

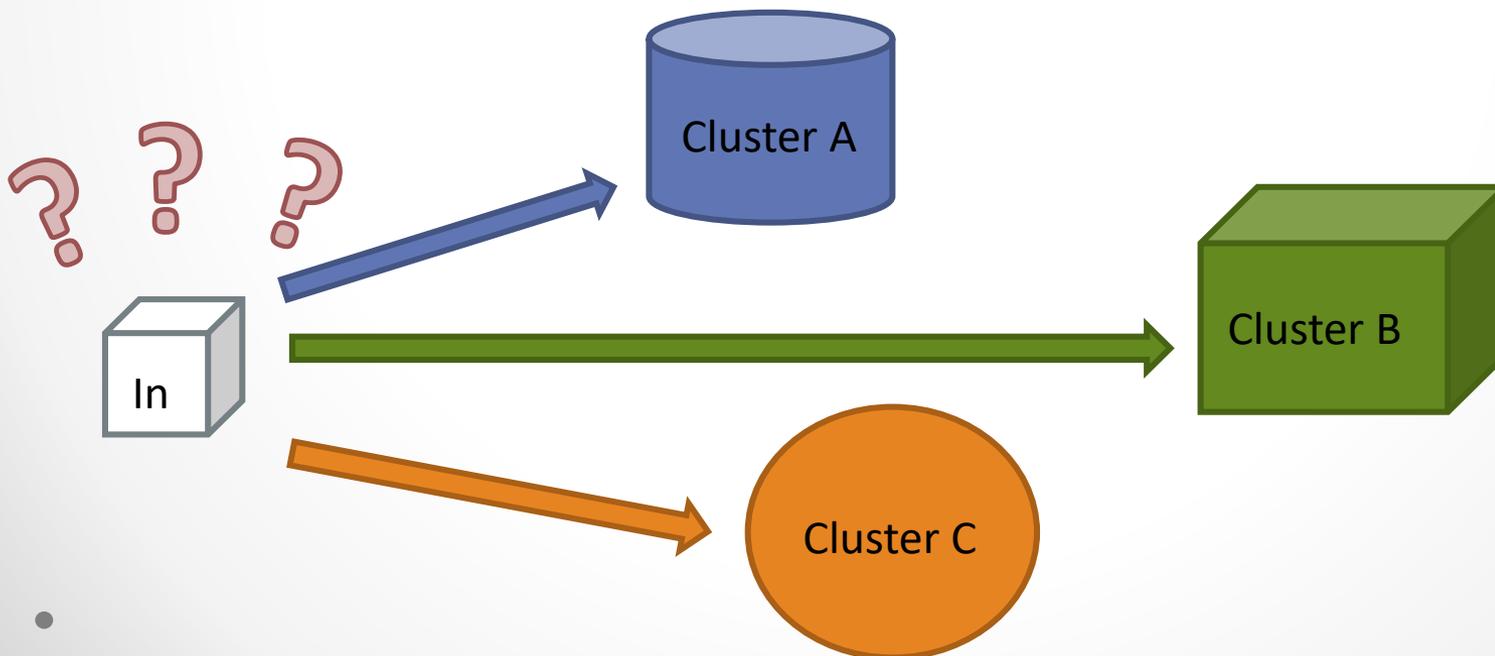
Francesco Gargano **PART 1: Theoretical Background**
Marco Cinà **PART 2: Concrete Applications**

Palermo 13 Dicembre 2016



Clustering

La risoluzione di un problema di clustering consiste nel decretare se un dato input I , in ingresso ad un sistema, appartiene o meno ad un determinato insieme costituente il possibile output O ad esso associato.

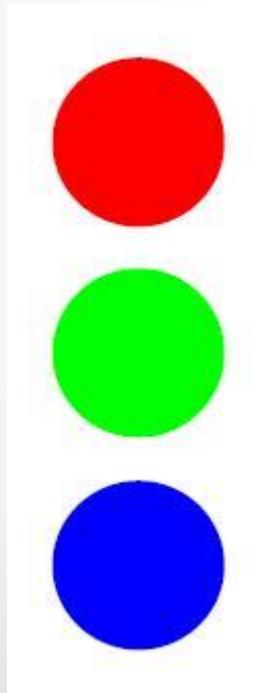
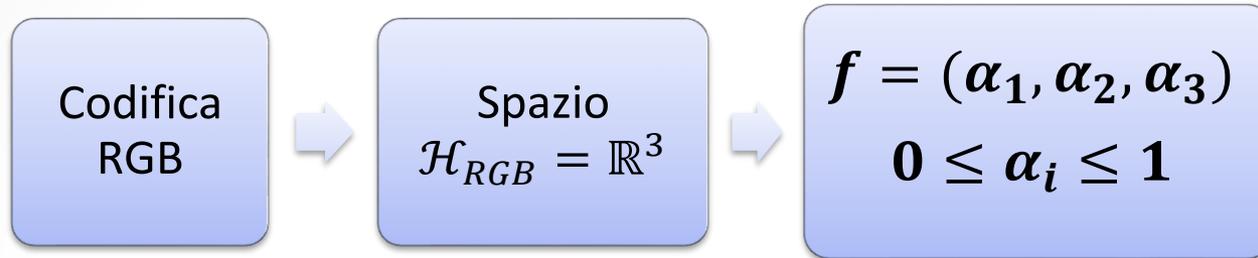




