

Università degli Studi di Bologna

---

**Facoltà di Ingegneria**  
Corso di Laurea in Ingegneria Elettronica  
Geometria e Algebra

**CODIFICA SINESTETICA  
DEI COLORI MEDIANTE SUONI  
IN UN DISPOSITIVO PER VIDEOLESI**

Tesi di Laurea di:  
**Ludovico Ausiello**

Relatore:  
**Prof. Massimo Ferri**

Correlatori:  
**Prof. Luigi Di Stefano**  
**Prof.ssa Nicoletta Caramelli**  
**Ing. Marco Mattioli**

Sessione Autunnale

---

**Anno Accademico 2003/2004**



# Indice

<b>1</b>	<b>Il progetto EAV</b>	<b>7</b>
<b>2</b>	<b>Implementazione</b>	<b>13</b>
2.1	Linguaggio di programmazione e formato dei dati . . . . .	13
2.2	Parallelismo . . . . .	14
2.3	Rappresentazione del colore . . . . .	15
2.3.1	Una codifica esatta . . . . .	16
2.3.2	Una codifica associativa . . . . .	20
2.4	Il Bianco ed il Nero . . . . .	23
2.5	Lenti . . . . .	24
2.6	Software per la sperimentazione . . . . .	24
<b>3</b>	<b>Sperimentazione</b>	<b>27</b>
3.1	Metodo . . . . .	27
3.1.1	Partecipanti . . . . .	27
3.1.2	Materiali . . . . .	28
3.1.3	Procedura . . . . .	30

3.1.4	Analisi dei dati e discussione dei risultati . . . . .	31
3.2	Valutazioni e possibili sviluppi futuri . . . . .	33
<b>A</b>	<b>Materiali</b>	<b>37</b>
A.1	Sequenze di apprendimento . . . . .	37

# Introduzione

E' realmente possibile pensare di compensare le disabilità con la tecnologia? Una risposta accettabile non può provenire solo dall'ingegneria intesa come aumentata potenza di calcolo o miniaturizzazione, bensì da un atteggiamento progettuale attento al fattore umano del problema. Il progetto di un ausilio elettronico per non vedenti, per esempio, dovrà tener conto di aspetti quali invasività (fisica o sensoriale), portabilità, istintività d'uso, tempi di apprendimento. In questa direzione si è mosso il progetto **EAV** promosso, in Spagna, dall'università La Laguna insieme con l'istituto nazionale dei ciechi (**ONCE**). All'appenza, infatti, l'**EAV** è costituito soltanto da un paio di occhiali da sole e due auricolari. Gli occhiali nascondono due microcamere: la scena ripresa viene convertita acusticamente sotto forma di un gocciolio che dà forma all'ambiente.

Con questa tesi vorremmo colmare una lacuna dell'**EAV**: la mancanza del colore. Presentiamo e confrontiamo, da un punto di vista della istintività d'uso e della resa cromatica, due soluzioni che codificano la varietà dei colori in varietà timbrica.

Dopo un primo capitolo, in cui si descrive il progetto **EAV**, si espongono i vincoli e si motivano le scelte progettuali delle due codifiche. Il terzo capitolo descrive la metodologia sperimentale che ha permesso di validare il lavoro svolto e che ha consentito di effettuare un confronto di prestazioni. La tesi è conclusa da due appendici contenenti i materiali di apprendimento e dei test e i dati sperimentali.



# Capitolo 1

## Il progetto EAV

### Creazione di uno spazio acustico virtuale

In questo capitolo ci proponiamo di descrivere la maggior parte degli aspetti, da quelli più strettamente matematici ed ingegneristici fino a quelli psicologici e neurofisiologici, che si vengono via via a toccare parlando di **EAV**;[?],[?] in questo modo potremo capire quali passi siano già stati compiuti nella realizzazione di questo dispositivo, quali si vogliano implementare con questa tesi e quanto resti ancora da affrontare in futuro.

Il limite di molti dispositivi per non vedenti che sfruttano il tatto, come ad esempio il **videt** [?], è che impongono una esplorazione ‘seriale’ dello spazio; per ottenere un quantitativo consistente di informazione è necessario un tempo lungo. Verrebbe da chiedersi, quindi, se vi sia un senso che, come la vista, sia intrinsecamente ‘parallelo’ e consenta così di veicolare più informazione in meno tempo: l’udito offre questa possibilità. Studi effettuati su piccoli mammiferi (gatti) parzialmente o totalmente privi della vista hanno evidenziato l’attivazione di una particolare zona della corteccia visiva (*EAV Cortex*) da parte di stimoli sonori [?],[?]. Queste ricerche, unitamente alla descrizione di **immagini sonore** evocate dalla pioggia, [?] hanno portato, nel 1996 in Spagna, alla formazione di un gruppo di ricerca presso l’università La Laguna con lo scopo di progettare un dispositivo di ausilio per non vedenti che, sfruttando l’udito, potesse ricreare uno ‘spazio acustico virtuale’, per offrire a persone non vedenti un mezzo di esplorazione per ambienti chiusi ed aperti.

Il dispositivo è in sostanza una catena di acquisizione ed elaborazione di informazione. L'anello iniziale è rappresentato da due microcamere nascoste in paio di occhiali da sole: la scena è ripresa in stereovisione ed è rappresentata nella memoria di un elaboratore sotto forma di *pixel* sterici. Un elaboratore (un palmare Pentium166) estrapola da questa mappa discreta un ambiente virtuale formato da piani orizzontali e verticali: non si deve pensare che in questa rappresentazione vi sia una schematizzazione troppo forte, perchè quasi tutti gli ambienti di vita quotidiana si possono rappresentare correttamente (superfici curve si possono approssimare con alcuni piani tangenti, a seconda della disponibilità di risorse di calcolo).

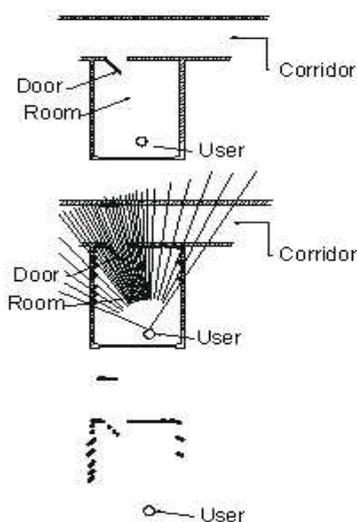


Figura 1.1: Esempio bidimensionale del comportamento del sistema **EAV**.

Sopra queste superfici il dispositivo crea tanti piccoli suoni impulsivi, come tante 'gocce di pioggia'; una volta 'ricoperte' dalle gocce, quindi, esse appaiono nel nostro campo visivo. Il suono è trasmesso al non vedente mediante un semplice paio di cuffie aperte, che permettono cioè anche ai normali rumori esterni di raggiungere l'orecchio. Un primo problema a cui si pensa subito è la capacità di risoluzione di un sistema così realizzato; ci sono diversi vincoli di progetto: il primo è la capacità di calcolo necessaria a rendere nella immagine virtuale i tanti particolari di una immagine tridimensionale reale. Dobbiamo pensare di suddividere lo spazio circostante in un certo numero di *pixel* sterici, ma quanti? La fisiologia umana suggerisce un sistema non lineare di codifica dell'immagine, con una risoluzione più alta al centro, cui

corrisponda un maggior numero di *stereopixel*, e molto più bassa ai margini, esattamente come avviene nell'analisi che compie la fovea dell'occhio umano.

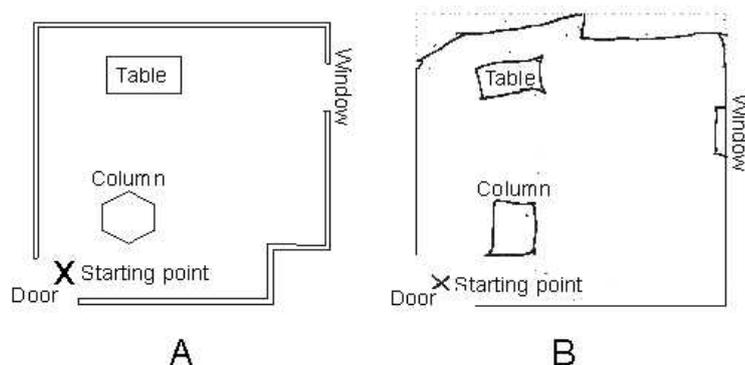


Figura 1.2: A. Rappresentazione schematica del laboratorio. B. Disegno fatto da un non vedente dopo una esplorazione molto breve, utilizzando un prototipo di **EAV**, senza uso del tatto.

