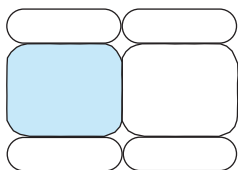




# STATO DELL'ARTE DELLA FOTOGRAFIA DIGITALE BASI TEORICHE E SVILUPPI TECNOLOGICI

Massimo Mancuso  
Stefano Tubaro  
Guido Vannucchi



La fotografia digitale, grazie alle tecniche ICT, ha fatto in questi ultimi anni enormi progressi e ha ormai soppiantato l'acquisizione di immagini con la classica pellicola fotografica. Un ruolo primario in questa evoluzione è stato giocato dai sensori per l'acquisizione di immagini e dagli importanti sviluppi dell'elaborazione numerica dei segnali (DSP) nella fotocamera. Scopo del presente articolo è quello di fornire una panoramica delle basi teoriche della rappresentazione dei colori e delle conseguenti tecnologie che sono alla base della fotografia digitale nella fase di cattura delle immagini.

## 1. INTRODUZIONE

Il campo dell'ICT è sempre più pervasivo estendendosi ormai ai più svariati settori. In particolare, la fotografia digitale, grazie alle tecniche ICT, ha realizzato forti innovazioni quasi completamente soppiantando l'acquisizione di immagini con la tradizionale pellicola. Questo risultato è stato possibile grazie ai sostanziali avanzamenti compiuti dai sensori microelettronici per l'acquisizione d'immagini e nel processamento dei relativi segnali elettrici con tecniche digitali. Inoltre, l'introduzione sempre più spinta di sensori a basso costo sta permettendo di dotare una larga serie di dispositivi personali di comunicazione ed elaborazione (*cellulari, palmari, computer portatili ecc.*) di dispositivi per la

cattura di immagini (fisse ed in movimento). Contrariamente a quanto si potrebbe pensare, la ripresa di un'immagine in movimento pone meno problemi di criticità di un'immagine fissa perché il movimento riduce la sensibilità dell'occhio al dettaglio spaziale ed alle componenti di rumore ad alta frequenza. Pertanto, mentre nella fase iniziale gli apparecchi per la fotografia digitale sono risultati una derivazione della tecnologia degli apparecchi "camcorder" (modificati per la cattura di immagini fisse), oggi invece - con una qualità digitale ormai paragonabile a quella della migliore fotografia a pellicola - alcuni sviluppi tecnologici per la cattura d'immagini fisse influenzano gli apparecchi per la ripresa d'immagini in movimento.

Scopo del presente articolo è pertanto quello

Nota - Il presente articolo rielabora e sintetizza alcuni degli argomenti trattati in una Giornata di Studio (dal titolo: "Le frontiere della fotografia digitale ed il mondo ICT") organizzata nell'Aprile del 2007 dall'Associazione AICT della Federazione AEIT in collaborazione con il Politecnico di Milano. Si ringraziano tutti i partecipanti ed in particolare il dr. Alfredo Roma per alcuni spunti delle conclusioni.

Si ringrazia anche l'Ing. Paolo Talone della FUB che, attraverso un'accurata rilettura del testo, ha dato utili suggerimenti di revisione ed integrazione.

di dare una panoramica tecnica delle basi teoriche della rappresentazione dei colori e delle tecnologie alla base della fotografia digitale all'atto della **fase di cattura** (acquisizione).

La **rappresentazione del colore** è sempre stata un elemento essenziale della fotografia. La luce, ossia la radiazione elettromagnetica che riusciamo a percepire con i nostri occhi è caratterizzata da lunghezze d'onda dello spettro elettromagnetico comprese fra 400 e 700 nm. La sensazione percepita dal nostro senso della vista sembrerebbe dover essere legata ad un numero elevato di parametri che descrivono lo spettro della luce che colpisce l'apparato visivo umano: dalla nostra esperienza quotidiana sappiamo invece che, attraverso tre soli parametri (per esempio le ampiezze di tre colori primari della luce oppure le informazioni di intensità, tinta e saturazione), è possibile dare un'ottima descrizione (in termini percettivi) di una larga porzione delle "luci" osservate dall'occhio umano che fisiologicamente ha una caratterizzazione "tricanale" che si adatta bene ad eccitazioni di tipo "tristimolo".

Questa **rappresentazione a tre valori** delle sensazioni di colore viene spesso data per scontata e l'argomento viene spesso affrontato con superficialità, senza approfondire in alcun modo la teoria e le esperienze fatte per avvalorare questo approccio. A tale scopo è infatti essenziale partire dai concetti fondamentali necessari alla descrizione del sistema visivo umano e della relativa percezione del colore, le cui caratteristiche e limiti sono essenziali per lo sviluppo di adeguate tecnologie di cattura delle immagini. Si è ritenuto pertanto di approfondire nella prima parte di quest'articolo tale importante tematica come necessaria introduzione alle importanti conseguenze pratiche che ne derivano nel campo della fotografia digitale.

Gli altri aspetti importanti che verranno trattati nella seconda parte dell'articolo sono la **tecnologia dei sensori** ed il **processamento dell'immagine** dopo la cattura, passando in rassegna i sensori CCD e CMOS e le più importanti tra le tecniche di elaborazione dell'immagine immediatamente successive alla fase di scatto.

**Ulteriori aspetti della fotografia digitale** che esulano tuttavia dalle finalità del presente articolo sono quelle della stampa, della proiezione, della trasmissione a distanza delle immagini catturate e, particolarmente importan-

ti, quelle delle operazioni di elaborazione e ritocco delle immagini acquisite. Quest'ultime, a differenza del caso d'impiego di pellicola fotografica, sono facilmente alla portata di tutti gli amanti della fotografia e ciò costituisce una grande potenzialità aggiuntiva.

Si può anzi affermare che, a differenza del passato, la fotografia digitale offre, al giorno d'oggi, **due momenti creativi**: l'istante dello scatto e quello della successiva elaborazione dell'immagine. Ciò non deve "scandalizzare" (com'era nella mentalità classica per le modifiche "post-cattura" delle immagini): infatti sia nelle macchine tradizionali che in quelle digitali il parametro della resa cromatica dipende a priori quasi esclusivamente dalle scelte tecnologiche operate nel primo caso dai produttori di pellicola e, nel secondo caso, dai costruttori delle macchine con le loro elaborazioni proprietarie. Ma nel caso delle macchine digitali il fotografo si riappropria della possibilità di agire sul colore agendo su una sua particolare sensibilità alla resa cromatica, attraverso le successive elaborazioni che è in grado di decidere personalmente. Nasce di conseguenza oggi, per molti fotografi, il **nuovo grado di libertà** di poter incidere a posteriori su tutta una serie di parametri che possono modificare sostanzialmente il risultato ottenuto all'atto dello scatto, risultato che, molto più del passato, può essere considerato come una particolare e soggettiva "interpretazione" del soggetto catturato attraverso lo scatto.

## 2. ELEMENTI DI COLORIMETRIA E TEORIA TRICROMATICA

### 2.1. Generalità

Il colore di una superficie o di un oggetto dipende da come lo spettro della luce incidente (bianca o colorata) viene modificata dalle caratteristiche superficiali dell'oggetto stesso.

Se una persona, con un sistema visivo che rientra nella media, cerca di ottenere una corrispondenza di percezione cromatica fra una particolare luce (direttamente generata da una fonte di luce o riflessa da una superficie) e la risultante dalla somma di più luci, scoprirà che, in generale, sarà necessario mescolare non meno di tre luci con una fissata composizione spettrale ed intensità opportuna per ottenere la cercata corrispondenza croma-

tica. Tale equivalenza rende palese la natura tricromatica del sistema visivo umano per ciò che attiene alla percezione del colore e ciò è suffragato dall'esistenza di tre differenti cellule sulla retina (coni) con diverse sensibilità spettrali. In altre parole, i due stimoli (luce di test e primari) anche se caratterizzati da differenti composizioni spettrali possono riprodurre la stessa sensazione di colore purché esse suscitino gli stessi livelli di attività nei coni e ciò dipende dall'integrale del prodotto della loro composizione spettrale per la curva di sensibilità di ciascun tipo di coni.

La colorimetria è la scienza che ha per oggetto lo studio e la misura delle "sensazioni di colore" che l'osservatore umano percepisce. L'approccio tricromatico per la rappresentazione della sensazione umana del colore si basa su alcuni importanti assiomi stabiliti a metà dell'800 da Grassman e su *test* accurati effettuati alla fine degli anni '20 per il "color matching". Per qualche maggiore dettaglio su questo argomento si rimanda all'**Appendice 1** a p. 36.

In definitiva, pur con alcune limitazioni, un osservatore che punti la sua attenzione su una luce colorata [C] (caratterizzata da una sua composizione spettrale o, come caso particolare, monocromatica) **percepisce la stessa impressione** visiva della somma di tre luci primarie [ $P_1$ ], [ $P_2$ ], [ $P_3$ ] (a loro volta con proprie distribuzioni spettrali o monocromatiche come il caso di figura A1.1 nell'Appendice 1) purché esse abbiano opportune **pesature**  $A_1^{[C]}$ ,  $A_2^{[C]}$ ,  $A_3^{[C]}$  (valori di *matching*). Ovviamente, anche a parità di [C], le pesature sono differenti a seconda della natura e caratteristiche spettrali dei primari scelti.

