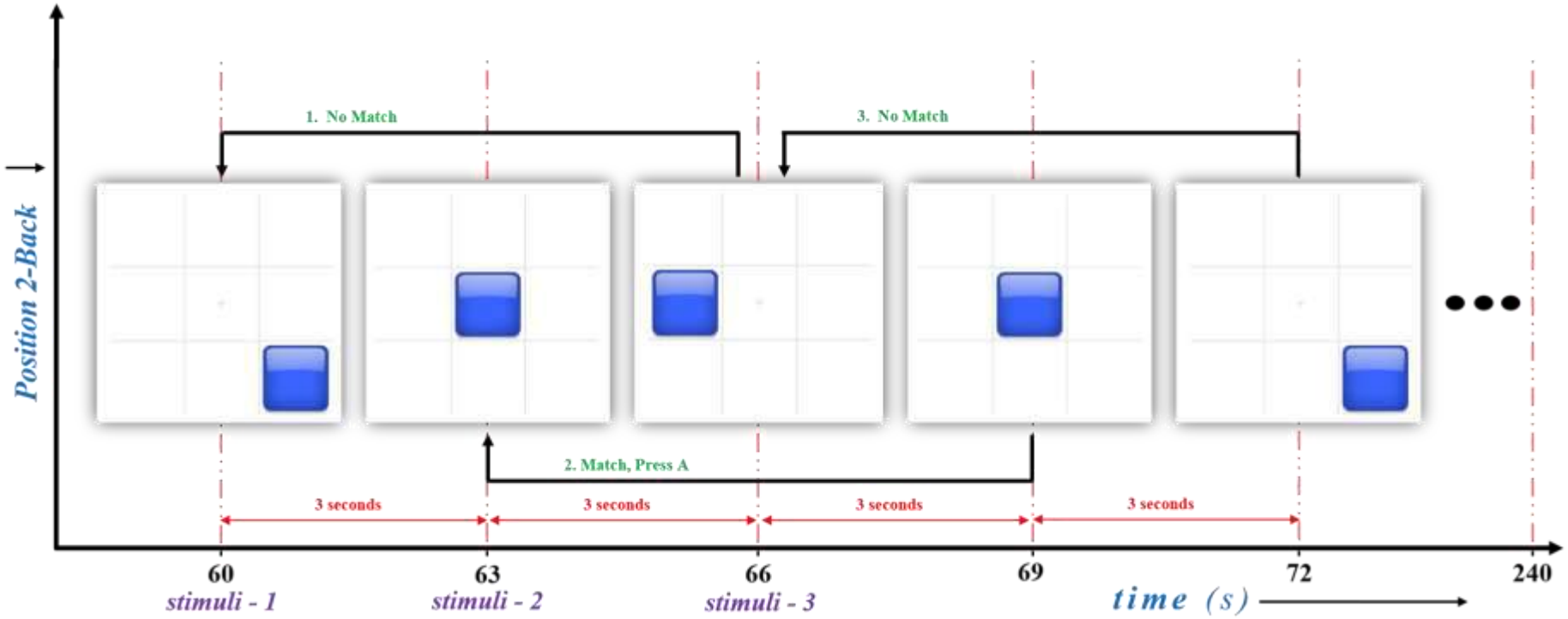
- Cognitive Workload (**CWL)** is a demand placed upon humans for mental recourses while performing a task.
- Mental resources include working memory, ability to process, etc.
- **Working memory (WM)** is a cognitive system with a limited capacity to hold a small amount of information and process it.