RICONOSCIMENTO DI COLORI IN SCALA RGB

Riconoscimento di colori

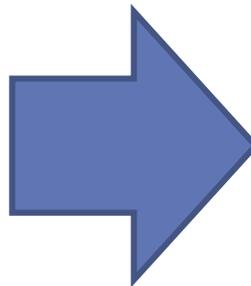
Operatori di proiezione



$$e_R = (1, 0, 0)$$

$$e_G = (0, 1, 0)$$

$$e_B = (0, 0, 1)$$



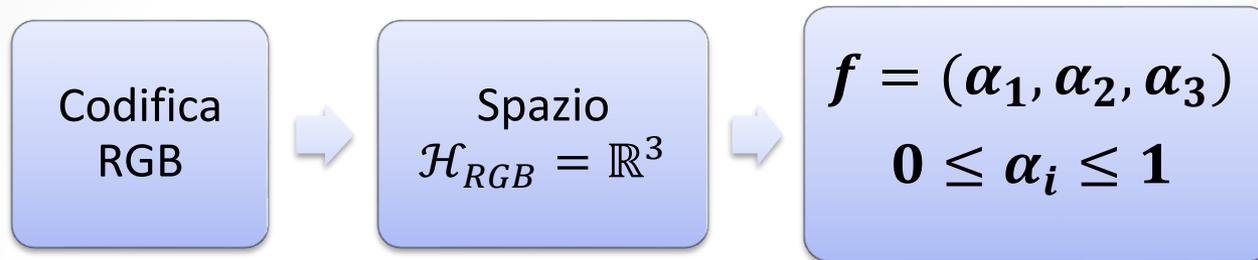
$$\circ P_R(f) = Q_R(f) = \langle e_R, f \rangle e_R$$

$$\circ P_G(f) = Q_G(f) = \langle e_G, f \rangle e_G$$

$$\circ P_B(f) = Q_B(f) = \langle e_B, f \rangle e_B$$

Riconoscimento di colori

Operatori di proiezione



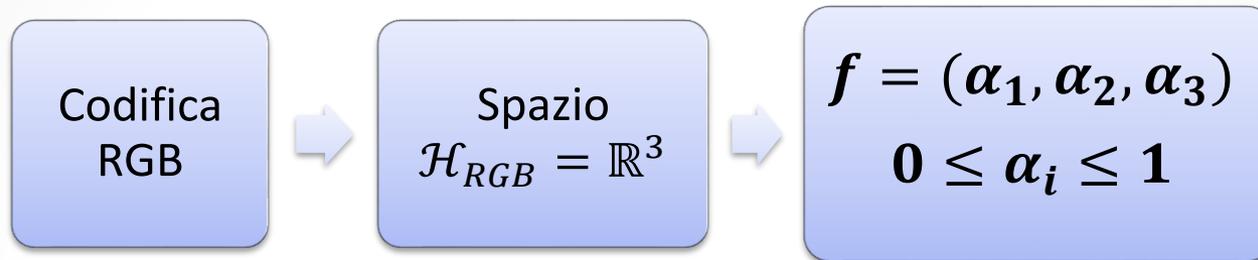
- $q_R(f) = \|Q_R(f)\|^2 = |\langle \mathbf{e}_R, \mathbf{f} \rangle|^2$

- $q_G(f) = \|Q_G(f)\|^2 = |\langle \mathbf{e}_G, \mathbf{f} \rangle|^2$

- $q_B(f) = \|Q_B(f)\|^2 = |\langle \mathbf{e}_B, \mathbf{f} \rangle|^2$

Riconoscimento di colori

Operatori di proiezione



$$q_R(f) \gg \max\{q_G(f), q_B(f)\}$$

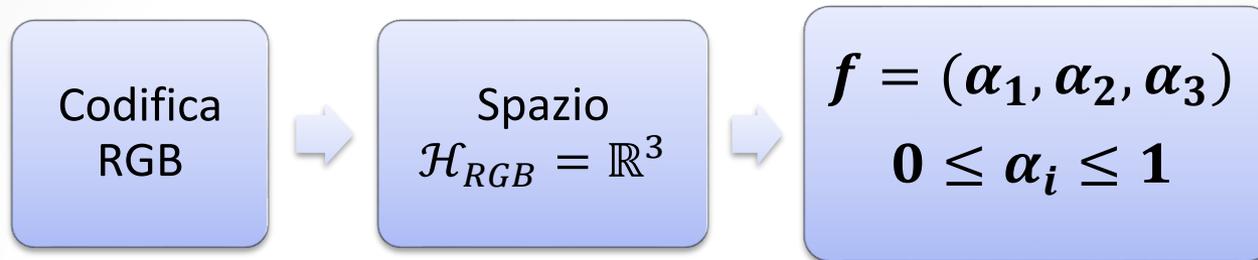


***f** è più vicino al Reference Point
Red, piuttosto che **Green** o **Blue***

*Selezionato un opportuno valore ε (abbastanza piccolo),
definiamo il cluster centrato su R di raggio ε , $K_\varepsilon(R)$*

Riconoscimento di colori

Operatori di proiezione



$$f \in K_\varepsilon(R) \Leftrightarrow q_R(f) \in [\|f\|^2 - \varepsilon, \|f\|^2]$$

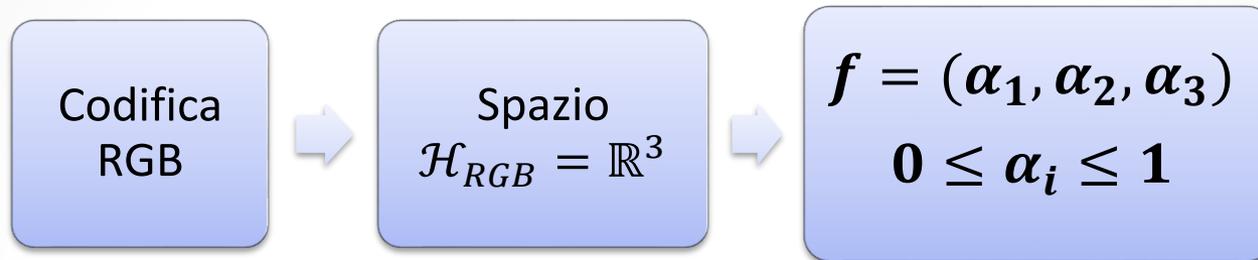
$$\hookrightarrow \|f\|^2 - \varepsilon \leq q_R(f) \leq \|f\|^2$$

$$\left\{ \begin{array}{l} q_R(f) \leq \|f\|^2 \quad \longleftarrow \text{(disuguaglianza di Schwarz)} \\ \|f\|^2 - \varepsilon \leq q_R(f) \quad \longrightarrow q_G(f) + q_B(f) \leq \varepsilon \end{array} \right.$$

$$q_G(f), q_B(f) \notin [\|f\|^2 - \varepsilon, \|f\|^2]$$

Riconoscimento di colori

Operatori di proiezione



ATTENZIONE !