La **tipica espressione colorimetrica generalizzata** è pertanto la seguente:

$$[C] \equiv \sum_{j=1}^3 A_j^{[C]} [P_j] = \sum_{j=1}^3 T_j^{[C]} [P_j] A_j^{[W]} \quad (1)$$

L'ultima parte dell'equazione esprime la normalizzazione dei pesi rispetto a quelli necessari, con le stesse tre luci primarie, a "riprodurre" una luce bianca [W] di riferimento a massima intensità: essi sono indicati con  $A_1^{[W]}$ ,  $A_2^{[W]}$ ,  $A_3^{[W]}$  e rappresentano solo costanti di normalizzazione.

Sono così definiti i coefficienti  $T_1^{[C]}$ ,  $T_2^{[C]}$ ,  $T_3^{[C]}$ , detti "**valori di tristimolo**", relativi al colore [C].

I valori di *matching*  $A_1^{[C]}$ ,  $A_2^{[C]}$ ,  $A_3^{[C]}$  (e quindi anche  $T_1^{[C]}$ ,  $T_2^{[C]}$ ,  $T_3^{[C]}$ ) dipendono dalla natura dei primari e dallo spettro della luce [C] ed, in linea di principio, possono essere valutati sperimentalmente per ciascuna combinazione di luce di test e di primari.

Nel caso che la luce [C] sia monocromatica e di intensità unitaria, al variare della sua lunghezza d'onda caratteristica ( $\lambda$ ) i valori di tristimolo  $T_1^{[C]}$ ,  $T_2^{[C]}$ ,  $T_3^{[C]}$  descriveranno tre funzioni di  $\lambda$  indicate col nome di *Color Matching Function* associate ai particolari primari selezionati (ed al bianco di riferimento scelto). Se, per esempio, si scegliessero i tre primari monocromatici RGB del CIE (vedi Appendice 1), le *Color Matching Function* sarebbero quelle riportate nella figura A.1.1, a meno di fattori di scala legati alle normalizzazioni utilizzate per tracciare i grafici.

Si indicheranno con  $T_1(\lambda)$ ,  $T_2(\lambda)$ ,  $T_3(\lambda)$  i **valori di tristimolo per colori monocromatici** per differenti lunghezze d'onda della luce di test (*Color Matching Function*). Una volta note le *Color Matching Function* (relative ai primari generici ed al bianco di riferimento scelti) le componenti di tristimolo di una generica luce [C] con uno spettro qualsiasi potranno essere valutate analiticamente, e non più solo sperimentalmente, in virtù della legge di addittività di Grassman. Più precisamente, si avrà che:

$$T_j^{[C]} = \int C(\lambda) T_j(\lambda) d\lambda \quad j = 1, 2, 3 \quad (2)$$

in cui  $C(\lambda)$  indica specificatamente lo spettro della luce [C]. I valori di tristimolo per il colore [C] così ottenuti possono anche essere negativi (cfr. per esempio Figura A1.1 nell'Appendice 1). Dal momento che essi rappresentano unità di energia, l'interpretazione fisica di questo risultato matematico è quella di una corrispondenza di sensazione cromatica fra la somma della luce in analisi con il primario caratterizzato da valore di tristimolo con segno negativo (la cui intensità sarà associata al modulo di tale valore) e la somma degli altri due primari. Comunque, da un punto di vista pratico, non essendo fisicamente realizzabili valori di tristimolo negativi, ciò implica, in conseguenza, l'esistenza di colori che non possono essere sintetizzati come somma pesata (con pesi positivi) delle tre luci primarie scelte. Fortunatamente è possibile selezio-

nare dei primari con i quali si è in grado di riprodurre la maggior parte dei colori più comuni e naturali e per maggiori informazioni sugli spazi cromatici e le varie standardizzazioni si rimanda all'**Appendice 2** a p. 37.

È possibile anche scegliere primari (vedi Figura A2.1 nell'Appendice 2) per cui le *Color Matching Function* sono sempre positive, ma ciò implica che tali primari non siano fisicamente realizzabili.

I valori di tristimolo possono essere considerati come le coordinate di un punto di uno spazio di colore tridimensionale. Un particolare colore può essere descritto come un punto in questo spazio. Bisogna tuttavia ricordare che la sensazione di colore non è descritta dalla lunghezza del vettore che unisce il punto rappresentativo del colore con l'origine, ma piuttosto dal rapporto fra le coordinate (valori di tristimolo) di tale punto.

L'insieme di colori riproducibili da una terna di luci primarie, ad esempio il Rosso, Verde e Blu con determinate distribuzioni spettrali di potenza, viene chiamato "**color gamut**" del sistema di primari definiti. I colori che rispetto a tale sistema non possono essere riprodotti esattamente sono **esterni al gamut** (gamma) del sistema scelto.

## 2.2. Colorimetria e sistemi di acquisizione della fotocamera digitale

A riguardo dei moderni **sistemi di acquisizione digitale di immagini**, va ricordato che essi si basano, in generale, sull'uso di un sistema ottico tradizionale nel cui piano immagine sono presenti uno o più sensori che effettuano la conversione fra energia ottica ed elettrica. Tipicamente nelle fotocamere digitali, come di vedrà in seguito, si utilizza un solo sensore, mentre per le telecamere professionali spesso ne vengono utilizzati tre, attraverso l'uso di opportune geometrie ottiche. In ogni caso fra ottica e sensore/i sono disposti opportuni filtri colorati (vedi più avanti Figura 3).

Per ogni elemento d'immagine acquisita, il sensore/i fornirà un vettore di 3 elementi i cui valori, indicati con  $U_R, U_G, U_B$ , sono dati, in analogia con la eq. (2), da:

$$U_X = \int C(\lambda) F_X(\lambda) h(\lambda) d\lambda \quad X = R, G, B \quad (3)$$

dove  $C(\lambda)$  è la densità di energia della luce

incidente,  $F_X(\lambda)$  è la risposta spettrale dei filtri utilizzati (vedi Figura 4), mentre  $h(\lambda)$  è la risposta spettrale dell'elemento fotosensibile comune a tutti i *pixel*. Chiaramente nell'eq. (3) le funzioni di  $\lambda$  sono sempre positive e quindi anche  $U_R, U_G, U_B$  saranno scalari sempre positivi indipendentemente dalla luce/colore considerata. Ciò porta ad affermare che i primari su cui si basa il sistema di acquisizione delle informazioni di colore sono dei primari non fisicamente realizzabili (come i ben noti CIE-XYZ di cui si parla nell'Appendice 2). Quindi le informazioni di colore acquisite in questo modo non possono essere utilizzate, così come sono, come *input* di un sistema di visualizzazione a tre primari fisicamente realizzabili, ma dovranno essere opportunamente trasformate (*rimappate*).

I dispositivi di acquisizione danno quindi in uscita le informazioni di colore non nel formato nativo (ossia il **segnale "Raw"** che "esce" dal sensore/i), ma trasformano le informazioni acquisite riferendole ad una terna di primari che possano essere fisicamente riprodotti. Tipicamente per le fotocamere digitali viene utilizzato il cosiddetto **spazio colore sRGB** (vedi Appendice 2 relativa agli spazi cromatici).

Il passaggio fra formato "Raw" e formato sRGB avviene attraverso una trasformazione lineare (moltiplicazione per una matrice  $3 \times 3$ ). Non è assicurato che all'uscita di tale trasformazione le componenti siano sempre positive. In caso si generino componenti negative esse vengono opportunamente rimappate su valori positivi.

Nel corso degli ultimi anni, data la sempre maggiore diffusione di software per la post-elaborazione digitale delle immagini, un numero sempre maggiore di fotocamere digitali permette di conservare l'immagine anche in formato *Raw*. Così facendo, per l'elaborazione del segnale, si possono usare dei programmi molto sofisticati avendo a disposizione la potenza di calcolo di un PC, senza le restrizioni di costo, consumo di potenza, tempo di processamento (per ridurre al minimo il tempo tra due scatti consecutivi) e quantità di memoria dinamica a cui devono sottostare le macchine fotografiche digitali. Ovviamente è necessario disporre di una notevole memoria per l'immagazzinamento delle immagini catturate in tale formato nativo.

Lo scopo di questa prima parte dell'articolo è stata quella di illustrare come risultino essere due cose completamente diverse le curve di sensibilità dei filtri utilizzati per l'acquisizione tricromatica nelle fotocamere digitali e la composizione spettrale (e conseguentemente i *gamut*) dei primari utilizzati per la rappresentazione delle immagini acquisite.