n-Back Task



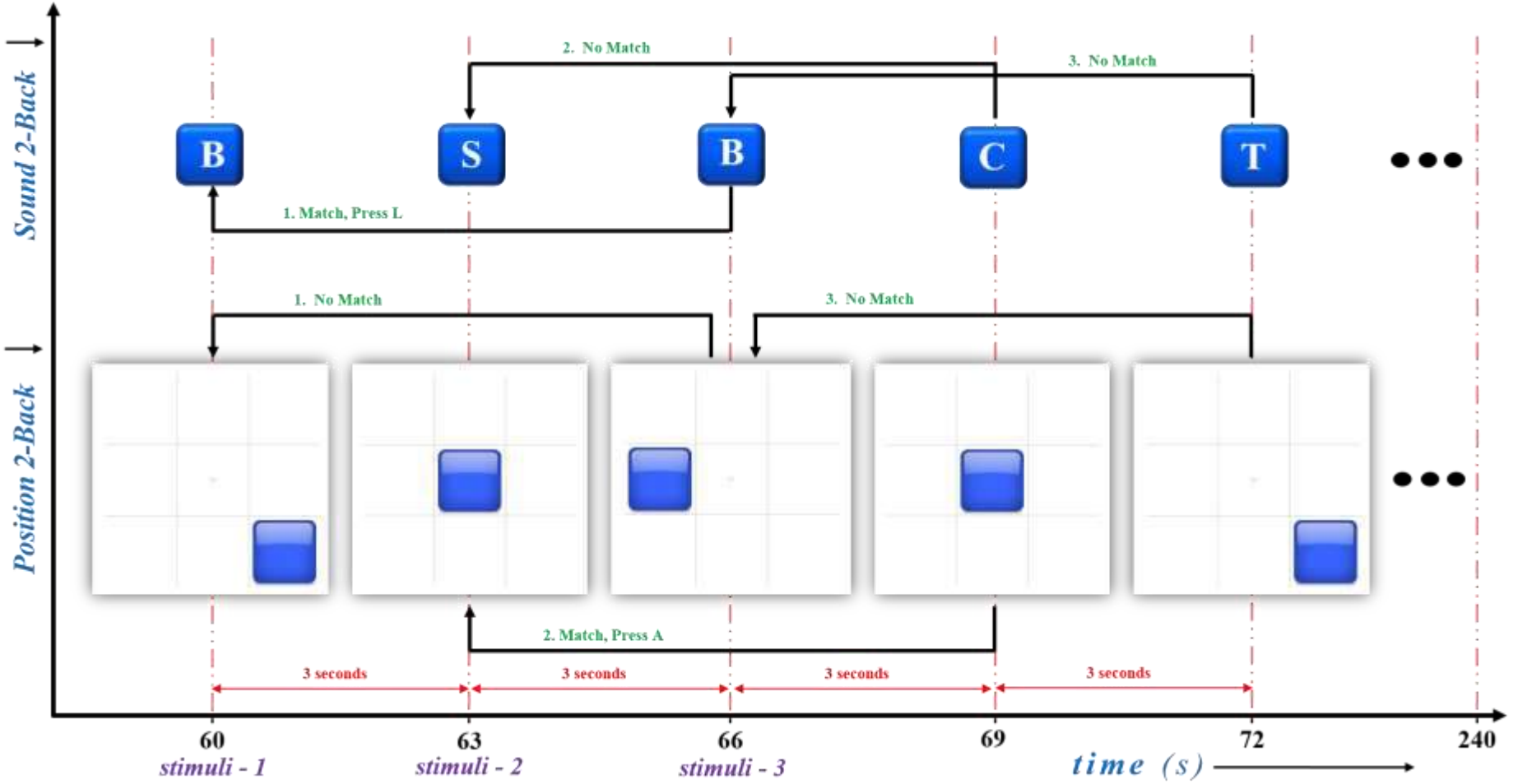
Sound *2*-Back

n-Back Task



Position *2*-Back

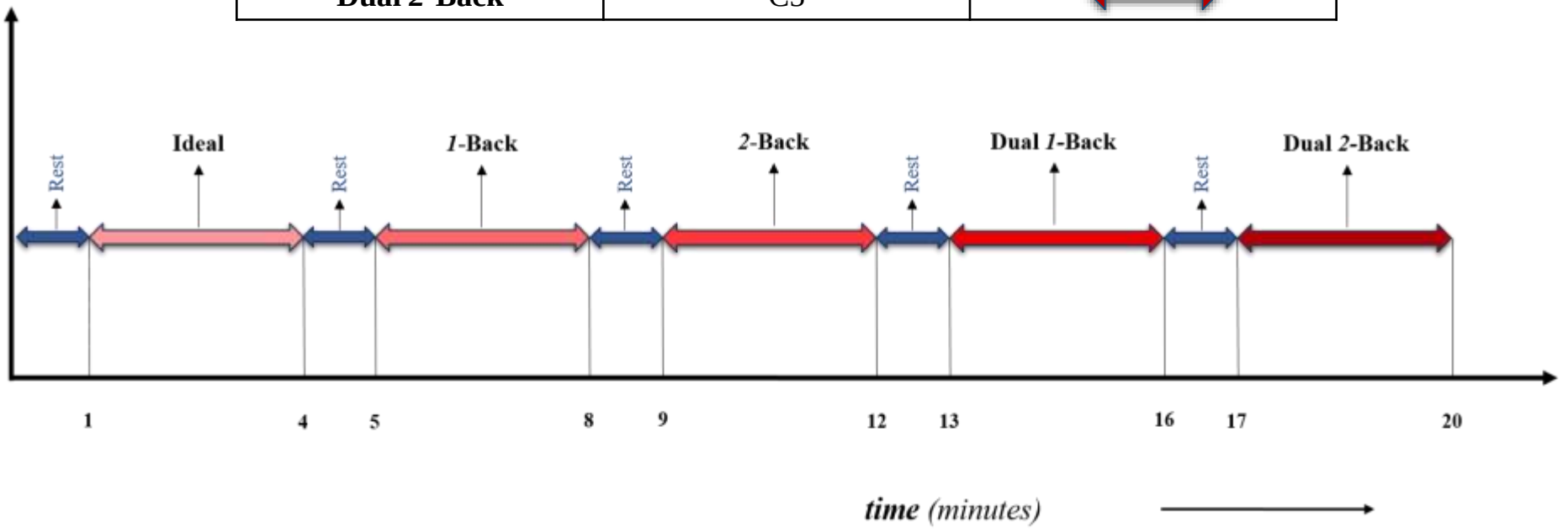
n-Back Task



Sound & Position *2*-Back

Data Epoching

Task	Cognitive Workload Class	Colour Code depicting Load Variation
Idle	C1	
***1*-Back**	C2	
***2*-Back**	C3	
Dual *1*-Back	C4	
Dual *2*-Back	C5	



Thank you!!

Spatial Filtering

Spatial Filtering

- Spatial filtering techniques take as input brain signals recorded from several different locations (or “channels”) and transform them in one of several ways.
- Possible goals include
  - enhancing local activity
  - reducing noise that is common across channels,
  - decreasing the dimensionality of the data,
  - finding projections that maximize discrimination between different classes

Bipolar

- Extract bipolar signals

$$\widetilde{s_{i,j}} = s_i - s_j$$

- Highlight the **electrical potential differences** between the two electrodes of interest (i and j).

Laplacian

- *Laplacian filtering*, extracts local activity at electrode *i* by subtracting the average activity present in the four orthogonal nearest neighboring electrodes

$$\tilde{s} = s_i - \frac{1}{4}\sum_{i\in\theta} s_i$$

Common Average Referencing

- *Common average referencing* (CAR), enhances the local activity at electrode *i* by subtracting the average over all electrodes

$$\widetilde{s_i} = s_i - \frac{1}{N}\sum_{i=1}^{N} s_i$$

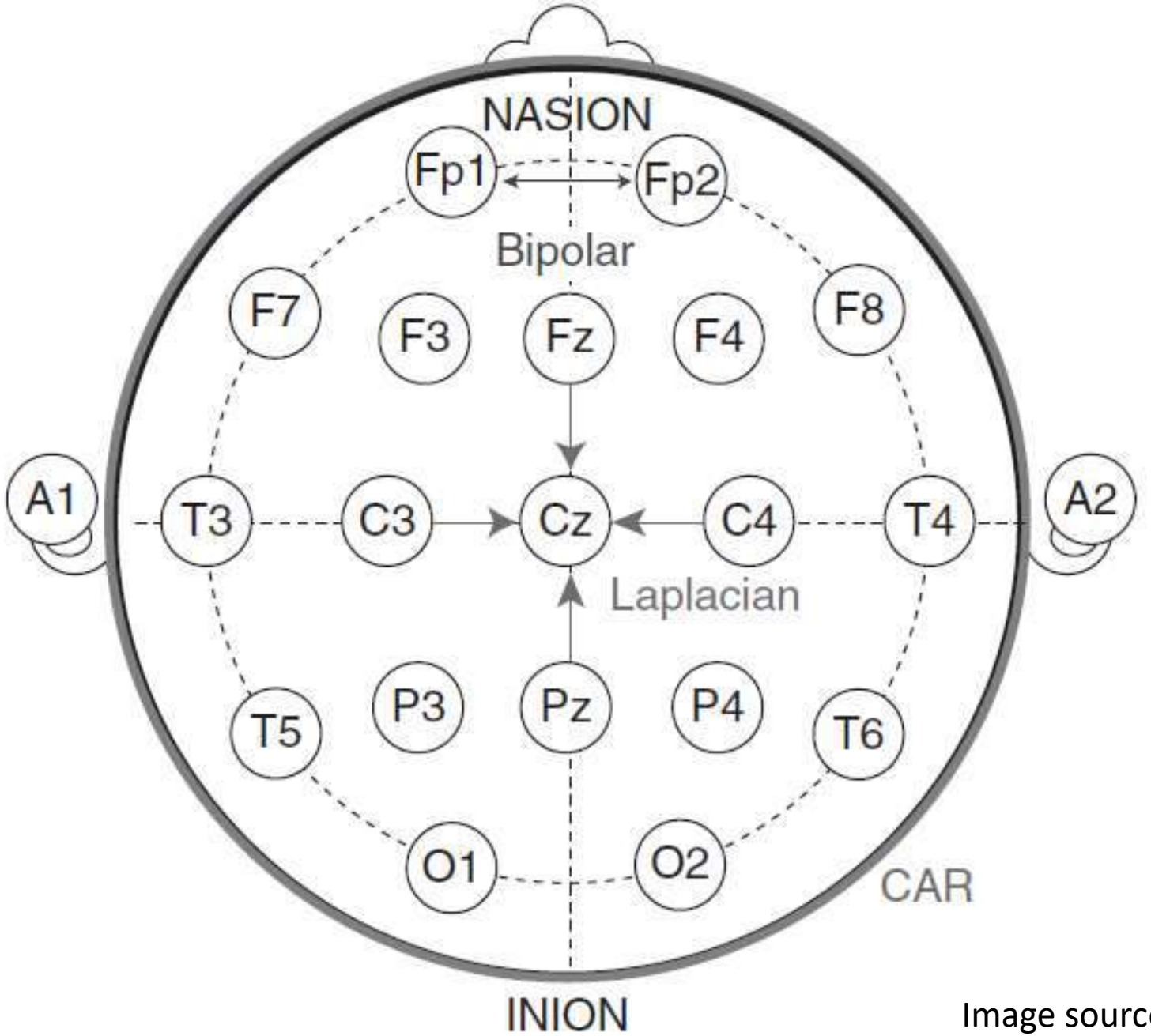


Image source: Rajesh P.N, Rao- Brain Computer Interfacing: An Introduction

Principal Component Analysis

- The goal in *principal component analysis* (PCA) (also called the *Hotelling transform*) is to discover the underlying statistical variability in the data and reduce the data's dimensionality from D to a much smaller number of dimensions L ($L << D$).