$$q_j(f) = 1 \quad \not\Rightarrow \quad q_j(f) \in K_\varepsilon(j)$$

$$f_{white} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \longrightarrow \quad q_j(f_{white}) = 1, \quad j = 1,2,3;$$

$$q_j(f_{white}) \geq \|f\|^2 - \varepsilon = 3 - \varepsilon, \quad j = 1,2,3; \quad \textbf{FALSE !}$$

Riconoscimento di colori

Operatori di proiezione

Analizziamo gli operatori $q_j(f)$ in altra chiave :

$$\|f\|^2 = q_R(f) + q_G(f) + q_B(f).$$

Ne consegue che , se $f \in K_\varepsilon(R)$, allora

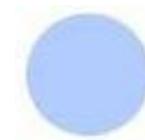
$$\|f - R\|^2 = q_R(f - R) + q_G(f - R) + q_B(f - R) \cong 0$$



$$f_1 = (0.8, 0.1, 1)$$



$$f_2 = (0.3, 0.6, 0.1)$$



$$f_3 = (0.7, 0.8, 1)$$



$$P_p = (0.6, 0, 0.6) \Rightarrow \|P_p - f_j\|^2 \cong \begin{cases} 0.21 & j = 1 \\ 0.49 & j = 2 \\ 0.81 & j = 3 \end{cases}$$



$$P_g = (0, 0.8, 0.2) \Rightarrow \|P_g - f_j\|^2 \cong \begin{cases} 1.76 & j = 1 \\ 0.13 & j = 2 \\ 1.12 & j = 3 \end{cases}$$

Riconoscimento di colori

Operatori di proiezione

$$F[f_1, f_2] := \frac{\|f_1\| \|f_2\| - |\langle f_1, f_2 \rangle|}{\|f_1\| \|f_2\|}$$





$f_1 = (0.8, 0.1, 1)$
 $f_2 = (0.3, 0.6, 0.1)$
 $f_3 = (0.7, 0.8, 1)$


 $P_p = (0.6, 0, 0.6)$

 $F[P_p, f_j] \cong \begin{cases} 0.01 & j = 1 \\ 0.34 & j = 2 \\ 0.22 & j = 3 \end{cases}$





 $P_g = (0, 0.8, 0.2)$

 $F[P_g, f_j] \cong \begin{cases} 0.78 & j = 1 \\ 0.06 & j = 2 \\ 0.36 & j = 3 \end{cases}$






RICONOSCIMENTO DI NOTE MUSICALI

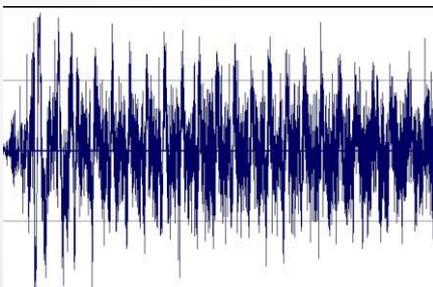
Acquisizione degli input

Acquisizione dei campioni audio in formato .wav

Transizione dal dominio del tempo al dominio delle frequenze

Spettro delle frequenze delle note campionate

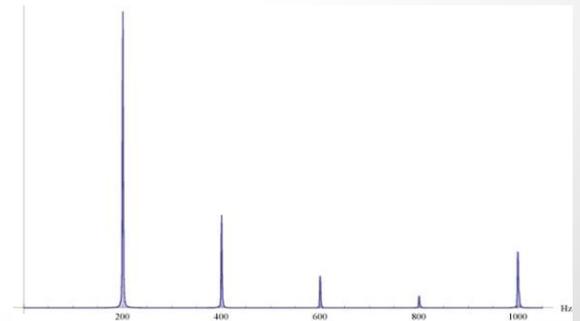
Rappresentazione del segnale f_s nel dominio di t



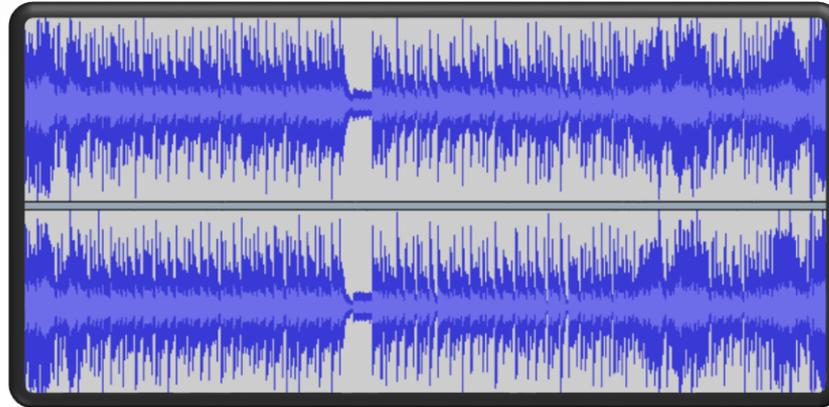
Trasformata di Fourier discreta

$$\widetilde{f}_s = \sum_{n=0}^{N-1} f_{s_n} e^{-i2\pi \frac{n}{N}}$$

Rappresentazione del segnale \widetilde{f}_s nel dominio di f



Acquisizione degli input

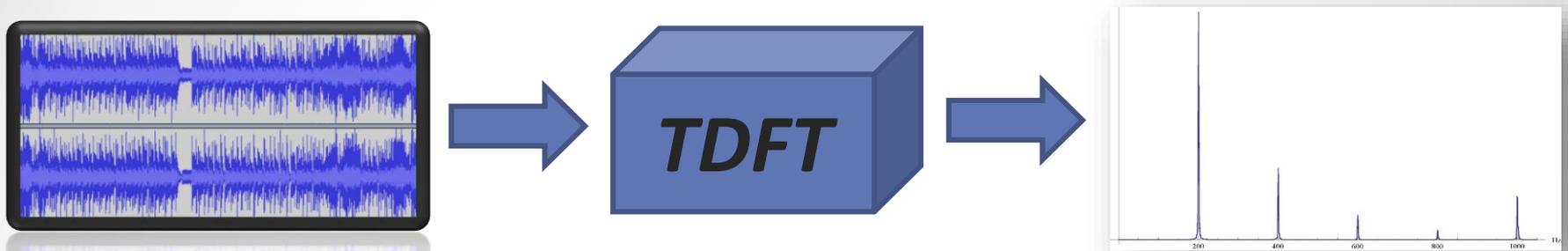


I file .WAV restituiscono una forma d'onda che riproduce l'andamento del suono campionato, di durata t_s , ad una frequenza $f_c = 44$ KHz (samples/sec), di conseguenza il modello matematico dell'input è un vettore f_s di numeri reali la cui dimensione è pari a

$$\mathit{dim}(f_s) = f_c * t_s$$

La durata dei campioni audio è, in tutti i nostri casi, pari ad un secondo, quindi la dimensione di ciascun vettore f_s è 44100.