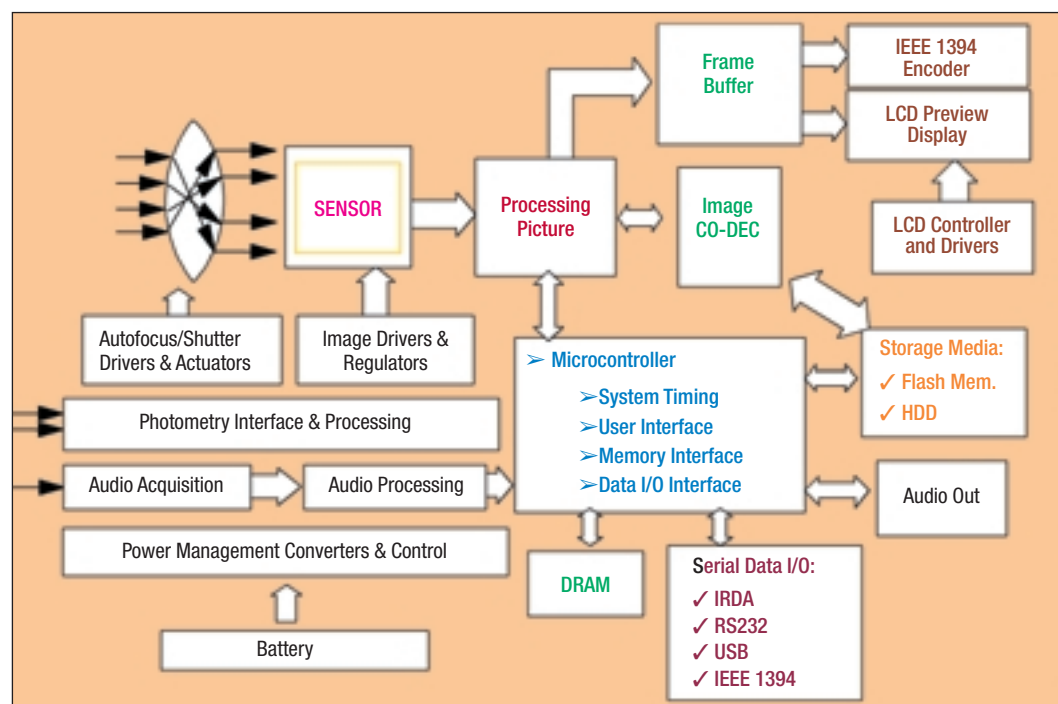
### 3. LA FOTOCAMERA DIGITALE

Lo schema a blocchi di una moderna fotocamera digitale è riportata nella figura 1. Nello schema si identifica un **sistema ottico** del tutto corrispondente a quello di una fotocamera tradizionale. Il secondo blocco che rappresenta il cuore della fotocamera digitale è il **sensore** cui spetta il compito di trasformare i fotoni (luce) in elettroni (cariche elettriche) per ciascuna porzione elementare dell'immagine. La quantità di carica accumulata in ogni elemento del sensore viene poi letta e convertita in informazioni digitali che subiscono di seguito una **serie di processi** assai importanti per garantire la qualità finale dell'immagine e per la sua rappre-

sentazione nel formato desiderato per la memorizzazione.

Al termine della catena, i dati processati vengono codificati ed elaborati numericamente per convertirli nei noti **formati compressi di visualizzazione** di immagini (JPEG, TIFF ecc. a seconda delle esigenze, come si vedrà meglio più avanti) che ne riducono drasticamente le dimensioni e ne permettono una più agevole memorizzazione. Sono questi i dati finali che vengono immagazzinati in una **memoria** (*storage media*) estraibile a stato solido di tipo "flash", con dimensioni e caratteristiche standard o proprietarie, che è in grado di immagazzinare un numero considerevole di immagini il cui valore esatto dipende dal numero di *pixel* del sensore e dal grado di compressione adottato.

L'immagine può, tuttavia, essere anche memorizzata (a spese di un ingombro notevole di memoria) in puri **dati numerici grezzi** (RAW), così come forniti dal sensore ed elaborata successivamente in un vero e proprio computer con opportuni programmi software di notevole complessità. Tale prestazione è oggi sempre più tipicamente fornita, oltre che da macchine professionali, anche da apparecchi



**FIGURA 1**

Schema a blocchi di una fotocamera digitale. Con le scritte a colori sono indicate le funzioni della fotocamera digitale analizzate nel seguito (Fonte: STM)



*prosumer* (fusione dei due termini *professional* e *consumer*) e da alcuni tipi di compatte.

## 4. SENSORI E “COSTRUZIONE” DELL’IMMAGINE

### 4.1. Principi dei sensori e caratteristiche di base

Le macchine fotografiche digitali per catturare l’immagine impiegano dispositivi fotosensibili a semiconduttori (**sensori**) in grado di trasformare l’informazione luminosa, messa a fuoco sul sensore, in cariche elettriche distribuite spazialmente su elementi sensibili che costituiscono la superficie del sensore medesimo. Nello studio dei sensori è utile adottare la distinzione concettuale tra “*photodetector*” (elemento unitario fotosensibile che nei sensori registra solo livelli di luce acromatica), “*photosite*” (luogo fisico del sensore dove si catturano i dettagli dell’immagine e che può essere costituito da un solo, o due o tre *photodetector* ed il “*pixel*” che rappresenta la più piccola porzione del singolo elemento di una matrice ideale con le relative informazioni di colore. Il *pixel*, anche se ha un preciso significato fisico, rappresenta spesso un dato elaborato - cioè un’entità *software* piuttosto che *hardware* - in quanto, per il colore di un singolo punto della matrice, si sfrutta l’interpolazione con le informazioni cromatiche dei punti adiacenti.

Qualora in un *photosite* si impieghino tre *photodetector* su ciascuno dei quali si deposita un opportuno filtro colorato, si è in grado di rivelare le tre componenti primarie (rossa, verde, blu) del colore della luce incidente nel punto considerato. Nel caso invece in cui per un *photosite* si adotta un solo *photodetector* (ed è la situazione più tipica) è necessario impiegare metodi di interpolazione con le informazioni adiacenti. Nel primo caso citato il *photosite* coincide anche con la definizione di *pixel*, mentre nel secondo caso il *pixel* nasce come risultato dell’elaborazione *software* con utilizzo delle informazioni dei *photosite/photodetector* adiacenti.

In realtà, nelle fotocamere digitali, vengono normalmente impiegate strutture diverse per garantire un migliore uso dell’area del sensore. In ogni caso il livello di entità di carica accumulata da ogni *photodetector* viene poi

convertita (tramite un convertitore A/D) in un numero binario con livello opportuno di *bit* (10, 12, 14) che determinano la precisione con cui il sensore è in grado di acquisire le informazioni di luminosità cromatica. Maggiore è il numero di livelli, minore sarà il rumore di quantizzazione (ossia l’approssimazione nella conversione A/D). Va ricordato come il rumore di quantizzazione possa portare nel caso delle immagini alla creazione di falsi colori od a falsi contorni (*contouring*) in particolare nelle regioni con passaggi a lievi sfumature. Agli inizi della fotografia digitale la risoluzione della fotocamera, e quindi dei sensori, è stato uno dei fattori più determinanti per la qualità e, pertanto, dal punto di vista del *marketing*, si era adottata l’equazione: “più *pixel* = migliore qualità”, in quanto risultava semplice associare ad un unico parametro l’indice di qualità del dispositivo.

Per queste ragioni storiche, nella presentazione delle macchine fotografiche i costruttori quasi sempre usano i termini *photodetector* e *pixel* come acronimi. Ciò in linea di principio non è corretto, ma consente di indicare valori numerici maggiori (di *pixel* che tali in realtà non sono), mentre dal punto di vista della **risoluzione** effettiva la situazione è peggiore di quella indicata. Occorrerebbero, infatti, tre diversi *photodetector* per *photosite* per descrivere con precisione un *pixel* nelle sue effettive informazioni di colore.

La corrente di pensiero sopra menzionata ha scatenato una corsa alla miniaturizzazione (per offrire più “*pixel*” a parità di area) che, in aggiunta alla migliore qualità, ha permesso una maggiore flessibilità (per esempio stampe di maggiori dimensioni senza perdita di risoluzione o stampe di porzione dell’immagine anch’esse senza perdita di risoluzione), nonché l’utilizzo della maggiore risoluzione per migliorare altri aspetti che influenzano la qualità. È importante tuttavia ricordare che accrescere eccessivamente la risoluzione del sensore senza accrescere la sua superficie, riduce l’area del singolo *photodetector* e di conseguenza la sensibilità del sensore. In conseguenza di ciò, oggi si sta facendo un pò di marcia indietro su tale aspetto.

Peraltro non è consigliabile realizzare sensori che occupino aree fisicamente molto grandi,

perché questo introduce costi e limitazioni sulle ottiche. Basta a tale proposito ricordare, nella classica fotografia su pellicola, le differenze tra le ottiche per fotocamere a 35 mm e quelle per macchine 6 × 6 cm.