- PCA achieves this goal by
  - Finding the directions of maximum variance in the D-dimensional data
  - Rotating the original coordinate system to align with these directions of maximum variance

Principal Component Analysis

- Most natural signals, including brain signals are redundant
- In the case of EEG measurements from N electrodes
  - Measurements from nearby electrodes may be correlated
  - Underlying rhythms across multiple electrodes.
- PCA attempts to find the dominant directions of variability in the data.
- New data points can be projected along the “principal” directions. Each projection is called a “principal component”
- The resulting *L*-dimensional vector can be used as a feature vector for classification or other purposes in BCI applications

PCA - Steps

- Suppose we are given $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_M$ (N x 1) vectors

N: # of features

M: # data

Step 1: compute sample mean

$$\bar{\mathbf{x}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{x}_i$$

Step 2: subtract sample mean (i.e., center data at zero)

$$\Phi_i = \mathbf{x}_i - \bar{\mathbf{x}}$$

Step 3: compute the sample covariance matrix Σ_x

$$\Sigma_{\mathbf{x}} = \frac{1}{M}\sum_{i=1}^{M}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T = \frac{1}{M}\sum_{i=1}^{M}\Phi_i\Phi_i^T = \frac{1}{M}AA^T$$

where $A=[\Phi_1\ \Phi_2\ ...\ \Phi_M]$

i.e., the columns of A are the Φ_i

(N x M matrix)

PCA - Steps

Step 4: compute the eigenvalues/eigenvectors of Σ_x

$$\Sigma_x = \lambda_i u_i$$

where we assume $\lambda_1 > \lambda_2 > \ldots > \lambda_N$

Note : most software packages return the eigenvalues (and corresponding eigenvectors) is decreasing order – if not, you can explicitly put them in this order)

Since Σ_x is symmetric, $\langle u_1, u_2, \ldots, u_N \rangle$ form an orthogonal basis in R^N and we can represent any $\mathbf{x} \in R^N$ as:

$$\mathbf{x} - \bar{\mathbf{x}} = \sum_{i=1}^{N} y_i u_i = y_1 u_1 + y_2 u_2 + \ldots + y_N u_N$$

$$y_i = \frac{(\mathbf{x} - \bar{\mathbf{x}})^T u_i}{u_i^T u_i} = (\mathbf{x} - \bar{\mathbf{x}})^T u_i \qquad if \quad \| u_i \| = 1$$

i.e., this is just a "change" of basis!

$$\mathbf{x} - \bar{\mathbf{x}}: \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ . \\ . \\ x_N \end{bmatrix} \rightarrow \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ . \\ . \\ y_N \end{bmatrix}$$

Note : most software packages normalize u_i to unit length to simplify calculations; if not, you can explicitly normalize them)

PCA - Steps

Step 5: dimensionality reduction step – approximate **x** using only the first K eigenvectors (K<<N) (i.e., corresponding to the K largest eigenvalues where K is a parameter):

$$\mathbf{x} - \bar{\mathbf{x}} = \sum_{i=1}^{N} y_i u_i = y_1 u_1 + y_2 u_2 + ... + y_N u_N$$

approximate $\mathbf{x}$ *by* $\hat{\mathbf{x}}$
using first K eigenvectors only

$$\hat{\mathbf{x}} - \bar{\mathbf{x}} = \sum_{i=1}^{K} y_i u_i = y_1 u_1 + y_2 u_2 + ... + y_K u_K$$

(reconstruction)

$$\mathbf{x} - \bar{\mathbf{x}}: \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ . \\ . \\ x_N \end{bmatrix} \rightarrow \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ . \\ . \\ y_N \end{bmatrix} \rightarrow \hat{\mathbf{x}} - \bar{\mathbf{x}}: \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_K \end{bmatrix}$$

note that if K=N, then $\hat{\mathbf{x}} = \mathbf{x}$
(i.e., zero reconstruction error)

What is the Linear Transformation implied by PCA?

- The linear transformation y = Tx which performs the dimensionality reduction in PCA is:

$$\hat{\mathbf{x}} - \bar{\mathbf{x}} = \sum_{i=1}^{K} y_i u_i = y_1 u_1 + y_2 u_2 + \ldots + y_K u_K$$

$$(\hat{\mathbf{x}} - \bar{\mathbf{x}}) = U \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_K \end{bmatrix}$$

$$where\ \ U = [u_1\, u_2 \ldots u_K]\ \ \ N\,x\,K\ \text{matrix}$$

i.e., the columns of U are the the first K eigenvectors of Σ_x

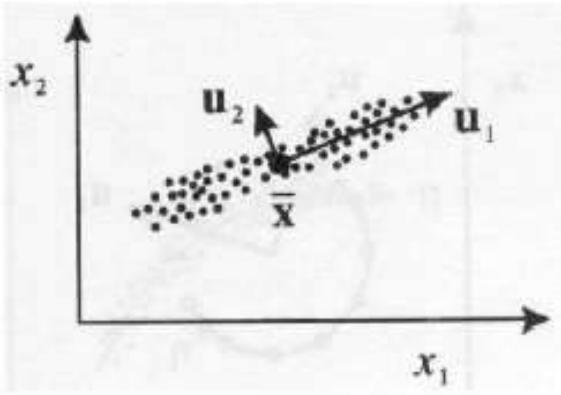
$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_K \end{bmatrix} = U^T (\hat{\mathbf{x}} - \bar{\mathbf{x}})$$

$T = U^T$ K x N matrix

i.e., the rows of T are the first K eigenvectors of Σ_x

Interpretation of PCA

- PCA chooses the eigenvectors of the covariance matrix corresponding to the largest eigenvalues.
- The eigenvalues correspond to the variance of the data along the eigenvector directions.
- Therefore, PCA projects the data along the directions where the data varies most.
- PCA preserves as much information in the data by preserving as much variance in the data.



u_1: direction of max variance
u_2: orthogonal to u_1

How do we choose K ?

- *K* is typically chosen based on how much information (variance) we want to preserve:

Choose the smallest K that satisfies the following inequality:

$$\frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{N} \lambda_i} > T \quad where\ T\ is\ a\ threshold\ (e.g., 0.9)$$

- If T=0.9, for example, we "preserve" 90% of the information (variance) in the data.

- If K=N, then we "preserve" 100% of the information in the data (i.e., just a "change" of basis and $\hat{\mathbf{x}} = \mathbf{x}$)