Acquisizione degli input



Sul segnale f_s così acquisito viene eseguito l'algoritmo di TDFT (Time Discrete Fourier Transform), il segnale risultante è un vettore di numeri complessi di dimensione 44100 i cui coefficienti sono così ottenuti:

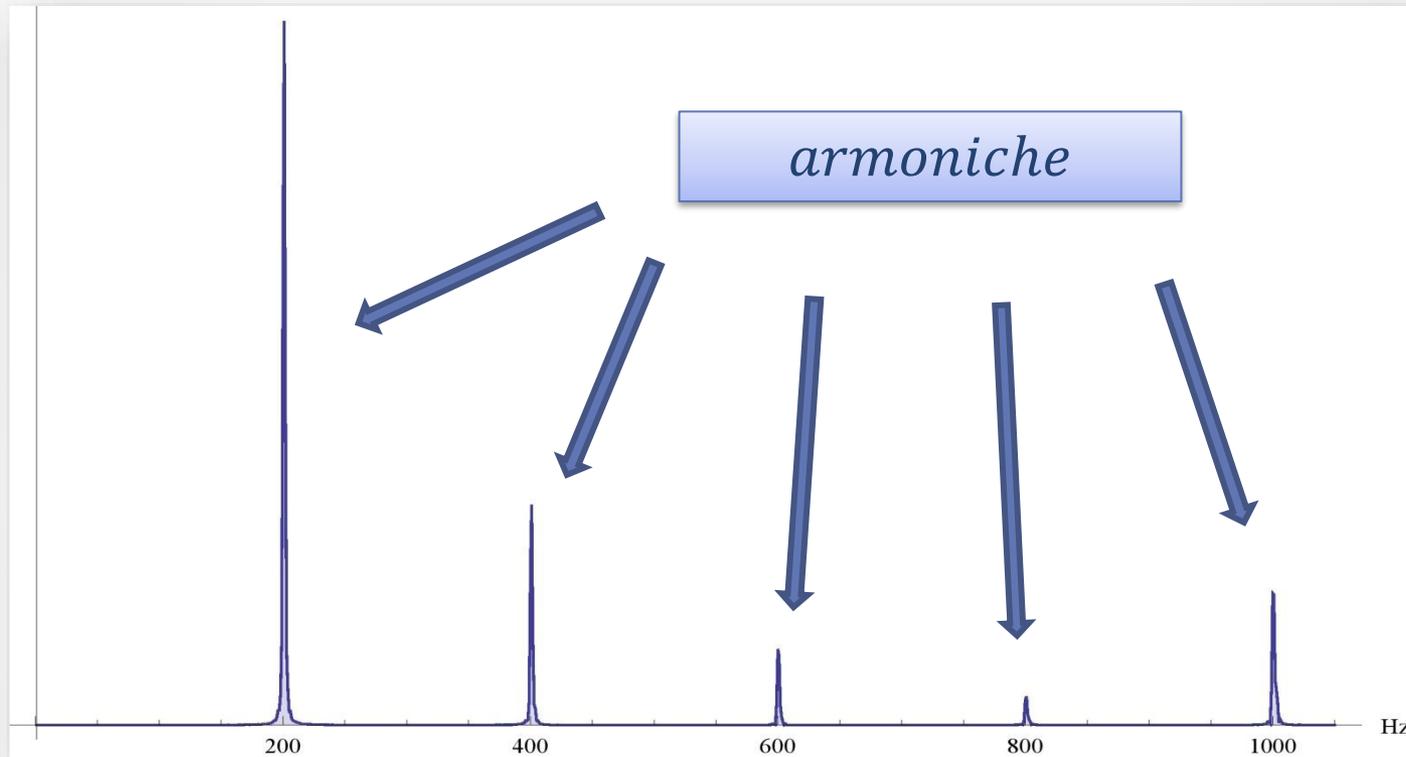
$$f_n = F_D[f_s](n) = \sum_{k=0}^{N-1} f_{s_k} e^{\frac{-ik2\pi n}{N}};$$