Il secondo vincolo è certamente la capacità dell'orecchio di saper distinguere la provenienza esatta di un suono; sono stati fatti molti studi negli ultimi 20 anni sia in America che in Europa [?],[?],[?],[?],[?],[?],[?],[?],[?] nell'intento di scoprire come l'uomo riesca a localizzare nello spazio una percezione sonora; i risultati di queste ricerche sono assolutamente necessari al nostro progetto. Come per la vista anche l'udito ha una risoluzione minima oltre la quale non può andare, risoluzione che si aggira intorno ai 2–3 gradi sul piano orizzontale e 5–10 gradi in elevazione.

Le cifre trovate dai ricercatori rispecchiano un fatto importante, la soggettività di queste capacità. Dobbiamo conoscere quindi una sorta di **funzione di trasferimento individuale** con cui ogni singolo utente, destinatario di un dispositivo reale, percepisca un suono come proveniente da destra o da sinistra, dall'alto o dal basso.[?],[?],[?],[?] Una volta nota questa funzione (*Head Related Transfer Function*) possiamo ricostruire questi meccanismi, al tempo stesso fisici e psicologici, con l'elaboratore ed 'ingannare' chi ascolta in cuffia il gocciolio della pioggia.[?],[?] In un ascolto in cuffia, infatti, si dis-

tingue correttamente la provenienza di tutti i suoni, ma questi sono sempre percepiti come all'interno della testa.

In questo dispositivo, invece, le gocce debbono certamente cadere su un tavolo o su di una parete che è sicuramente fuori dalla testa! Un passo ulteriore del progetto prevede, quindi, di filtrare opportunamente le gocce per assegnare loro il corretto attributo di spazialità e rendere così corretta la percezione soggettiva. Va notato come la *HRTF* debba essere specificatamente misurata per ciascun utente; volendo presentare comunque solo immagini percepite idealmente con la vista si può restringere il campo di misurazione alla metà frontale dell'angolo giro.

A questo punto cosa manca alla nostra rappresentazione sonora della realtà? Porte, tavoli, finestre, tante forme della vita di tutti i giorni sembrano poter apparire di fronte a noi .. ma di che colore sono? Ecco, finalmente, l'obiettivo di questa tesi: riuscire a trasmettere l'informazione relativa al colore mediante il suono, o sue variazioni. Si è detto informazione relativa al colore con un intento specifico; vi sono, infatti, almeno due diverse categorie di utenti possibili per l' **EAV**, ovvero o persone non vedenti dalla nascita o persone che abbiano perso la vista per motivi legati a traumi o malattie avvenute in età più o meno avanzata: la differenza fondamentale tra questi due soggetti sta nell'avere o meno la nozione di colore, almeno sotto forma di ricordo della percezione del colore.

Il fatto è che, a prescindere dall'esperienza sensoriale dovuta al colore, ognuno di noi ha appreso una serie vasta di nozioni ad esso correlate: come dire, un rosso non vuol dire nulla in se, ma un semaforo rosso vuol dire '*non attraversare*'! Questa è comunque informazione e, di nuovo, in presenza di oggetti difficilmente distinguibili mediante il solo **EAV**, come ad esempio due figure umane, il fatto che queste possano indossare indumenti di colori diversi può essere un mezzo semplice per percepirli come diversi.

Lo scopo di questa tesi, dunque, è realizzare un'associazione tra colore e suoni che consenta al non vedente di percepire il mondo esterno in modo sufficientemente articolato da permettergli di esplorare ambienti complessi in un lasso temporale adeguato e permetta quindi di ricevere più informazione in meno tempo. Possiamo fissare alcune priorità di lavoro che dovranno essere affrontate per ottenere un risultato quanto più realistico e pratico possibile:

- minima complessità computazionale
- minima necessità di apprendimento
- buona resa cromatica
- personalizzazione da parte dell'utente finale



# Capitolo 2

## Implementazione

### 2.1 Linguaggio di programmazione e formato dei dati

Poichè l'**EAV** utilizza un elaboratore palmare (Pentium166) si deve pensare a come scrivere il software di basso livello necessario al corretto interfacciamento con l'hardware: questo problema suggerisce la scelta di un insieme di primitive per la gestione dei dati audio (in formato *.wav*) basato su piattaforma Microsoft, che offre il pacchetto Directsound.

Per quanto riguarda la scelta del formato di dati (all'inizio del progetto non esisteva ancora il formato *.mp3*) è comunque preferibile continuare ad utilizzare file *.wav*; questa scelta è motivata da una caratteristica intrinseca del timbro degli stimoli acustici (le nostre 'gocce di pioggia'). Dovendo usare infatti suoni impulsivi abbiamo a che fare con segnali con spettro in frequenza assai esteso, con bande fino a 10 **KHz**, che verrebbero quindi fortemente distorti dagli algoritmi di compressione. Una scelta ulteriore potrebbe essere anzi quella di sostituire al singolo impulso di una goccia di pioggia un piccolo '*treno d'acqua*', un pacchetto di impulsi ravvicinati, scelta che sarà suggerita anche da altre problematiche sorte durante lo sviluppo di questa tesi. [?]

Queste considerazioni suggeriscono che, nell'associazione tra colori e timbri di strumenti a percussione che intendiamo stabilire, la scelta degli strumenti avvenga in base al loro contenuto spettrale ed in base all'immediatezza di riconoscimento da parte dei partecipanti alla sperimentazione.

## 2.2 Parallelismo

Nella prospettiva di realizzare un software che si integri con un dispositivo portatile (un sistema di tipo embedded) è importante valutare la complessità computazionale richiesta dai vari elementi del progetto, in modo da poter evidenziare vincoli sul processore da impiegare piuttosto che sulla quantità di memoria da utilizzare; il nostro algoritmo, infatti, dovrebbe essere utilizzato con un elevato parallelismo, richiedendo una grande quantità di risorse, per rendere l'effetto sonoro di una vera pioggia a partire dal suono di una goccia singola. Il nostro apparato uditivo però ha una capacità di elaborazione limitata: esiste un fenomeno molto importante e complesso che prende il nome di mascheramento; si può avere mascheramento in frequenza, in intensità ed infine mascheramento temporale. [?]

Il mascheramento in frequenza è diretta conseguenza della conformazione fisica dell'orecchio interno e fa sì che un suono, sia esso puro, a banda stretta o a banda larga, percepito in un istante di tempo  $t_0$  alteri la soglia uditiva per un certo tempo  $t_{Mf}$  in una banda  $\Delta f$  intorno alle frequenze del primo impulso. Maggiore è la banda del primo impulso, più esteso risulta lo span  $\Delta f$  delle frequenze la cui soglia è aumentata (un impulso di Dirac altera quindi la soglia di udibilità in tutto lo spettro udibile); le gocce di pioggia devono avere uno spettro ampio, per poter esibire le caratteristiche di impulsività temporali necessarie a facilitare la valutazione della loro provenienza nello spazio che si vuole esplorare.

Il secondo tipo di mascheramento è quello in intensità: in presenza di due suoni  $s_1$  ed  $s_2$ , aventi analoghe caratteristiche spettrali, se la pressione acustica generata da  $s_1$  supera di **10 dB** quella generata da  $s_2$  allora il nostro apparato uditivo non è in grado di percepire il suono di intensità minore.

Il mascheramento temporale, infine, fa sì che il nostro orecchio non riesca a riconoscere come distinti due suoni se tra di essi non vi è un certo intervallo di tempo minimo  $t_{Mt}$ , ma li interpreti invece come un primo suono seguito da un riverbero o da un'eco, a seconda dell'entità del ritardo tra i due. Ognuno di questi fenomeni, soprattutto se vogliamo ottenere dei risultati efficaci senza sprecare risorse, ci induce a riflettere attentamente su alcuni punti:

- nell'evidenziare un oggetto in primo piano, caratterizzato da un certo volume sonoro, la pioggia 'nasconde' altri aspetti dell'ambiente, a prescindere dalla loro ubicazione.