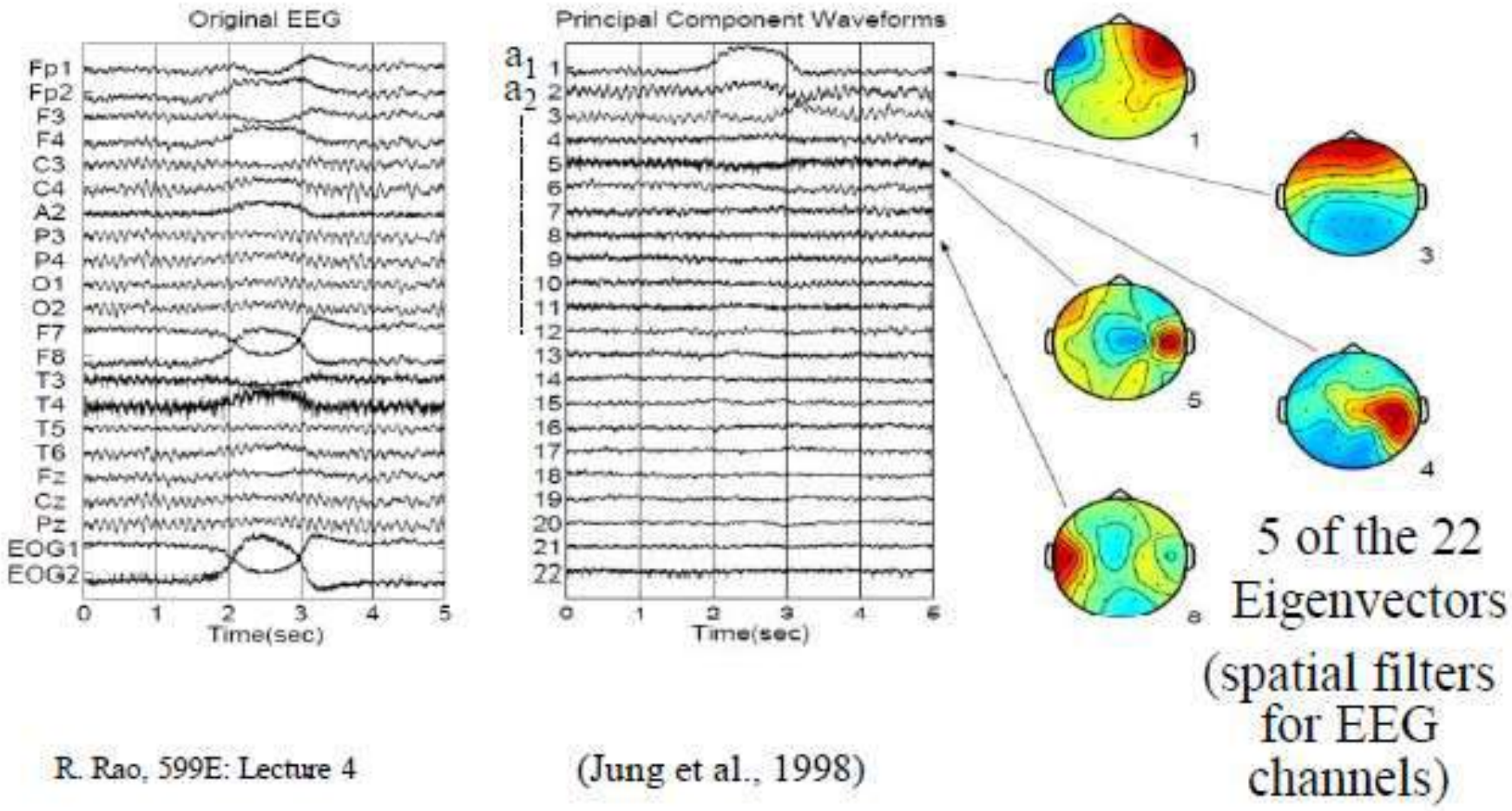
Data Normalization

- The principal components are dependent on the *units* used to measure the original variables as well as on the *range* of values they assume.
- Data should always be normalized prior to using PCA.
- A common normalization method is to transform all the data to have zero mean and unit standard deviation:

$$\frac{x_i - \mu}{\sigma}$$

where μ and σ are the mean and standard deviation of the i-th feature x_i

PCA applied to EEG



R. Rao, 599E: Lecture 4

(Jung et al., 1998)

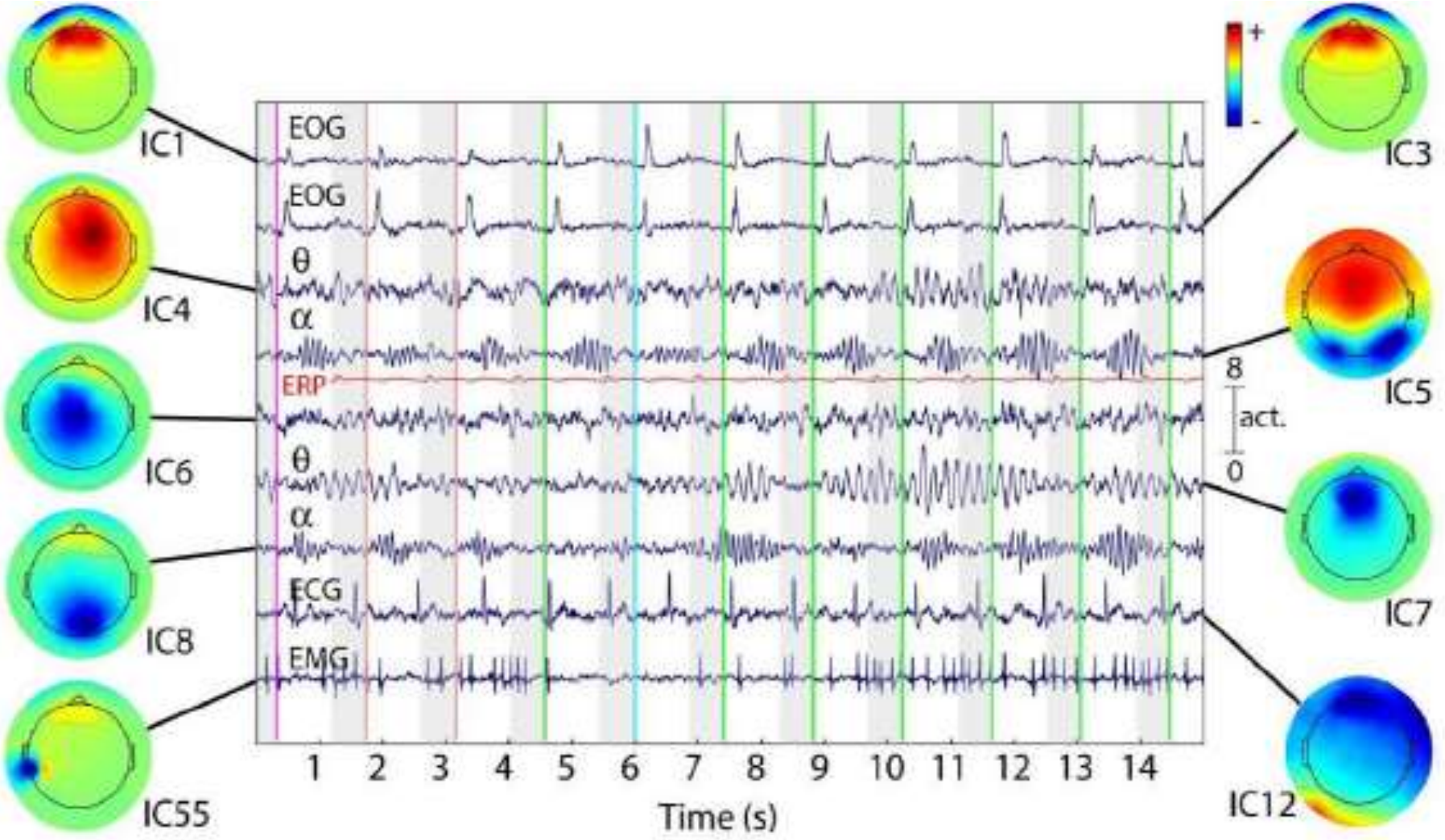
Independent Component Analysis

- PCA finds a matrix **V** that decorrelates the inputs but the resulting feature vector $\boldsymbol{a}$ may still retain higher order statistical dependencies
- There may be a possibility that the variables are independent.
- ICA tries to find a matrix W of filters (columns of W) such that the output **a** is **statistically independent**:

$$\boldsymbol{a} = \boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{x} \text{ such that } \boldsymbol{P}(\boldsymbol{a}) \approx \prod_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{D}} \boldsymbol{P}(\boldsymbol{a}_{\boldsymbol{i}})$$