$$f = \sum_{n=1}^{22050} |f_n| e_n;$$

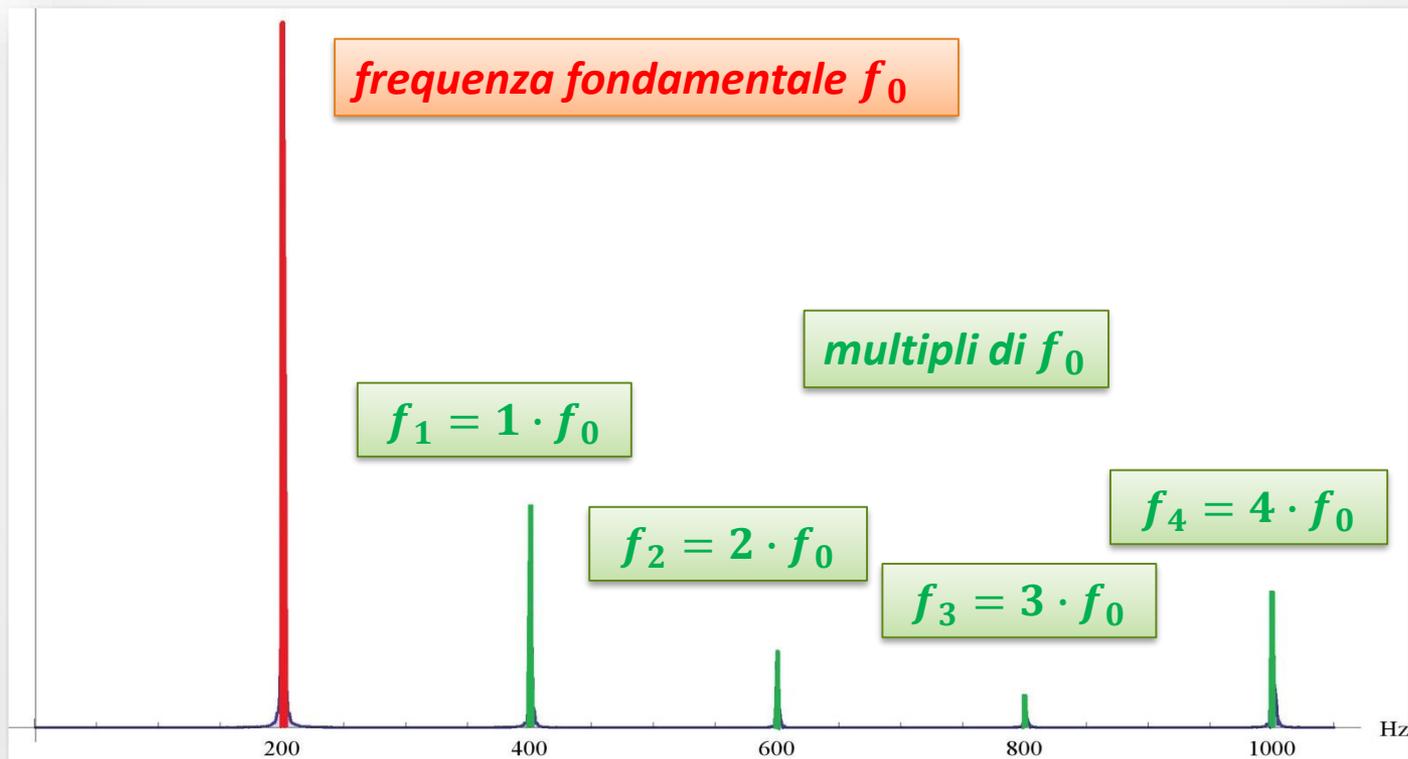
e_n : elementi della base canonica di $\mathcal{H}_{sound} = \mathbb{R}^{22050}$;

$\|f\| = 1$ ($\|\cdot\|$ è la norma in $\mathcal{H}_{sound} = \mathbb{R}^{22050}$);

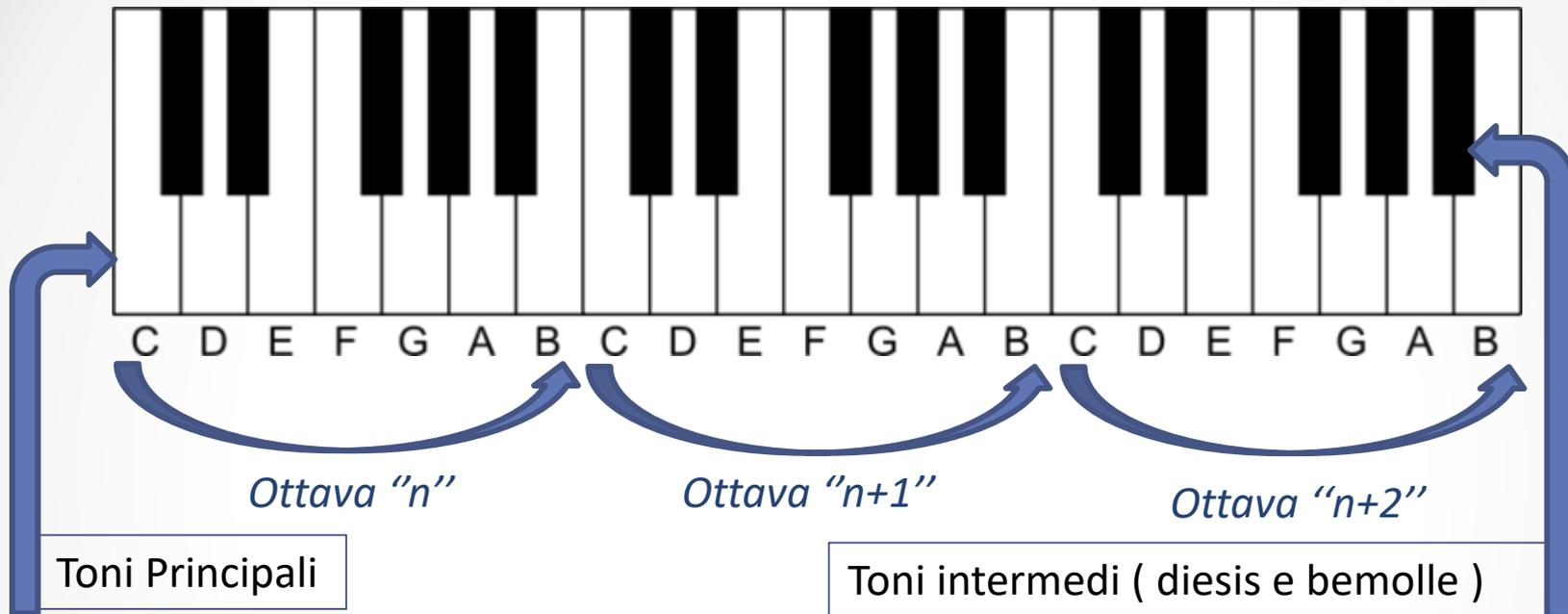
Acquisizione degli input



Acquisizione degli input



Costruzione dei Reference Points



Ogni ottava consta, in totale, di 12 semitoni, tra principali ed intermedi.

Considerando le ottave numero 2,3,4 e 5 di un pianoforte si ha un totale di 48 semitoni, i quali costituiscono l'insieme dei nostri Reference Points.

Ciascun RP è caratterizzato da una frequenza fondamentale eccitata e, a seconda delle caratteristiche dello strumento utilizzato, un certo numero di armoniche.

Costruzione dei Reference Points

Il generico RP " O^t " sarà, quindi, un vettore del tipo :

$$O^t = \frac{1}{\sqrt{n_h + 1}} (e_{k_t} + e_{2k_t} + e_{3k_t} + \dots + e_{(n_h+1)k_t})$$

e_j : vettore della base o.n. considerata;

k_t : frequenza fondamentale del semitono considerato;

n_h : numero di armoniche imposto.

