

# Dai segnali multimediali ai segnali fisiologici: come l'analisi del segnale ci svela l'emozione



Francesca Gasparini

## **MULTI MEDIA SIGNAL PROCESSING LABORATORY**

[www.mmsp.unimib.it](http://www.mmsp.unimib.it)

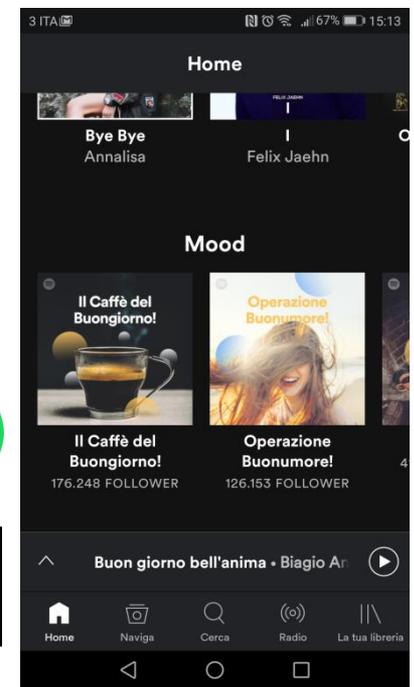


Image, Video and Audio Processing  
Analysis and Classification  
Machine Learning



Siamo quotidianamente «aggrediti» da stimoli multimediali: immagini, audio, video, testo.

Questi stimoli sono in grado di cambiare il nostro «mood», il nostro stato d'animo.



Pubblichiamo dati multimediali: immagini, audio, video, testo per comunicare il nostro stato d'animo.



Si può prevedere quale emozione sia legata ad un segnale multimediale?

Come possiamo  
MISURARE l'emozione  
legata ad un segnale?



### MISURE SOGGETTIVE:

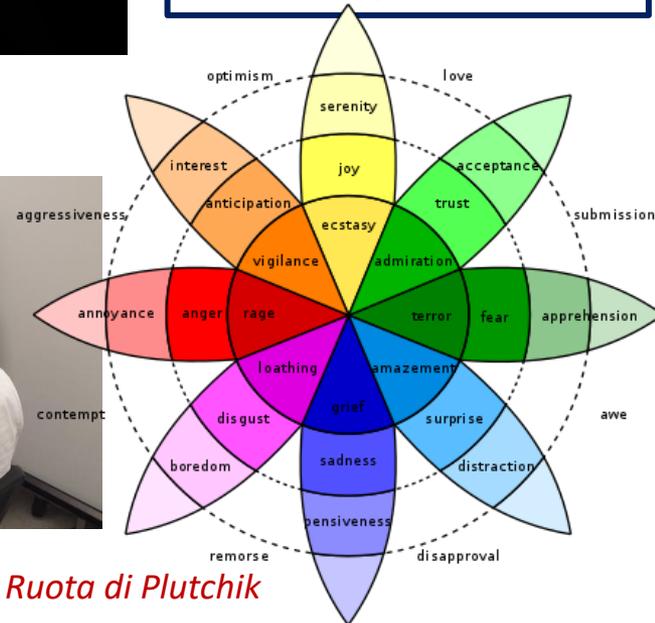
Indotte dal segnale  
sul soggetto:

- Dati fisiologici
- Dati psicovisuali

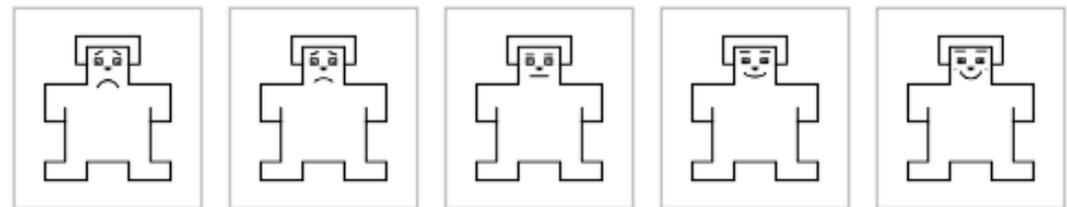
### MISURE OGGETTIVE:

Caratteristiche del segnale:

- Analisi del colore
- Analisi dei contenuti frequenziali
  - Analisi dei bordi
  - Analisi dei pattern
  - Analisi delle forme
- Valutazione dell'estetica
  - Presenza di persone
  - Presenza di oggetti
    - .....



*Ruota di Plutchik*



negative



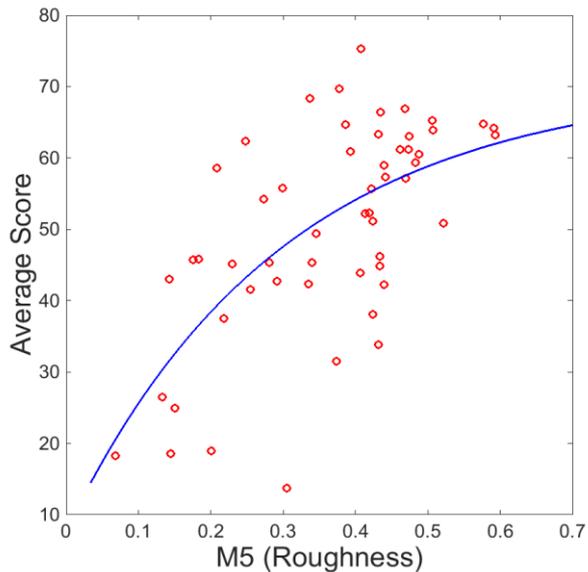
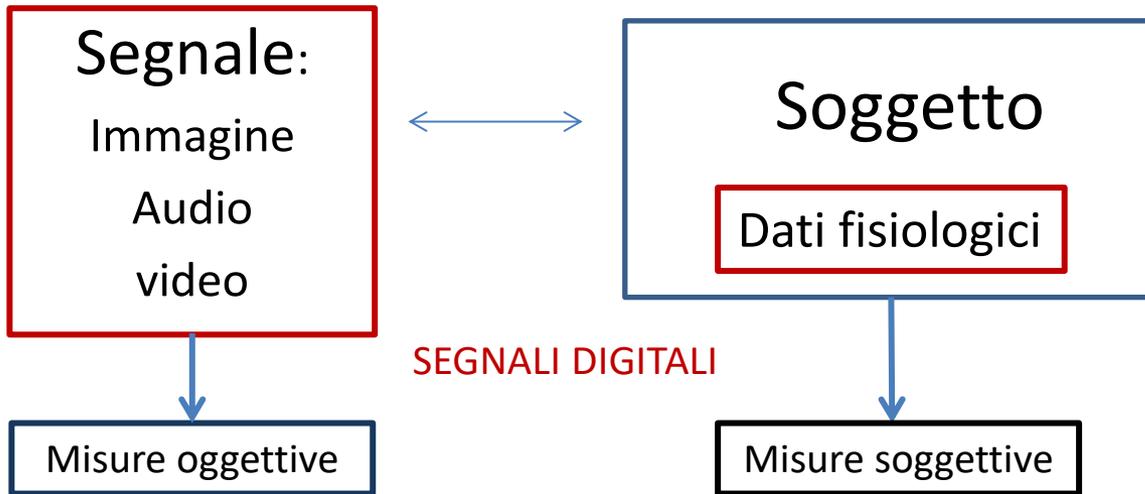
neutral



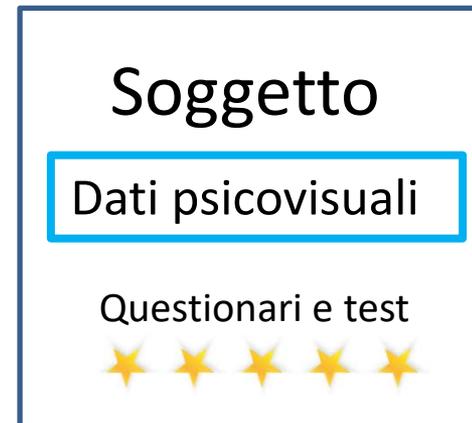
positive

*SAM self assessment manikin*

# APPROFONDIAMO

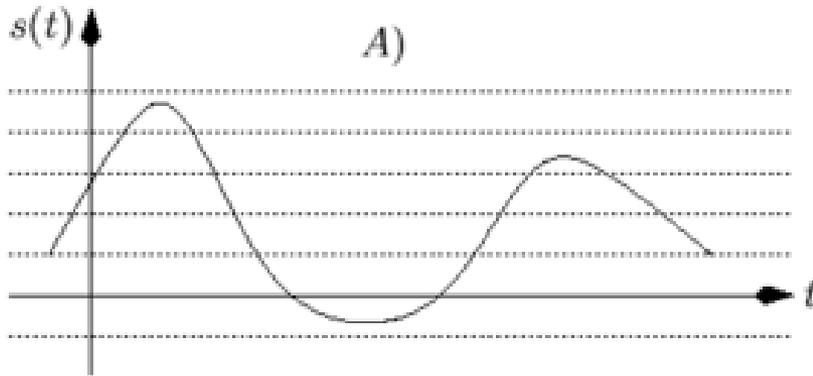


Le misure oggettive o soggettive vengono messe in correlazione con la ground truth: la verità. In questo modo si può capire quale misura descrive meglio l'emozione indotta nel soggetto

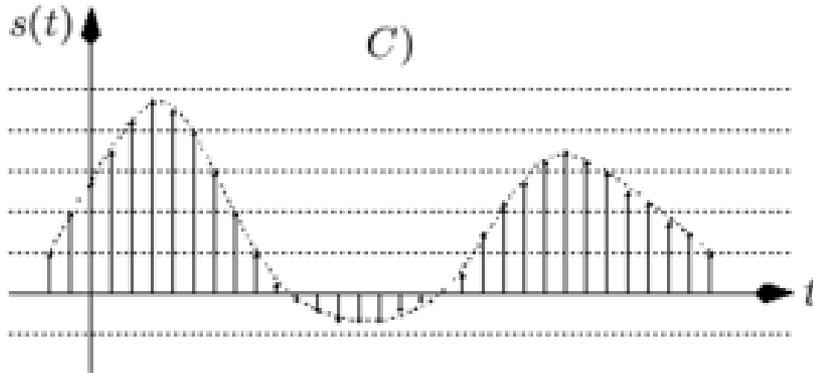
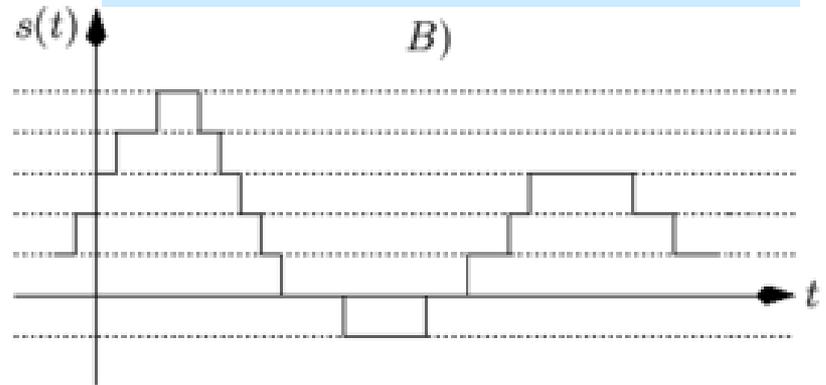


# CLASSIFICAZIONE DEI SEGNALI

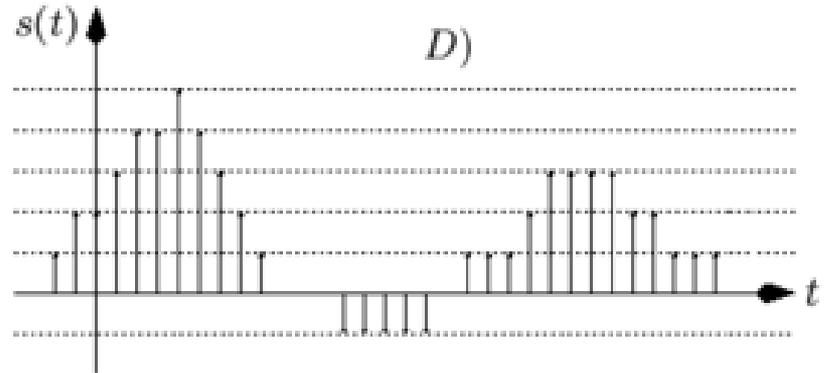
$D=R, C=R$  segnale analogico



$D=R, C=K$  segnale analogico quantizzato



$D=K, C=R$ ; segnale campionato



$D=K, C=K$ ; segnale digitale.