Premesso quanto sopra, le principali caratteristiche a cui, in sintesi, un sensore deve soddisfare possono così riassumersi:

❑ **adeguato numero di photodetector**, per garantire un elevato dettaglio di immagine e per soddisfare la qualità di eventuali ingrandimenti, senza tuttavia far nascere eccessivi problemi di velocità nel trasferimento dei dati al processore d'immagine;

❑ **elevato rapporto segnale-rumore**, in particolare nella cattura d'immagini a bassa luminosità, evitando segnali spuri derivanti da rumori di fondo degli elementi fotosensibili;

❑ **elevata gamma dinamica**, ossia ampiezza d'intervallo di luminosità registrabile prima che il photodetector vada in saturazione;

❑ **capacità di non trattenere "ombre"**, relative cioè alla persistenza di immagini precedenti (vale in particolare per i sensori di tipo CMOS). Se la dimensione fisica del sensore è elevata a parità di numero di *pixel* denunciati dal costruttore, questo comporta ovviamente una maggiore dimensione fisica dei *photodetector*. Tale fatto rende maggiormente sensibile il sensore garantendone un miglior rapporto segnale/rumore.

I sensori di alcune costose fotocamere REFLEX professionali hanno il sensore CCD con una dimensione 24 × 36 mm, pari cioè a quella del fotogramma della pellicola. Con queste dimensioni, oltre ad avere un basso rumore, la lunghezza focale delle ottiche usate per le fotocamere tradizionali non viene alterata nel loro impiego in macchine digitali ad obiettivo intercambiabile. Fuori da questi casi le dimensioni dei sensori possono variare da macchina a macchina (pur mantenendo sempre 2/3 come rapporto dei lati) e la focale di un obiettivo di una fotocamera tradizionale si riduce dello stesso rapporto dei lati del sensore rispetto agli omologhi del fotogramma della pellicola. Tuttavia, come già accennato, per le fotocamere digitali è invalso l'uso di esprimere le focali in termini equivalenti a quelle delle fotocamere tradizionali, mentre in realtà il valore effettivo è più piccolo e questo spiega la compattezza degli obiettivi.

## 4.2. Tecnologie CCD e C- MOS impiegate per i sensori

I sensori possono essere di tipo **CCD** (*Charge Coupled Devices*) o di tipo **CMOS** (*Complementary Metal Oxide Semiconductor*). Senza entrare in eccessivi dettagli, si riportano di seguito le maggiori differenze tra i due tipi di sensori.

Entrambi i sensori CCD e CMOS sono costituiti da una matrice di milioni di celle sensibili alla luce che, per ambedue i tipi, sono realizzati attraverso *photo-diodes* oppure *photo-gates*. Ciascuna di queste celle converte in cariche elettriche i fotoni che incidono sulla relativa porzione d'immagine: le celle a foto-diode usano la giunzione per immagazzinare la carica generata, mentre le celle a foto-gate usano le barriere di potenziale della loro capacità. Tali celle costituiscono i *photodetector* o i *photo-site* (nel caso di un solo *detector*) già definite poco sopra e tanto maggiore è l'intensità della luce incidente tanto maggiore è la carica elettrica che viene accumulata nel relativo sito.

Il passo successivo è quello di leggere i valori delle cariche accumulate in ciascuna cella. In un **CCD** la carica è "estratta" ad un angolo estremo del sensore e le cariche di ciascun elemento fotosensibile debbono, attraverso opportune tecniche di indirizzamento che generano i necessari campi elettrici, essere trasferite nelle diverse celle contigue per giungere alla porta di uscita (da cui il nome del dispositivo). Ciò significa che la carica deve essere "spinta" da un *pixel* a quello adiacente senza che parte di essa rimanga confinata nel *pixel* da cui è stata generata o in cui transita. Tale metodo di scansione dei CCD richiede tensioni nel campo 2.5 ~ 10 V e comporta un maggiore consumo di potenza rispetto ai sensori di tipo CMOS. Le cariche trasferite, alla fine del processo di scansione, sono trasformate in tensione attraverso l'uso di un'opportuna capacità. Si genera in tal modo, attraverso la lettura, una successione di campioni analogici sequenziali che vengono tradotti in numeri binari per mezzo di un convertitore analogico-digitale.

Lo scopo principale della progettazione e dei processi di costruzione dei CCD è quello di massimizzare il parametro *Charge Transfer Efficiency* (CTE), che misura il rapporto tra la carica trasferita e quella che rimane nel *pixel* che l'ha generata. Ciò si traduce in speciali processi manifatturieri e nella necessità, come già ac-

cennato, di tensioni di alimentazione più elevate con un conseguente **maggiore consumo di potenza rispetto al caso dei C-MOS**.

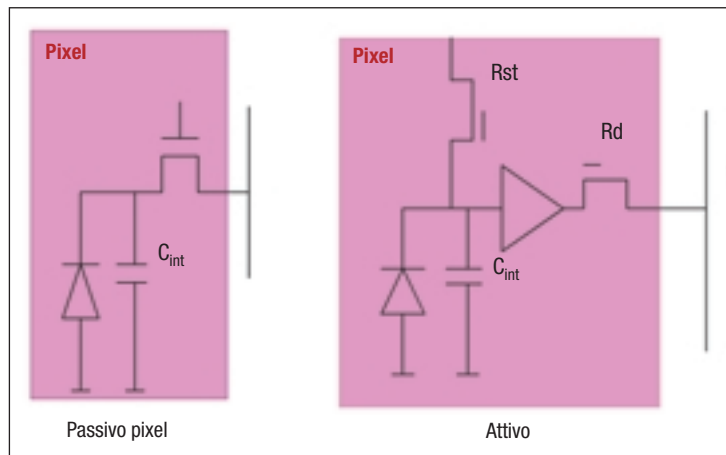
I vantaggi del CCD consistono nell'aver maggiore sensibilità, un trasferimento di carica immune da rumore e, a causa delle barriere di potenziale per il trasferimento di carica, una minore *dark-current* (ossia un flusso di cariche generate dal dispositivo in assenza di segnale). In sintesi si può dire che i CCD hanno una maggiore fedeltà e sensibilità alla luce incidente, con la possibilità di creare immagini con basso rumore e di alta qualità.

I sensori indicati col nome **C-MOS** si basano sullo stesso principio di trasformazione della luce in carica elettrica, ma il segnale generato in ciascuna cella viene trasformato in tensione prima del trasferimento: ne consegue la necessità di introdurre *in loco* un condensatore che trasformi la carica in tensione e, spesso, l'opportunità di introdurre un circuito di amplificazione a livello di *photodetector*, riducendo la percentuale dell'area fotosensibile (tale percentuale è nota come *Fill-Factor*). Il segnale generato viene trasportato all'uscita attraverso percorsi altamente conduttivi e questa è un'altra ragione per un consumo di potenza ridotto rispetto al CCD. Aspetto negativo risulta invece il fatto che l'assenza di barriere di potenziale produce, nei sensori C-MOS una maggiore *dark-current*.

Per quanto detto sopra, come illustrato nella figura 2, i sensori di tipo CMOS possono essere classificati in sensori con *photodetector* di tipo passivo o attivo (APS: *Active Pixel Sensor*). Nei primi il *photodetector* è costituito esclusivamente dal fotodiode, un interruttore ed un condensatore. Nei sensori APS si ha invece anche la presenza di un amplificatore per ciascun *photodetector* ottenendo quindi un maggiore SNR, ma riducendo il *fill-factor*.

Il **grande vantaggio dei sensori a C-MOS** è quello di usare le stesse tecnologie a semiconduttore utilizzate in grande serie per i dispositivi logici, ciò che permette integrazioni sullo stesso *chip* con altra circuitistica.

In sintesi si può affermare che i vantaggi dei sensori CMOS, rispetto ai CCD, risiedono in un minor costo di sistema (dovuto ai più semplici criteri di indirizzamento ed alla integrazione del convertitore analogico-digitale sullo stes-



so *chip*), un consumo di potenza più basso e nell'uso di tecnologie a semiconduttore standard che permettono di integrare più funzionalità (per esempio la parte di conversione A/D e di elaborazione all'interno dello stesso *chip*. Ciò tuttavia a scapito di una maggiore *dark-current*, di un *fill-factor* meno buono e quindi, a parità di area, di una minore sensibilità rispetto ai sensori CCD (per cui si può ricorrere all'uso di microlenti su ciascun photodetector come illustrato più avanti).

Le precedenti considerazioni spiegano il perché della grande diffusione nel campo della telefonia cellulare di fotocamere realizzate con tecnologia CMOS. Il grande miglioramento dei processi produttivi di questo tipo di sensori ha portato al loro uso nelle fotocamere digitali denominate SLR (*Single Lens Reflex*) ossia di fascia alta.

Non vi è dubbio che, nel tempo, i sensori CMOS miglioreranno al punto da essere paragonabili ai sensori CCD per la massima parte delle applicazioni, ma ancora questo traguardo non è stato raggiunto.

Poiché l'architettura ed il *layout* fisico è essenzialmente indipendente dalla tecnologia, per il seguito di questo articolo i due tipi di sensori possono essere considerati sostanzialmente coincidenti.

### 4.3. Tecniche per la cattura del colore nei sensori

Il sensore è costituito da una matrice di elementi fotosensibili a tutte le lunghezze d'onda della radiazione elettromagnetica visibile (ed anche all'infrarosso). Per rendere possibile l'acquisizione di informazioni di colore in modalità tri-

**FIGURA 2**

Sensori con celle passive od attive



cromatica, ciascuno elemento della matrice viene reso sensibile ad una particolare intervallo di lunghezze d'onda con la deposizione di un filtro appropriato su ciascun *photodetector*.