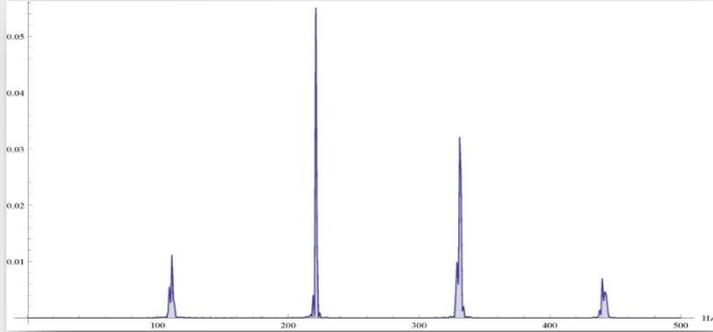
Le componenti diverse da zero nei RP sono unicamente quelle relative alla posizione k_t ed alle successive n_h armoniche nelle posizioni ik_t ($i = 1, \dots, n_h + 1$).

ATTENZIONE !

Il numero di armoniche n_h è importante ai fini del corretto riconoscimento del semitono.

Il numero di armoniche dei RP " O^t " dev'essere simile a quello dei vettori d'input " f "!

Riconoscimento di note mediante operatori di proiezione



Input (f_{A2}) : semitono A2 con $k_t = 110$ Hz.

Armoniche considerate nei RP : $n_h = 0$.

O^t	2nd	3rd	4th	5th
C	1.999	1.993	1.997	1.998
C#	1.998	1.992	1.995	1.970
D	1.998	1.992	1.998	1.994
D#	1.997	1.994	1.987	2.000
E	1.998	1.995	1.127	1.984
F	1.993	1.989	1.985	1.995
F#	1.996	1.988	1.994	1.996
G	1.984	1.981	1.999	1.999
G#	1.977	1.962	1.994	1.995
A	1.491	0.8482	1.524	1.922
A#	1.976	1.978	1.991	1.999
B	1.989	1.993	1.998	1.975

$$\|f_{A2} - O^t\|^2$$

O^t	2nd	3rd	4th	5th
C	0.9997	0.9967	0.9987	0.9991
C#	0.9988	0.9961	0.9975	0.9849
D	0.9992	0.9960	0.9990	0.9970
D#	0.9984	0.9969	0.9934	0.9998
E	0.9992	0.9974	0.5630	0.9919
F	0.9967	0.9947	0.9924	0.9978
F#	0.9980	0.9938	0.9970	0.9980
G	0.9919	0.9904	0.9995	0.9994
G#	0.9887	0.9811	0.9971	0.9974
A	0.7460	0.4240	0.7620	0.9608
A#	0.9881	0.9892	0.9955	0.9995
B	0.9946	0.9964	0.9989	0.9873

$$F[f_{A2}, O^t]$$

Riconoscimento di note mediante operatori di proiezione

Riconoscimento errato dovuto al differente numero di armoniche tra l'input (3 armoniche eccitate) e i RP (0 armoniche eccitate)

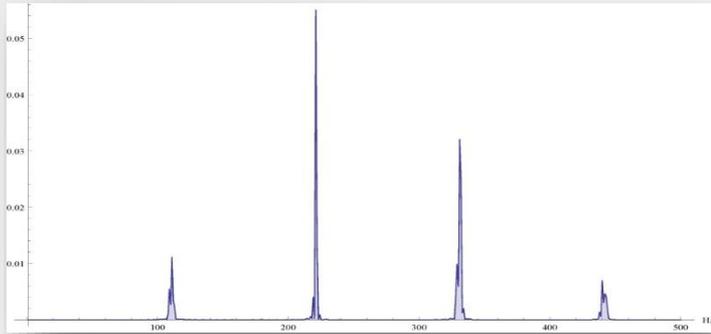
C#	1.998	1.992	1.995	1.970
D	1.998	1.992	1.995	1.994
D#	1.997	1.994	1.997	2.000
E	1.998	1.995	1.992	1.984
F	1.993	1.989	1.985	1.995
F#	1.996	1.988	1.994	1.996
G	1.984	1.985	1.999	1.999
G#	1.977	1.992	1.994	1.995
A	1.491	0.8482	1.524	1.922
A#	1.976	1.978	1.991	1.999
B	1.989	1.993	1.998	1.975

$$\|f_{A2} - O^t\|^2$$

C#	0.9988	0.9961	0.9975	0.9849
D	0.9992	0.9960	0.9990	0.9970
D#	0.9984	0.9969	0.9934	0.9998
E	0.9992	0.9974	0.5630	0.9919
F	0.9967	0.9947	0.9924	0.9978
F#	0.9988	0.9938	0.9970	0.9980
G	0.9919	0.9904	0.9995	0.9994
G#	0.9887	0.9811	0.9971	0.9974
A	0.7460	0.4240	0.7620	0.9608
A#	0.9881	0.9892	0.9955	0.9995
B	0.9946	0.9964	0.9989	0.9873

$$F[f_{A2}, O^t]$$

Riconoscimento di note mediante operatori di proiezione



Input (f_{A2}) : semitono A2 con $k_t = 110$ Hz.

Armoniche considerate nei RP : $n_h = 2$.

