

CAPITOLO 1 - Le immagini e la rappresentazione del loro contenuto.

In prima istanza possiamo definire un'immagine come “una metodica di rappresentazione secondo coordinate spaziali indipendenti di un oggetto o di una scena. Un'immagine contiene informazioni descrittive riferite all'oggetto, alla scena che rappresenta: un'immagine è quindi una distribuzione, che può essere bi o tri-dimensionale, di un'entità fisica”. Questa prima definizione è di carattere enciclopedico e quindi comprensibile a chiunque. Per gli scopi di questa tesi non è però sufficiente limitarsi a fornire questo unico concetto, poiché il nostro interesse è rivolto alla rappresentazione delle immagini in un computer, mediante la descrizione del loro contenuto, ed al loro confronto. Dobbiamo quindi approfondire quali sono i concetti che permettono la riproduzione di un'immagine in un dispositivo digitale e, successivamente, analizzare in che modo può essere effettuato un confronto tra le immagini stesse.

In questo capitolo verranno affrontati a tal fine i seguenti argomenti:

- 1.1 - La rappresentazione digitale di un'immagine.
- 1.2 - Il colore e la definizione dei modelli di colore.
- 1.3 - I modelli di colore RGB e HSI.
- 1.4 - La rappresentazione del contenuto di un'immagine: l'istogramma di colore e la matrice di similarità.
- 1.5 - La mappatura di un'immagine ed i colori della mappa.

1.1 - La rappresentazione digitale di un'immagine.

Per poter rappresentare un'immagine in un dispositivo digitale, l'immagine deve essere sottoposta ad un processo di discretizzazione nello spazio, detto campionamento, e nel colore, detto quantizzazione. Il campionamento è il processo che ha come obbiettivo quello di dividere l'asse orizzontale (larghezza immagine) e

quello verticale (altezza immagine) in elementi discreti detti pixel. La quantizzazione, invece, è il processo che associa ad un pixel un valore che codifica il colore dell'immagine in quel determinato punto spaziale. La gamma dei colori possibili è stabilita dal numero di bit con i quali viene effettuata l'operazione di quantizzazione: con 8 bit avremo 256 colori, con 24 bit ne avremo invece ben oltre 16 milioni! Quando la gamma dei colori possibili, determinati come appena visto dal numero di bit, è inferiore al numero di colori presenti nell'immagine originale, il processo di quantizzazione introduce una certa perdita di informazione. Anche il campionamento può essere più o meno accurato: maggiore è il numero di pixel in cui viene suddivisa l'immagine e maggiore sarà la risoluzione alla quale l'immagine originale verrà rappresentata. Ovviamente, il prezzo da pagare, per avere un'immagine discretizzata che sia il più possibile uguale a quella originale, risiede nella maggiore quantità di informazioni, espressa in numero di bit, che devono essere memorizzate per l'archiviazione dell'immagine in un dispositivo digitale. Siamo quindi nella seguente situazione: abbiamo un'immagine, caratterizzata da un certo numero di pixel, dato dal prodotto dei pixel che costituiscono la larghezza e l'altezza dell'immagine, dove ad ogni pixel, identificato dalle coordinate spaziali X e Y, è associato un valore espresso tramite un prestabilito numero di bit, che codifica il colore del pixel stesso. Il problema successivo, quindi, è quello di capire come effettuare la codifica di questi valori che rappresentano il colore. A questo scopo è necessario introdurre il concetto di colore, per poter procedere successivamente all'introduzione dei modelli di colore.

1.2 - Il colore e la definizione dei modelli di colore.

Il colore è l'effetto percepito quando un'onda elettromagnetica colpisce alcune strutture del nostro sistema visivo poiché, di fatto, la luce è un'onda elettromagnetica. Un essere umano è in grado di percepire la luce quando l'onda elettromagnetica della luce stessa ha una frequenza compresa tra 384THz e 857THz. Solitamente la luce che vediamo tutti i giorni non corrisponde mai ad

un'unica onda magnetica ad una certa frequenza, ma è una combinazione di più onde a frequenze diverse. La rappresentazione dei contributi delle varie componenti delle onde a diversa frequenza viene detto “spettro”.

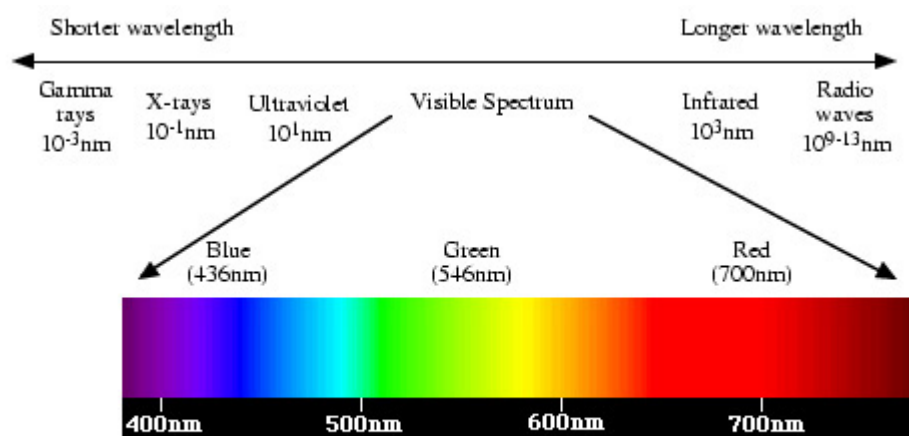


Figura 1.1 - Spettro del visibile per l'occhio umano

La luce bianca pura è il risultato della somma di tutte le componenti frequenziali presenti nello spettro del visibile, del quale abbiamo già indicato la frequenza inferiore e superiore che ne determina il range. Invece, un colore corrispondente ad una radiazione elettromagnetica ad una sola frequenza, viene detto colore puro.

Adesso è importante descrivere con quali meccanismi l'essere umano percepisce il colore. Vedremo a breve che esiste una certa relazione tra questi meccanismi e le modalità con le quali sono concepiti alcuni dei modelli di colore, ovvero le codifiche utilizzate per tradurre i colori in dati numerici.

La percezione visiva nasce grazie allo stimolo luminoso che, attraverso la cornea ed il cristallino, giunge alla retina, dove sono localizzati i fotorecettori. I fotorecettori sono di due tipi: i bastoncelli e i coni. I bastoncelli provvedono alla visione scotopica acromatica, fornendo informazioni sulla sola intensità luminosa, non sul colore, ed in presenza di poca luce. In presenza di forte intensità luminosa il funzionamento dei bastoncelli è inibito. I coni sono invece responsabili della visione fotopica cromatica, fornendo informazioni in presenza di maggiore intensità luminosa. Nella retina esistono tre tipi di coni, ognuno dei quali è sensibile ad una

diversa lunghezza d'onda, approssimativamente associabili alle frequenze del rosso (Red), verde (Green) e blu (Blue).