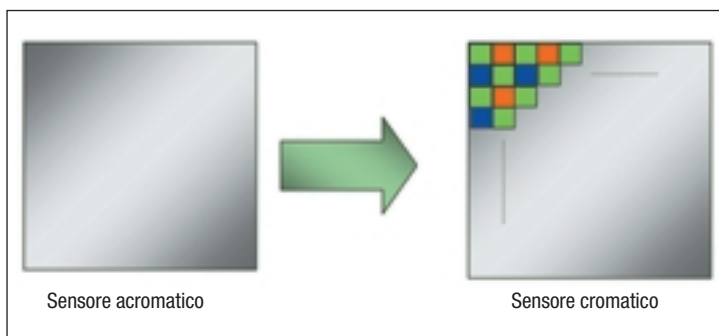
Sono normalmente utilizzate **due classi di filtri**: quelli "trasparenti" a bande spettrali centrate intorno al rosso, verde e blu (colori primari) o quelli "trasparenti" a bande centrate intorno a Ciano, Magenta e Giallo (colori complementari). La prima classe è quella normalmente preferita poiché permette una maggiore fedeltà nell'acquisizione delle informazioni di colori e durante il seguito dell'articolo si farà riferimento solo ad essa. Le camere digitali più popolari utilizzano un solo sensore per l'acquisizione dell'immagine

e su di esso è disposto un filtro a mosaico denominato "**Color Filter Array**" (*CFA*). I filtri sono disposti secondo opportuni mosaici di cui il più utilizzato è noto come "**Bayer pattern**" (dal nome di chi lo ha proposto), caratterizzandosi in realtà con tale nome una famiglia con varianti nel numero dei colori filtrati (3 o 4) e nel tipo di disposizione dei colori sul mosaico. Il più comunemente usato è quello denominato GRGB che ha il 50% dei *pixel* che catturano il Verde (**G**), il 25% che catturano il Rosso (**R**) ed il rimanente 25% che catturano il Blu (**B**) (Figura 3). Questa particolare distribuzione tiene conto della maggiore sensibilità dell'occhio umano alle lunghezze d'onda prossime al Verde. Nella figura 4 sono riportate le risposte spettrali tipiche di una terna di filtri colore: si noti che in realtà i filtri utilizzati sono quattro, poiché deve essere utilizzato anche un filtro IR anti-infrarosso per evitare l'acquisizione di energia ottica relativa all'infrarosso.

Di recente sono state introdotte delle soluzioni alternative al *pattern* di Bayer dove il secondo *pixel* sensibile al Verde è stato sostituito da uno sensibile allo Smeraldo (Sony) nel tentativo di migliorare la qualità dei colori. Sono state anche introdotte particolari strutture di *pixel* per una maggiore definizione e campo dinamico (Fujifilm) in modo da evitare la perdita di dettagli dovuta alla saturazione. Infine sono stati anche proposti nuovi *pattern* con l'introduzione di *pixel* non mascherati con filtri colorati (Kodak) per migliorare entrambi gli aspetti di cui sopra.

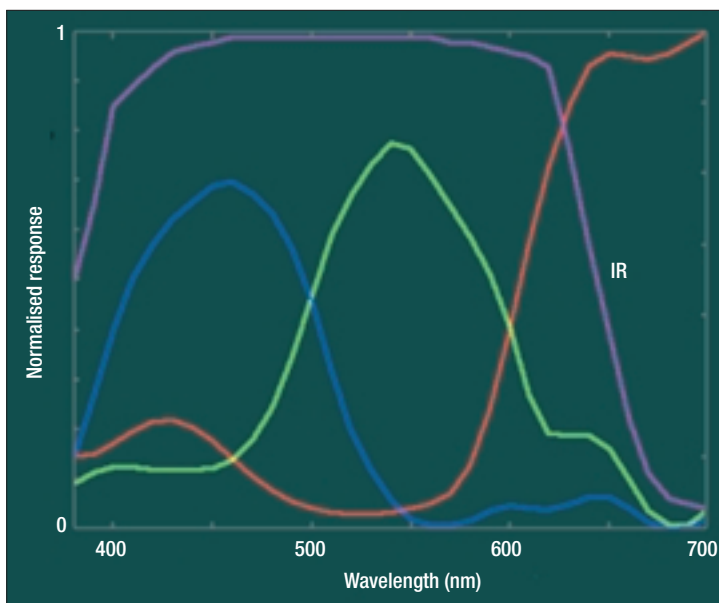
Poiché per la struttura adottata, ogni *photodetector* cattura solo una delle componenti primarie RGB, se si vuole avere in uscita un'immagine con un numero di *pixel* pari a quello dei *photodetector* le altre due componenti cromatiche che concorrono alla formazione di un *pixel* vengono calcolate dal processore d'immagine con l'algoritmo di "**demosaicing**" che consiste nell'applicare un opportuno metodo di **interpolazione** utilizzando anche le informazioni dei valori cromatici dei *photodetector* adiacenti. Con tale procedura si ottengono buoni risultati anche in virtù della minore risoluzione dell'occhio umano alle differenze cromatiche rispetto a quelle di luminosità.

Il processo di "ricostruzione" del colore ed i metodi di bilanciamento del bianco di cui si parlerà poco avanti, assumono un ruolo deter-



**FIGURA 3**

Un sensore fotosensibile acromatico (sensibile cioè a tutte le lunghezze d'onda della luce visibile) si trasforma in un sensore a colori attraverso l'aggiunta sui *photodetector* di filtri di colore RGB (in questo caso secondo il *pattern* di Bayer)



**FIGURA 4**

Risposte dei filtri depositati sui *photodetector* del sensore di immagine

minante nell'ottenere immagini di buona qualità e vengono di conseguenza utilizzati algoritmi di elevata complessità e software proprietari che i costruttori mantengono riservati. Val la pena citare che in commercio esiste un solo sensore, il FOVEON, in grado di catturare tutte e tre le componenti RGB su un unico photosite e che, in quanto tale, è più simile all'occhio umano. La sua diffusione è tuttavia finora abbastanza più ridotta rispetto ai sensori dotati di *Color Filter Array*.

Tecnologicamente legato al deposito dei filtri di colore è l'accrescimento di microlenti su ciascun *photodetector* in modo da focalizzare i fotoni sulla parte sensibile del *photodetector* (Figura 5). Si compensa in tal modo la perdita di sensibilità del sensore conseguente a fill-factor significativamente minori di 1.

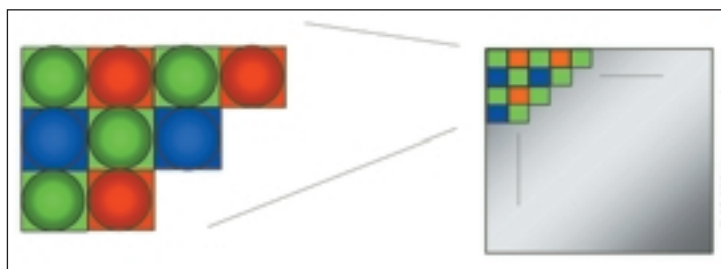
## 5. PROCESSAMENTO DEI DATI FORNITI DAL SENSORE

La figura 6 descrive una tipica catena di elaborazione digitale dei dati forniti dal sensore. Alcuni di essi sono attivi durante la fase di pre-cattura: viene infatti analizzata l'immagine (o il flusso di immagini acquisite dal sensore) estraendone caratteristiche come l'istogramma dei tre canali colore, valori medi

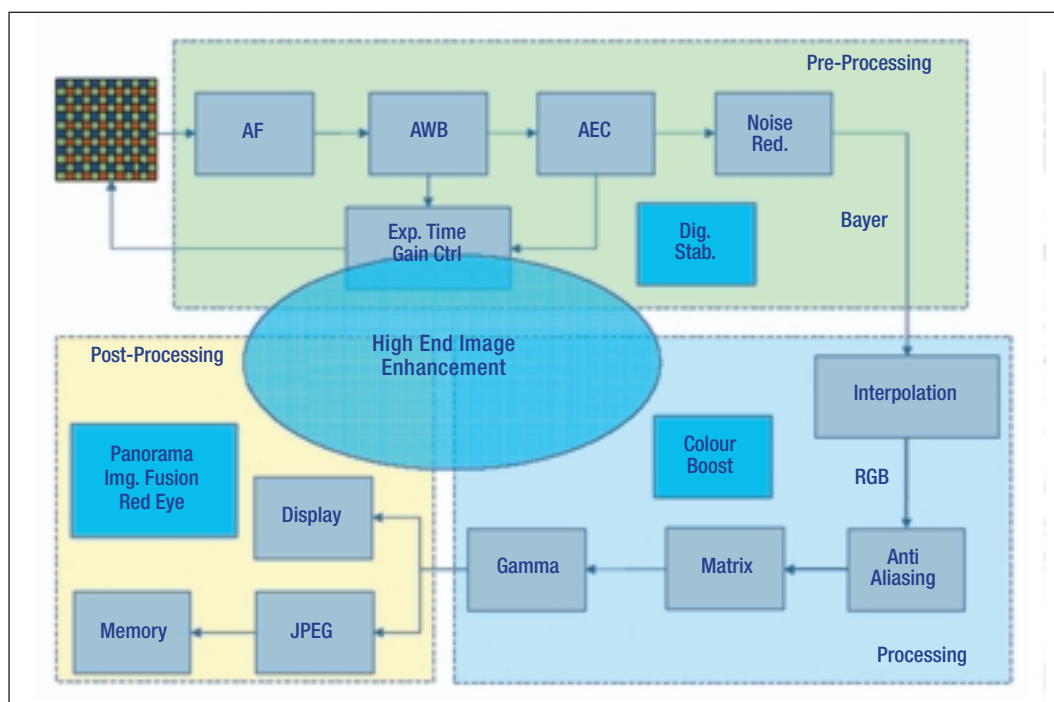
di luminosità, contenuto di alta frequenza ed energia media per canale.