# Segnale digitale-immagine a colori

**Immagine a colori:** combinazione di 3 segnali bidimensionali R,G,B, campionato nello spazio, in una griglia di pixel, es. 1024x1024, e quantizzato nell'ampiezza: 8 bit per pixel per canale R G B, valori assunti tra 0-255

221 214 222 228 227 225 224...  
221 223 223 222 220 221 222...  
223 222 223 222 223 225 227...  
223 220 221 220 222 224 225...  
222 217 216 218 218 221 222...  
...

**R**

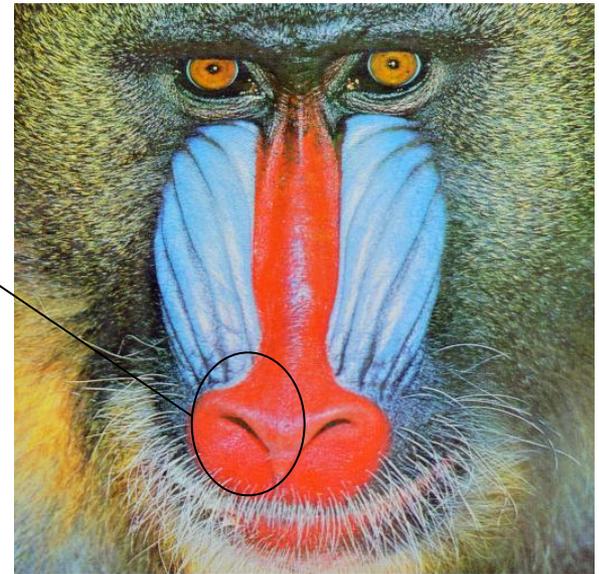


97 91 99 106 108 105 102...  
93 96 100 100 101 102 103...  
91 94 96 99 100 103 105...  
87 88 91 94 95 100 102...  
83 80 83 86 88 93 94...  
...

**G**

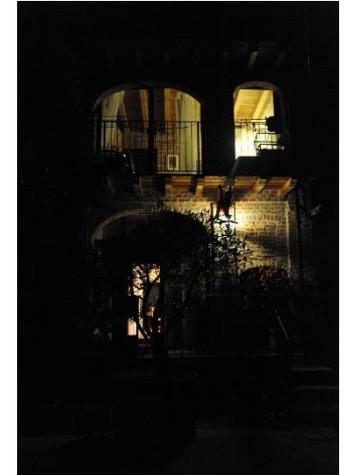
84 89 94 103 104 104 101...  
82 87 93 95 97 98 99...  
78 83 87 91 95 98 100...  
71 75 78 82 88 92 94...  
64 64 68 73 75 82 83...  
...

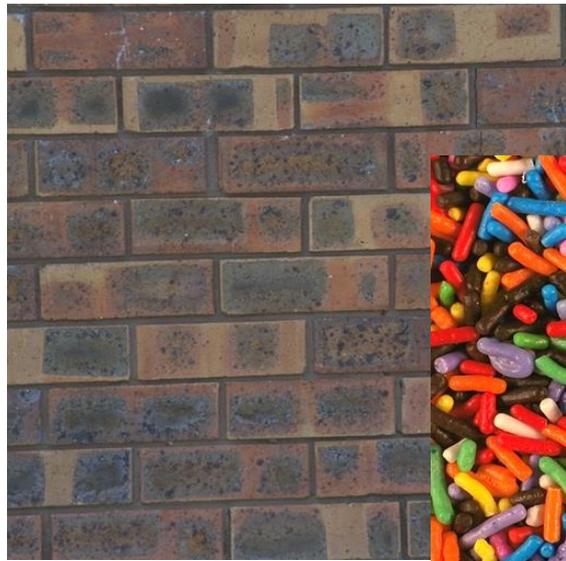
**B**



# QUALI SONO LE MISURE OGGETTIVE SUL SEGNALE? (immagine digitale)

- Misure di contrasto
  - Analisi del colore
- Analisi dei contenuti frequenziali
  - Analisi dei contorni
  - Analisi della texture
  - Analisi delle forme
- Valutazione dell'estetica
  - Presenza di persone
  - Presenza di oggetti
    - .....





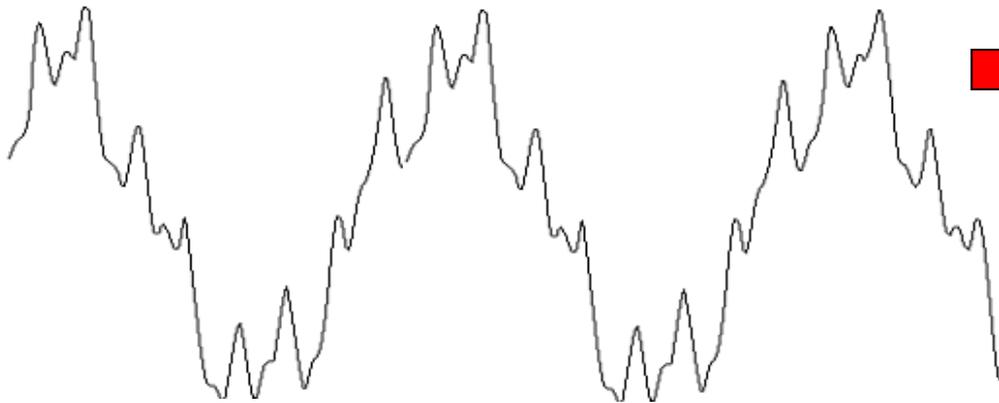
# Analisi di Fourier



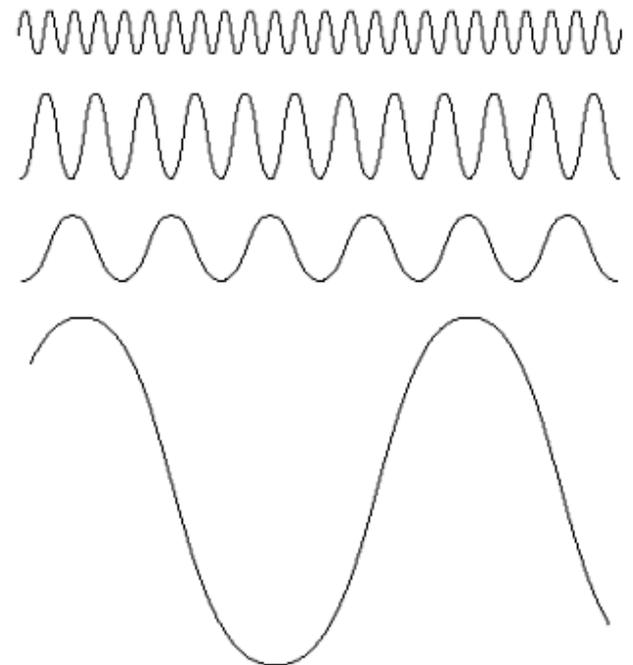
# Analisi di Fourier

- Decomporre il segnale in costituenti sinusoidali di differenti frequenze.
- Essa consente quindi di vedere il segnale non più nel dominio del tempo/spazio ma nel dominio delle frequenze.

*tempo/spazio*



*frequenze*



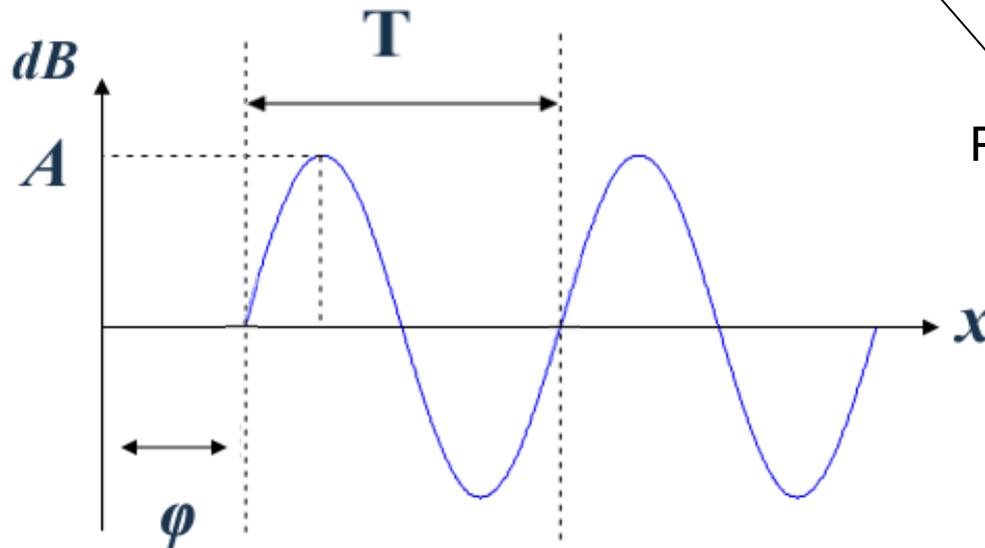
# Analisi di Fourier

- Una funzione armonica è una funzione periodica del tipo:

$$y = A \sin(\omega x + \varphi)$$

$$y = A \cos(\omega x + \varphi)$$

Ampiezza



Fase

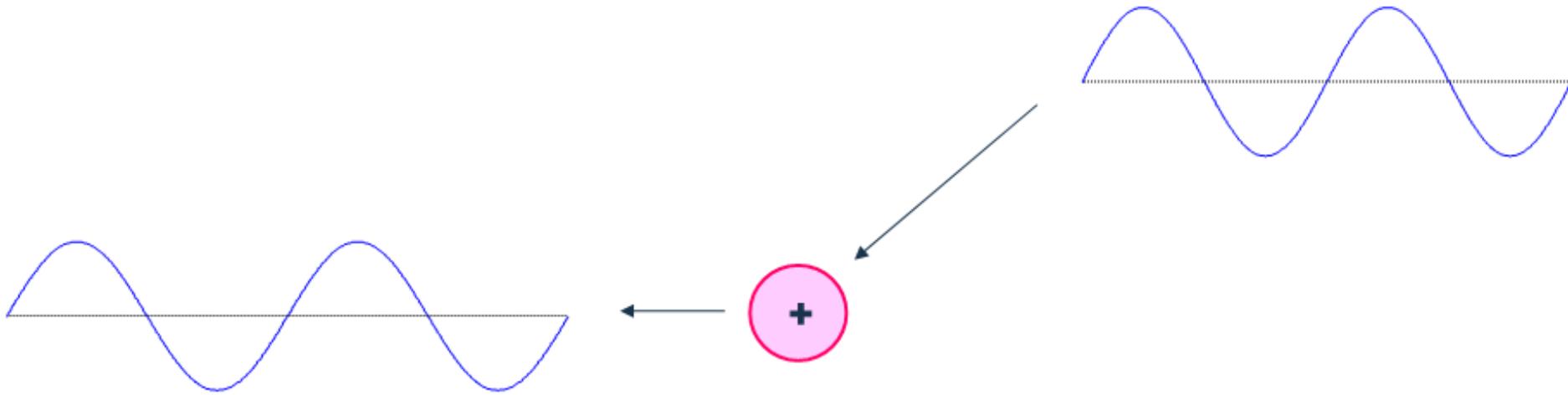
Pulsazione

$$\omega = 2\pi / T$$

$1/T = \text{Frequenza}$

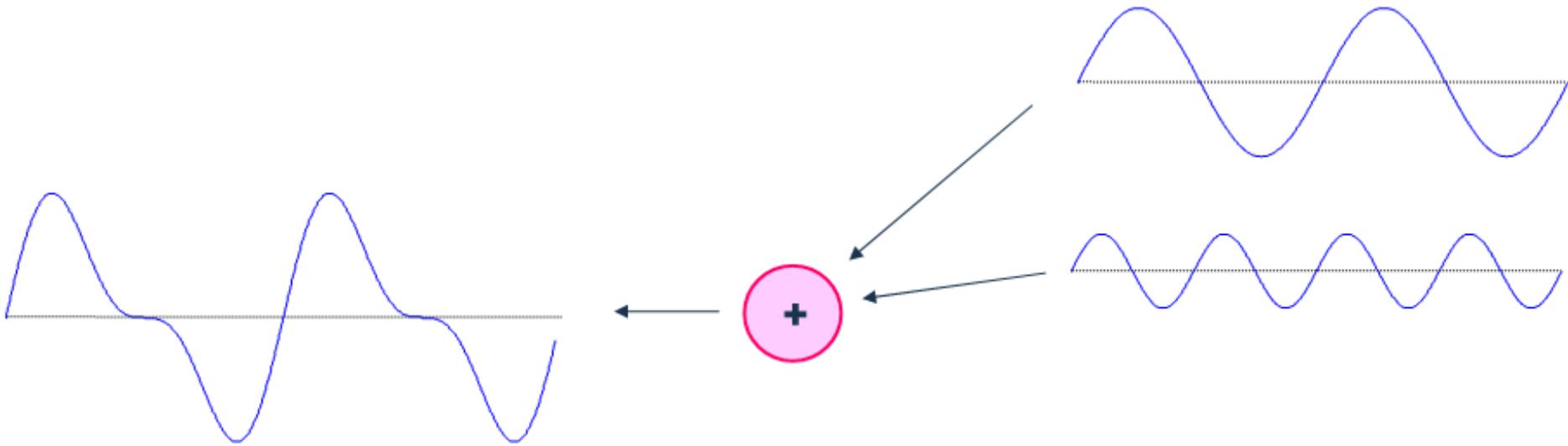
# Analisi di Fourier

- Combinazione delle armoniche



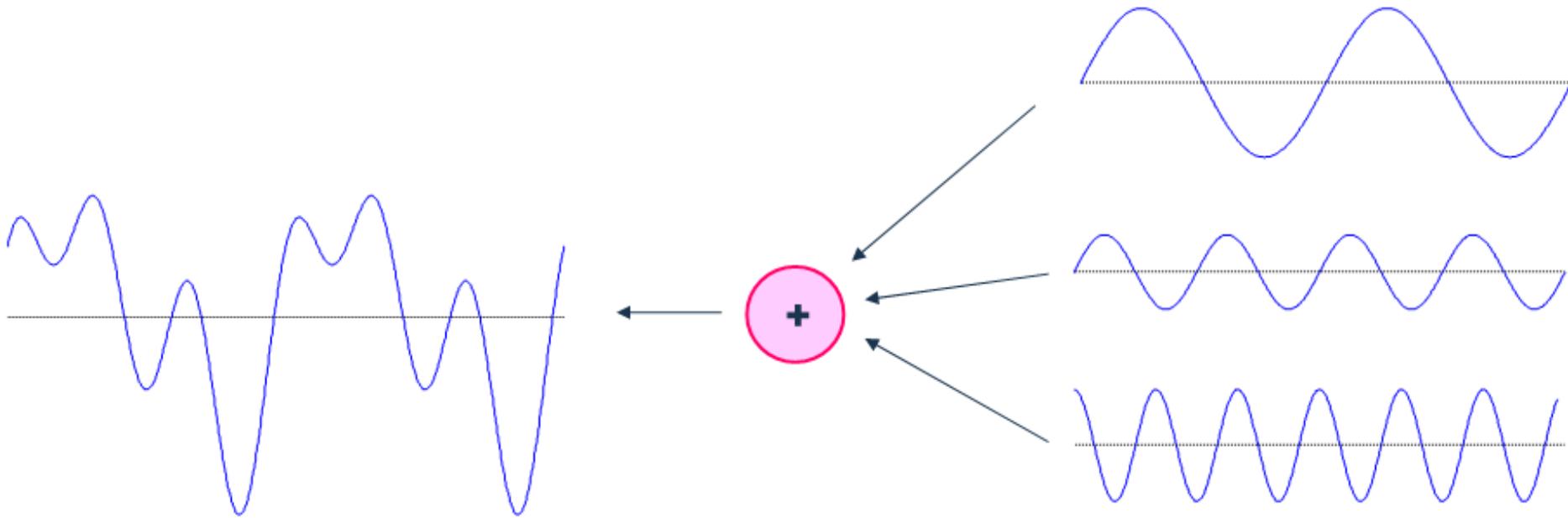
# Analisi di Fourier

- Combinazione delle armoniche



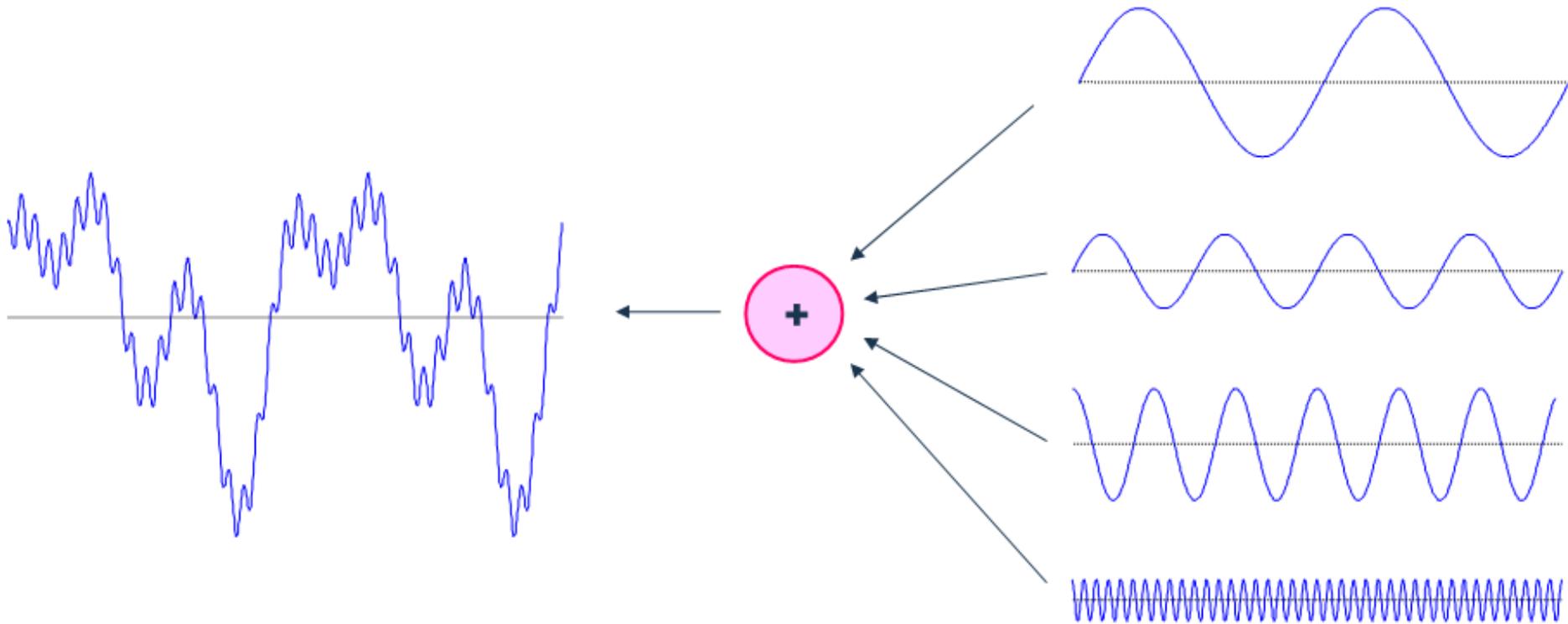
# Analisi di Fourier

- Combinazione delle armoniche



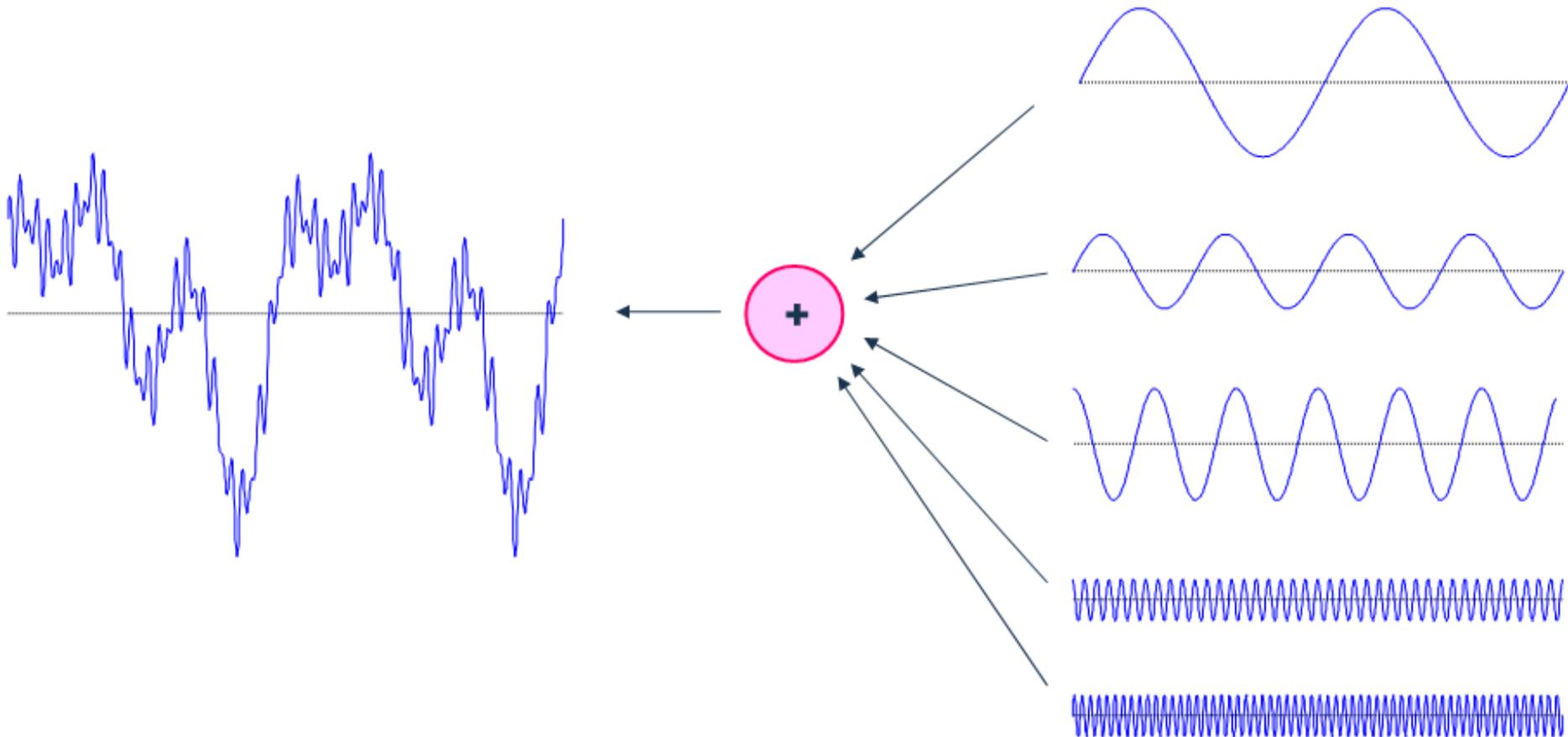
# Analisi di Fourier

- Combinazione delle armoniche



# Analisi di Fourier

- Combinazione delle armoniche



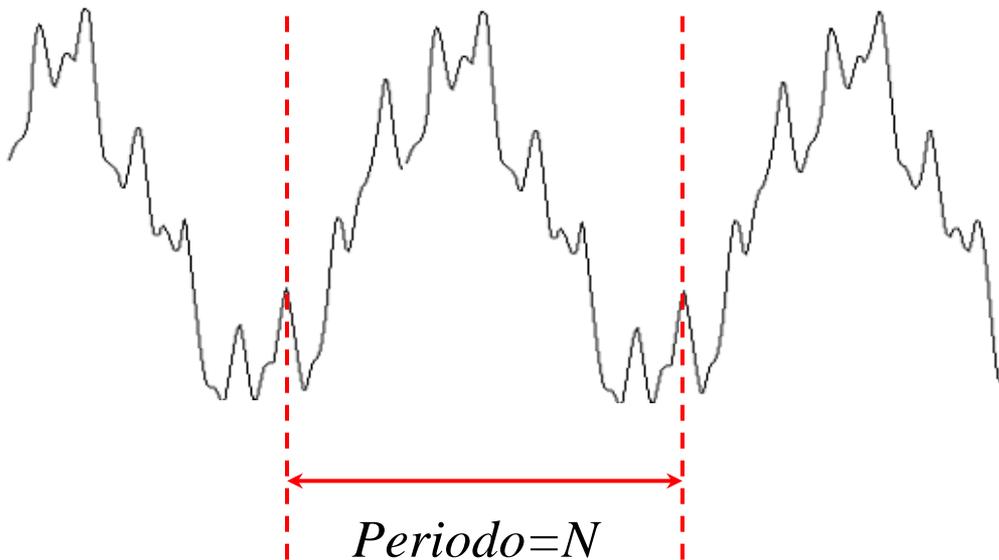
# SERIE DI FOURIER

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi}{N} kx\right) + b_k \sin\left(\frac{2\pi}{N} kx\right)$$