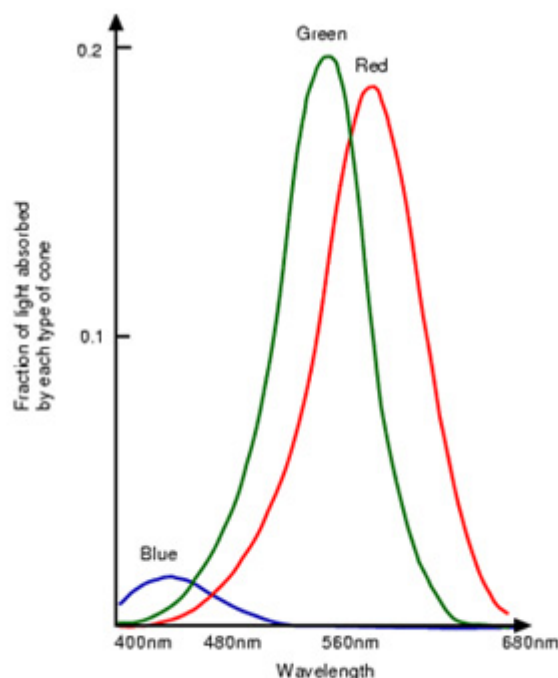


Figura 1.2 - Risposta spettrale dei tre coni

Tuttavia, questo non deve far pensare che i coni trasferiscano al cervello separatamente le intensità alle varie lunghezze d'onda. Infatti, una sorgente monocromatica appare all'occhio dotata di un colore che può in genere ottenersi in modo perfetto combinando le emissioni di tre altre sorgenti monocromatiche sufficientemente spaziate nello spettro, ciascuna di opportuna intensità. Questa proprietà è nota come “natura tricromatica della visione” ed essa sarà concettualmente molto importante in seguito, poiché, come già accennato, questo modo di vedere dell'occhio umano è una delle caratteristiche basilari per la costruzione di alcuni modelli di colore.

Siamo quindi arrivati nella situazione di aver caratterizzato il colore; inoltre, abbiamo appena appreso come l'essere umano, tramite l'occhio, lo percepisce e come il suo apparato visivo trasmette al cervello le informazioni relative ad esso. Pertanto possiamo riassumere dicendo che la percezione cromatica dipende sia

dall'intensità delle componenti di luce alle diverse frequenze, che dalla sensibilità dei tre tipi di coni .

Adesso dobbiamo distinguere quali sono le diverse modalità con cui può essere fatta una sintesi cromatica, fatto ciò avremo in mano tutti gli strumenti per iniziare a parlare dei modelli di colore nei termini necessari per lo sviluppo di questo lavoro di tesi. Le possibilità sono due e ciò dipende dal fatto se l'oggetto è una sorgente di luce oppure è illuminato da una sorgente di luce esterna. Nel primo caso la luce percepita è quella irradiata dall'oggetto. Nel secondo caso, la luce percepita, è invece quella riflessa dall'oggetto. Il colore che noi attribuiamo all'oggetto, di fatto, corrisponde al colore della luce che lo illumina a cui si sottrae il colore assorbito dall'oggetto stesso. In pratica, se si illumina un oggetto con una luce pura (con tutte le componenti spettrali) e lo si vede rosso, significa che tale oggetto assorbe tutte le componenti della luce bianca eccetto quelle associate al rosso.

Un modello di colore può essere quindi costruito cercando di seguire dei criteri che tengano conto della “natura tricromatica della visione” e, allo stesso tempo, rispettino le modalità con le quali può essere effettuata la sintesi cromatica del colore. Diretta conseguenza di tutto ciò, sono i modelli di colore additivi ed i modelli sottrattivi. Nei primi, un colore è definito dalla somma di altri colori, mentre nei secondi, dalla differenza. La definizione di un colore come somma o sottrazione dei diversi contributi forniti da un insieme prestabilito di colori, benché sussista il parallelismo descritto con il sistema visivo umano, non è però il metodo più naturale per descrivere un colore. Nel prossimo paragrafo saranno introdotti due modelli di colore, utili ai fini di questa tesi: il modello RGB ed il modello HSI. Il primo è un modello additivo, caratterizzato secondo i criteri appena descritti, mentre il secondo descrive il colore in un linguaggio che possiamo definire “naturale”, del quale parleremo a breve.

1.3 - I modelli di colore RGB e HSI

Come logica conseguenza di quanto appurato fino a questo momento, possiamo

definire in maniera precisa un modello di colore (o spazio di colore o sistema di colore) come uno strumento con il quale si può specificare, creare e visualizzare un colore. In sostanza, un modello di colore è una specifica di un sistema di coordinate e di un sottospazio al suo interno, dove ogni colore è rappresentato da un singolo punto.

I modelli di colore principalmente in uso oggi, oltre alle distinzioni fatte nel paragrafo precedente, possono essere catalogati anche in base al loro orientamento di utilizzo:

- Hardware di acquisizione e di restituzione (RGB, CMY, CMYK).
- Elaborazione ed analisi di immagini (RGB, HSI, HSV).
- Trasmissione (YUV (PAL), YIQ (NTSC)).

Come abbiamo già detto i modelli di colore di interesse sono RGB e HSI. Il perché sarà immediatamente chiarito nei paragrafi successivi, quando parleremo della descrizione del contenuto di un'immagine ed in particolare della sua mappatura. In questo paragrafo, quindi, ci limiteremo ad introdurre brevemente le caratteristiche di questi due modelli di colore, con il fine di dare una base di conoscenze sufficienti per comprendere i motivi delle scelte che verranno fatte in seguito.

Nel modello RGB (Red-Green-Blue), ogni colore è rappresentato dalla somma delle tre componenti spettrali relative al rosso, verde e blu. Il modello è basato su un sistema di coordinate cartesiane, mentre il sottospazio di interesse è un cubo di spigolo unitario sui cui vertici sono disposti i colori primari, quelli secondari, il bianco (0,0,0) ed il nero (1,1,1). I livelli di grigio sono disposti sul segmento che congiunge (0,0,0) con (1,1,1). Vedere in proposito la figura 1.3.

Un colore rappresentato nel modello RGB è perciò dato dalla somma di ogni singolo contributo fornito da ognuno dei tre distinti canali (modello additivo) e, come già evidenziato, è principalmente utilizzato per la descrizione dei colori nei dispositivi hardware: il monitor di un pc utilizza la rappresentazione RGB.