- utilizzare troppe gocce di pioggia è inutile, perché l'orecchio 'sommerebbe' i contributi di tutte quelle gocce che cadessero in successione troppo serrata l'una dopo l'altra (il tempo di integrazione varia con l'età, da un minimo di 50ms per le persone giovani fino a 200ms per persone anziane).
- una goccia che cada su di un oggetto di un certo colore maschera in modo più o meno intenso gli altri oggetti a seconda del colore di questi ultimi.

Queste considerazioni ci portano quindi verso un numero massimo di gocce, oppure ci suggeriscono nuovamente di sostituire il suono elementare di una singola goccia con un pacchetto di impulsi. Ai fini dell'elaborazione del colore, inoltre, vedremo che a seconda del tipo di codifica potremo avere vincoli di progetto diversi, ovvero potremmo aver bisogno di più o meno gocce per evitare troppa 'confusione cromatica'.

## 2.3 Rappresentazione del colore

L'uomo mostra una capacità sensoriale nello spettro elettromagnetico caratterizzato da lunghezze d'onda comprese tra 380 nm e 780 nm; in questo range di frequenze si percepiscono i colori come luce riflessa da oggetti illuminati da una sorgente. Se un oggetto riflette tutta la luce incidente esso ci appare bianco, mentre se non riemette radiazione ci sembra nero. [?] L'occhio umano è in grado di cogliere piccole differenze cromatiche grazie alla sensibilità della retina ed è capace di confrontare tonalità ed intensità con una tavolozza standard fornita come riferimento. Vi sono anche in questo caso delle limitazioni alla capacità di elaborazione che si riflettono principalmente in:

- Influenza della distanza dell'oggetto da osservare; nel confronto tra due campioni, senza linea di separazione, le capacità dell'occhio di registrare differenze di tono e di intensità si indeboliscono rapidamente non appena i campioni vengano allontanati.
- Influenza dell'ambiente; un oggetto circondato da colori diversi viene percepito in modo differente.
- Il ricordo del colore è soggetto a rapido decadimento

- Metamerismo; fissata una condizione di illuminazione, due oggetti con spettri di riflessione diversi possono dare la stessa impressione cromatica ad una valutazione visiva, mentre appaiono diversi in altre condizioni di luce.

La stima visiva è inoltre difficile da comunicare verbalmente, rendendo molto complicato anche solo il confronto tra i giudizi di due soli soggetti. E' necessaria quindi una codifica standard per poter rappresentare lo spazio dei colori in modo univoco.

Oltre che a dipendere dalla sensibilità umana, come si è detto, il colore dipende anche dalla sorgente luminosa e dalla natura cromatica della superficie da osservare. Di queste tre variabili quella che interessa ai fini di un'analisi comparativa e all'associazione con i suoni è l'ultima, per cui partiremo dal presupposto di fissare le altre due, standardizzando cioè sia la distribuzione energetica della fonte illuminante sia la risposta dell'occhio dell'osservatore medio. Possiamo a questo punto estrapolare tre attributi fondamentali del colore:

1. Il *tono* dipende dalla lunghezza d'onda della luce diffusa e fa apparire gli oggetti variamente colorati.
2. La *saturazione* definisce l'intensità di un colore, poiché indica quanto è stato diluito con luce bianca
3. La *luminanza*, infine, rappresenta la quantità totale di luce riflessa dall'oggetto

### 2.3.1 Una codifica esatta

Nella retina esistono tre diversi centri di stimolo il cui massimo di sensibilità spettrale è situato nel rosso, nel verde e nel blu. Possiamo definire i valori **X** (red), **Y** (green), **Z** (blue) come i valori cromatici di base, che indicano quanto intensamente i tre centri di stimolo sono sollecitati dalla radiazione incidente. La codifica più semplice per il colore, dunque, ricalca esattamente il funzionamento dell'occhio ed è costituita da tre vettori (tristimolo) mediante i quali possiamo rappresentare ogni singolo elemento del nostro spazio vettoriale; nella figura 2.1 possiamo osservare le componenti dei tre vettori in funzione dello spettro elettromagnetico. [?]

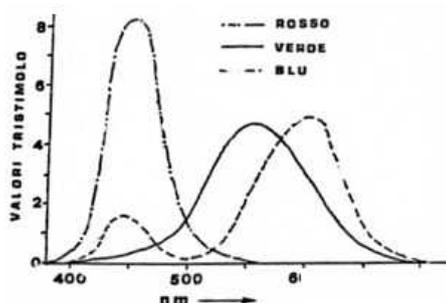


Figura 2.1: Valori del tristimolo in funzione della lunghezza d'onda.

Poichè un colore non è facilmente intuibile attraverso i valori tristimolo, si può effettuare un cambiamento di base (noto come Yxy Color System, CIE 1931) che lega alla terna originale

$$\begin{bmatrix} X & Y & Z \end{bmatrix}$$

una nuova tripletta di valori

$$\begin{bmatrix} x & y & z \end{bmatrix}$$

mediante le seguenti relazioni

$$x = \frac{X}{X+Y+Z} \quad y = \frac{Y}{X+Y+Z} \quad z = 1 - x - y$$

Lo spazio ottenuto in questo modo può essere rappresentato in un diagramma spettrale costituito da una curva a campana all'interno della quale si trovano tutti i colori. [?],[?]

Al centro della figura 2.2 vi è una zona dove tutti e tre gli stimoli raggiungono i valori massimi, determinando la tonalità cromatica del bianco; questa codifica è un ottimo punto di partenza per la nostra associazione tra colore e suono: potremmo infatti, rispettando l'esattezza geometrica di questa rappresentazione, associare a ciascun valore x, y e z un diverso timbro di

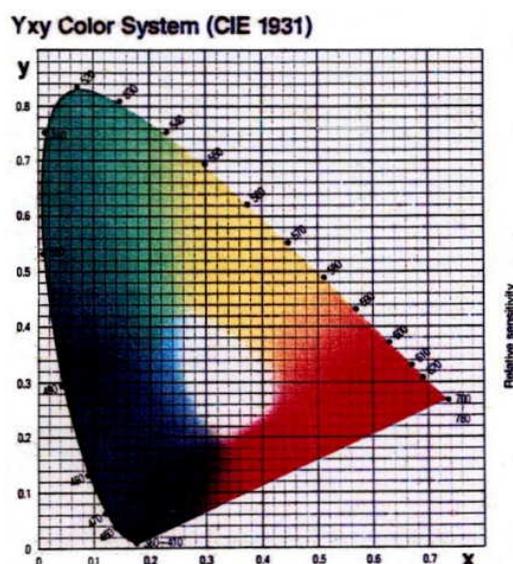


Figura 2.2: Diagramma di cromaticità CIE1931.

strumento (considerazioni riguardanti la scelta dei timbri saranno presentate più avanti). Il volume relativo di ciascuno di essi rappresenterà l'intensità relativa dello stimolo, lasciando la possibilità di rappresentare tutto lo spettro cromatico. Due domande sorgono immediate: quanto fine è la capacità dell'orecchio di distinguere i diversi volumi relativi? che influenza ha il fenomeno del mascheramento in questa possibile realizzazione? Vi è inoltre un piccolo difetto di partenza, che complica le cose: con questo cambiamento di base le distanze relative tra i colori non sono identiche, sarebbe a dire, ad esempio, che tra il viola ed il blu il passo è più breve che tra il blu ed il verde.

Nelle rappresentazioni del colore questo problema si risolve con un ulteriore cambiamento di base che porta ad ottenere sempre tre vettori linearmente indipendenti che, però, hanno valori sia positivi che negativi (Hunter Color System e  $L^*A^*B$  Color System, CIE 1976). [?] Questo aspetto è difficile da affrontare, quando associamo ad ogni vettore del colore un timbro, poiché non possiamo avere volumi negativi! Se vogliamo adottare una relazione geometricamente corretta tra suono e colore, quindi, dobbiamo inevitabilmente utilizzare una scala non lineare, con le probabili distorsioni ad essa collegate.