Con tali parametri si opera una serie di operazioni di "pre-processing" per impostare l'**autofocus (AF)**, l'**esposizione automatica (AE)** ed il **bilanciamento automatico del bianco (AWB)** (di cui solo quest'ultimo è tipico delle fotocamere digitali) nonché la **riduzione del Rumore (Noise Reduction)**.

Si ha poi una catena di ulteriori processamenti (escluso il formato RAW cui sopra si è accennato) relativi specificatamente al colore quali la **ricostruzione dell'informazione colore (Pixel Interpolation)**, la **correzione dei contorni (Anti-Aliasing)** per modificare falsi colori, lo **sharpening** per migliorare la nitidezza e la trasformazione "**gamma**" per tener conto della non-linearità dei *monitor*.



**FIGURA 5**  
Micro-lenti disposte sui photodetector del pattern del sensore per aumentarne la sensibilità



**FIGURA 6**  
Tipica catena di elaborazione digitale dei dati forniti dal sensore (Fonte: ST)

Le informazioni relative ad ogni *pixel* vengono infine registrate in un file che possiede le informazioni di tutte e tre le componenti RGB, per subire infine i processi di *compressione* e *memorizzazione*.

### 5.1. Equalizzazione automatica di immagini a colori (bilanciamento del bianco)

Le informazioni di colore associate alle immagini acquisite con tecniche tradizionali (pellicola fotografica) e con tecniche digitali sono influenzate sia dalla riflettanza delle superfici inquadrature, sia, in maniera altrettanto significativa, dal tipo di illuminazione della scena. Uno stesso oggetto appare di colore diverso se osservato in piena luce solare o sotto una luce fluorescente o ad incandescenza.

Il nostro sistema percettivo è tuttavia in grado di compensare, in buona misura, queste variazioni e pertanto la nostra percezione del colore risulta, entro certi limiti, indipendente dalle caratteristiche spettrali della sorgente luminosa che illumina la scena. (un foglio bianco, ad esempio, verrà percepito grossomodo dello stesso colore se visto sotto luce solare o sotto una lampada con filamento al tungsteno, pur essendo gli spettri delle due luci profondamente diversi).

Ovviamente queste capacità di “*chromatic adaptation*” del nostro sistema visivo si attivano quando si osserva una scena reale. Osservando, invece una fotografia, queste differenze ci appaiono in tutta la loro importanza. È quindi importante che le moderne fotocamere digitali ed i programmi di fotoritocco rendano disponibili delle tecniche che simulino, in qualche modo, il comportamento del nostro sistema visivo.

La tecnica più semplice a questo riguardo è il cosiddetto “*bilanciamento del bianco*” che, in ultima analisi, non è altro che l’aggiustamento dei colori di una fotografia in modo da evitare dominanti poco gradevoli. Ovviamente il risultato gradevole può essere anche un fatto soggettivo legato alla creatività e sensibilità del fotografo. Facendo riferimento a quanto illustrato nella prima parte dell’articolo sulla colorimetria, ciò significa riscaldare le componenti di colore in modo che le aree che si ritengono bianche siano rappresentati con valori R, G, B uguali fra loro (e di ampiezza massima se la superficie considerata può essere assunta come la più luminosa possibile).

Un modo oggettivo e semplice di procedere per il bilanciamento è quello di effettuare la **calibrazione a priori del bianco** della fotocamera inquadrando una superficie bianca di riferimento. Il processore della camera legge i valori R, G, B medi dell’immagine e trova i fattori di scala da applicare ad ogni canale di colore per ottenere il colore associato alla superficie bianca osservata sotto la luce di riferimento (per esempio luce solare indiretta). Questi fattori di scala saranno poi automaticamente applicati a tutte le immagini in seguito acquisite. Oltre alla procedura qui menzionata, le fotocamere hanno poi la possibilità di calibrazione del bianco per i tipi di luce più comuni (solare, incandescenza, fluorescente, flash,...). La tecnica è semplice ed interessante, ma richiede di eseguire correttamente questa fase di calibrazione a priori. Se ad esempio si sbaglia il *pre-set* di calibrazione con macchina impostata per luce solare, ma con foto ripresa sotto una luce ad incandescenza, tutto apparirà comunque rosso-giallastro.

Al posto della pre-calibrazione si può adottare una tecnica di **bilanciamento automatico del bianco**, ma in questo caso il problema diviene molto più complesso da risolvere. Infatti, se si fa riferimento alla precedente eq. (3) è facile intuire come una stessa superficie venga acquisita in modo diverso se osservato sotto luci diverse, ma al contempo può avvenire che superfici diverse osservate sotto luci diverse vengono acquisite dalla fotocamera come se fossero dello stesso colore. In genere, perciò, le tecniche di bilanciamento automatico del bianco devono operare sulla base di alcune ipotesi/assunzioni a priori.

A seconda delle ipotesi fatte si hanno tecniche diverse di bilanciamento automatico del bianco, e fra queste le più note vanno sotto i nomi di: “*Grey World*”, “*White Patch*”, “*Gamut Match*”, “*Retinex*”.

Nella tecnica “**Grey World**” si assume che la media dei colori di un’immagine (acquisita con un’illuminazione “canonica”) sia il grigio, da cui il nome dell’approccio. La deviazione da questa situazione è attribuita ad un’illuminazione diversa da quella di riferimento. I fattori di scala (costanti per tutta l’immagine) applicati alle componenti rossa, verde e blu di ogni *pixel* vengono calcolati per riportarsi quanto meglio possibile al “*Grey World*”.

La tecnica del **"White Patch"** assume invece che i *pixel* più luminosi dell'immagine si riferiscano alla ripresa di una superficie bianca (e quindi corrispondano al colore dell'illuminante). In questo caso i fattori di scala applicati ai tre canali di colore sono calcolati in modo da riportare la porzione di immagine considerata ad essere del colore bianco "canonico".

Nel caso del **"Gamut Match"** invece si crea l'istogramma tridimensionale delle componenti RGB di ciascun punto immagine. Si trovano poi i fattori di scala per le componenti colore che portano tale istogramma ad avere la forma più "canonica" possibile.

Infine l'approccio **"Retinex"** (Retina + Cortex) si basa, almeno concettualmente, sulle esperienze di Edwin Herbert Land (fondatore della società Polaroid, 1909 – 1991) sulle modalità di funzionamento del sistema visivo umano. Attraverso di esse Land ipotizzò, infatti, che i meccanismi della visione nascano da una forte interazione fra retina e corteccia cerebrale. Questa tecnica di equalizzazione automatica dei colori si basa sull'ipotesi che l'immagine possa essere vista come il prodotto fra una componente con lente variazioni spaziali (campo di illuminazione) ed una a veloce variazione spaziale (campo delle riflettanze degli oggetti). Le veloci variazioni sono dovute alle brusche transizioni associate al passaggio fra due oggetti diversi. Attraverso opportune tecniche statistiche si cerca di stimare (quindi separare) queste due componenti, si procede poi ad una equalizzazione del campo di illuminazione ed ad una successiva ricostruzione delle immagini. L'uso dell'operatore logaritmo semplifica la procedura trasformando i prodotti (campo di illuminazione per campo di riflettività) in somme. La tecnica *Retinex* viene spesso utilizzata anche come tecnica di equalizzazione delle intensità delle immagini e non solo delle informazioni cromatiche per mettere in risalto particolari altrimenti poco o per nulla visibili.

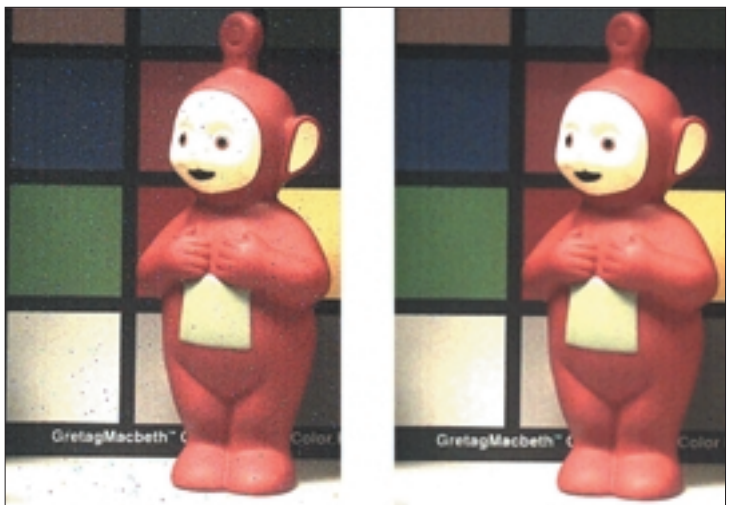
Nella figura 7 è riportato un esempio di correzione cromatica per compensazione dell'illuminante.