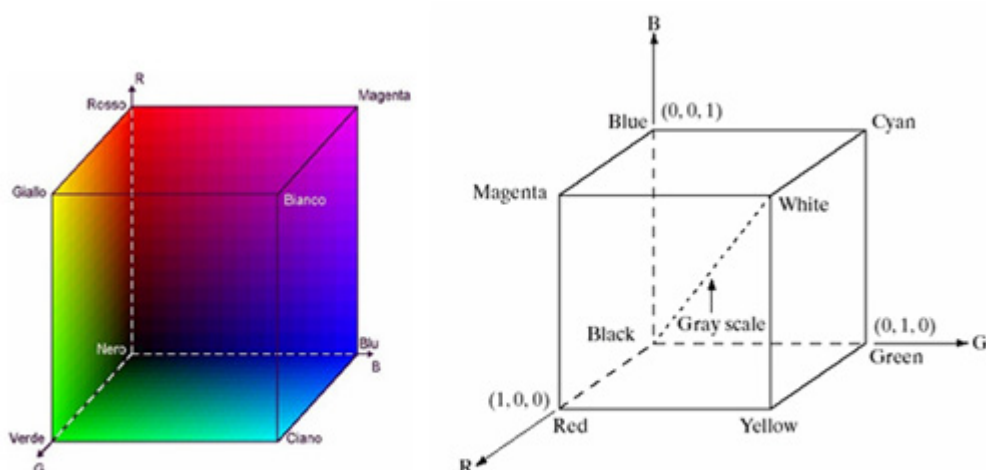


Figura 1.3 - Lo spazio di colore RGB

Abbiamo visto che questo modello è costruito cercando di seguire fedelmente alcune delle modalità tramite le quali l'apparato visivo umano acquisisce informazioni riguardo un colore. Questo però non significa che il modello RGB sia quello in cui i colori vengono descritti nella maniera più naturale. Nel modello RGB, ad esempio, rispetto alla luminosità, una caratteristica naturale tramite la quale caratterizzare un colore, le singole componenti risultano essere correlate, perciò la descrizione del colore, quando questo è rappresentato tramite l'utilizzo del modello RGB, si discosta evidentemente da quella naturale, in quanto, dal valore delle tre distinte componenti di rosso, verde e blu non è possibile ricavare alcuna informazione sulla luminosità. Esistono perciò dei modelli di colore basati sulle caratteristiche naturali con cui usualmente viene definito un colore: tinta (Hue), saturazione (Saturation) e luminosità o intensità (Brightness or Intensity). Diamo adesso una descrizione di cosa effettivamente rappresentano queste grandezze caratteristiche:

- Tinta: attributo legato alla lunghezza d'onda dominante. Rappresenta ciò che un osservatore definisce “il colore dominante”.
- Saturazione: attributo che si riferisce alla purezza della tinta; si ha una bassa saturazione quando nel colore è presente un'elevata quantità di luce bianca mescolata alla tinta.

- Luminosità (intensità): attributo che si riferisce alla quantità di luce presente.

L'insieme della tinta e della saturazione definiscono la cromaticità, ovvero, ciò che caratterizza il colore indipendentemente dall'intensità luminosa presente. Uno dei modelli di colore fondati sulla definizione di tre coordinate, che rappresentano le grandezze caratteristiche appena descritte, è proprio il modello HSI (Hue-Saturation-Intensity).

Generalmente lo spazio di colore HSI è descritto da due coni, di cui uno rovesciato e sovrapposto all'altro, come riportato in figura 1.4. Possiamo notare che la componente H, relativa alla tinta, descrive sempre il valore di un angolo. Diversamente, il valore della saturazione S, rappresenta la distanza del punto che identifica il colore nello spazio dal centro dello spazio stesso. Infine, la luminosità I, definisce il livello della superficie circolare nella quale identificare il punto stesso. In maniera analoga a quanto visto nello spazio RGB, il segmento che congiunge le punte dei due coni, unisce i punti che definiscono il bianco (0,0,255) ed il nero (0,0,0); su tale segmento sono ovviamente collocate le varie tonalità di grigio.

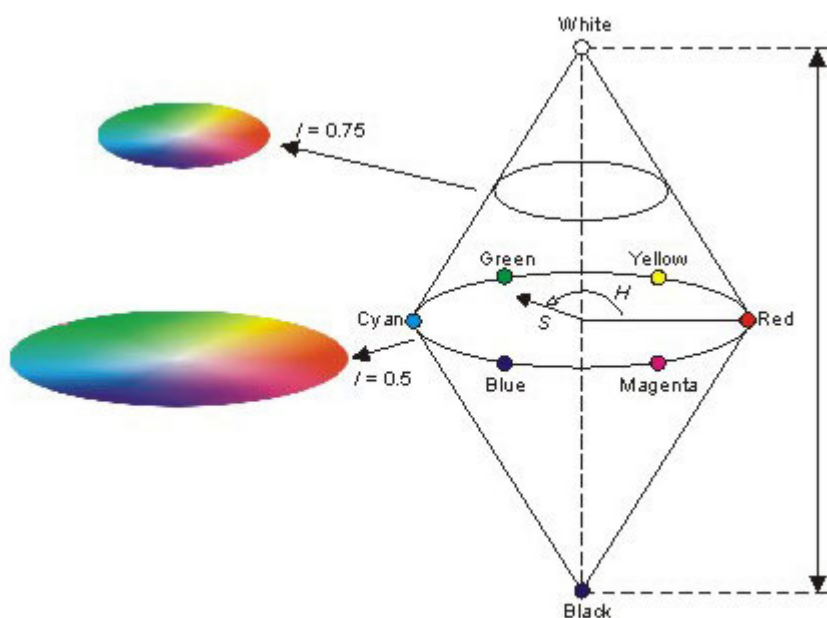


Figura 1.4 - Lo spazio di colore HSI

Sono stati quindi introdotti gli spazi di colore RGB e HSI. Per gli scopi di questa

tesi non sono necessari ulteriori approfondimenti, mentre può risultare senz'altro utile fissare alcuni concetti esposti fino ad ora, poiché sarà a questi che faremo riferimento nel proseguo della trattazione.