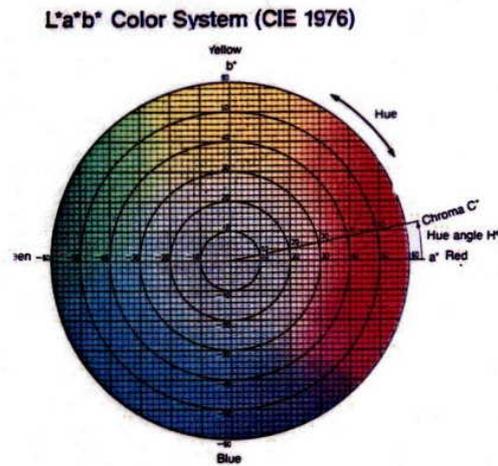


Figura 2.3: Diagramma Hunter- $L^*A^*B$  1976.

Una volta applicati appropriati cambiamenti di base, comunque, possiamo ottenere una figura di forma regolare in cui sia facile esprimere i rapporti tra i volumi dei timbri primari. La figura 2.4 mantiene una rappresentazione sufficientemente fedele del Gamut del RGB e non pone il problema di avere volumi negativi per gli stimoli acustici.

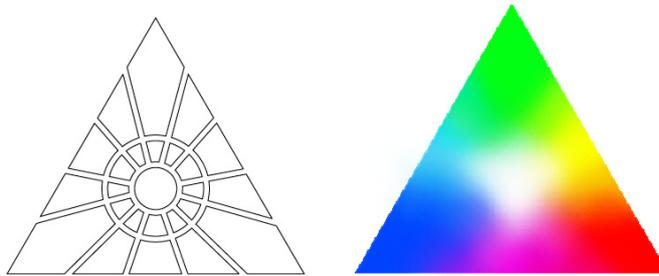


Figura 2.4: Gamut del RGB e mappa discreta dei volumi relativi dei timbri associati ai colori primari.

Dal punto di vista della complessità computazionale, però, questa associazione richiede elaborazione in tempo reale al nostro sistema; una volta ricostruito l'ambiente da esplorare, infatti, ad ogni stereopixel devono essere associati i tre valori della terna  $[x, y, z]$  ed in base a questi devono essere calcolati i volumi relativi dei tre timbri che, di volta in volta, costituiscono

un nuovo suono di goccia elementare da collocare nello spazio da parte del sistema. [?],[?],[?]

Viene da chiedersi se non vi sia un sistema più semplice e soprattutto meno oneroso da un punto di vista del calcolo per associare ai colori dei timbri; facendo un parallelo con il mondo dei microprocessori sarebbe bello avere una sorta di cache, di memoria associativa, da cui poter reperire velocemente, mediante una semplice lettura, i timbri delle gocce di pioggia che colpiscono oggetti di colore diverso. Un sistema di questo tipo, infatti, richiederebbe al sistema solamente di riconoscere il colore, per effettuare così la scelta corretta del suono da riprodurre.

### 2.3.2 Una codifica associativa

Nel momento in cui si scelga una codifica di tipo puramente associativo tra colori e suoni bisogna subito precisare come l'algoritmo di codifica sia soggettivo; ciascuno degli utenti finali del sistema, infatti, potrà scegliere un suo personale accostamento tra timbri e colori. Non vedenti dalla nascita in linea di principio non dovrebbero avere alcuna preferenza verso una specifica associazione, a meno di casi particolari di cui discuteremo più avanti, mentre persone che avevano la vista possono avere ricordi legati ai colori e di conseguenza potranno voler effettuare delle scelte precise nelle associazioni. Si tratta a questo punto di stabilire quanto grande dovrà essere la nostra memoria associativa, ovvero quanti colori avremo a disposizione nella tavolozza, e quanti gradi di libertà poter offrire all'utente finale per personalizzare il sistema.

Un primo vincolo sui possibili colori rappresentabili con questa codifica è posto proprio dai vari timbri che si possono utilizzare: abbiamo già accennato al fatto che il nostro suono di goccia di pioggia non può essere qualsiasi, ma che deve possedere una caratteristica di impulsività, o dualmente una larghezza di banda ben definita; ciò perché tanto più breve è il nostro impulso tanto meglio riesce ad apparire localizzato nello spazio, di conseguenza meglio descrive l'ambiente circostante. Un suono con caratteristiche impulsive ideali dovrebbe avere banda più estesa possibile, fermo restando che l'udito di un soggetto normale va da 20 **Hz** a 20 **KHz**.

Questa prima considerazione esclude in partenza tutti i fiati (legni, ottoni etc. etc.) e gli strumenti ad arco, perché presentano o un timbro troppo

concentrato sulla fondamentale o un attacco troppo lungo; rimangono tutti quelli strumenti le cui note sono prodotte da una parte in movimento, percussivi dunque, siano essi a corda, a pelle o più semplici ancora. Non bisogna poi dimenticare che qualunque strumento suonato da un musicista con un minimo di esperienza è in grado di fornire una grande varietà di sfumature timbriche che dovremo certamente sfruttare per rendere la grande continuità cromatica dei colori. Potremmo pensare, quindi, di partire da una tavolozza relativamente semplice di 16 colori di base, cui poi aggiungere un certo numero di sfumature; ciascuno dei 16 elementi iniziali della nostra tabella associativa può essere sostituito da una sottotabella contenente le sfumature (per questo lavoro iniziale abbiamo scelto delle sottomemorie da 4 elementi).

Un sistema intelligente per disegnare la nostra tavolozza è ricalcare, piuttosto che la codifica Yxy, la codifica di Hunter, in cui i colori sono equidistanti tra loro e sono disposti su di una circonferenza: possiamo così suddividere la circonferenza in 16 settori circolari, uno per ogni timbro, e rappresentare le diverse sottomemorie dividendo a loro volta i singoli settori in sottosettori radiali.

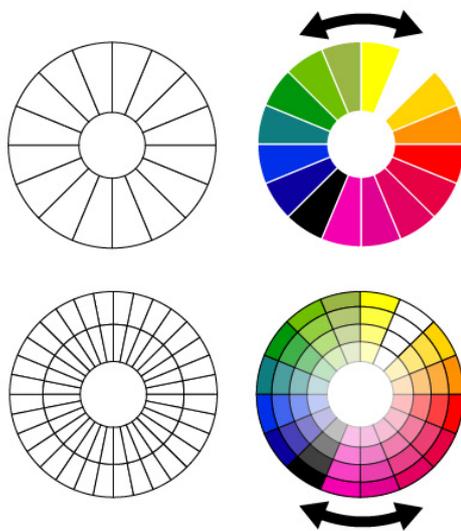


Figura 2.5: Tavolozza associativa a 16 e 64 colori. Le frecce indicano il grado di libertà offerto, ovvero l'associazione iniziale tra un timbro ed un colore.

A questo punto non ci resta che affrontare il problema della personaliz-

zazione: come può l'utente finale interagire con il sistema e modificare la tabella di associazione tra colori e suoni? per risolvere questo problema non si può prescindere dalla considerazione di un fenomeno molto interessante relativo al sistema sensoriale umano: la sinestesia. [?],[?],[?],[?] Normalmente usiamo questa parola per indicare quel fenomeno della lingua per cui vocaboli relativi ad una certa modalità sensoriale vengono utilizzati per descrivere sensazioni prodotte da una modalità sensoriale diversa come quando, ad esempio, si dice 'un suono alto' in cui la modalità uditiva viene descritta attraverso la modalità visiva o un 'colore freddo' in cui la modalità visiva viene descritta attraverso la modalità tattile.

Questo fenomeno, che trova espressione anche nel linguaggio, in realtà ha una matrice nell'integrazione dell'informazione sensoriale. Nel nostro continuo rapporto con il mondo esterno, infatti, i cinque sensi funzionano in parallelo nell'elaborare l'informazione sensoriale: vediamo gli oggetti che tocchiamo, ne sentiamo l'odore ecc. E' questa unitarietà crossmodale dell'elaborazione delle sensazioni che comporta l'instaurarsi di associazioni tra una modalità sensoriale e l'altra: così un timbro ricco di toni alti spesso è associato ad un colore chiaro, mentre uno con toni gravi è associato ad un colore scuro. Queste associazioni, inoltre, sono più vivide in certe persone e solo raramente sono condivise. Attualmente tra i ricercatori non c'è consenso su di una spiegazione univoca di questo fenomeno. [?] Infatti, se per alcuni la sinestesia è solo un fatto socio-culturale, e non biologico, in quanto basata su somiglianze emotive, per altri, invece, è da ricondurre all'integrazione multisensoriale mediata da processi fisiologici flessibili di carattere adattivo e per altri ancora alla cross-attivazione di aree specializzate della corteccia cerebrale.