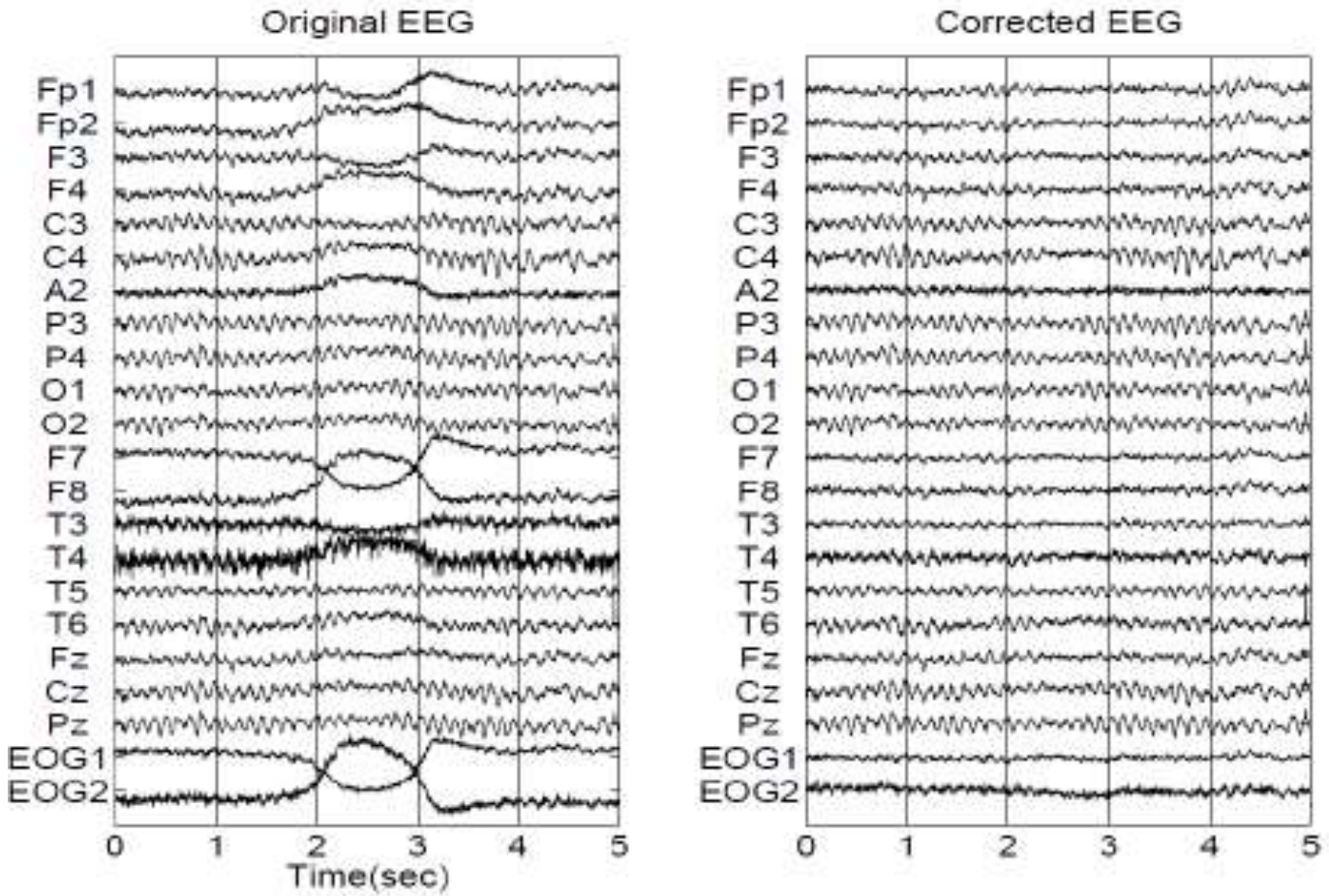
Independent Component Analysis

- ICA assumes sources are linearly mixed to produce x
- The feature vector dimension in ICA can be lesser than, equal to, or greater than the number of input dimensions.
- ICA has proved useful in a variety of settings in BCI applications, ranging from the use of the output vector **a** as a feature vector in classification.



Application of ICA to EEG data for isolating electrooculographic (EOG) (eye-related), electromyographic (EMG) (muscle-related) and electrocardiographic (ECG) (heart-related) artifacts, and unmixing putative source signals in the brain. Image (adapted from Onton and Makeig, 2006)

ICA for Artifact Removal in EEG



Common Spatial Pattern

- Supervised Technique
- Data is labeled with class to which each data vector belongs
  - E.g., EEG obtained for right versus left hand imagery
- CSP finds a matrix of spatial filters
  - the variance of the filtered data for one class is maximized
  - variance of the filtered data for the other class is minimized
- CSP filters can significantly enhance discrimination ability between the two classes

CSP applied to EEG for Right/Left Hand Imagery

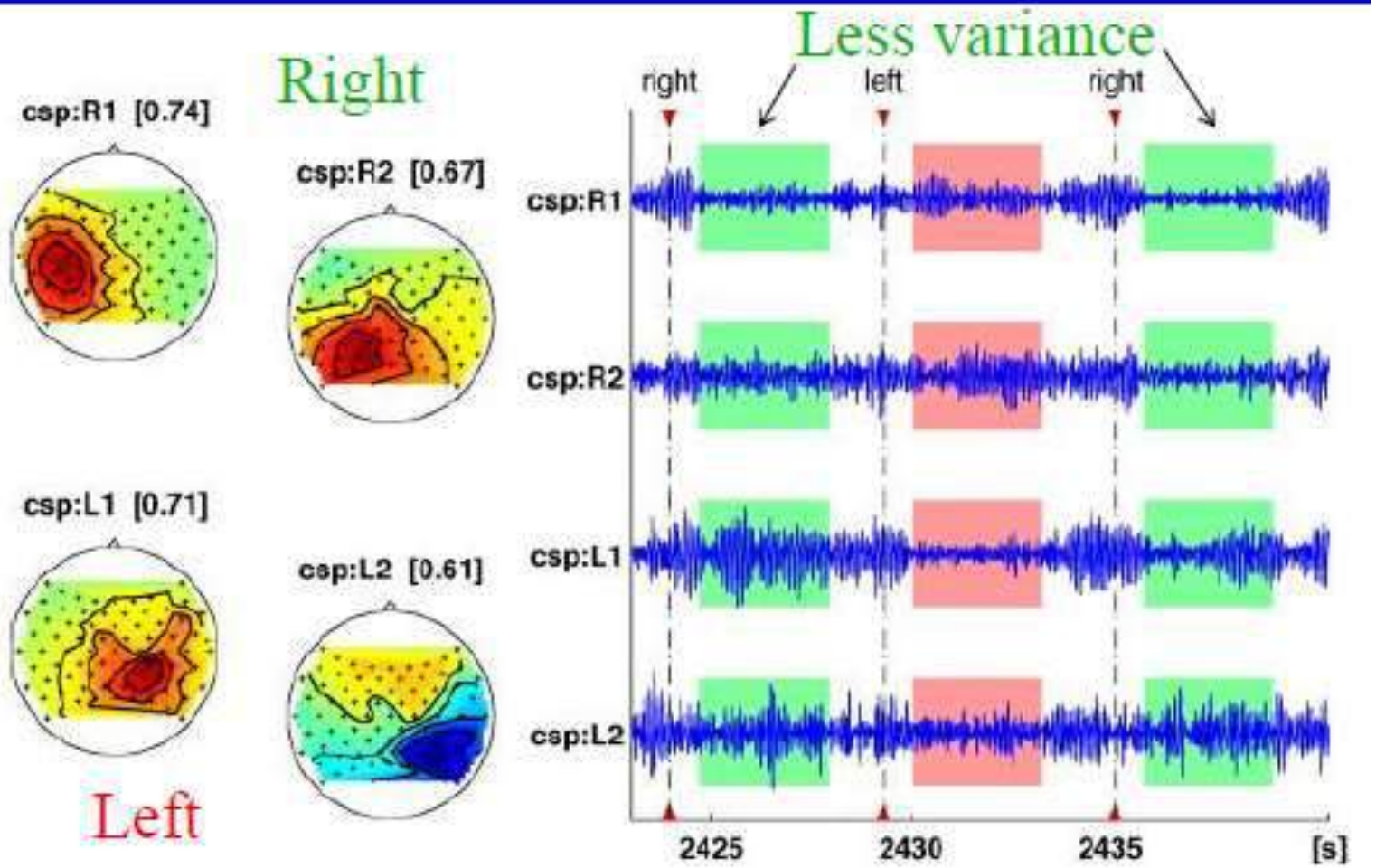
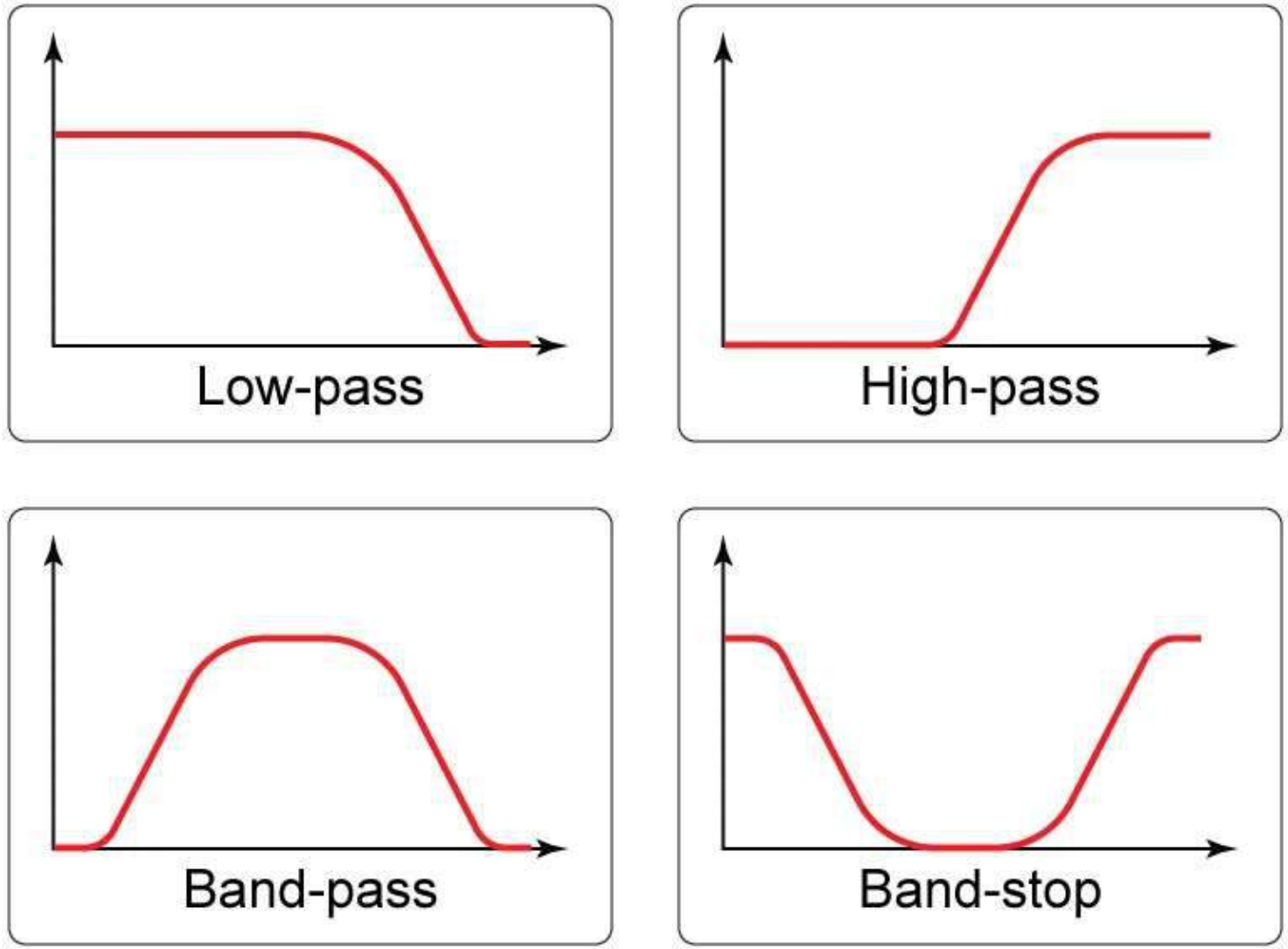