Abbiamo detto che il modello di colore RGB è un modello additivo, nel quale ogni colore è definito dalla somma di tre contributi derivanti da tre distinti canali, ognuno dei quali individua un colore primario: rosso, verde e blu. Abbiamo visto quindi che questo modello cerca di rappresentare i colori in maniera conforme al metodo con cui il sistema visivo umano trasmette i colori al cervello. Infine è stato evidenziato che, nonostante sussista quest'ultima analogia, tale modello non è quello che più si avvicina alla descrizione del colore in linguaggio naturale, ovvero tramite la caratterizzazione del colore in termini di tinta, saturazione e luminosità. A tale scopo, per avere cioè una descrizione del colore in linguaggio naturale, è stato introdotto il modello HSI, che descrive un colore attraverso le grandezze di tinta, saturazione e intensità luminosa.

A questo punto, tenendo in dovuta considerazione quanto detto riguardo ai modelli di colore RGB e HSI, poiché a breve sarà necessario aver ben chiare le nozioni ed i concetti esposti su questi modelli, passiamo all'argomento successivo: la scelta del modo con cui rappresentare il contenuto di un'immagine.

1.4 - L'istogramma di colore e la matrice di similarità.

Dovrebbe essere acquisito il fatto che una qualsiasi immagine memorizzata in un supporto digitale è codificata da una sequenza di bit, nella quale sono mantenute le informazioni necessarie a caratterizzare i singoli pixel che costituiscono l'immagine. Dalla codifica di ogni singolo pixel dell'immagine è quindi sempre possibile risalire alla sua collocazione spaziale ed al suo contenuto di colore, espresso tramite un qualsiasi modello. Disponendo di tali informazioni possiamo quindi riprodurre esattamente il contenuto di un'immagine tramite un calcolatore, sempre in conformità ai limiti descritti nel paragrafo 1.1, imposti dai processi di campionamento e quantizzazione, ai quali è sottoposta un'immagine per essere

digitalizzata, che come abbiamo visto possono introdurre una certa perdita d'informazione. Puntualizzato questo fatto, facciamo adesso un piccolo passo indietro, cerchiamo infatti di ricordare quali premesse e quali scopi ci hanno indotto ad introdurre i concetti riguardanti le immagini, il colore e i modelli di colore. Lo scopo inizialmente dichiarato era quello di realizzare un'applicazione web che, con l'ausilio di un database, potesse archiviare delle immagini ed effettuare dei confronti di similarità tra le medesime. A questo punto, per quanto appurato sulla memorizzazione delle immagini, dovrebbe sorgere spontanea una domanda: come possono essere confrontate due immagini archiviate in un supporto digitale? Consideriamo due immagini, come quelle riportate in figura 1.5, entrambe di dimensione 400x300 pixel e codificate in RGB, poiché i formati supportati in ambito web (gif, jpg, jpeg e png) utilizzano sempre questa codifica. Guardando le due immagini si può notare senza ambiguità che esse descrivono due paesaggi simili; come può essere quindi possibile quantificare quanto questi due paesaggi raffigurati nelle immagini sono simili? Oppure, come possiamo fare per valutare se queste due immagini sono perfettamente identiche? Ovviamente dobbiamo sempre ricordarci che a queste domande deve rispondere l'applicazione che vogliamo realizzare e non il nostro occhio!

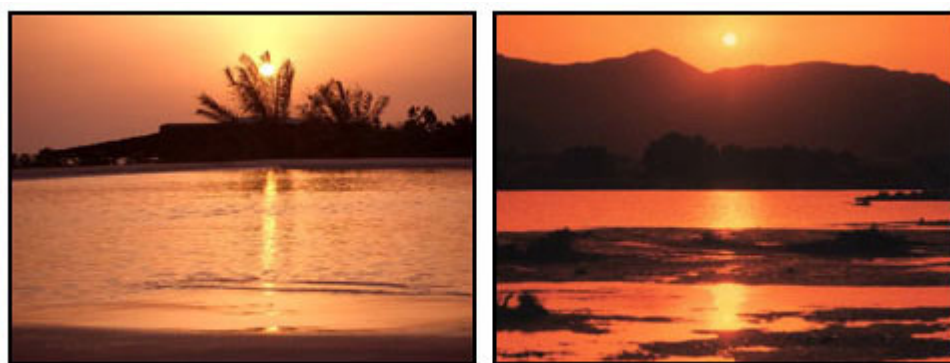


Figura 1.5 - Due immagini che possono essere considerate simili

Abbiamo ricordato ad inizio paragrafo che, un calcolatore, vede queste due immagini come sequenze di bit che forniscono informazioni riguardo ai pixel che le

costituiscono; detto ciò non può essere comunque ipotizzabile, come potrebbe essere lecito pensare istintivamente, valutare la similarità tra le due immagini andando a confrontare le codifiche binarie di ogni singolo pixel estratto dall'una e dall'altra immagine. Questo genere di confronto, infatti, non avrebbe nessuna valenza generale: basti pensare che, adottando tale strategia, confrontare due immagini di diverse dimensioni non sarebbe possibile, o almeno, qualsiasi operazione di confronto fornirebbe un risultato non attendibile. Risulta quindi indispensabile identificare un metodo che, per confrontare due o più immagini, non prenda in esame il loro codice sorgente, ma bensì effettui i confronti avvalendosi di una qualche forma di rappresentazione del loro contenuto che, possibilmente, risulti anche indipendente dalle dimensioni. In particolare, nell'obiettivo di questo lavoro di tesi, è esplicitamente richiesto di ricercare, e quindi confrontare, le immagini per contenuto cromatico. In ragione di ciò, è naturale individuare un metodo per caratterizzare il contenuto di un'immagine in base alle sue caratteristiche cromatiche. La rappresentazione del contenuto di un'immagine, facendo riferimento alle sole caratteristiche cromatiche, non tiene ovviamente in considerazione della sua forma, ovvero delle informazioni legate alla distribuzione spaziale dei singoli pixel. Questo comporta che due immagini, con layouts molto differenti, possono comunque avere due rappresentazioni molto simili. D'altra parte, una rappresentazione del contenuto che tiene conto delle sole informazioni cromatiche di un'immagine, ha un aspetto molto positivo: la facilità con la quale può essere implementata. A tal fine, possiamo quindi introdurre una struttura idonea a fornire una rappresentazione del contenuto di un'immagine, basata solo sulle caratteristiche cromatiche: “l'istogramma di colore di un'immagine”. L'utilizzo di un “istogramma”, detto anche “descrittore”, è una delle varie possibilità esistenti per descrivere il contenuto di un'immagine. Come già evidenziato, questo tipo di rappresentazione non è sicuramente la più efficiente, poiché non prende in considerazione alcuna informazione legata alla forma dell'immagine; ha il vantaggio, però, di poter essere realizzata in modo abbastanza semplice. La definizione di un istogramma di un'immagine può essere fatta principalmente in due