Ci si deve chiedere, allora, quali implicazioni comporti per il nostro lavoro questo 'fenomeno irriducibile'; la risposta è che dobbiamo lasciare un margine di personalizzazione più ampio o, in altri termini, ricorrere ad una codifica associativa a più vie. Un modo immediato di codifica associativa che presenti comunque un grado di libertà di personalizzazione si potrebbe realizzare così: da un lato abbiamo il cerchio di Hunter dei colori, dall'altro abbiamo i nostri settori circolari con i relativi timbri; lo sfumare dei colori l'uno nell'altro si può rendere in modo efficace scegliendo le posizioni di timbri simili in settori adiacenti, in modo da non creare stacchi bruschi nelle gocce di colori vicini. All'utente rimane la possibilità di scegliere quale sia la associazione di partenza, potendo in pratica far ruotare la corona circolare con i diversi timbri sulla corona dei diversi colori. La presenza della sinestesia, invece, ci obbliga a lasciare libera ogni possibile associazione tra ciascun timbro e

ciascun colore, fermo restando che le diverse sfumature ‘di interpretazione’ possano essere stabilite a priori.

## 2.4 Il Bianco ed il Nero

Abbiamo analizzato in modo accurato la rappresentazione dei colori, senza però soffermarci su alcuni problemi. Poniamo di avere di fronte a noi una parete bianca e di utilizzare la codifica geometrica: ciascuno dei tre timbri avrà volume relativo massimo, ed il volume totale varierà in funzione della distanza. Se al contrario avessimo di fronte una parete nera cosa sentiremmo? semplicemente assenza di gocce, come se fossimo di fronte alla soglia di una stanza buia? ma se fosse un ostacolo di colore nero ad occupare il nostro campo visivo? Non facendo suonare la pioggia in corrispondenza dell’assenza di colore perderemmo una grande quantità di informazione, rischiando di compromettere completamente la funzionalità esplorativa del nostro dispositivo! Questa è una diretta conseguenza della base scelta per la rappresentazione del colore, una base additiva, che, a partire dal nero (colore nullo), addizionando frazioni energetiche diverse per ciascuna delle tre componenti del tristimolo (**RGB**), ricrea tutto lo spettro visibile.

Potremmo scegliere, allora, una base sottrattiva, come il **CMYK** (Ciano, Magenta, Giallo, Nero), che partendo da uno sfondo bianco e sottraendo i diversi colori primari ottenga di nuovo lo spettro visibile nella sua totalità. A questo punto però, è intuibile, il bianco presenterà il medesimo problema del nero del caso precedente: vettori rappresentativi di intensità nulla, quindi assenza di suono. A seconda della base scelta, additiva o sottrattiva, vi sarà sempre un punto, corrispondente al vettore nullo, cui verrà associato il silenzio; è importante aver sempre presente queste considerazioni, soprattutto nella fase di apprendimento all’uso del dispositivo.

Una soluzione pratica immediata per la codifica geometrica sarebbe utilizzare una goccia di diverso timbro per il nero: avremmo una base non ortogonale e quindi una codifica ridondante, però potremmo gestire tutti gli eventi di colore nello stesso modo, senza dover modificare software o hardware. Così, nella codifica associativa dovremo riservare uno dei 16 settori circolari al nero, utilizzando la relativa sottomemoria per rappresentare 4 livelli di grigio.

## 2.5 Lenti

Nel capitolo precedente abbiamo descritto il dispositivo **EAV** ed abbiamo visto come le due microcamere, l'anello iniziale di tutta la nostra catena di acquisizione dati e di elaborazione dell'informazione, siano nascoste dietro le lenti di un paio di occhiali da sole; in questo modo il dispositivo è assolutamente non invasivo sia per gli utenti finali che per le persone con cui questi ultimi si possano relazionare, quindi la soluzione è ottima. Però la presenza di una lente scura, per non pensare ad una lente volutamente colorata, altera il contenuto spettrale della luce che verrà analizzata dai sensori. In questa tesi non affronteremo questo problema e ci limiteremo a descrivere immagini non modificate cromaticamente; un'ulteriore direzione della ricerca, però, potrà essere quella di compensare questa aberrazione cromatica utilizzando tecniche di filtraggio digitale.

## 2.6 Software per la sperimentazione

In questa ultima sezione descriveremo i due programmi sviluppati per iniziare il vero percorso sperimentale di associazione sinestetica: si chiamano **Colorama** e **Georama** e sono rispettivamente l'implementazione pratica della codifica associativa e geometrica. Per semplicità sono stati programmati in Flash, un tool professionale che permette di creare pagine HTML multimediali, rendendone così facile l'uso e la divulgazione.

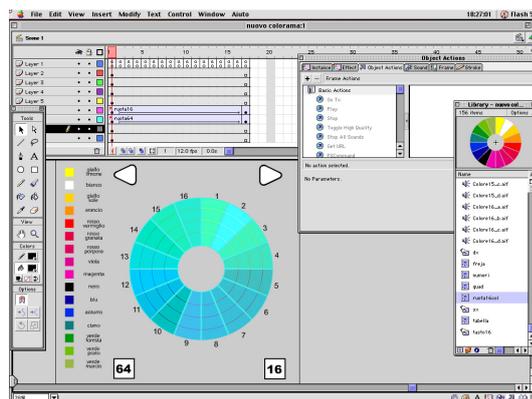


Figura 2.6: Programmando Colorama.

**Colorama** è una corona circolare che riproduce la codifica Hunter con una tavolozza di 16 colori principali; premendo un tasto si possono visualizzare le sottomemorie, ottenendo così 64 settori, che presentano un grado di saturazione via via decrescente. Per simulare l'aumento di confusione cromatica al diminuire della saturazione abbiamo mixato i suoni campione originali con del rumore gaussiano bianco (prefiltrato con un involuppo impulsivo), mantenendo però l'energia dei campioni risultanti pari al loro valore iniziale; l'intento di questa normalizzazione è quello di non fornire una chiave di distinzione tra i diversi gradi di saturazione basata sull'aumento della pressione sonora associata al mixaggio. La discretizzazione dello spazio dei colori è rispecchiata dal timbro dei diversi campioni di strumenti percussivi che, sebbene scelti in base a criteri di somiglianza, rimangono comunque discontinui.

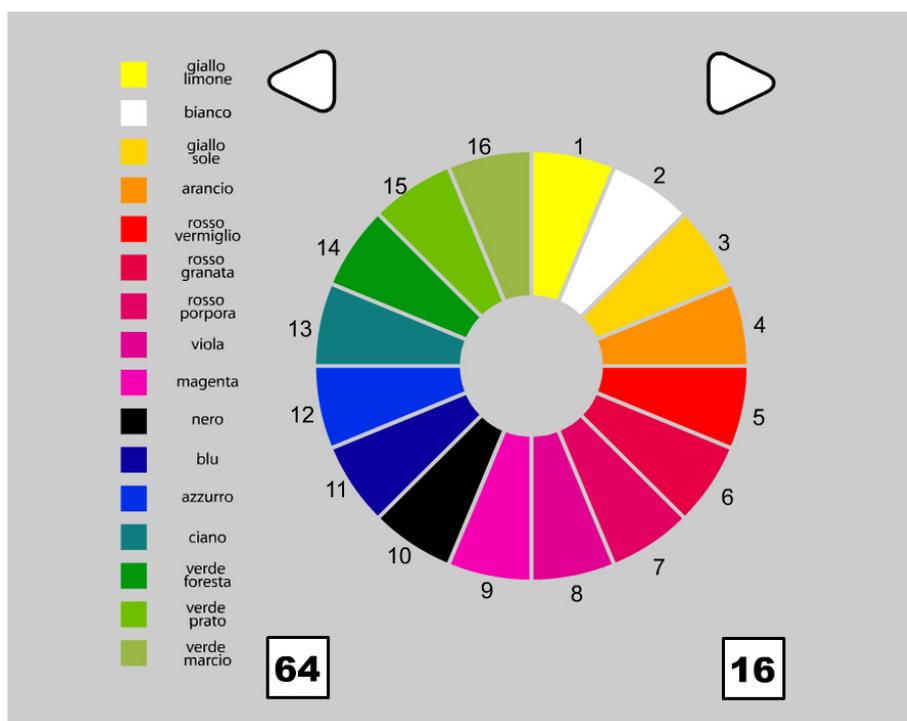


Figura 2.7: Software Colorama (versione a 16 colori).