All'atto pratico nessuna delle tecniche descritte, e delle molte altre presenti in letteratura, riesce a dare risultati "piacevoli" all'occhio umano in tutte le circostanze. Le moderne fotocamere digitali, dotate di potenti processori, sulla base di *preset* impostati ed attraverso tecniche di ana-



**FIGURA 7**

Effetto dell'applicazione di una tecnica di compensazione dell'illuminante



**FIGURA 8**

Elaborazione dell'immagine per eliminare difetti dei pixel

lisi di immagine, classificano le tipologie di foto che si possono scattare (volto in primo piano, ripresa di gruppo, panorama di mare, panorama con predominanza di prati, foto notturne ecc.). A seguire, caso per caso, si utilizzano tecniche di equalizzazione diverse per ottenere le immagini più gradevoli possibili.

## 5.2. Eliminazione dei pixel difettosi e riduzione del rumore

Le tecniche di **"Defect Correction"** hanno lo scopo di correggere la modalità di funzionamento dei *pixel* che deterministicamente hanno una sensibilità minore o maggiore della media rispetto a quelli con funzionamento "normale". Tali tecniche si basano su algoritmi legati all'analisi del contesto "locale" (e cioè sull'informazione presente in un *pixel* e nel *pixel* immediatamente vicino). Le tecniche utilizzate si basano su una tabella pre-calcolata che contiene la posizione dei *pixel* difettosi oppure su algoritmi che determinano se alcuni *pixel* risultano difettosi durante la cattura dell'immagine (Figura 8).



Le tecniche di eliminazione dei *pixel* difettosi utilizzano massicciamente filtri non-lineari di tipo "mediano". Nel caso più semplice un filtro mediano sostituisce il valore di intensità (o di componente di colore) di un *pixel* corrente con il valore mediano fra tutti quelli associati ai *pixel* di una regione che circonda il *pixel* corrente. Tale regione (che comprende anche il *pixel* corrente) ha spesso la dimensione di  $3 \times 3$  *pixel*.

Il passo successivo è rappresentato dalla riduzione del rumore statistico (**Noise Reduction**). Tale tecnica viene anch'essa, come la precedente, applicata direttamente nel dominio del Bayer pattern perché così facendo i passi successivi possono beneficiare di un'immagine meno rumorosa.

Il rumore può avere distribuzioni statistiche diverse e cambia a seconda della luminosità della scena. Le tecniche di riduzione del rumore sono di tipo non lineare in quanto devono riuscire a distinguere le zone omogenee da quelle contenenti i contorni degli oggetti. Nelle prime il rumore risulta evidente e quindi deve essere eliminato senza rimuovere informazione significativa, in particolare i contorni. Questi ultimi devono invece essere preservati quanto più è possibile perché rappresentano il contenuto di informazione presente nell'immagine.

## 6. COMPRESSIONE E FORMATI DI SALVATAGGIO DELLE IMMAGINI

Come già accennato, i dati che formano le immagini sono codificati ed ordinati sequenzialmente per poi essere compressi in modo da salvarli agevolmente su una memoria. La semplice scrittura dei dati in aritmetica binaria con ordinamento sequenziale (formato non compresso) richiederebbe infatti uno spazio in memoria troppo grande.

Se, ad esempio, desideriamo ottenere, in formato 3:2, un'immagine di  $3.072 \times 2.048$  pixel è necessario adottare un sensore da circa 6 Megapixel. Volendo memorizzare l'immagine in RGB e adottando 24 bit/pixel di dinamica complessiva (ossia 8 bit/pixel e quindi 256 livelli di dinamica per ciascuna delle tre componenti) si richiede, senza compressione, un'occupazione di memoria di circa 18 Mbyte. Tale

occupazione sale a 24 Mbyte se si adottano per la codifica binaria 11 bit, come spesso viene richiesto in immagini professionali destinate alla post elaborazione per rappresentare un numero più elevato di gradazioni di colore. Per ridurre l'ammontare di memoria richiesta da ciascuna foto si ricorre ad **algoritmi di compressione** che riducono la correlazione dei dati. A seconda della qualità desiderata gli algoritmi di compressione possono essere **loss-less**, **lossy** e **visually loss-less**.

Le tecniche di codifica **loss-less** prendono anche il nome di codifiche entropiche e sono intrinsecamente senza perdita (ovvero dai dati codificati è sempre possibile la ricostruzione integrale dell'immagine originale). Una buona codifica entropica si basa su una rappresentazione ed ordinamento dei dati nel modo più efficiente possibile, sfruttando le proprietà statistiche che i dati medesimi possiedono in un'immagine reale. Su una foto reale ciò porta ad una compressione dei dati di circa 2 volte. Con gli algoritmi **lossy**, dopo la decompressione, si ottiene invece un'immagine diversa dall'originale, ma che risulta "accettabile" all'utente finale. A questa classe appartengono anche gli algoritmi di compressione **visually loss-less** con i quali si ottengono delle differenze tra originale e decodificato che non sono visibili attraverso l'ispezione visiva, ma solo attraverso comparazione numerica. Un esempio di **visually loss-less** è costituito dal passaggio da uno spazio di colore RGB a quello delle luminanze e crominanze, seguito dal sottocampionamento delle crominanze perché queste vengono percepite in maniera ridotta dal sistema visivo umano. A seconda delle codifiche adottate si perviene ad una compressione dei dati pari a 1,5 volte (nel cosiddetto formato 4:2:2) od a 2 (nel formato 4:2:0).

Le tecniche universalmente usate, con molte differenti varianti, per una forte compressione sono però quelle basate su opportune "trasformate" seguite da quantizzazione e codifica entropica, processo che, nel suo complesso, elimina le irrilevanze e le ridondanze dell'immagine. Queste tecniche portano ad un risparmio pari ad un fattore di 20-30, pur mantenendo una buona qualità. Un esempio di risultato di compressione **lossy** è dato nella figura 9 per la fotografia di un





FIGURA 9

Esempio di compressione (30 volte) della fotografia di un francobollo.

(Si può notare che l'informazione persa con la compressione JPEG riguarda essenzialmente i contorni degli "oggetti". Tuttavia la qualità finale rimane più che accettabile per un uso amatoriale)

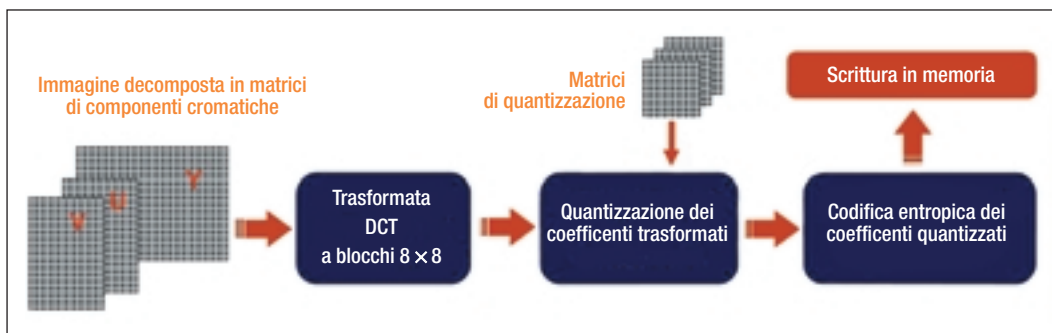


FIGURA 10

Schema di elaborazione digitale operato dalla compressione JPEG

francobollo ed ove è utilizzata una compressione di tipo JPEG.

Il **formato di compressione** più noto ed adottato nel campo della fotografia digitale è il **JPEG**, basato sulla DCT (*Discrete Cosine Transform*), ormai universalmente utilizzato nel campo del *visually loss-less* e che adotta lo schema di elaborazione riportato in modo schematico nella figura 10.

Lo standard JPEG ha rivoluzionato, dagli anni '80 in poi, il trattamento delle immagini aprendo la strada alla diffusione della fotografia digitale. Regolando il rapporto di compressione si può ottenere una qualità *visually loss-less*, rappresentando una foto reale con  $2 \div 4$  bit/pixel. Se, si aumenta il rapporto di com-

pressione (pur mantenendo l'immagine accettabile) si ottiene una qualità *lossy* che può essere rappresentata con 1 bit/pixel o meno.

È opportuno notare che, a parità di fattore di compressione, si può ottenere una qualità diversa per tipi di immagini differenti. In particolare possono comparire, in alcuni casi, i cosiddetti "artefatti" tipici della compressione a blocchi basata sulla trasformata DCT, quale l'effetto di "*blocchettamento*" dell'immagine ed il cosiddetto *ringing* attorno ai contorni. Le prestazioni delle tecniche di compressione sono, infatti, fortemente dipendenti dall'entropia dei dati da comprimere.