modi, ovvero facendo uso di una o due distribuzioni. Nel primo caso, viene creata una sola distribuzione, di dimensione equivalente al numero di colori distinti presenti nell'immagine. Ogni singolo elemento della distribuzione è associato ad un colore e il valore di ognuno indica la percentuale di pixel dell'immagine che presenta quel colore. Nel secondo caso, invece, vengono utilizzate due distribuzioni indipendenti, una riferita alle informazioni cromatiche ed un'altra a quelle di luminosità. In questo caso, gli elementi delle due distribuzioni, sono rispettivamente associati alle varie coppie tinta-saturazione ed ai possibili valori d'intensità luminosa. La prima distribuzione avrà quindi un numero di elementi pari al numero delle distinte coppie di tinta e saturazione potenzialmente presenti nell'immagine, mentre la seconda avrà un numero di elementi pari ai possibili valori di intensità luminosa che l'immagine può presentare. Anche in questo caso, gli elementi di entrambe le distribuzioni, indicano in percentuale quanti pixel dell'immagine presentano le caratteristiche che un determinato elemento della distribuzione rappresenta. Per questo lavoro di tesi utilizzeremo il primo tipo di implementazione dell'istogramma di colore, quale struttura atta alla rappresentazione del contenuto cromatico di un'immagine. In definitiva, questa struttura può essere rappresentata in un calcolatore tramite un vettore, caratterizzato da un numero di posizioni pari al numero dei possibili colori che l'immagine può avere, ovvero, ogni posizione del vettore identifica un colore distinto, dove, il numero di questi ultimi è determinato dal modello di colore scelto per la codifica dell'immagine, e dal numero di bit utilizzati per codificare ogni singolo pixel. Data un'immagine, come abbiamo appena appreso, il suo istogramma di colore viene costruito attribuendo ad ogni posizione del vettore un valore che indica, in percentuale, quanto quel determinato colore è presente nell'immagine. Nell'istogramma di colore riportato in figura 1.6, sono presenti solo i valori percentuali associati ai colori che hanno una presenza non nulla nell'immagine. In realtà, se l'immagine utilizza una codifica per la quale essa può presentare un numero di colori superiore, rispetto a quelli effettivamente visibili, l'istogramma avrà sempre delle posizioni nulle, le quali indicano che il colore riferito a quella posizione è assente dall'immagine.

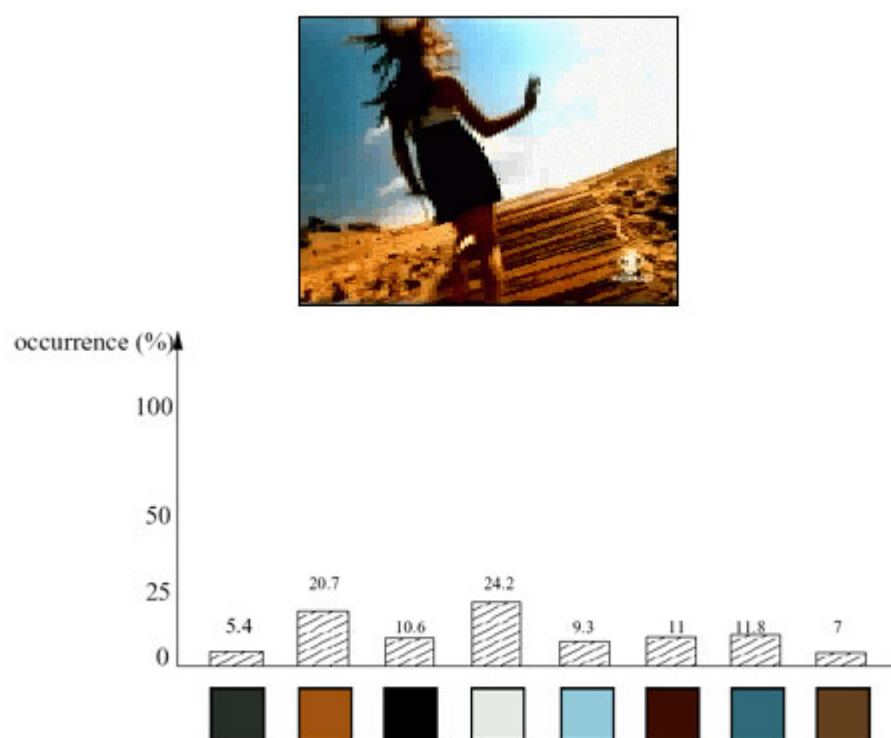


Figura 1.6 - L'istogramma di colore di un'immagine

Con questo tipo di rappresentazione del contenuto, può quindi accadere che due istogrammi associati ad immagini diverse siano uguali. Come premesso, tale evenienza è conseguenza diretta del fatto che, con questa rappresentazione del contenuto, non è presente nessuna informazione sulla forma dell'immagine. Le immagini riportate in figura 1.7, seppur diverse, hanno due istogrammi di colore uguali.



Figura 1.7 - Immagini diverse che presentano istogrammi uguali

Quello appena esposto rappresenta sicuramente un caso limite, poiché le immagini prese in considerazione non rappresentano alcuna scena, oggetto o paesaggio, ma bensì due scacchiere, fatte appositamente con caselle di uguali colori disposte in ordine diverso, proprio per mettere in risalto il concetto espresso in precedenza, riguardo al fatto che, non disponendo nell'istogramma di nessuna informazione sulla distribuzione spaziale dei pixel, può accadere di considerare uguali due immagini che in realtà sono completamente diverse. Come avremo modo di vedere e di discutere nell'ultimo capitolo, questo caso limite, quando il contenuto dell'immagine rappresenta scene o paesaggi realistici, raramente si verifica in modo così evidente. Una situazione sicuramente più rilevante, quando si fanno dei confronti di similarità tra immagini, è rappresentata dal caso in cui si confrontano quattro immagini monocolori come quelle riportate in figura 1.8.