**Georama** riflette, invece, la codifica geometrica: aprendo l'applicazione si nota una immagine rappresentante il Gamut della codifica RGB, ovvero quel particolare sottoinsieme dei colori codificati con la base RGB realmente visibili con un monitor. Premendo il mouse su un punto della figura i tre timbri sono mixati con volumi differenti, per offrire in tempo reale un timbro che con continuità rappresenti il colore scelto. A questo punto sarebbe molto interessante poter aggiungere questo modulo software all'**EAV** e poter presentare oggetti ed ambienti di diversi colori. I passi logoci sono questi: un primo algoritmo ricostruisce, a partire da una immagine del Gamut RGB, le corrette coordinate RGB e le passa ad una funzione che esegua il mixaggio di tre timbri precedentemente scelti; il file *.wav* così ottenuto viene utilizzato dal software di generazione dell'**EAV** in tempo reale.

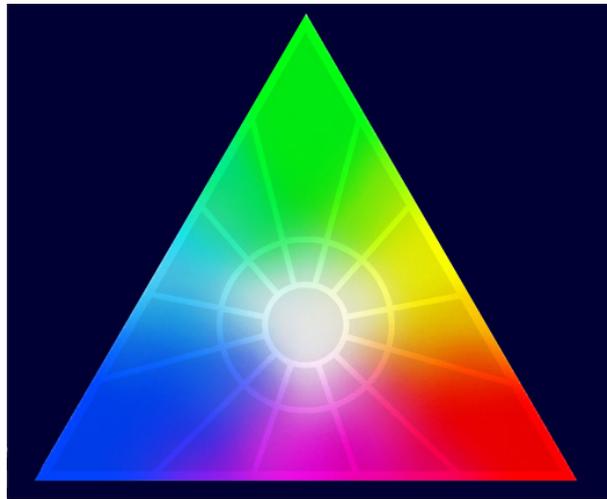


Figura 2.8: Software Georama.

# Capitolo 3

## Sperimentazione

### Ipotesi della ricerca

In vista di un successivo sviluppo della ricerca volto alla realizzazione di un dispositivo portatile di **pioggia virtuale**, è parso opportuno confrontare con uno studio pilota la validità ergonomica dal punto di vista cognitivo delle due codifiche sinestetiche precedentemente presentate ed implementate: **Colorama**, basata sulla codifica associativa, e **Georama**, basata sulla codifica geometrica. E' così possibile stabilire quale delle due codifiche costituisca lo strumento maggiormente 'user friendly' da utilizzare nella futura realizzazione del dispositivo. L'ipotesi alla base dell'esperimento è che le due codifiche, a parità di quantità di apprendimento svolto, producano nei partecipanti prestazioni diverse, ovvero che differiscano nell'elaborazione cognitiva richiesta.

### 3.1 Metodo

#### 3.1.1 Partecipanti

Hanno preso parte all'esperimento 20 partecipanti di cui 10 vedenti con età media di 27 anni e 10 non vedenti con età media di 45 anni. L'opportunità di condurre l'esperimento con partecipanti vedenti e non vedenti è stata suggerita.

ta dalla necessità di controllare se la conoscenza della continuità cromatica propria ai vedenti possa influenzare la prestazione. Analogamente è stata controllata e bilanciata tra i gruppi la competenza musicale dei partecipanti.

### 3.1.2 Materiali

Gli stimoli da presentare ai partecipanti erano sequenze di suoni costruite utilizzando i 2 software **Colorama** e **Georama**. Per ciascuno di essi sono stati costruiti 2 tipi di sequenze la prima delle quali si ipotizza essere più ‘facile’ della seconda, più ‘difficile’, in quanto basata su un minor numero di alternative possibili<sup>1</sup>. In ciascuna sequenza ogni suono è costituito da un campione audio registrato in mono con risoluzione dinamica a 16 bit e frequenza di campionamento pari a 44,1 KHz. La durata dei campioni selezionati è stata normalizzata a 500 ms, così come è stato normalizzato il massimo livello di pressione sonora ad essi associato.

Con il software **Colorama** le sequenze ‘facili’ sono state costruite a partire dai 16 colori di base da cui sono state scelte 4 sequenze di 8 colori contigui; le 4 sequenze sperimentali ‘difficili’ sono state costruite a partire dai 64 colori tra cui sono stati scelti 4 colori casuali che presentassero tutti e 4 i gradi di desaturazione. Per simulare l’aumento di confusione cromatica al diminuire della saturazione abbiamo mixato i suoni campione originali con del rumore gaussiano bianco (prefiltrato con un involuppo impulsivo), mantenendo l’energia delle sequenze risultanti pari al loro valore iniziale. Obiettivo di questa normalizzazione era evitare di fornire una chiave di distinzione tra i diversi gradi di saturazione basata sull’aumento della pressione sonora associata al mixaggio. Ogni tipo di sequenza sonora così ottenuta è stata denominata con una lettera dell’alfabeto in funzione del suo grado di desaturazione: ‘a’ indicava una desaturazione nulla, ‘b’ una desaturazione al 25%, ‘c’ al 50% e ‘d’ al 75%.

Per quanto riguarda la costruzione delle sequenze sperimentali con il software **Georama**, invece, siamo partiti dai tre timbri associati al rosso, verde e blu ed abbiamo discretizzato il triangolo del Gamut RGB in 25 celle. In ogni cella ognuno dei tre timbri ha un volume (o più correttamente viene riprodotto con un livello di pressione sonora) pari all’intensità del colore ad esso associato. In questo modo le sequenze-stimolo derivate dalle celle esterne sono sempre il risultato del mescolamento di due soli timbri (muovendosi dal-

---

<sup>1</sup>Vedi Appendici A e B

la cella 1 alla 5, infatti, si passa dal timbro associato al verde puro a quello associato al rosso puro, e il blu ha sempre valore nullo).

Il bianco è stato ottenuto mixando i tre timbri con volumi identici. Inoltre, per evitare che alcune celle suonassero più intensamente di altre, i valori ottenuti sono stati tutti normalizzati in energia. Una volta ottenuto il bianco si è costruita una corona interna mixando i suoni della corona esterna con il bianco. Ognuno dei suoni della corona interna, così, è sempre riconoscibile da quelli della corona esterna perché presenta sempre tutti e tre i timbri. Ragionando per sottrazione si potrebbe riconoscere facilmente il giallo chiaro (corona interna) dal giallo scuro (corona esterna) proprio perché presenta pochissimo blu.

Le sequenze sperimentali ‘facili’ della versione a 12 colori di Georama presentavano i 12 suoni della corona esterna (mixaggi di due timbri alla volta); quelle ‘difficili’ della versione a 26 colori sono state ottenute scegliendo i 15 stimoli compresi tra due colori primari ed il bianco. Anche in questo caso ogni tipo di sequenza sonora così ottenuta è stata denominata con una lettera dell’alfabeto: ‘a’ indica uno stimolo appartenente alla corona esterna, ‘b’ indica uno stimolo della corona interna ‘c’ è il bianco; per la sua natura particolare (nella codifica RGB dei colori è infatti un concetto limite, prodotto dai valori massimi di ciascun valore del tristimolo) si è preferito usare una lettera da sola, non accompagnata da un numero. E’ in pratica una terza corona circolare che degenera in un elemento solo.

In totale le sequenze stimolo utilizzate nell’esperimento erano:

- Per Colorama:
  1. sequenze di 8 stimoli consecutivi, presentati in modo casuale (sequenze ‘facili’)
  2. sequenze di 4 stimoli non consecutivi con tutti e 4 i gradi di desaturazione, presentati in modo casuale (sequenze ‘difficili’)
- Per Georama:
  1. sequenze di 12 stimoli consecutivi, presentati in modo casuale (sequenze ‘facili’)
  2. sequenze di 15 stimoli compresi tra due colori primari ed il bianco, presentati in modo casuale (sequenze ‘difficili’)

### 3.1.3 Procedura

Ogni partecipante indossava cuffie di tipo aperto (Beyerdynamics 990 Pro) con risposta in frequenza lineare tra 5 e 35 KHz e livello di sensibilità di 96 dB. Prima della fase sperimentale, ogni partecipante veniva sottoposto ad una fase di apprendimento.