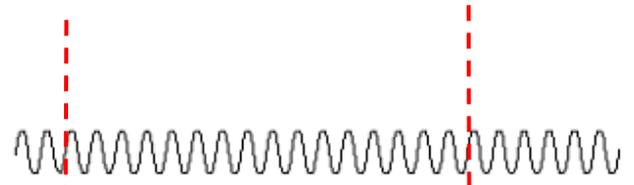
$N$  → periodo

$(1/N)$  → frequenza fondamentale  $f_0$

$(1/N) k$  → frequenze multiple  $kf_0$



$f = kf_0$



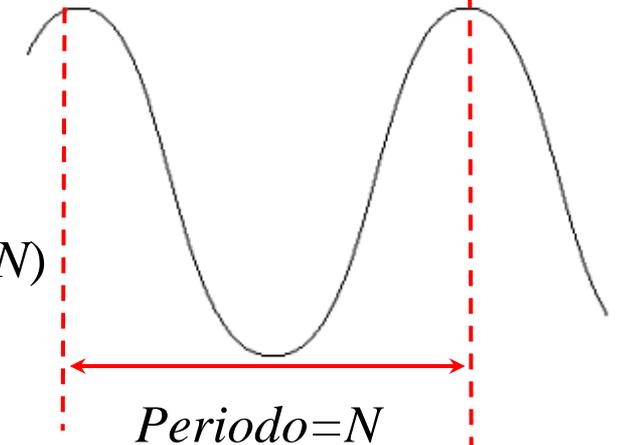
$f = 8f_0$



$f = 4f_0$

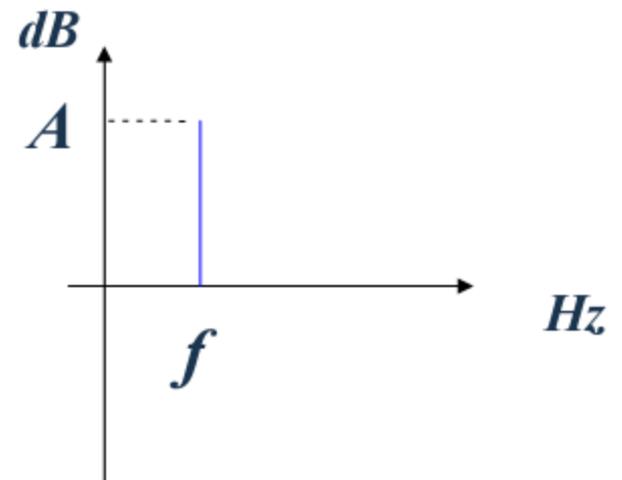
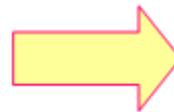
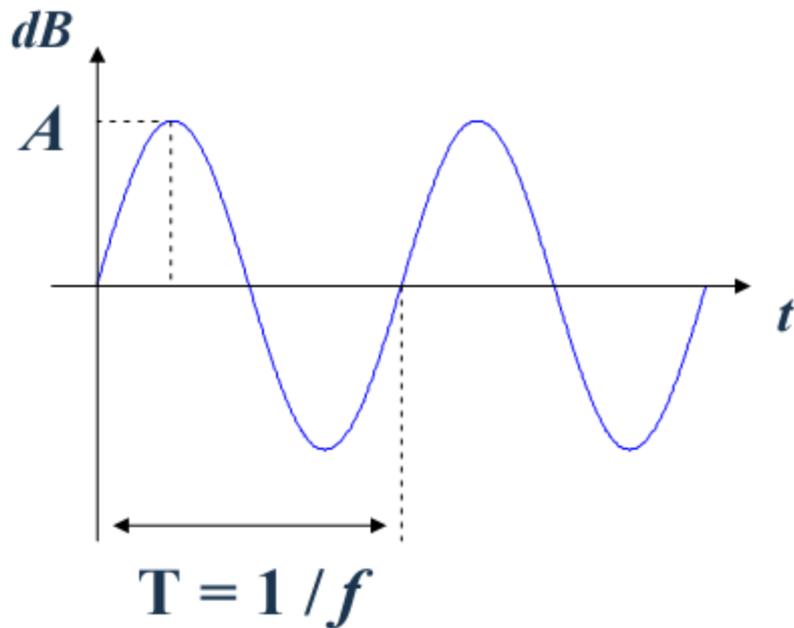


$f_0 = (1/N)$



# Serie di Fourier

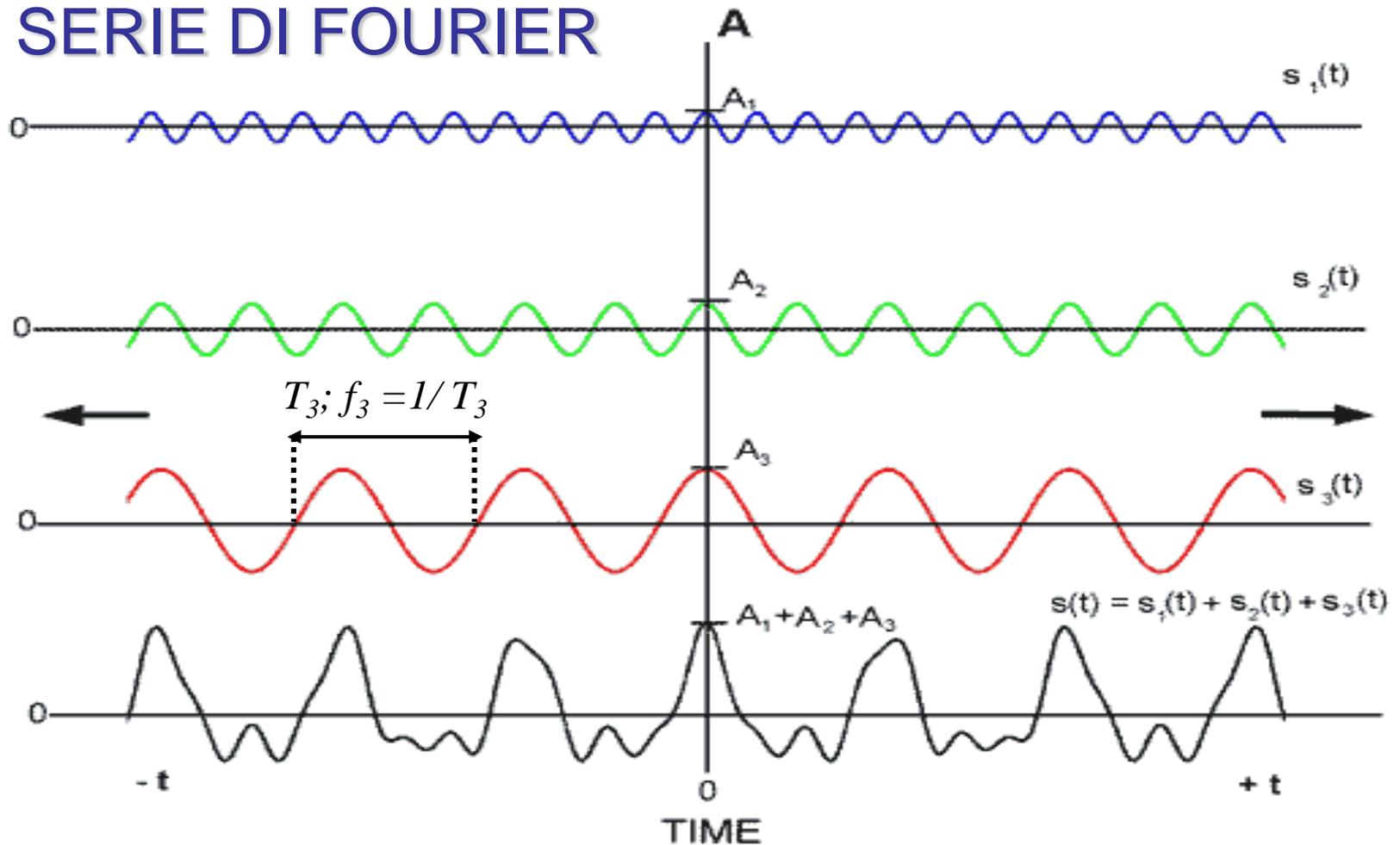
- Armoniche e dominio della frequenza



**Dominio del tempo -  $t$**

**Dominio della frequenza -  $Hz$**

# SERIE DI FOURIER



$$s(t) = s_1(t) + s_2(t) + s_3(t) =$$

$$= \underbrace{A_1}_{\text{Coefficienti}} \cos(2\pi f_1 t) + \underbrace{A_2}_{\text{Coefficienti}} \cos(2\pi f_2 t) + \underbrace{A_3}_{\text{Coefficienti}} \cos(2\pi f_3 t)$$