Figura 1.8 - Immagini monocolori distinte

Se vengono costruiti i quattro istogrammi, relativi ad ogni singola immagine, è evidente che ognuno di essi presenterà tutte posizioni nulle, eccetto quella riferita all'unico colore presente nell'immagine, la quale assumerà valore uno, indicando appunto che il 100% dei pixel dell'immagine sono di quel colore. Se ora andiamo a confrontare tra loro le immagini, cioè confrontiamo i loro istogrammi, sicuramente otteniamo come risultato che tutte le immagini sono assolutamente diverse e questo è un dato di fatto inconfutabile. Però, è altrettanto realistico il fatto che, noi umani, senza dubbio, percepiamo molta differenza tra la prima immagine e la terza e la quarta, mentre ne percepiamo molta meno tra la prima e la seconda. In ragione di ciò, mettere sullo stesso piano la differenza che intercorre tra la prima immagine e

le ultime due, e la differenza tra la prima e la seconda, non è il metodo che quantifica nel miglior modo possibile quella che è la nostra naturale sensazione di differenza cromatica. Risulta perciò necessario, introdurre un qualcosa, che possa notificare al calcolatore il fatto che, due immagini, comunque diverse, possono essere più o meno simili. Questo è necessario poiché, utilizzando il solo istogramma di colore, ci siamo appena resi conto che in un eventuale confronto, l'applicazione, valuterebbe nella stessa misura la diversità tra tutte e quattro le immagini, fornendo così un risultato insensibile alla sensazione percettiva che l'uomo ha guardando e mettendo a confronto le quattro immagini.

Per fare in modo che il calcolatore possa ordinare delle immagini, caratterizzate da istogrammi diversi come quelli prodotti dalle immagini di figura 1.8, secondo una scala di similarità, è quindi opportuno introdurre un'apposita struttura ausiliaria: “la matrice di similarità”. Questa matrice è una matrice di tipo simmetrico, i cui coefficienti forniscono una misura sulla differenza percettiva che intercorre tra due colori distinti. Ipotizziamo di essere nel caso in cui i pixel di un'immagine sono codificati in RGB e supponiamo, per semplicità, che ogni immagine possa presentare solo quattro colori distinti. Siano quindi questi quattro colori quelli utilizzati nelle singole immagini monocolori di figura 1.8:

1. RGB(0,0,255);
2. RGB(0,100,200);
3. RGB(255,0,0);
4. RGB(0,255,0);

La matrice di similarità sarà in questo caso una matrice 4×4 , dove ogni coefficiente $M_{i,j}$ fornisce una misura della differenza percettiva tra il colore i ed il colore j , con $i,j=1,2,3,4$. Risulta immediato che $M_{i,j}$ è uguale a $M_{j,i}$. Inoltre è evidente che i coefficienti sulla diagonale $M_{i,i}$ ($i=j$), dovranno assumere un valore che indica che tra i due colori non c'è differenza, essendo questi esattamente gli stessi. Esistono vari metodi per calcolare i coefficienti della matrice di similarità; qui di seguito

sarà esposto quello che effettivamente verrà utilizzato nell'applicazione web per effettuare il confronto tra immagini (tra istogrammi). Comunque sia, tutti i metodi di calcolo dei coefficienti della matrice di similarità, sono volti ad assicurare che la distanza tra due istogrammi risulti in generale non negativa e nulla quando gli istogrammi sono perfettamente identici. Nel secondo capitolo verrà illustrato che il criterio di confronto tra istogrammi opera calcolando una distanza tra essi, ragion per cui è importante assicurare il fatto che questa distanza risulti sempre maggiore o uguale a zero. Il metodo di calcolo scelto per i coefficienti è il seguente:

$$M_{i,j} = \left(1 - \frac{D_{i,j}}{D_{\max}} \right)$$

$$D_{i,j} = \sqrt{|R_i - R_j|^2 + |G_i - G_j|^2 + |B_i - B_j|^2}$$

Prospetto 1.1 - Formule per il calcolo della matrice di similarità

Con $D_{i,j}$ viene indicata la distanza euclidea tra il colore i ed il colore j , mentre con D_{\max} ci si riferisce alla massima distanza euclidea rilevata tra due colori distinti appartenenti all'insieme considerato. Con questo metodo, la matrice di similarità, riferita alle ipotesi fatte, assume i seguenti valori:

	RGB(0,0,255)	RGB(0,100,255)	RGB(255,0,0)	RGB(0,255,0)
RGB(0,0,255)	1	0.68353	0	0
RGB(0,100,255)	0.68353	1	0.05954	0.29853
RGB(255,0,0)	0	0.05954	1	0
RGB(0,255,0)	0	0.29853	0	1

Prospetto 1.2 - Esempio di calcolo della matrice di similarità

Per maggiore chiarezza sono state inserite una riga ed una colonna d'intestazione,

dove sono riportati i valori RGB utilizzati per il calcolo dei $D_{i,j}$ e degli $M_{i,j}$. Come era lecito aspettarsi, osservando la formula per il calcolo dei coefficienti, possiamo notare che: i valori sono normalizzati nell'intervallo $[0,1]$, il valore uno indica l'uguaglianza tra colori, mentre il valore zero è indice del fatto che i due colori sono alla massima distanza percettiva possibile nell'insieme considerato. In generale, quindi, valori prossimi all'unità indicano colori simili, mentre valori tendenti allo zero indicano colori diversi. Dagli altri valori, diversi da uno e zero, possiamo quindi constatare che il colore in posizione uno è molto più simile a quello in posizione due, che non a quelli in posizione tre e quattro. Questo era di fatto il risultato auspicato, possiamo perciò affermare che, con l'utilizzo della matrice di similarità, abbiamo individuato un sistema per dare una misura della differenza percettiva che intercorre tra due colori distinti. In particolare, introducendo attivamente questi coefficienti nel criterio di confronto tra istogrammi, sarà sempre possibile effettuare un confronto, tra un'immagine di riferimento ed un'insieme di immagini distinte, ottenendo come risultato l'ordinamento delle immagini confrontate, da quella più simile a quella meno, secondo una scala di similarità percettiva determinata dai colori che le immagini stesse contengono.

In conclusione, in questo paragrafo, abbiamo introdotto l'istogramma di colore di un'immagine quale struttura idonea alla rappresentazione del suo contenuto cromatico. Abbiamo anche intuito chiaramente che l'istogramma (rappresentato nel calcolatore tramite un vettore) è il dato strutturato che deve essere utilizzato, per effettuare i confronti tra immagini, nel predicato da richiamare all'interno delle query SQL. Infine, per ottenere dei risultati consistenti dai confronti tra istogrammi, è stata introdotta la matrice di similarità, della quale abbiamo chiaramente esposto il significato e che, come già detto, utilizzeremo attivamente nell'implementazione del criterio di confronto, descritto nel prossimo capitolo.

Resta adesso da discutere di un ultimo aspetto, sempre legato al contenuto cromatico di un'immagine, prima di entrare nel merito di come realizzare l'applicazione web che archivi ed effettui le ricerche tra immagini attraverso l'utilizzo di un database.

1.5 - La mappatura di un'immagine ed i colori della mappa.