L'apprendimento con **Colorama** nella versione a 16 colori prevedeva l'ascolto per 4 volte della sequenza di 8 stimoli contigui numerati, prima in ordine crescente poi in ordine decrescente. Si procedeva poi con un ascolto di 4 sequenze degli stessi 8 stimoli in ordine casuale. In entrambe queste fasi insieme a ciascuno degli stimoli veniva comunicato il loro numero identificativo. Successivamente si effettuava un 'riconoscimento guidato' di una sequenza casuale degli 8 stimoli. Si proponevano gli stimoli al partecipante e si chiedeva di fornire i numeri identificativi, che non gli erano stati comunicati. Ad ogni risposta del partecipante si comunicava l'esito del riconoscimento; se era positivo, si procedeva alla presentazione dello stimolo successivo, se era negativo, si annotava la risposta errata in modo da creare una lista dei tentativi effettuati per riconoscere gli stimoli presentati; al partecipante si ripresentava quindi lo stimolo chiedendo una nuova risposta. In questo modo si creavano delle liste di risposte (frutto di tentativi e correzioni) associate ad ogni stimolo della sequenza; l'analisi di queste liste permette di evidenziare timbri difficilmente memorizzabili.

Il test con **Colorama** nella versione a 16 colori prevede l'ascolto di 4 sequenze in ordine casuale degli 8 stimoli riportando le risposte del partecipante si possono misurare le distanze con cui questi riconosce i diversi suoni utilizzando una norma Manhattan: se venisse proposto lo stimolo numero 2, ad esempio, ed il partecipante fornisse la risposta numero 5 allora la sua distanza di riconoscimento sarebbe pari a 3.

L'apprendimento con **Colorama** nella versione a 64 colori prevedeva la presentazione di 4 sequenze di 4 stimoli non contigui identificati da un numero e da una lettera, prima ordinati per saturazione decrescente poi per saturazione crescente e, poi, due volte in ordine casuale. Senza effettuare il riconoscimento guidato si passava direttamente alla fase sperimentale che consisteva nella presentazione in ordine casuale di 4 sequenze di 16 stimoli ciascuna in cui erano presenti tutti e 4 i gradi di desaturazione. In questo caso la norma Manhattan si specializza così: la distanza angolare, associata al numero degli stimoli proposti, resta invariata, mentre quella radiale si calcola attribuendo una distanza pari a 0,25 tra ciascun grado di desaturazione.

Se venisse proposto lo stimolo 4a, ad esempio, ed il partecipante fornisse la risposta 4c allora la sua distanza di riconoscimento sarebbe pari a 0,5.

L'apprendimento con **Georama** nella versione a 12 colori prevedeva l'ascolto per tre volte della sequenza di 12 stimoli contigui numerati, prima in ordine crescente poi in ordine decrescente. Si procedeva poi alla presentazione di 3 sequenze degli stessi 12 stimoli in ordine casuale. Successivamente si effettuava il riconoscimento guidato di un'ulteriore sequenza casuale dei 12 stimoli. La fase sperimentale prevedeva l'ascolto di 3 sequenze in ordine casuale dei 12 stimoli. Anche in questo caso la misura raccolta era la distanza tra stimoli proposti e risposte fornite dal partecipante.

L'apprendimento con Georama nella versione a 26 colori prevedeva l'ascolto di 2 sequenze di 15 stimoli identificati da un numero e da una lettera, prima ordinati a partire dagli elementi della corona esterna fino al bianco (corona interna degenera) e poi in senso inverso. Anche in questo caso non si è proceduto al riconoscimento guidato, ma si è passati direttamente alla fase sperimentale che consisteva nella presentazione di 2 sequenze di 15 stimoli presentati in ordine casuale. Di nuovo la misura registrata era la distanza tra stimoli proposti e risposte fornite dal partecipante.

Le sequenze di apprendimento, sia per Colorama che per Georama, sono state assegnate ai diversi partecipanti in modo bilanciato.

### 3.1.4 Analisi dei dati e discussione dei risultati

Per poter essere analizzate, le misure ottenute con le due diverse codifiche, **Colorama** e **Georama**, sono state normalizzate rispetto alla distanza massima misurabile su ciascuna sequenza; sui dati così ottenuti è stata eseguita una ANOVA  $2 \times 2 \times 2$  le cui variabili indipendenti erano il fattore vedenti-non vedenti tra gruppi, il fattore tipo di codifica e il fattore difficoltà entrambi 'nested'. Possiamo, infatti, rappresentare le condizioni sperimentali in una struttura ad albero che evidenzia in modo chiaro come i fattori tipo di codifica (**Colorama** e **Georama**) e facilità (rispettivamente 16 o 64 colori e 12 o 26 colori) siano autoinclusi.

L'analisi ha mostrato che sia il fattore tipo di codifica che il fattore facilità sono significativi (rispettivamente  $F(2, 72) = 6,879$ ,  $Mse = 0,507$ ,  $p < 0,0018$ ;  $F(4, 72) = 5,492$ ,  $Mse = 0,507$ ,  $p < 0,0006$ ), mentre non lo è il fattore vedenti-non vedenti. Da tale analisi si può concludere che la diversa

conoscenza della continuità cromatica dei vedenti rispetto a quella dei non vedenti non ha influito sui risultati. Ciò che, invece, ha influito è sia la condizione di facilità relativa al numero dei colori sia quella relativa ai due tipi di codifica.

Condizione facile	0,860
Condizione difficile	0,997

Tabella 3.1: Medie normalizzate delle distanze tra stimoli proposti e risposte dei partecipanti nelle due condizioni ‘facile’ e ‘difficile’

Come si vede dalla tabella 3.1, infatti, anche indipendentemente dal tipo di codifica, la prestazione dei partecipanti è stata migliore nella condizione ‘facile’ (rispettivamente con 16 e 12 colori) che nella condizione ‘difficile’ (rispettivamente con 64 e 26 colori).

Codifica con Colorama	1,028
Codifica con Georama	0,830

Tabella 3.2: Medie normalizzate delle distanze tra stimoli proposti e risposte dei partecipanti nelle due codifiche

Inoltre, come si vede nella tabella 3.2, la prestazione dei partecipanti è stata migliore con la codifica realizzata con il software **Georama** che con la codifica realizzata con il software **Colorama**.

	Cond. Facile	Cond. Difficile
Codifica con Colorama	0,670	1,387
Codifica con Georama	1,051	0,608

Tabella 3.3: Medie normalizzate delle distanze tra stimoli proposti e risposte dei partecipanti nelle due condizioni di facilità nei due tipi di codifica

Un risultato particolarmente interessante emerge dalla tabella 3.3. La codifica associativa di **Colorama** presenta, nella versione a 16 colori, una distanza media di riconoscimento nettamente inferiore, quasi il 35%, a quella della codifica geometrica di **Georama**. Ciò pare indicare che, data una tavolozza di colori netti l’associazione tra un colore e un timbro è più immediata con la codifica con **Colorama**. Tuttavia, se si aumenta la risoluzione della codifica, aumentando così la sensibilità del sistema nelle associazioni tra

colori e timbri la codifica associativa, con **Colorama**, determina un peggioramento nella prestazione dei partecipanti che non percepiscono il cambiamento nella risoluzione in modo così chiaro come nella codifica geometrica; infatti la distanza media aumenta sensibilmente quando, invece, dovrebbe diminuire data la distanza reale minore degli stimoli presentati.

La codifica geometrica, invece, seppur più difficile da interpretare nella versione base con 12 colori, mantenendo la medesima logica nella versione a più alta risoluzione, dà luogo a prestazioni migliori. La distanza media di riconoscimento durante la prova migliora in circa il 50% dei partecipanti.

## 3.2 Valutazioni e possibili sviluppi futuri

Inizialmente abbiamo proposto una dicotomia tra una codifica associativa che non richiede elaborazione in tempo reale e una codifica geometrica esatta dall'altro. I risultati dello studio pilota che abbiamo condotto ci consentono di concludere che, almeno da un punto di vista cognitivo, la codifica geometrica di **Georama** produce prestazioni migliori soprattutto ad alti livelli di risoluzione. Tale codifica, infatti, ci consente, di presentare una grande continuità cromatica a chi usufruisca del dispositivo di pioggia virtuale senza che l'acquisizione della capacità di distinguere molti colori simili comporti un carico cognitivo eccessivo neppure nella fase del primo apprendimento.