# TRASFORMATA DI FOURIER

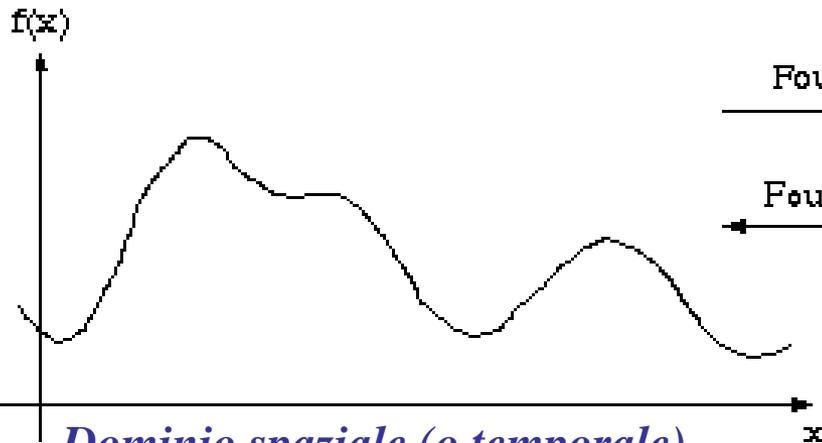
*Antitrasformata (SINTESI)*

$$f(x) = F^{-1}[F(u)]$$



*Trasformata (ANALISI)*

$$F(u) = F[f(x)]$$



*Dominio spaziale (o temporale)  
o dominio diretto*

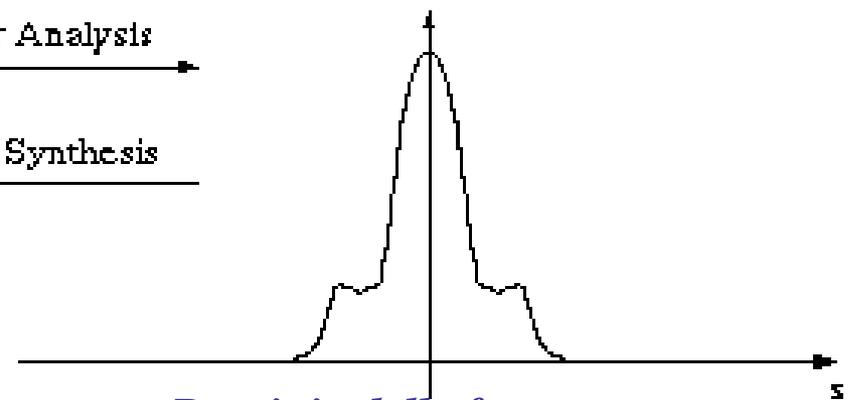
Fourier Analysis



Fourier Synthesis



$(F(f))(s)$

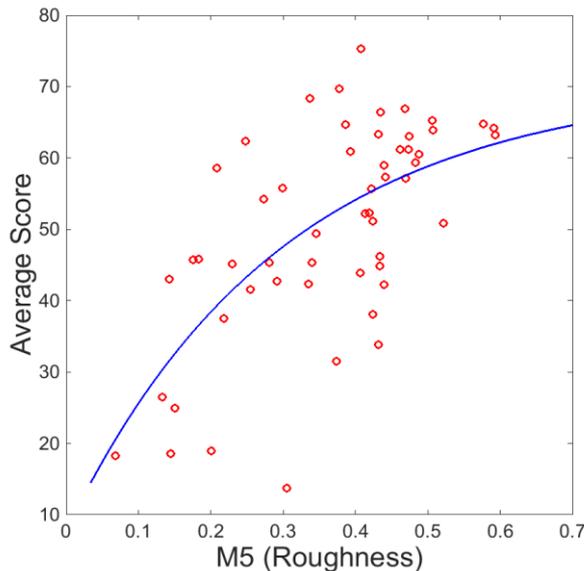
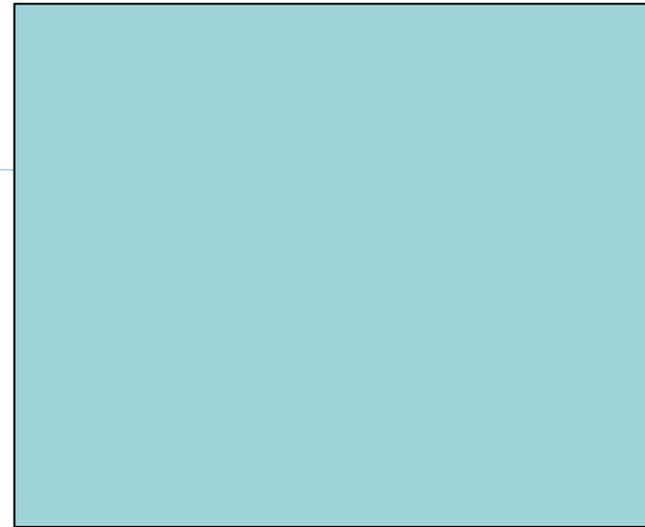


*Dominio delle frequenze  
o dominio trasformato*

# APPROFONDIAMO

Segnale:  
Immagine  
Audio  
video

Misure oggettive



Le misure oggettive o soggettive vengono messe in correlazione con la ground truth: la verità. In questo modo si può capire quale misura descrive meglio l'emozione indotta nel soggetto

## Soggetto

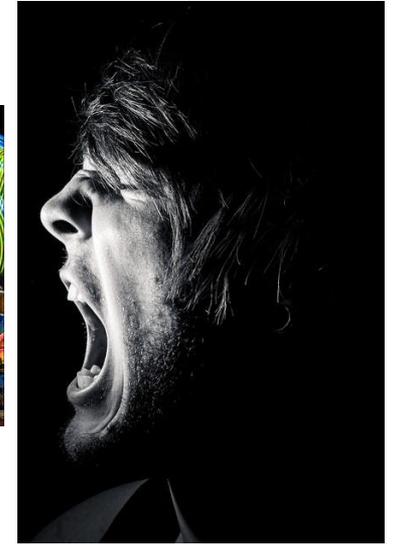
Dati psicovisuali

Questionari e test



# MACHINE LEARNING

20500  
immagini da  
flicker con  
testo  
associato

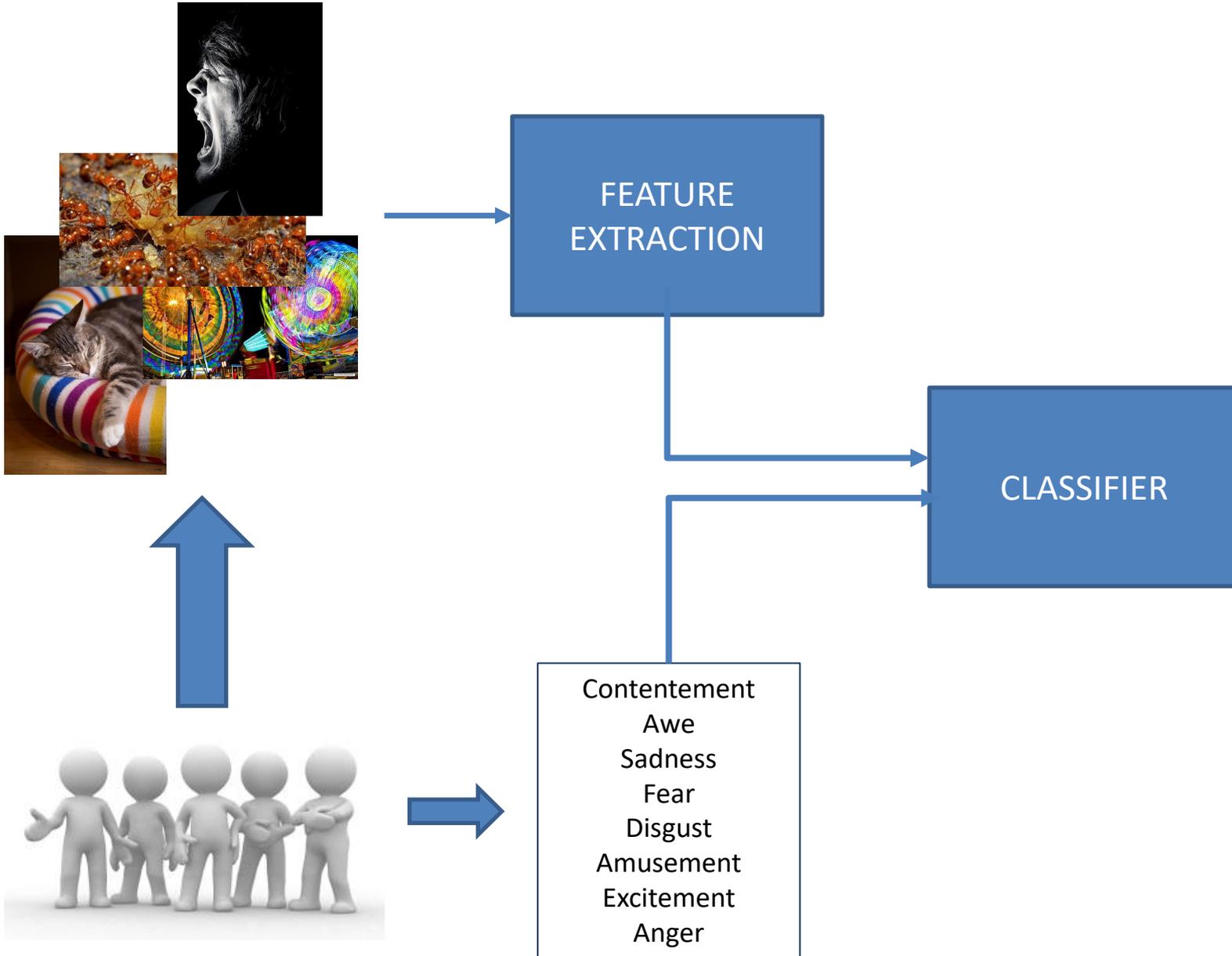


Classificate rispetto a 8 diverse emozioni:

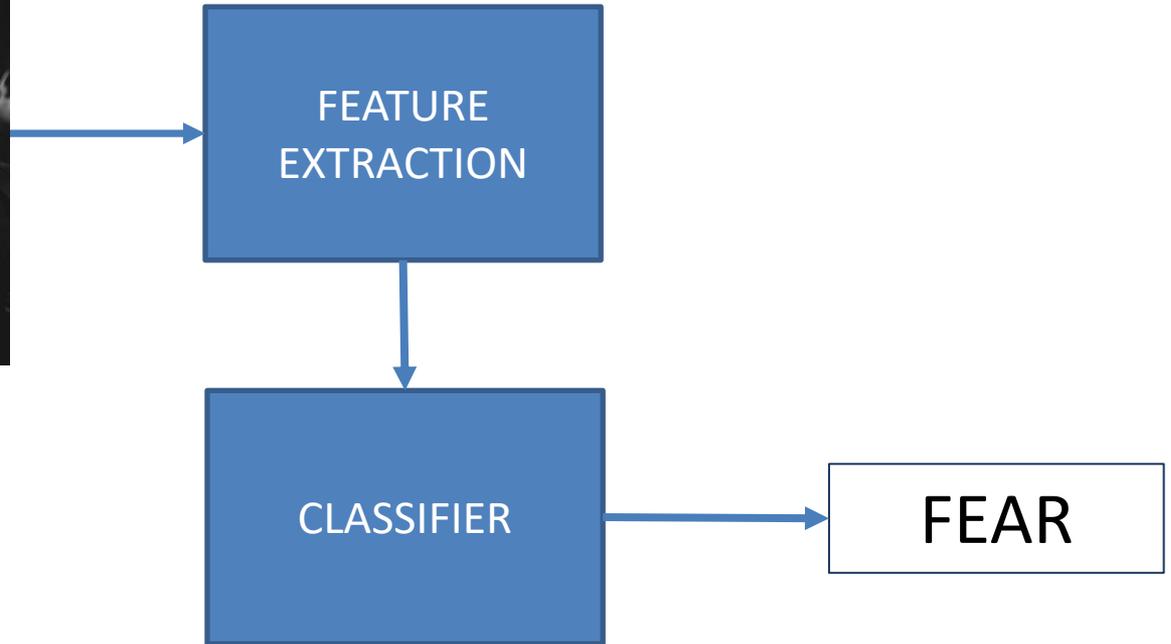
Contentment  
Awe  
Sadness  
Fear  
Disgust  
Amusement  
Excitement  
Anger