O^t	2nd	3rd	4th	5th
C	1.984	1.992	1.992	1.991
C#	1.984	1.985	1.993	1.953
D	1.988	1.883	1.930	1.957
D#	1.980	1.980	1.989	1.957
E	1.991	1.699	1.577	1.891
F	1.984	1.990	1.991	1.984
F#	1.986	1.987	1.982	1.989
G	1.985	1.993	1.988	1.987
G#	1.971	1.992	1.993	1.976
A	1.152	1.479	1.836	1.938
A#	1.970	1.992	1.994	1.960
B	1.985	1.988	1.848	1.817

$$\|f_{A2} - O^t\|^2$$

O^t	2nd	3rd	4th	5th
C	0.9933	0.9972	0.9990	0.9977
C#	0.9935	0.9961	0.9974	0.9997
D	0.9974	0.9433	0.9670	0.9797
D#	0.9970	0.9915	0.9955	0.9843
E	0.9981	0.8540	0.8870	0.9486
F	0.9949	0.9984	0.9967	0.9930
F#	0.9991	0.9957	0.9952	0.9988
G	0.9959	0.9982	0.9959	0.9969
G#	0.9949	0.9975	0.9966	0.9880
A	0.7140	0.8480	0.9517	0.9797
A#	0.9873	0.9971	0.9979	0.9805
B	0.9950	0.9980	0.9246	0.9186

$$F[f_{A2}, O^t]$$



RICONOSCIMENTO DI PATOLOGIE MEDICHE: CELIACHIA

Riconoscimento della celiachia

Si considera un dataset di “*training*” composto da 300 vettori $I_j \in \mathcal{H} = \mathbb{R}^{35}$, ciascuno dei quali rappresenta un paziente e per i quali si conosce il corretto output : O_0 per pazienti non affetti da celiachia, O_1 per pazienti affetti da celiachia.

La dimensione di questi vettori è dovuta al numero di parametri a disposizione per determinare se un paziente sia celiaco o meno (16 sintomi, 11 evidenze fisiche, 8 indagini di laboratorio). Ciascun parametro può assumere un valore compreso tra 0 e 10.

Di conseguenza questo dataset \mathcal{S}_{train} è l’unione tra il sub-dataset \mathcal{S}_0 , insieme degli input aventi O_0 come output, e \mathcal{S}_1 , insieme degli input aventi O_1 come output.

$$\mathcal{S}_{train} = \mathcal{S}_0 \cup \mathcal{S}_1, \quad \mathcal{S}_0 \cap \mathcal{S}_1 = \emptyset.$$

Riconoscimento della celiachia

Ciascun paziente viene modellizzato come un vettore $f = (f_1, f_2, f_3, \dots, f_{35})$, per stabilire se sia celiaco o meno viene applicato a questo vettore l'operatore $F[f, I_j]$, distinguendo i risultati in base al sub-dataset di appartenenza del vettore I_j , e viene estratto il valore minimo per ciascuno dei due insiemi ottenuti:

$$d_0 = \min_{f_0 \in \mathcal{S}_0} F[f, f_0]$$

$$d_1 = \min_{f_1 \in \mathcal{S}_1} F[f, f_1]$$

$$d_0 > d_1$$



NOT CELIAC

$$d_0 < d_1$$



CELIAC !

Riconoscimento della celiachia

Risultati su un nuovo dataset di 30 pazienti (output inizialmente sconosciuto, poi rivelato ai fini della verifica)

$$F[f_{new}, I_j]$$

$$\|f_{new} - I_j\|^2$$



26 diagnosi corrette su 30

Decision Tree



23 diagnosi corrette su 30



FRAMES E SEGNALI ALTERATI DA RUMORE

Frames

Frame

Insieme di vettori φ_i non linearmente indipendenti tali che:

$$A\|f\|^2 \leq \sum_i |\langle \varphi_i, f \rangle|^2 \leq B\|f\|^2$$

Tight
Frame

Frame in cui $A = B$, per cui vale la relazione :

$$\|f\|^2 = A \sum_i |\langle \varphi_i, f \rangle|^2$$

Frames

- E' possibile costruire, a partire da un frame \mathcal{F}_φ , il relativo frame duale $\mathcal{F}_{\tilde{\varphi}}$ tale che:

$$f = \sum_i \langle \varphi_i, f \rangle \tilde{\varphi}_i = \sum_i \langle \tilde{\varphi}_i, f \rangle \varphi_i$$

- Per un tight frame è valida la relazione :

$$\tilde{\varphi}_i = \alpha \varphi_i$$

- Un esempio di tight frame in uno spazio di Hilbert a dimensione "n" è l'insieme di vettori \mathcal{F}_φ :

$$\mathcal{F}_\varphi = \left\{ \varphi_1 = e_1, \varphi_2 = \frac{1}{2} e_1, \dots, \varphi_{2n-1} = e_n, \varphi_{2n} = \frac{1}{2} e_n \right\}$$

Frames



Distanza in norma

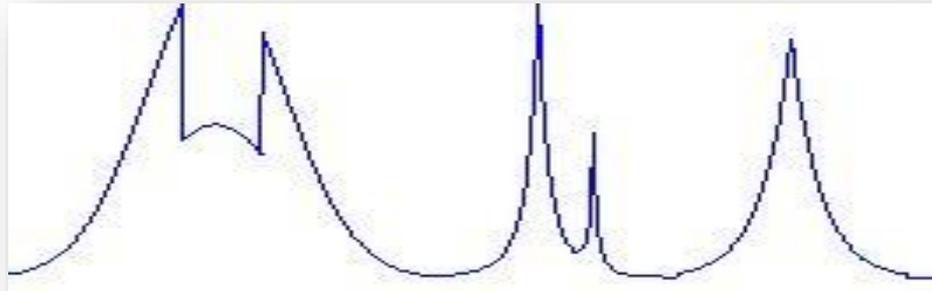
$$\|f - g\|^2 = \sum_{k=0}^n (f_k - g_k)^2$$

Distanze estese

$$\Delta(f, g) = \max\{\sup_i | \langle \varphi_i, f - g \rangle |, \sup_i | \langle \tilde{\varphi}_i, f - g \rangle | \}$$

$$\nabla(f, g) = \min\{\sup_i | \langle \varphi_i, f - g \rangle |, \sup_i | \langle \tilde{\varphi}_i, f - g \rangle | \}$$