Nel paragrafo precedente è stato appurato che, ai fini di questo lavoro di tesi, il contenuto cromatico di un'immagine sarà completamente rappresentato tramite un istogramma di colore. Abbiamo appreso che tale struttura è un vettore caratterizzato da un numero di posizioni equivalente al numero di colori distinti che i pixel di un'immagine possono assumere. Sempre nel precedente paragrafo è stato detto che un'applicazione di tipo web, alla quale un'utente si interfaccia mediante l'utilizzo di un browser, può manipolare e visualizzare solo certi tipi d'immagini, in particolare quelle nei formati gif, jpg, jpeg e png. In relazione a questi formati è stato anche asserito che essi utilizzano per la codifica dei colori il modello RGB, del quale, sempre in questo capitolo, è stato definito lo spazio di colore e le caratteristiche principali (paragrafi 1.2 e 1.3). Se adesso andiamo ad analizzare una qualsiasi delle due immagini riportate in figura 1.5, estraendo da essa la codifica di colore di un'insieme qualunque di pixel, potremmo constatare che in generale ognuno di essi è codificato con 8 bit per ogni singolo canale RGB. Se poi volessimo ripetere l'esperimento con delle immagini scelte casualmente da un sito web, ci accorgeremmo del fatto che, eccetto casi particolari, il risultato è il medesimo: tutte le immagini, qualsiasi sia il loro formato, rappresentano i colori con una codifica RGB che utilizza 8 bit per canale. Tutto questo implica che i colori totali che una generica immagine può presentare sono oltre 16 milioni! Diretta conseguenza di tutto ciò è che ogni istogramma associato ad un'immagine, fruibile da un'applicazione web, è rappresentato da un vettore di oltre 16 milioni di posizioni. A sua volta consegue da questo fatto che definire una matrice di similarità, relativa ad un insieme di oltre 16 milioni di colori, benché questa sia simmetrica, comporterebbe il calcolo di un numero di coefficienti distinti elevatissimo.

A questo punto, ponendosi nella condizione operativa delineata dalla situazione appena esposta, pur non conoscendo ancora nel dettaglio il criterio di confronto tra gli istogrammi delle immagini, ma avendo già appreso che questo calcola una distanza, utilizzando sia i valori degli istogrammi che i coefficienti della matrice di

similarità, dovrebbe essere evidente il fatto che, con un numero di colori così elevato, ogni confronto tra immagini avrebbe un costo computazionale davvero troppo elevato. Dobbiamo quindi trovare un sistema per ridurre il numero dei colori distinti presenti in un'immagine, facendo però in modo che questa operazione non alteri eccessivamente il suo contenuto. Questo è possibile tramite un processo chiamato “mappatura”. Senza volerne dare una definizione rigorosa, possiamo dire che “dato un insieme di colori A ed uno di colori B, dove B è contenuto in A, si definisce mappatura di un'immagine il processo mediante il quale il colore di ogni pixel di un'immagine codificata con A colori viene sostituito con il colore appartenente all'insieme B ad esso cromaticamente più simile (o a distanza minima)”. In questo caso, l'insieme A, corrisponde agli oltre 16 milioni di colori definiti dai tre canali RGB che codificano un'immagine di tipo web, mentre resta da identificare chi è l'insieme B, o meglio, quanti colori appartengono a tale insieme e quali sono. Appare evidente dalla definizione data di mappatura che i colori dell'insieme B sono un sotto insieme dei colori dell'insieme A. Per determinare, quali colori dell'insieme A inserire nell'insieme B, si potrebbe pensare di selezionare un certo numero di valori equidistanti nei tre canali RGB, combinare questi valori, e definire l'insieme B con i colori risultanti dalle varie combinazioni possibili. Se, ad esempio, selezioniamo 5 valori per canale, sostanzialmente equidistanti, come i valori '0', '64', '128', '192' e '255', otteniamo 125 colori distinti, ovvero un numero di colori inferiore di ben cinque ordini di grandezza rispetto agli oltre 16 milioni di colori di partenza! Potrebbe quindi sembrare tutto risolto, poiché 125 colori, nell'ottica delle computazioni da effettuare per eseguire i confronti, è un numero più che accettabile, quale dimensione dell'istogramma. Purtroppo, rimanendo comunque valida quest'ultima affermazione, c'è un altro problema: i 125 colori identificati con questo metodo, saranno adeguati affinché il contenuto dell'immagine non risulti alterato dopo essere stata sottoposto al processo di mappatura? La risposta è no, o meglio, diciamo che così facendo c'è il rischio di un'alterazione del contenuto maggiore di quella inevitabile che in ogni caso un processo di mappatura generalmente introduce. Abbiamo infatti detto che la mappatura sostituisce un

colore originale dell'insieme A, con quello dell'insieme B a distanza minima, dove per distanza facciamo sempre riferimento alla distanza euclidea. La minore alterazione possibile del contenuto, dopo aver sottoposto un'immagine al processo di mappatura, si ottiene quando il colore a distanza minima, appartenente all'insieme B, che va a sostituire quello originale dell'insieme A, è di fatto quello che nell'insieme B è cromaticamente più simile a quello dell'insieme A, ottenendo così un'immagine mappata che fornisce una percezione visiva molto fedele a quella data dall'immagine originale. Con la scelta dei colori fatta nel modo esposto, questa garanzia non è però certa, e questo dipende dalle caratteristiche alla base del modello RGB. Abbiamo già discusso del fatto che il modello RGB non esprime i colori in linguaggio naturale, ovvero in termini di tinta, saturazione e luminosità, ma ricrea i colori come sovrapposizione di tre distinti canali riferiti al rosso, verde e blu. Abbiamo anche già evidenziato che, in un modello additivo di questo tipo, la componente di luminosità di un colore varia in maniera correlata al variare dei valori dei tre canali. Questo indica che una scelta dei colori per l'insieme B, definiti dalle combinazioni di valori equidistanti delle singole componenti dei tre canali, non è quella che mappa meglio l'insieme dei colori A. Possiamo riassumere questo concetto dicendo che, nello spazio RGB la distanza minima e la percezione naturale tra due colori non varia in modo uniforme e, di conseguenza, dobbiamo identificare i colori dell'insieme B in altra maniera, affinché il processo di mappatura non introduca nel contenuto di un'immagine delle alterazioni, tali da rendere molto dissimile la percezione visiva dell'immagine mappata rispetto a quella originale.