Image source: http://www.sciencedirect.com/science/article/pii/S0165027007004657

Artifact Reduction Techniques

- Thresholding
  - If the magnitude or some other characteristic of a recorded EOG or EMG signal exceeds a pre-determined threshold, the brain signals recorded during that epoch are deemed to be contaminated and rejected.
- Band-Stop and Notch Filtering
  - Band-stop filtering is a useful artifact reduction technique that attenuates the components of a signal in a specific frequency band and passes the rest of the components of the signal.
  - A notch filter set to the 59–61 Hz band (in the United States) for filtering out the 60 Hz power-line noise artifact.


Low-pass
High-pass
Band-pass
Band-stop

Artifact Reduction Techniques

- Linear Modeling
  - A simple way of modeling the effect of artifacts on a recorded brain signal is to assume that the effect is additive.
  - For example, if $EEG_i(t)$ is the EEG signal recorded from electrode *i* at time *t*, then a model of how the signal has been contaminated could be:

$$EEG_i(t) = EEG_i^{true}(t) + K \cdot EOG(t)$$

  - EEG_i^{true}(t) is the uncontaminated ("true") EEG signal from electrode *i* at time *t,* $EOG(t)$ is the recorded EOG signal at time *t* and *K* is a constant.
  - Given an estimated value for K, one can obtain an estimate of the true EEG signal using:

$$EEG_i^{true}(t) = EEG_i(t) - K \cdot EOG(t)$$

Brain Computer Interaction

Feature Extraction

Course Instructors

Dr. Sreeja S R

Assistant Professor

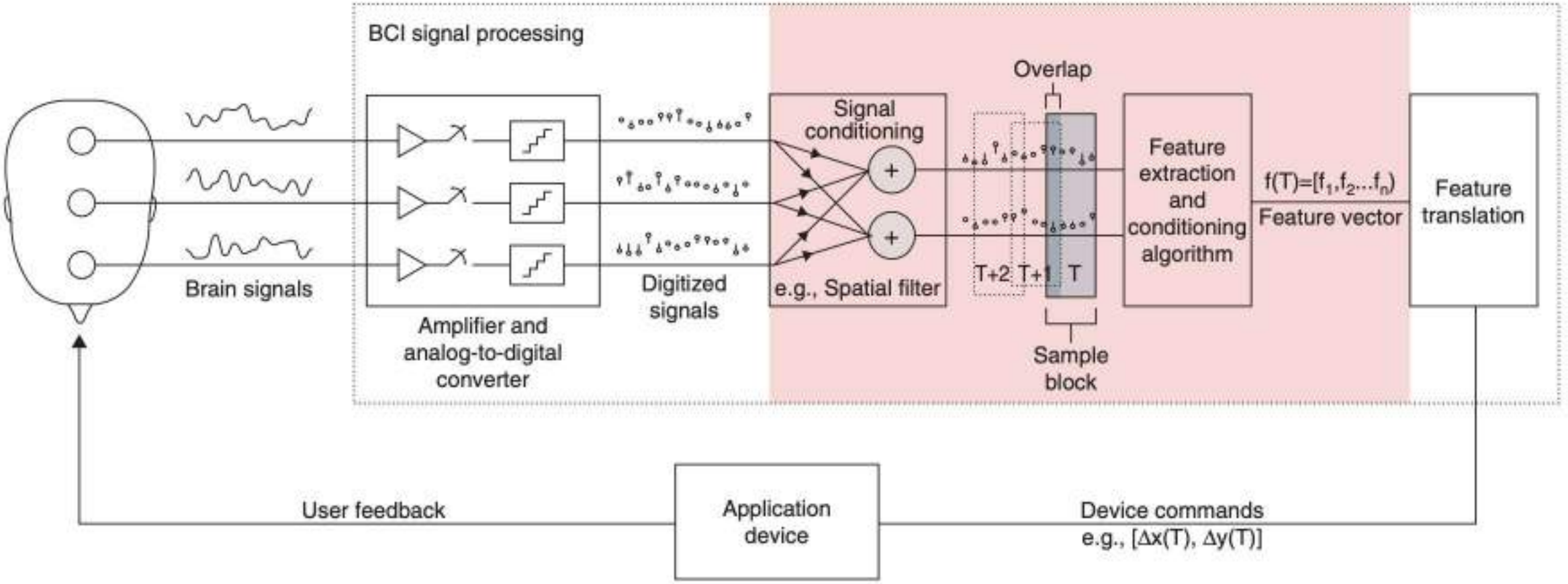
Indian Institute of Information Technology

IIIT Sri City

Features

- The purpose of a BCI is to **detect and quantify characteristics of brain signals** that indicate what the user wants the BCI to do, to translate these measurements in real time into the desired device commands, and to provide concurrent feedback to the user.
- The brain-signal characteristics used for this purpose are called ***signal features*, or simply *features***.
- ***Feature extraction*** is the process of distinguishing the pertinent signal characteristics from extraneous content and representing them in a compact and/or meaningful form, amenable to interpretation by a human or computer.