Dopo questo studio, il problema della necessità di calcolo in tempo reale può essere affrontato con la filosofia propria della codifica associativa: la discretizzazione. Appurato, infatti, che anche la codifica geometrica può essere mappata con una certa risoluzione, un immediato sviluppo di questa tesi è trovare la mappatura ottima del Gamut del RGB in modo tale da sfruttare al massimo la sensibilità dell'orecchio umano. Inoltre si potrà procedere alla personalizzazione della codifica cosicché i singoli utenti finali possano provare e scegliere i timbri da associare al Rosso, Verde e Blu e ciò permetterà loro di affinare le proprie capacità di valutazione.

Un'ulteriore prospettiva da esplorare sarà il confronto tra le prestazioni offerte da questa codifica sinestetica tra colori e suoni e altre possibili codifiche sinestetiche basate su gli altri sensi. Dal punto di vista ingegneristico, le applicazioni possono essere molteplici, da quelle in campo medico fino a quelle ludiche. Così, se potessimo associare univocamente un colore ad un elemento di un file musicale, ad esempio, potremmo utilizzare tecniche pro-

prie del riconoscimento di immagini per indicizzare ed esplorare database di musica. Allo stesso modo se, invece di passare da colori a suoni, eseguiamo il passaggio inverso potremmo pensare, di associare al Rosso, Verde e Blu tre ultrasuoni per eseguire ecografie a colori sfruttando il diverso scattering delle tre frequenze. Sempre restando nell'ambito di dispositivi di ausilio per persone non vedenti, potremmo pensare anche di crearne uno in grado di aiutare nella scelta dei colori dei vestiti da indossare, sfruttando l'eufonicit  dei timbri per riprodurre la gradevolezza degli accostamenti cromatici.

## Conclusioni

Abbiamo elaborato due proposte di resa acustica del colore come integrazione del sistema **EAV** per videolesi. **Colorama** associa ad una tavolozza di colori un insieme di timbri percussivi distinti; **Georama** combina con continuità tre timbri associati univocamente ai tre colori primari, Rosso, Verde e Blu. Sono stati condotti esperimenti su volontari vedenti e non vedenti che hanno messo in evidenza come la codifica **Georama** ottenga migliori prestazioni soprattutto dal punto di vista della facilità d'uso. Resta da affrontare il confronto con ogni eventuale altro approccio sinestetico al problema e debbono essere esplorate possibili applicazioni in altri ambiti.



# Appendice A

## Materiali

### A.1 Sequenze di apprendimento e test

Di seguito riportiamo esempi delle sequenze di apprendimento per **Colorama** nella versione ‘facile’, cioè a 16 colori. Gli stimoli hanno sempre un codice identificativo formato da numeri.

APP. 1		APP. 1B				APP. 1C		
[sequenza ordinata]		[ordine random, 1 volta ciascuna]				[ordine random, feedback]		
Numero		Numero	Numero	Numero	Numero	Numero	Liste associate	
1		2	1	5	7	3		
2		3	3	6	4	1		
3		1	2	3	8	2		
4		4	5	4	5	5		
5		7	8	1	6	8		
6		5	7	7	2	7		
7		8	4	2	1	4		
8		6	6	8	3	6		
APP. 2		APP. 2B				APP. 2C		
[sequenza ordinata]		[ordine random, 1 volta ciascuna]				[ordine random, feedback]		
Numero		Numero	Numero	Numero	Numero	Numero	Liste associate	
3		9	10	5	7	5		
4		3	3	6	4	4		
5		10	9	3	8	8		
6		4	5	4	5	3		
7		7	8	10	6	9		
8		5	7	7	9	10		
9		8	4	9	10	6		
10		6	6	8	3	7		



Questa è una sequenza di apprendimento per **Colorama** nella versione 'difficile', cioè a 64 colori. Gli stimoli hanno sempre un codice identificativo formato da numeri e lettere.

APP. N 1 [Saturazione crescente, decrescente, 2 volte random]				
Numero		Numero		Numero
1_a		1_d		1_c
1_b		1_c		1_a
1_c		1_b		1_d
1_d		1_a		1_b
5_a		5_d		5_b
5_b		5_c		5_d
5_c		5_b		5_a
5_d		5_a		5_c
11_a		11_d		11_b
11_b		11_c		11_c
11_c		11_b		11_a
11_d		11_a		11_d
16_a		16_d		16_a
16_b		16_c		16_d
16_c		16_b		16_c
16_d		16_a		16_b

Questo è il test relativo alla sequenza di apprendimento 1.

TEST N 1		[scrivere nelle caselle le risposte]							
Numero	Risposte	Numero	Risposte	Numero	Risposte	Numero	Risposte	Numero	Risposte
1_d		11_b		1_c		5_a			
11_a		5_c		11_a		16_a			
5_b		16_c		5_a		1_a			
1_a		1_a		11_d		16_d			
11_b		1_c		1_a		5_b			
16_a		5_b		5_d		1_c			
5_a		16_d		16_a		5_d			
11_c		11_a		5_b		16_b			
1_b		16_a		11_b		1_b			
11_d		5_a		16_b		11_a			
5_d		1_b		1_d		1_d			
16_d		11_d		16_c		5_c			
16_c		16_b		1_b		11_b			
1_c		5_d		11_c		11_d			
16_b		1_d		5_c		16_c			
5_c		11_c		16_d		11_c			





Questa è una sequenza di apprendimento per **Georama** nella versione 'difficile', cioè a 26 colori. Gli stimoli hanno sempre un codice identificativo formato da numeri e lettere.

APP. N 1	[due sequenze ordinate, due casuali]			
Numero	Numero	Numero	Numero	Numero
1_a	c	4_b	c	
1_b	5_b	3_b	5_a	
c	5_a	c	4_b	
2_a	c	2_a	2_a	
2_b	4_b	4_a	c	
c	4_a	c	4_a	
3_a	c	3_a	c	
3_b	3_b	5_a	5_b	
c	3_a	c	3_a	
4_a	c	2_b	1_b	
4_b	2_b	5_b	2_b	
c	2_a	c	c	
5_a	c	1_b	3_b	
5_b	1_b	1_a	1_a	
c	1_a	c	c	

Questo è il test relativo alla sequenza di apprendimento 1.

TEST N 1		[scrivere nelle caselle le risposte]	
Numero	Risposte	Numero	Risposte
c		2_b	
2_b		5_a	
1_b		3_b	
c		c	
3_a		1_a	
2_a		5_b	
1_a		c	
4_b		4_b	
c		3_a	
5_a		c	
3_b		2_a	
c		c	
4_a		4_a	
c		1_b	
5_b		c	

# Elenco delle figure

1.1	Esempio bidimensionale del comportamento del sistema <b>EAV</b> .	8
1.2	A. Rappresentazione schematica del laboratorio. B. Disegno fatto da un non vedente dopo una esplorazione molto breve, utilizzando un prototipo di <b>EAV</b> , senza uso del tatto. . . . .	9
2.1	Valori del tristimolo in funzione della lunghezza d'onda. . . . .	17
2.2	Diagramma di cromaticità CIE1931. . . . .	18
2.3	Diagramma Hunter-L*A*B 1976. . . . .	19
2.4	Gamut del RGB e mappa discreta dei volumi relativi dei timbri associati ai colori primari. . . . .	19
2.5	Tavolozza associativa a 16 e 64 colori. Le frecce indicano il grado di libertà offerto, ovvero l'associazione iniziale tra un timbro ed un colore. . . . .	21
2.6	Programmando Colorama. . . . .	24
2.7	Software Colorama (versione a 16 colori). . . . .	25
2.8	Software Georama. . . . .	26



## Elenco delle tabelle

3.1	Medie normalizzate delle distanze tra stimoli proposti e risposte dei partecipanti nelle due condizioni ‘facile’ e ‘difficile’ . . . . .	32
3.2	Medie normalizzate delle distanze tra stimoli proposti e risposte dei partecipanti nelle due codifiche . . . . .	32
3.3	Medie normalizzate delle distanze tra stimoli proposti e risposte dei partecipanti nelle due condizioni di facilità nei due tipi di codifica . . . . .	32