A tal fine, come spero possa essere già stato intuito, entra in gioco il modello di colore HSI. Sappiamo, infatti, che questo modello descrive i colori in linguaggio naturale, per cui possiamo identificare i colori dell'insieme B utilizzando delle opportune componenti di tinta, saturazione e luminosità in questo spazio di colore e, successivamente, convertire i colori individuati dalle possibili combinazioni di queste componenti, dalla codifica HSI a quella RGB. Cosa rappresentano i valori H, S ed I nello spazio HSI è già stato spiegato nel paragrafo 1.3, possiamo quindi procedere ad elencare le singole componenti scelte per definire i colori:

- H: {0°, 30°, 60°, 90°, 120°, 150°, 180°, 210°, 240°, 270°, 300°, 330°}.
- S: {0, 33, 67, 100}.
- I: {0, 64, 128, 192, 255}.

Prospetto 1.3 - Componenti HSI scelte per generare la Mappa dei Colori

Con questa scelta si ottengono 149 colori distinti nello spazio HSI. Infatti, per $I=0$, abbiamo un solo colore possibile (il nero), qualunque siano i valori di H ed S. Per $S=0$, ne abbiamo invece 5, dovuti ai 5 livelli distinti di luminosità I, qualunque siano i valori della tinta H; uno di questi cinque però, il nero, è già stato considerato. Infine, quando luminosità e saturazione non sono nulle, abbiamo $12 \times 3 \times 4 = 144$ combinazioni di colori possibili, dovute alle dodici tinte considerate, che si combinano con i tre livelli di saturazione e i quattro livelli di luminosità selezionati. Resta ora da compiere l'ultimo passo, convertire questi colori dallo spazio HSI allo spazio RGB, ed avremo così definito l'insieme B dei colori della mappa. Per effettuare questa operazione sono state utilizzate le formule riportate di seguito:

$$h = H \cdot \pi / 180 ; \quad s = S / 100 ; \quad i = I / 255$$

$$x = i \cdot (1 - s)$$

$$y = i \cdot \left[1 + \frac{s \cdot \cos(h)}{\cos(\pi/3 - h)} \right]$$

$$z = 3i - (x + y);$$

$$\text{Se: } h < 2\pi/3, \quad b = x; \quad r = y; \quad g = z.$$

$$\text{Se: } 2\pi/3 \leq h < 4\pi/3, \quad h = h - 2\pi/3, \quad r = x; \quad g = y; \quad b = z.$$

$$\text{Se: } 4\pi/3 \leq h < 2\pi, \quad h = h - 4\pi/3, \quad g = x; \quad b = y; \quad r = z.$$

$$r = r \cdot 255; \quad \text{se } r > 255 \quad \text{allora } r = 255.$$

$$g = g \cdot 255; \quad \text{se } g > 255 \quad \text{allora } g = 255.$$

$$b = b \cdot 255; \quad \text{se } b > 255 \quad \text{allora } b = 255.$$

Prospetto 1.4 - Formule di conversione da HSI a RGB

Osservando le formule di conversione, si può notare che, in alcuni casi, può succedere che una terna HSI produca un colore RGB nel quale una componente, o più, ecceda il massimo valore consentito, pari a 255. In tal caso, come descritto nelle medesime formule, il valore del canale in eccesso viene posto a 255, poiché la codifica RGB non può accettare valori superiori. Questa situazione si è manifestata anche nella conversione dei 149 colori HSI selezionati, dalla quale di fatto sono risultati 122 colori distinti nello spazio RGB e non 149. Questo risultato è proprio il frutto di quanto appena detto, ovvero che in più di un caso due colori distinti HSI hanno generato lo stesso colore RGB.

I 122 colori individuati nello spazio RGB costituiscono quindi i colori dell'insieme B e verranno chiamati nel seguito “Colori della Mappa”. Facendo un breve excursus, possiamo anticipare che questi colori possono essere visualizzati dall'area “HELP” del menù dell'applicazione web “Image-Compare”. Per ogni colore vengono mostrate sia le componenti HSI che quelle RGB, oltre ad una piccola immagine monocolora che consente di vedere a video il colore al quale le due terne si riferiscono. Come sarà spiegato meglio nel terzo capitolo, dove parleremo dell'applicazione web e della sua interfaccia, la mappa dei colori (e quindi anche la matrice di similarità, poiché i suoi coefficienti sono direttamente dipendenti dai colori presenti nella mappa), viene generata ad ogni nuova apertura di sessione dell'applicazione. Questo modo di operare è determinato dal fatto che, i colori della mappa, vengono sempre calcolati prendendo come riferimento le possibili terne generate dalle componenti HSI definite nel file di configurazione dell'applicazione (setting.php). Per questa ragione, ogni qual volta viene fatta una scelta diversa dei valori HSI, rispetto a quelli indicati in questo paragrafo e attualmente impostati nel suddetto file di configurazione, l'applicazione può gestire questo cambiamento in tempo reale. Ovviamente, questo non deve suggerire al lettore l'idea che l'applicazione possa operare con immagini mappate su insiemi diversi, infatti, come avremo modo di vedere nel secondo capitolo, descrivendo il criterio di confronto tra istogrammi, non è possibile e non avrebbe alcun senso effettuare dei confronti tra istogrammi che presentano dimensioni diverse. Ciò comunque non esclude, come è

stato detto, che l'applicazione possa essere predisposta per operare correttamente utilizzando mappe dei colori diverse; l'importante, ribadiamo, è assicurarsi che tutte le immagini sulle quali opera l'applicazione siano state sottoposte ad un processo di mappatura che utilizzi la stessa mappa dei colori. In riguardo alla mappa dei colori possiamo fare infine un'ultima anticipazione: sempre dall'area “HELP”, nella sezione “Colori della Mappa”, è possibile accedere ad una rappresentazione tridimensionale della distribuzione dei colori della mappa nel cubo RGB. Questa visualizzazione verrà riportata nel capitolo quarto, dove verranno fatte alcune considerazioni volte a spiegare in quali modi le caratteristiche della distribuzione possono influenzare determinati risultati prodotti dall'applicazione.

Detto ciò, interrompiamo questa trattazione riguardo alle funzionalità dell'applicazione, poiché, per gli scopi di questo paragrafo e del capitolo in generale, non sono necessari ulteriori approfondimenti in merito.

Possiamo concludere adesso il paragrafo e l'intero capitolo mostrando due immagini, un'originale e la sua corrispettiva mappata, dove nel processo di mappatura sono stati utilizzati i 122 colori RGB generati dai valori HSI indicati in precedenza nel prospetto 1.3.



Figura 1.9 - Immagine originale e corrispettiva mappata

Le considerazioni, riguardo agli effetti prodotti su un'immagine dal processo di mappatura e in che modo questi effetti incidano sui risultati dei confronti tra immagini, saranno l'oggetto dell'ultimo capitolo.