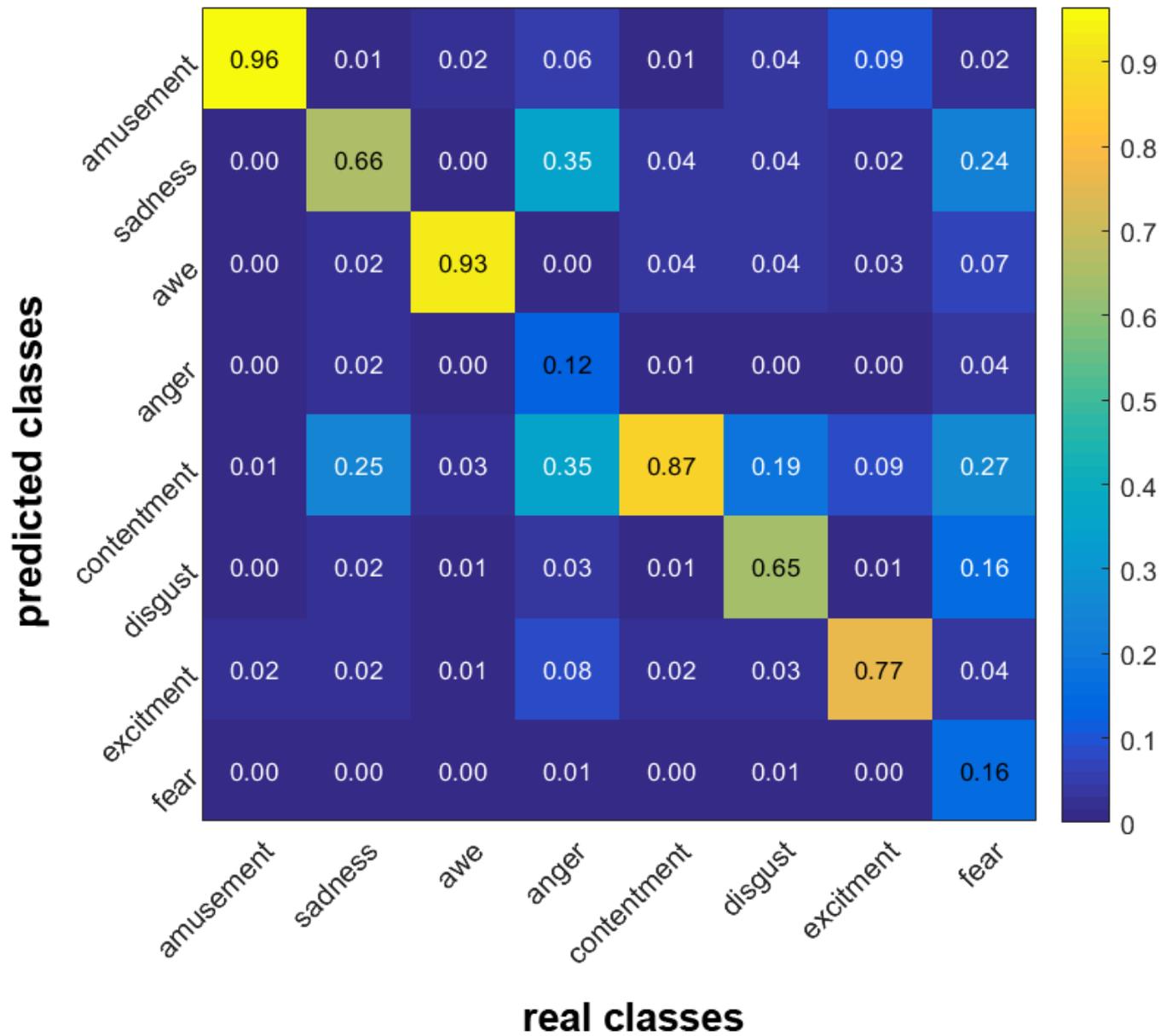


# TRAINING DEL CLASSIFICATORE

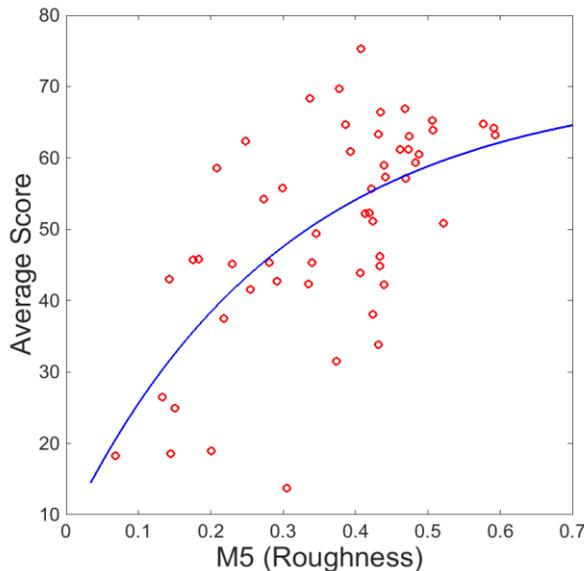
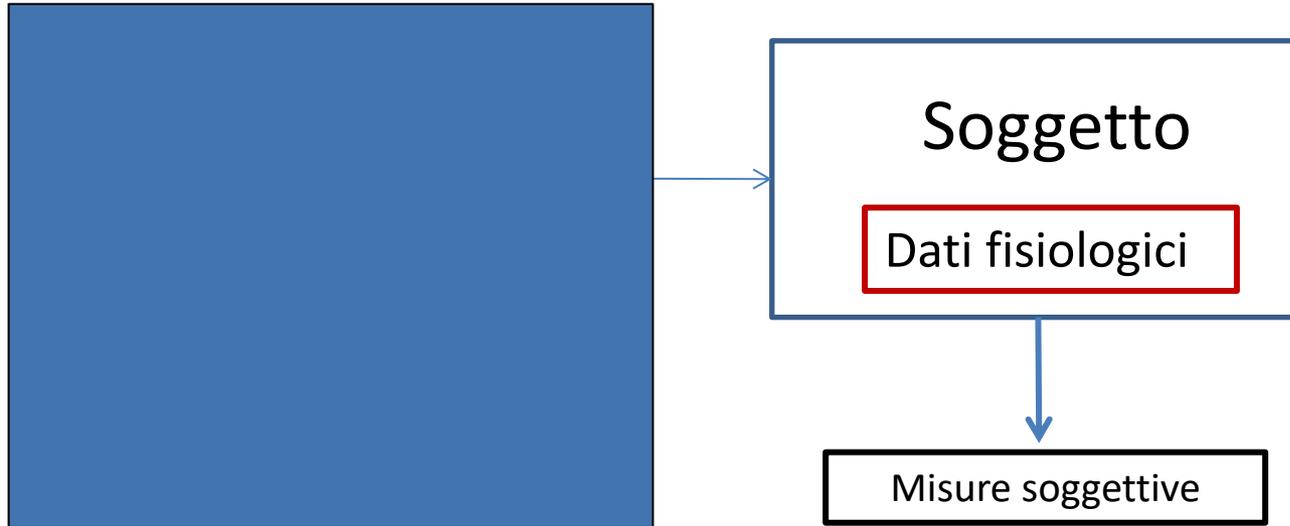


# CLASSIFICATORE

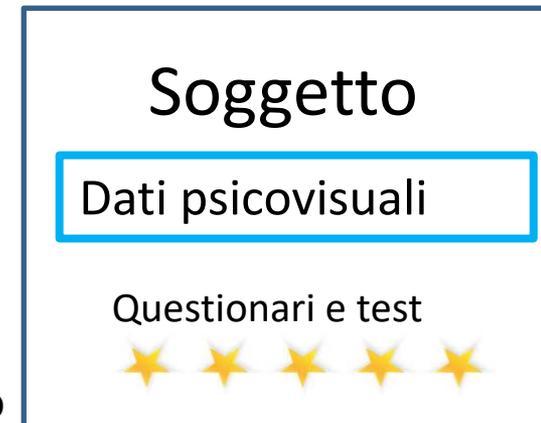




# APPROFONDIAMO

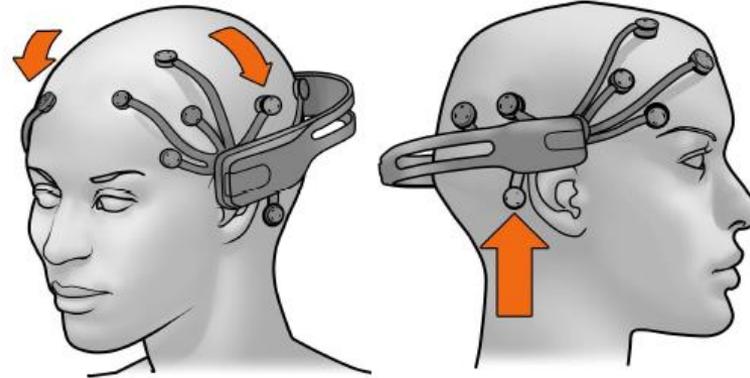
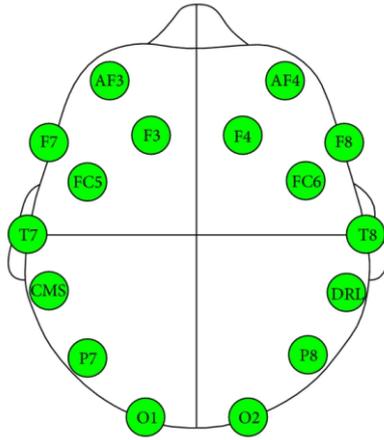


Le misure oggettive o soggettive vengono messe in correlazione con la ground truth: la verità.  
In questo modo si può capire quale misura descrive meglio l'emozione indotta nel soggetto



# DATI FISIOLGICI

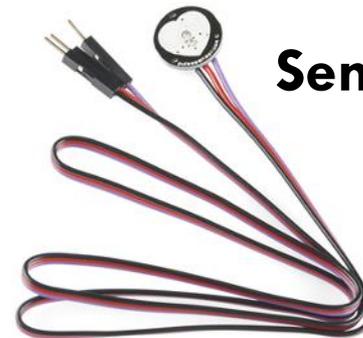
- EEG



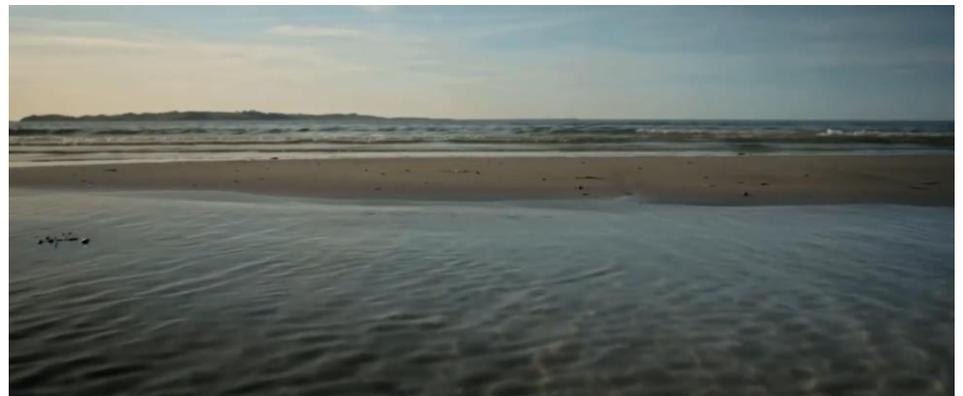
## Sensore GSR



## Sensore HR



# CLASSIFICARE VIDEO STRESSANTI E RILASSANTI

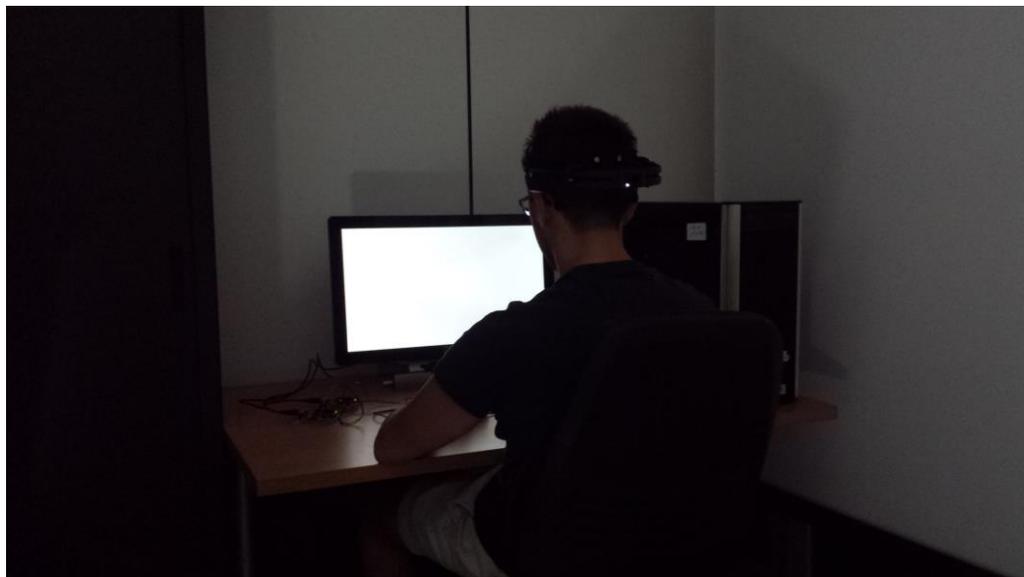


Esempi di scene stressanti.