Overall structure of a BCI



Feature vector

- A **fundamental signal feature** is simply a direct measurement of the signal. They usually provide limited relevant information about typically complex brain signals.
- Thus, it is more common for BCIs to use features that are **linear or nonlinear combinations, ratios, statistical measures, or other transformations of multiple fundamental features** detected at multiple electrodes and/or multiple time points.
- Such complex features, if selected appropriately, can reflect the user's desires more accurately than the fundamental features themselves.
- Most features used in BCI applications are based on **spatial, temporal, and/or spectral analyses** of brain signals or the relationships among them.
- Furthermore, in order to determine the user's wishes as accurately as possible, most BCIs extract a number of features simultaneously. This set of features is referred to as a ***feature vector***.

Feature vector

To be effective for BCI applications, a feature should have the following attributes:

- its spatial, temporal, spectral characteristics, and dynamics can be precisely characterized for an individual user or population of users
- it can be modulated by the user and used in combination with other features to reliably convey the user's intent
- its correlation with the user's intent is stable over time and/or can be tracked in a consistent and reliable manner

BCI SIGNAL PROCESSING - Fourier Analysis

- Much of signal-processing theory is rooted in Fourier analysis, which transforms a time-domain (i.e., time on the x-axis) signal into its equivalent frequency-domain (i.e., frequency on the -axis) representation.
- The primary utility of Fourier analysis is to decompose a signal into individual sinusoidal components that can be isolated and evaluated independently.
- Using Fourier analysis, practically any signal can be accurately represented as the sum of a number (possibly an infinite number) of amplitude-scaled and time-shifted sinusoids at specific frequencies.
- In order to model these signals, it is necessary to properly adjust the phase and the magnitude of each sinusoid.

Fourier Analysis

- For an arbitrary signal x(t), the magnitude (scale) and phase (shift) of the sinusoid at each frequency [ω(radians) = 2πf (Hz)] required to represent an arbitrary signal can be determined from the Fourier transform:

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt = \int_{-\infty}^{\infty} x(t)[\cos\omega t + j\sin\omega t]dt$$

$$= \underbrace{\int_{-\infty}^{\infty} x(t)\cos\omega t dt}_{a(\omega)} + j\underbrace{\int_{-\infty}^{\infty} x(t)\sin\omega t dt}_{b(\omega)}$$

$$= a(\omega) + jb(\omega)$$

The magnitude and phase for each sinusoidal component are given as:

$$Magnitude: \quad |X(\omega)| = \sqrt{a^2(\omega) + b^2(\omega)}$$

$$Phase: \quad \theta = \arg(X(\omega)) = \tan^{-1}\left(\frac{b(\omega)}{a(\omega)}\right)$$

THE THREE STEPS OF FEATURE EXTRACTION

The process of feature extraction is discussed here as a three-step procedure:

- signal conditioning to reduce noise and to enhance relevant aspects of the signals
- extraction of the features from the conditioned signals
- feature conditioning to properly prepare the feature vector for the feature-translation stage

FIRST STEP: SIGNAL CONDITIONING

- The first step of feature extraction is called signal conditioning or preprocessing.
- This step enhances the signal by preemptively eliminating known interference (i.e., artifacts) or irrelevant information, and/or by enhancing spatial, spectral, or temporal characteristics of the signal that are particularly relevant to the application.
- It is common to have some prior knowledge about the general signal characteristics relevant for a particular application, and this knowledge is used in conditioning.

FIRST STEP: SIGNAL CONDITIONING

Signal conditioning can include a number of different procedures that can primarily be categorized as:

- frequency-range prefiltering
- data transformation and normalization
- spatial filtering
  - Data independent spatial filtering (*common- average reference* and *surface Laplacian spatial filters)*
  - Data dependent spatial filtering (PCA, ICA and CSP)
- removal of environmental interference and biological artifacts

Frequency-Range prefiltering

- Digital filters are central to digital signal processing. They modify the frequency content of a digital signal by attenuating some frequencies (or frequency ranges) and amplifying others. Each successive sample of a digitized signal is passed through a digital filter to produce a new value as the output.

  - Low pass filter (higher frequencies are attenuated and lower frequencies are preserved)
  - High pass filter (very specific ranges of signal frequencies can be amplified, attenuated, preserved, and/or eliminated)
  - Bandpass filter (A band-pass filter preserves signal power within a specified continuous frequency range, while attenuating signal power outside of this range)
  - Notch filter/Band-stop filter (A notch filter is the converse of a bandpass filter; it attenuates signal power within a specified continuous frequency range, while preserving signal power outside of this range)

Data Transformation

Data transformation are needed to bring uniformity to the data. In addition, it can be used to scale the data to a preferred range. Following are few methods as data transformation.

- Change of origin
- Change of scale
- Change of origin and scale
- Decimal scaling
- Min-Max normalization
- Standard normalization

Data Transformation

1. Change of Origin:

- Arbitrarily choose a constant. If sample values are integers, an integer constant is preferred.
- Shift data point by subtracting (or adding) the chosen constant from each sample observation.
- This technique is useful when data values are large, and variability is not so large.

Lemma 6.1:

a) The range of original data is preserved by a change of origin transformation.

b) If $\bar{x}$ is the old mean and c is the chosen constant, then the new mean of the transformed data is $\bar{x} = c + \bar{x}'$

Data Transformation

Example:

Apply change of origin method to the following data and calculate the old and new means.

$$x = \{115, 128, 110, 104, 133\}$$

Let us arbitrarily chose the constant 120 and subtract this from each value in x to get the transformed data x' as

$$x' = \{-5, 8, -10, -16, 13\}$$

The new mean $\bar{x}' = \frac{\sum x_i'}{5} = -2$

The mean of the original data is

$$\bar{x} = c + \bar{x}' = 120 - 2 = 118$$

2. Change of scale:

- This method is used to shorten the range of large numbers or lengthen the range of very small numbers.
- Chose (arbitrarily) a non-zero constant c. If c is less then the minimum of the observation, each value will be transformed to a value greater than 1. On the other hand, if it is greater than the maximum of the observation, then each value will be transformed to a value less than 1.
- If the value is between min and max of the sample, then the transformed values lie on the real line (positive real line if all x_i's are positive)
- If all values are small fractions, we may multiply a large constant to scale them up and vice-versa.

3. Change of origin and scale:

- This is the most frequently used method to standardize data values. Depending upon the constants used to change the origin and scale, a variety of transformation intervals can be obtained.

Example: A sample in the range (a, b) can be transformed to a new interval (c, d) by the following transformation.
Let x is the original and y is the transformed value. Then

$$y = c + \frac{(d-c)}{(b-a)} \times (x-a)$$

[Prove that all values in the range (a, b) are mapped onto the range (c, d)].

4. Min-Max Normalization:

- Min-Max normalization performs a linear transformation on the original data.
- Suppose, min_A and max_A are the minimum and maximum values of an attribute A. Min-Max normalization maps a value, v of A to v' in the range $min_A{}'$ and $max_A{}'$ using the following transformation:

$$v' = \frac{(v - min_A)}{(max_A - min_A)} \times (max'_A - min'_A) + min_A{}'$$

If $[min_A{}' , max_A{}'] = [0,1]$, then it is a special case of Min-max normalization.

5. Standard Normalization:

- This transformation is so called because it is extensively used in statistics in standardizing normal scores.
- Here, the origin is changed using the mean of the sample, and the scale is changed using the standard deviation of the sample.

$$v' = \frac{(v - \bar{A})}{\sigma_A}$$

where $\bar{A}$ and σ_A are the mean and std for the attribute A.

- This method is also alternatively termed as z-score normalization (zero-mean normalization) and the transformed values are called z-scores.

5. Standard Normalization:

Example: Given X = {32, 80, 56, 75, 69, 26, 44, 50}. Apply the standard normalization.