Segnali alterati da rumore



Segnale originale f

$$\|P - f\| \leq \varepsilon$$



$$f \in K_{\varepsilon}^{\|\cdot\|}(P)$$

$$\Delta(P, f) \leq \varepsilon$$



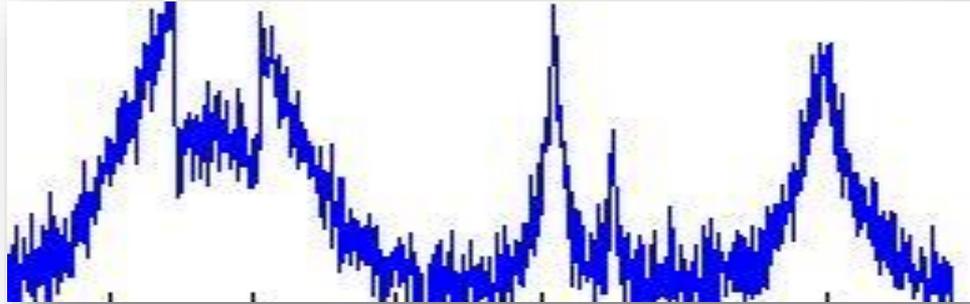
$$f \in K_{\varepsilon}^{\Delta}(P)$$

$$\nabla(P, f) \leq \varepsilon$$

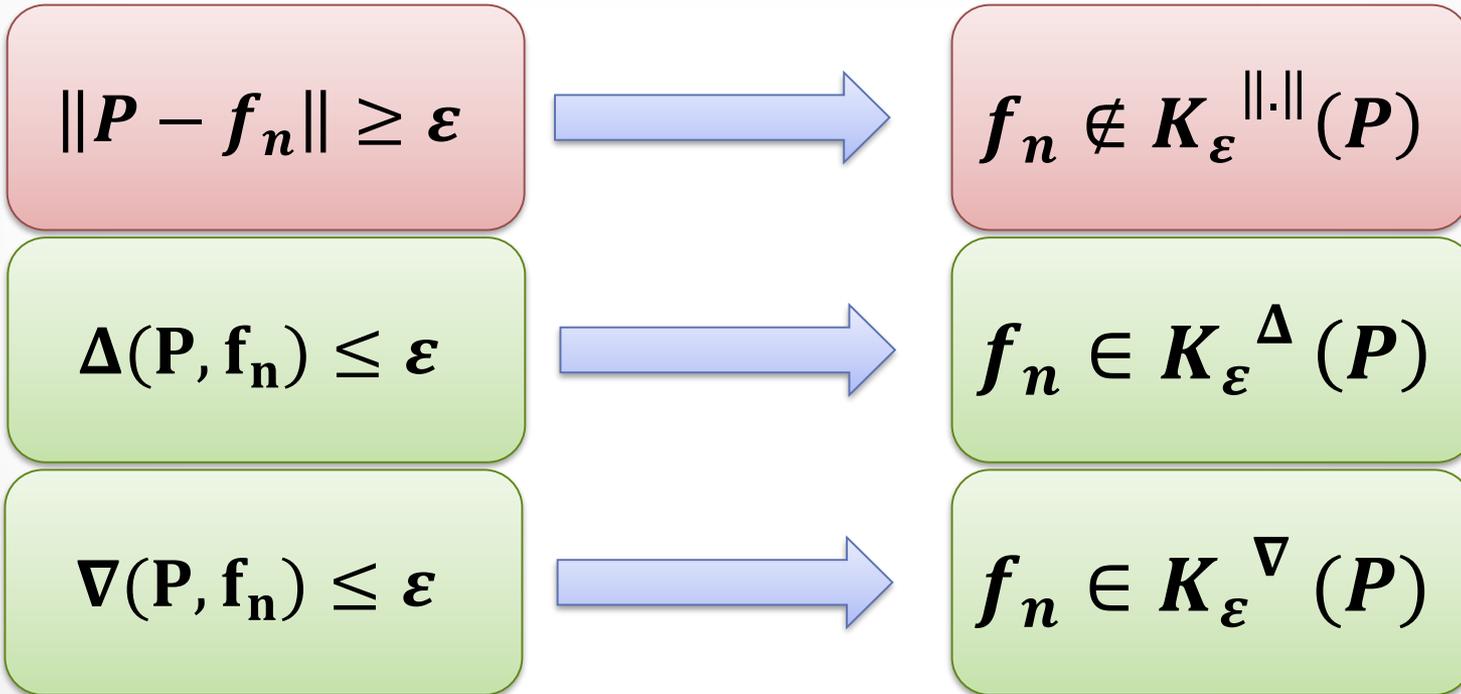


$$f \in K_{\varepsilon}^{\nabla}(P)$$

Segnali alterati da rumore



Segnale alterato f_n



Riconoscimento di suoni alterati da rumore

$$f = LA_{2nd}$$



$$P = e_{110}$$



$$LA_{2nd} \in K_{\varepsilon}^{\parallel \cdot \parallel}(e_{110})$$

$$LA_{2nd} \in K_{\varepsilon}^{\Delta}(e_{110})$$

$$LA_{2nd} \in K_{\varepsilon}^{\nabla}(e_{110})$$

Riconoscimento di suoni alterati da rumore

- Consideriamo il segnale alterato $f_n = \widetilde{L}A_{2nd}$, ottenuto modificando in maniera casuale le componenti del vettore di input, e valutiamo la sua appartenenza ai tre cluster prima definiti.

$$f = \widetilde{L}A_{2nd}$$

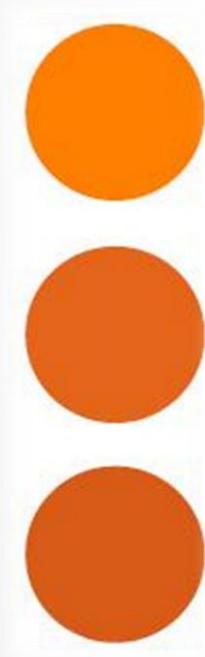


$$\widetilde{L}A_{2nd} \in K_{\varepsilon}^{\|\cdot\|}(e_{220})$$

$$\widetilde{L}A_{2nd} \in K_{\varepsilon}^{\Delta}(e_{110})$$

$$\widetilde{L}A_{2nd} \in K_{\varepsilon}^{\nabla}(e_{110})$$

Riconoscimento di colori alterati



$$P_o = (1, 0.5, 0)$$

$$f_o = (0.9, 0.4, 0.1)$$

$$\tilde{f}_o = (0.83, 0.36, 0.08)$$



$$\|P_o - \tilde{f}_o\| = 0.22$$

$$\Delta(P_o, \tilde{f}_o) = 0.15$$

$$\nabla(P_o, \tilde{f}_o) = 0.12$$

Grazie della vostra
attenzione!