Esempi di scene rilassanti.

# SETUP SPERIMENTALE

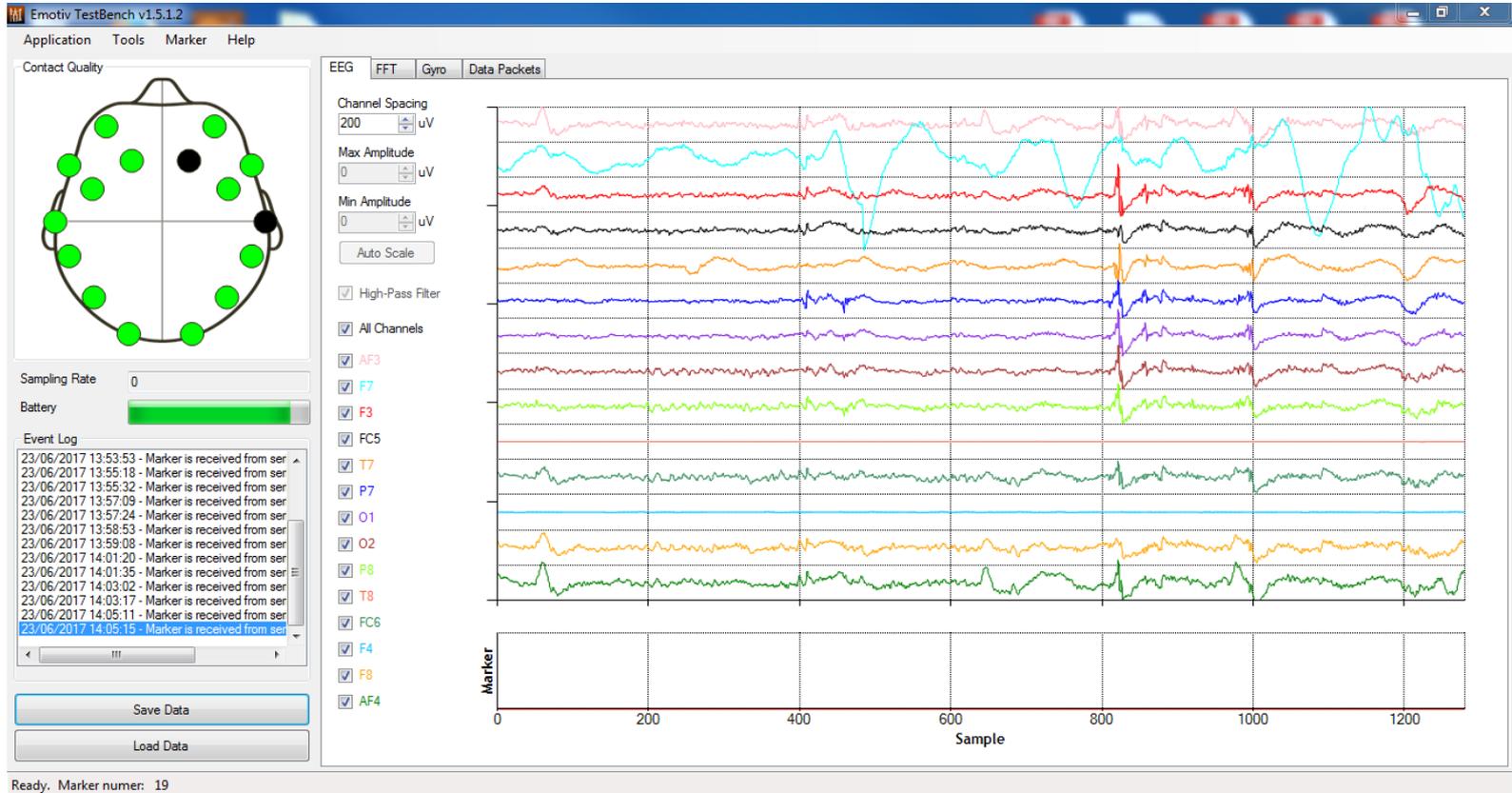




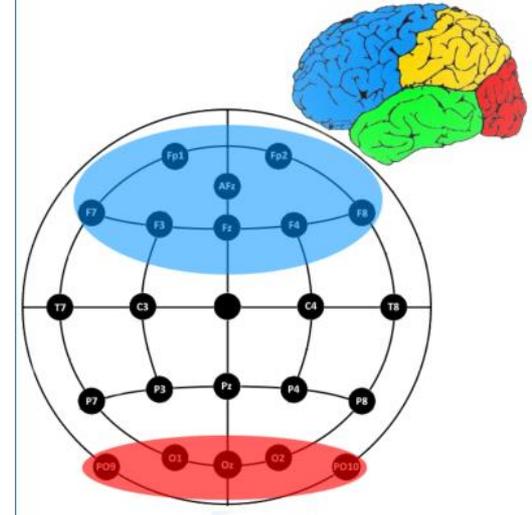
- 25 partecipanti coinvolti.
- 16 video in totale (8 rilassanti + 8 stressanti).
- Procedura:
  - 2 minuti di registrazione di baseline.
  - Presentazione di 8 stimoli (4 rilassanti + 4 stressanti in ordine casuale).
  - Valutazione soggettiva.