Here, the mean $\bar{A} = 54$, $\sigma_A^2 = 390$. Thus the z-scores are

$v' = \frac{(v-\bar{A})}{\sigma_A}$ = {-1.1140133 1.3165612 0.1012739 1.0633763 0.7595545 -1.4178351 -0.5063697 -0.2025479}

Note: It is very interesting to note that z-scores will always lie in the interval [-3, +3]

6. Decimal Scaling:

This data transformation method is same as the “change of scale” method by either scale up or scale down.

$$v' = \frac{v}{10^j}$$

Here j represents movement of decimal points. Decided based on the maximum value in the data.

SECOND STEP: EXTRACTING THE FEATURES

- ***BLOCK PROCESSING***
  - For most BCI applications, it is highly desirable for the processing to occur in real time. Prior to feature extraction, the incoming signal samples are commonly segmented into consecutive, possibly overlapping, sample blocks.
  - A feature vector is created from the signal samples within each individual sample block. The feature vectors from the successive sample blocks are then fed to the translation algorithm, which produces a device command or user feedback corresponding to each sample block or corresponding to sets of consecutive sample blocks.
  - For efficient online implementation, the length and overlap of these sample blocks should fit the relevant temporal dynamics of the signal, the feature-extraction method, the nature of the application, and the concurrent user feedback, as well as the available processing power.
    - E.g., BCI cursor control
    - P300 response

SECOND STEP: EXTRACTING THE FEATURES

- ***TIME (TEMPORAL) FEATURES***

1. Peak-Picking and Integration
   - Peak-picking simply determines the minimum or maximum value of the signal samples in a specific time block (usually defined relative to a specific preceding stimulus) and uses that value (and possibly its time of occurrence) as the feature(s) for that time block.
   - The signal can be averaged or integrated over all or part of the time block to yield the feature(s) for the block. Some form of averaging or integration is typically preferable to simple peak-picking, especially when the responses to the stimulus are known to vary in latency and/or when unrelated higher-frequency activity is superimposed on the relevant feature

SECOND STEP: EXTRACTING THE FEATURES

- ***TIME (TEMPORAL) FEATURES***

2. Correlation and Template-Matching
   - The similarity of a response to a predefined template might also be used as a feature.

SECOND STEP: EXTRACTING THE FEATURES

Statistical features

Sl. No	Features	Short description
1	MEAN	Mean value
2	STD	Standard deviation
3	MAX VALUE	Maximum positive amplitudes
4	MIN VALUE	Maximum negative amplitudes
5	SKEWNESS	a measure of asymmetry of the distribution
6	KURTOSIS	a measure of flatness of the distribution
7	MEDIAN	the middle value of a set of ordered data

SECOND STEP: EXTRACTING THE FEATURES

Interval or period analysis features.

Sl. No	Features	Short description
8	LINE LENGTH	Line length
9	MEAN VV AMPL	Mean of vertex-to-vertex amplitudes
10	VAR VV AMPL	Variance of vertex-to-vertex amplitudes
11	MEAN VV TIME	Mean of vertex-to-vertex times
12	VAR VV TIME	Variance of vertex-to-vertex times
13	MEAN VV SLOPE	Mean of vertex-to-vertex slope
14	VAR VV SLOPE	Variance of vertex-to-vertex slope
15	ZERO CROSSING	Number of zero crossings in a signal
16	MIN MAX NUMBER	Number of local minima and maxima
17	COEFF OF VARIATION	a statistical measure of the deviation of a variable from its mean, standard deviation divided by mean
18	AMPL RANGE	The difference between the maximum positive and maximum negative Amplitude values

SECOND STEP: EXTRACTING THE FEATURES

Features derived from the first and second derivative.

Sl. No	Features	Short description
19	1st DIFF MEAN	Mean value of the first derivative of the signal
20	1st DIFF MAX	Maximum value of the first derivative of the signal
21	2nd DIFF MEAN	Mean value of the second derivative of the signal
22	2nd DIFF MAX	Maximum value of the second derivative of the signal

SECOND STEP: EXTRACTING THE FEATURES

The Hjorth parameters

Sl. No	Features	Short description
23	HJORTH 1	Ability
24	HJORTH 2	Mobility $(\sigma x\,'/\sigma x)$
25	HJORTH 3	Complexity $\frac{(\sigma x''/\sigma x''}{(\sigma x\,'/\sigma x)}$

NOTE:

$\sigma x\,'$ is the standard deviation of the first derivative of the signal
$\sigma x''$ is the standard deviation of the second derivative of the signal

SECOND STEP: EXTRACTING THE FEATURES

- ***FREQUENCY (SPECTRAL) FEATURES***
- Much brain activity manifests itself as continuous amplitude- and frequency-modulated oscillations. Therefore, it is often advantageous to accurately track these changes in the frequency domain. Although the Fourier transform is the most common method for converting from the time domain to the frequency domain, there are several alternatives that have characteristics that are particularly desirable given specific constraints or specific objectives. These include:
  - band power
  - fast Fourier transform (FFT)
  - autoregressive (AR) modeling

SECOND STEP: EXTRACTING THE FEATURES

FFT-based features calculated from the EEG spectra.

Sl. No	Features	Short description
26	FFT DELTA	0.1 - 3 Hz
27	FFT THETA	3 - 7 Hz
28	FFT ALPHA	7 - 12 Hz
29	FFT BETA	12 - 30 Hz
30	FFT GAMMA	30 - 40 Hz
31	FFT WHOLE	0.1 - 40 Hz

SECOND STEP: EXTRACTING THE FEATURES

FFT-based Spectral Features.

Sl. No	Features	Short description
32	FFT DT RATIO	*DELTA / THETA*
33	FFT DA RATIO	*DELTA / ALPHA*
34	FFT TA RATIO	*THETA / ALPHA*
35	FFT DTA RATIO	*(DELTA + THETA) / ALPHA*
36	FFT SEF	Spectral edge frequency (95 % of the total spectral power resides)
37	FFT SP-ROLL OFF	The frequency below which 85 % of the total spectral power resides

SECOND STEP: EXTRACTING THE FEATURES

Wavelet based Features.

Sl. No	Features	Short description
38	MIN WAV VALUE	Minimum value
39	MAX WAV VALUE	Maximum value
40	MEAN WAV VALUE	Mean value
41	MEDIAN WAV VALUE	Median value
42	STD WAV VALUE	Standard deviation
43	SKEWNESS WAV VALUE	Skewness
44	KURTOSIS WAV VALUE	Kurtosis
45	WAV BAND	Relative energy

SECOND STEP: EXTRACTING THE FEATURES

Wavelet based Features.

Sl. No	Features	Short description
46	ENTROPY SPECTRAL WAV	The spectral entropy
47	1st DIFF WAV MEAN	Mean value of the 1*st* derivative
48	1st DIFF WAV MAX	Maximum value of the 1*st* derivative
49	2nd DIFF WAV MEAN	Mean value of the 2*nd* derivative
50	2nd DIFF WAV MAX	Maximum value of the 2*nd* derivative
51	ENERGY PERCENT WAV	Percentage of the total energy of a detail/approximation
52	WAV ZERO CROSSING	Zero crossing
53	WAV COEFF OF VARIATION	Coefficient of variation
54	WAV TOTAL ENERGY	Total Energy

SECOND STEP: EXTRACTING THE FEATURES

Other Features.

Sl. No	Features	Short description
55	ENTROPY SPECTRAL	The spectral entropy
56	ENTROPY SHANNON	The Shannon entropy
57	MAX ABS XCORR EEG-EEG	Maximum positive amplitude of auto-correlation or cross-Correlation function
58	MEAN ABS XCORR EEG-EEG	Mean value of auto-correlation or cross-correlation function

THIRD STEP: FEATURE CONDITIONING

- The distributions and the relationships among the features can have a significant effect on the performance of the translation algorithm that follows feature extraction. These effects depend on the characteristics of the particular translation algorithm.

  - ***NORMALIZATION***
  - ***LOG-NORMAL TRANSFORMS***
  - ***FEATURE SMOOTHING***
  - ***PCA AND ICA***

Thank You!