# SEGNALE EEG



# RITMI EEG



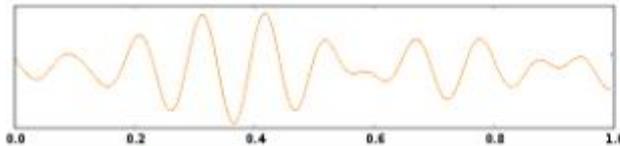
**Ritmo**

**Forma d'onda**

**Frequenze**

**Presenza**

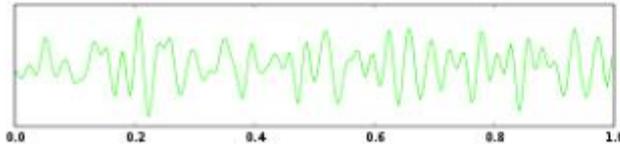
**Alfa**



**8 - 13 Hz**

**Veglia rilassata**

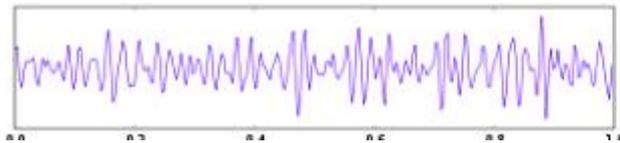
**Beta**



**13 - 30 Hz**

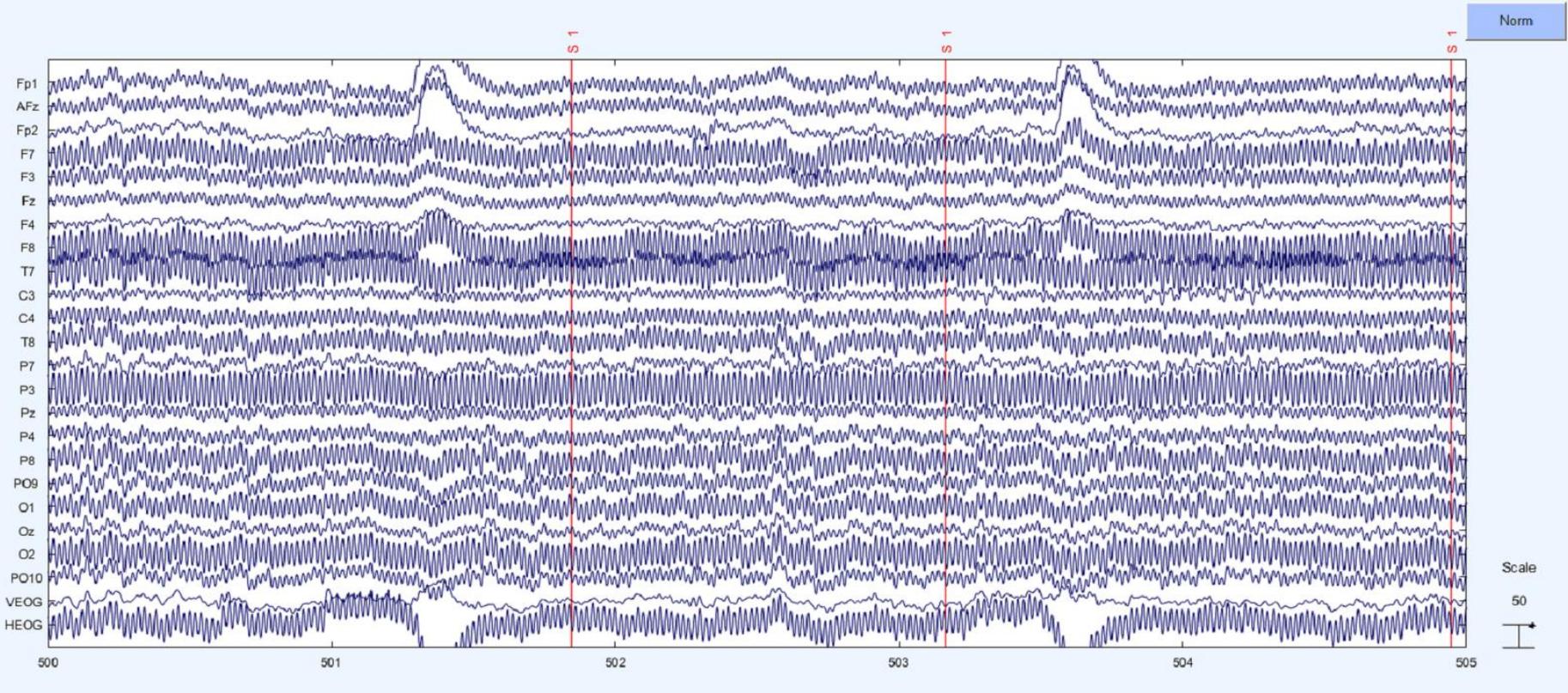
**Veglia attenta**

**Gamma**



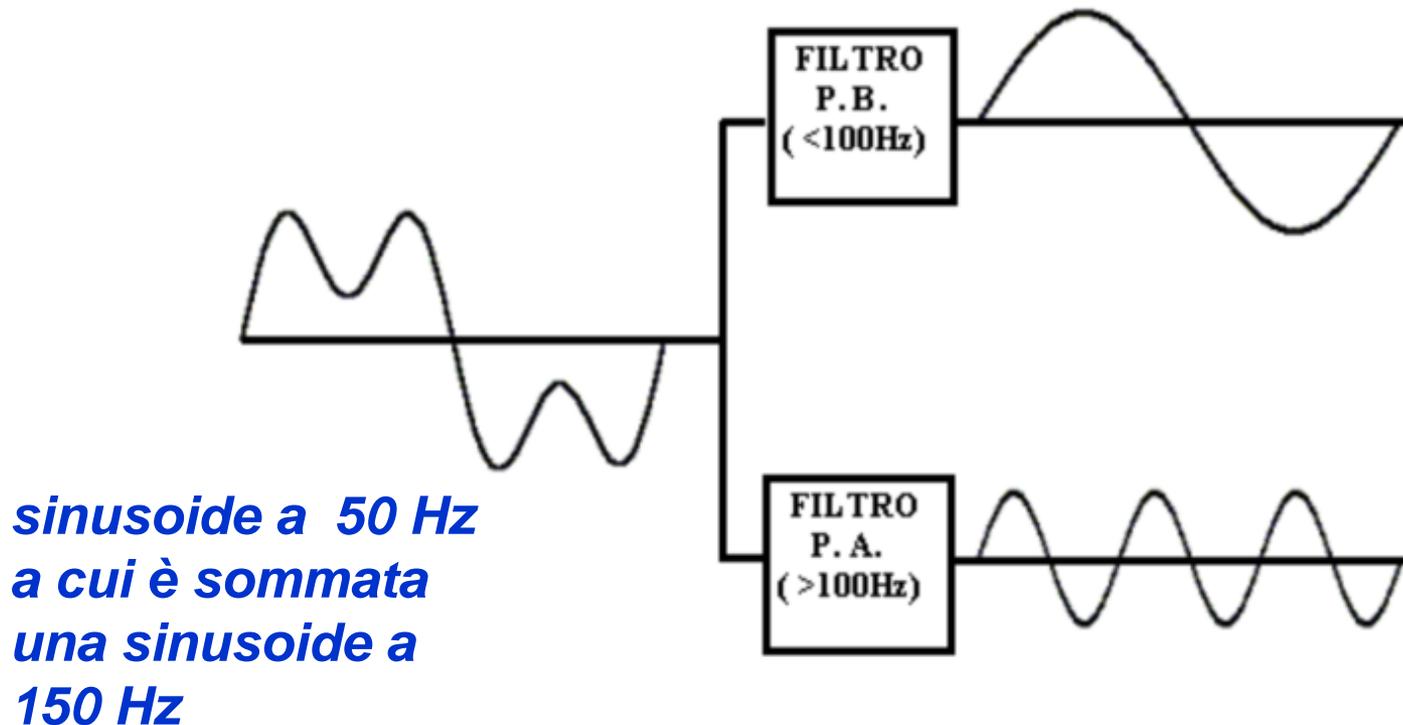
**30 - 100 Hz Problem Solving**

# SEGNALE GREZZO



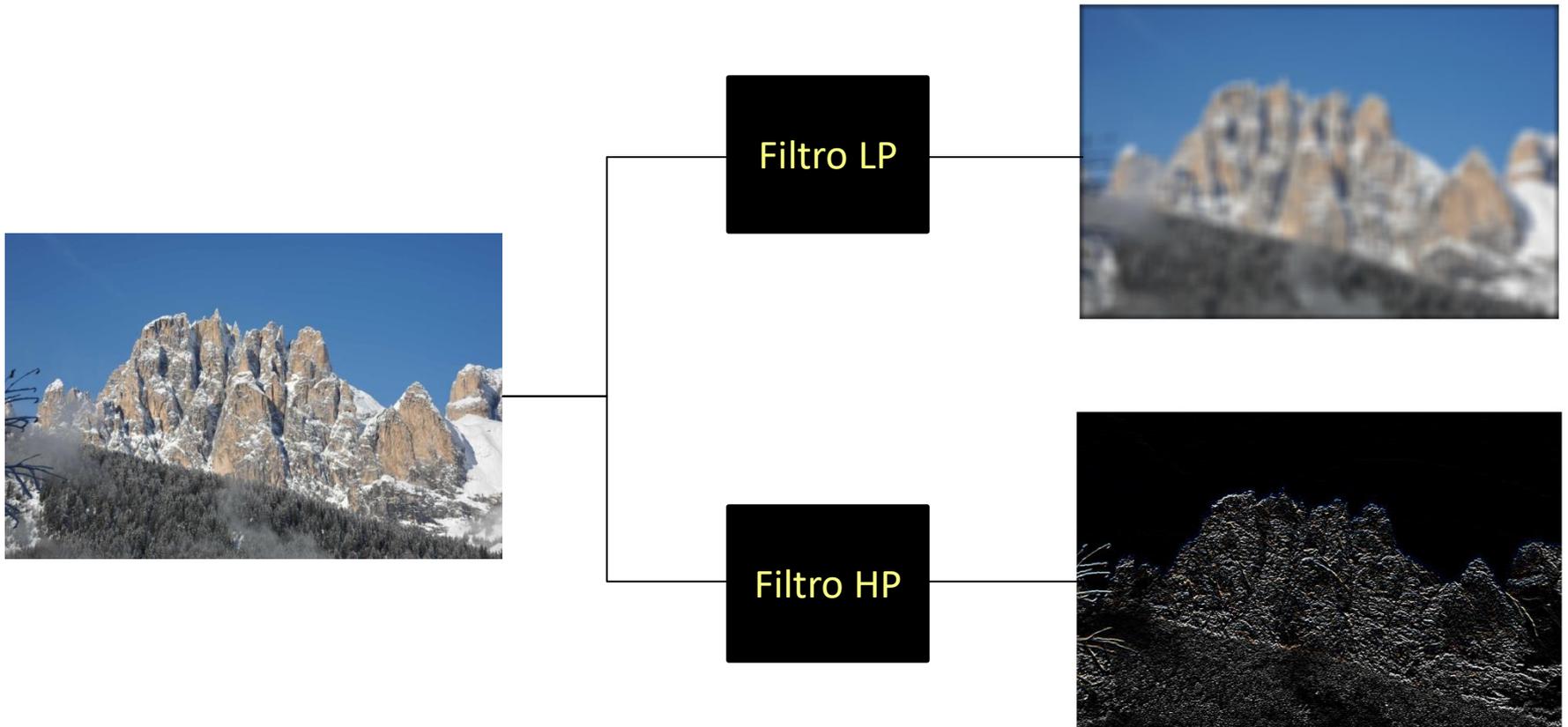
# Filtraggio del segnale

Un filtro passa-basso (P.B) lascia passare solo frequenze minori di un certo valore, un passa-alto (P.A) frequenze superiori ad un certo valore.

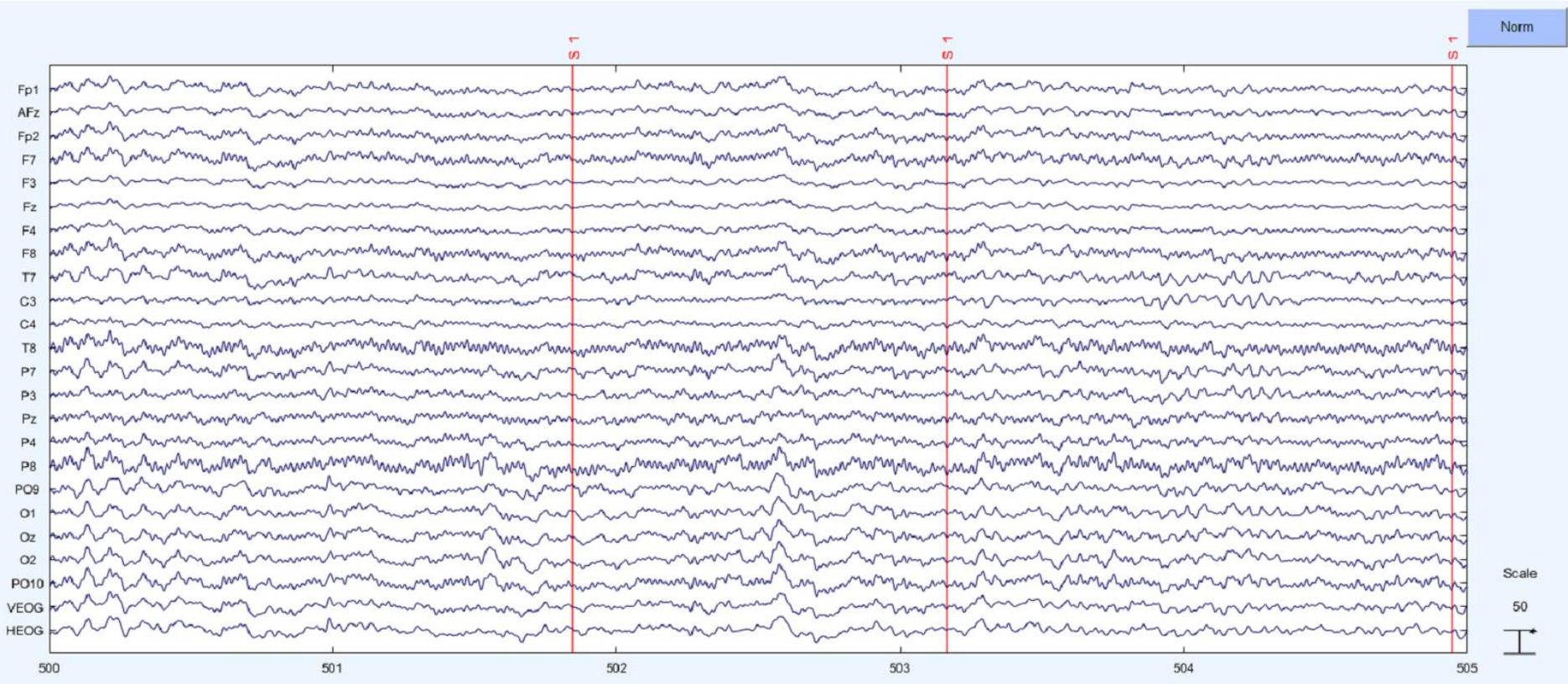


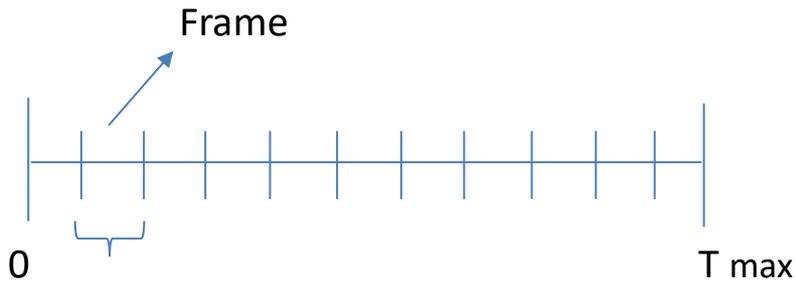
# Filtraggio del segnale

Un filtro passa-basso (LP) lascia passare solo frequenze minori di un certo valore, un passa-alto (HP) frequenze superiori ad un certo valore.



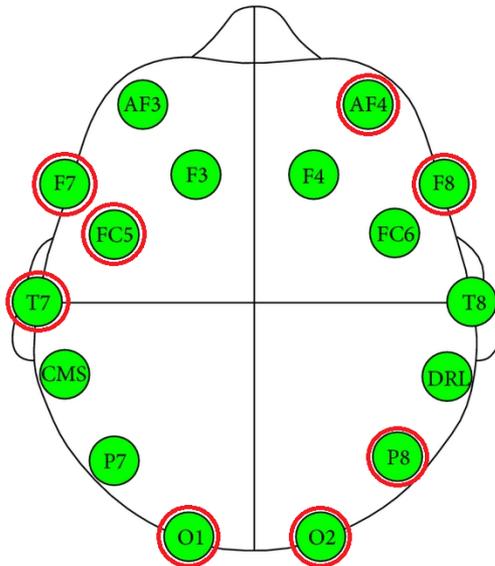
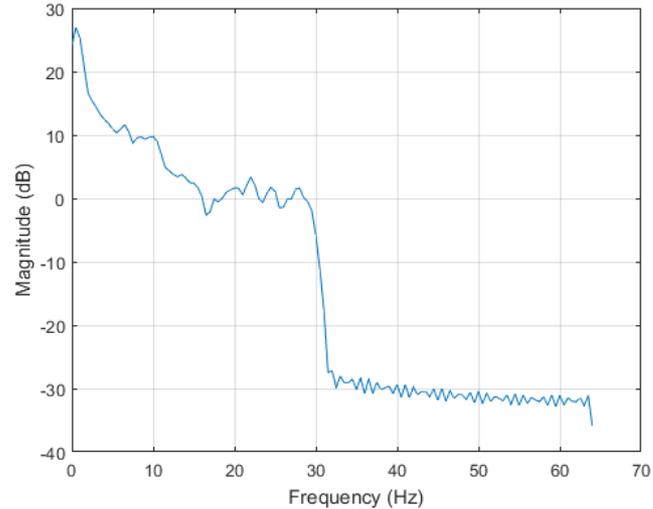
# SEGNALE PULITO





Per ogni soggetto, per ogni sensore, per ogni stimolo, la registrazione EEG viene divisa in intervalli di 5 secondi

Su ogni frame viene stimata la Power spectral density



Si confrontano la distribuzione di potenze spettrali nelle bande alpha beta e gamma, nelle condizioni di baseline, stress e reax e si cercano le aree dove vi è stata una variazione di Potenza statisticamente significativa

# Segnali gsr: stress VS relax