Per ovviare a tutto ciò ed ottenere una migliore qualità sono state sviluppate delle tec-

niche che permettono, a parità di fattore di compressione, di **selezionare i parametri dell'encoder** a seconda del tipo di immagine. In aggiunta allo standard JPEG è stato sviluppato e proposto il nuovo standard **JPEG-2000**, molto più flessibile (in cambio di una maggiore complessità realizzativa) del JPEG, che tuttavia al momento non ha ancora trovato la "killer application" sul mercato.

In sintesi i **formati normalmente usati nelle fotocamere digitali per il salvataggio** delle immagini sono:

❑ **BMP**: formato di salvataggio poco utilizzato poiché il file è di dimensioni piuttosto elevate in cui le immagini possono essere salvate a 16, 24 e 32 bit/pixel senza nessun tipo di compressione;

❑ **TIFF**: formato in grado di salvare immagini senza perdita d'informazioni in uno spazio relativamente contenuto di compressione;

❑ **JPEG**: il più usato nelle fotocamere anche se è del tipo *lossy* (e pertanto con perdita di informazioni);

❑ **RAW**: formato "grezzo" utilizzato dai professionisti che sta ad indicare che l'immagine catturata dal sensore CCD o CMOS della macchina fotografica non viene processata in nessun modo dalla fotocamera, a parte la conversione Analogico/Digitale. Le immagini vengono ricomposte e regolate in quanto a profondità colore con potenti programmi per PC ed infine convertite in uno qualsiasi dei formati conosciuti. I *file RAW*, partendo dai dati grezzi, consentono di apportare in un secondo momento dei miglioramenti significativi alla qualità dell'immagine scattata, riportandosi alle condizioni iniziali di scatto. Si può, per esempio, aggiustare il bilanciamento del bianco, ridurre eventuali aberrazioni cromatiche degli obiettivi, ottimizzare l'esposizione ecc.. I *file RAW* hanno anche una grande profondità cromatica (fino a 16 bit per *photodetector* ossia 48 bit per i tre canali di colore) che permettono "lavorazioni" anche sostanziali sulle immagini senza alterarne la qualità.

La scrittura e lettura dei *file* in formato RAW è molto più lenta rispetto a quella in formato JPEG per la maggiore quantità di dati da scrivere, rendendo difficoltosa la loro archiviazione e la loro successiva visione. Alcuni produttori di fotocamere digitali hanno inserito il doppio formato di registrazione negli appa-

recchi. Questo consente di leggere l'immagine registrata in formato JPEG con una buona velocità (per esempio, nelle operazioni di selezione ed archiviazione delle immagini), mentre è sempre possibile poter utilizzare il formato RAW per le elaborazioni complesse. Per elaborare un file RAW al computer occorre evidentemente un software adeguato che, tra l'altro, compie anche tutte le operazioni del processore d'immagine della fotocamera.

Nelle fotocamere, i file rappresentanti le immagini vengono immagazzinati in memorie EProm di tipo Flash, gestibili dall'utente.

Quelle oggi utilizzati dalle case costruttrici di fotocamere digitali sono principalmente:

- *Compact flash*;
- *Memory Stick*;
- *Microdrive*;
- *MultiMedia (MMC)*;
- *SD (secure digital)*;
- *SmartMedia*;
- e *Fuji*.

A conclusione di questo quadro è opportuno notare che il vantaggio di un formato standard per i file d'immagine è anche quello di promuovere l'interoperabilità di dispositivi diversi (Camere, PC, Lettori DVD, Stampanti ecc).

## 7. CONCLUSIONI

Per quanto riguarda i sensori, le moderne macchine fotografiche hanno raggiunto ormai dimensioni pari a molti *Megapixel* a confronto della fovea dell'occhio umano che ha "solo" 337.000 sensori. Ciononostante il confronto più significativo va fatto con la pellicola per permettere gli stessi ingrandimenti dei "particolari" di una fotografia. E solo recentemente i sensori hanno raggiunto, a costi accettabili, una definizione paragonabile a quella della pellicola.

I successivi processamenti dei dati fotografici applicati dai diversi costruttori di macchine fotografiche e così pure le elaborazioni con i programmi di fotoritocco portano, come già è stato accennato in precedenza, a quello che noi "vogliamo vedere" in termini di resa e saturazione cromatica e non corrispondono sempre alla realtà oggettiva.

Nel campo *prosumer*, le due grandi famiglie di sensori (CCD e CMOS) stanno man mano avvicinandosi come caratteristiche ed i secondi

Segmenti di mercato	Tipi di sensore	Risoluzione sensore	Ottica	Prestazioni	Range di costo
DSC Low-End	CCD Bayer Size 5.76 × 4.29 mm 3 μm pixel size	3-5 Mpixel	35 mm (equivalente) Optical zoom 3-4x	AF-TTL 5-Points VGA video Multi-Shot 1.5 FPS	< 200 \$
DSC High-End	CCD Bayer Size 7.6 × 5.7 mm 5 μm pixel size	5-12 Mpixel	35 mm (equivalente) Optical zoom 3-4x	AF-TTL 9-points XGA video Multi-Shot 2.5 FPS Image Stabilisation Face Tracking	300-500 \$
SLR Low-End	CCD / CMOS Bayer Size: 18 × 13.5 mm - - 23.7 × 15.5 mm 5.7 μm pixel size	10 Mpixel	Separate Optics (optical stabilisation)	AF Multi-Basis TTL, 9-p Advanced AWB & AEC Multi-Shot 3 FPS - RAW Sensor Dust Removal Low Noise	< 1.000 \$
SLR High-End	CMOS Bayer/ Multi-Layer Full size: 36 × 24 mm 6.4 μm pixel size	21 Mpixel	Separate Optics (Optical Stabilisation)	AF Multi-Basis, 45-P Advanced AWB & AEC Multi-Shot 8.5 FPS-RAW Cleaning System FULL Params Control Very Low Noise Wireless Connection	1.000-9.000 \$

Segmenti: DSC: "Digital Still Camera" - SLR: "Single Lens Reflex" - Low End: fascia bassa; High End: fascia alta  
Altri nomil: Bayer (vedi paragrafo 5.3); AF-TTL: "Auto Focus Through The Lens"; AWB: "Automatic White Balance"

**TABELLA 1**

Confronto di caratteristiche e costi nel mercato delle fotocamere digitali

permettono di trarre a costi minori ed ad una migliore integrazione monolitica.

Interessanti da seguire l'evoluzione del **sensore Foveon** che maggiormente assomiglia all'occhio umano, anche se ha ancora problemi di passaggio della luce attraverso gli strati di silicio.

Sempre nell'interessante segmento del mercato *prosumer* delle fotocamere sono molto interessanti le ottiche compatte e leggere che si riescono a realizzare, nate a seguito della introduzione dei sensori che normalmente hanno una minor dimensione rispetto al fotogramma della pellicola, anche se tali ottiche rimangono inferiori a quelle intercambiabili.

Nella tabella 1 è riportato un confronto delle caratteristiche di diversi segmenti di mercato delle moderne fotocamere digitali per avere un quadro comparativo delle prestazioni attuali (dimensioni, consumo, costo...).

Dal **punto di vista del mercato**, impressionante è la crescita della penetrazione dei sensori di immagine nel mercato dei *camera phones* con possibilità di acquisire foto e video. Si vuole foto-

grafare ogni attimo della nostra vita e della vita di tutti in tempo reale. Si vive nella civiltà dell'immagine reale, ma il rischio è di non immaginare più nulla e di vedere solo quello che telefoni o macchine fotografiche registrano.

A seguito dell'evoluzione della tecnologia, anche la macchina fotografica "seria" sta diventando un prodotto di consumo ad obsolescenza rapidissima, a causa dell'introduzione di sempre maggiori risoluzioni ed ad altre diverse funzionalità (panorama, riduzione occhi rossi, *antishake* ecc.).

Tutto ciò può lasciare disorientato il consumatore e solo una buona conoscenza della tecnologia delle fotocamere digitali e delle proprie esigenze (in termini di risoluzione e funzionalità) permette di fare delle scelte oculute ed oggettive.

La fotografia digitale offre anche immense possibilità che però non vengono ancora sfruttate nella maggioranza dei casi perché è tuttora limitato il numero di persone che, pur essendo i programmi *software* da impiegare assai *friendly*, si dedicano alla successiva fa-

se della elaborazione che è importante per creare un'immagine.

A tale proposito, si può infine aggiungere una considerazione sulla fotografia artistica in quanto la moderna tecnologia ICT di procesamiento delle immagini offre possibilità che la fotografia analogica non offriva. Ora il fotografo, attraverso il fotoritocco, ha la possibilità di creare immagini intervenendo sul cromatico, l'inquadratura, la saturazione dei

colori, il contrasto e realizzando così quello che la sua cultura, le sue emozioni, la sua visione degli eventi e delle situazioni gli suggeriscono. Può abbandonare vecchi schemi mettendosi continuamente in discussione e sperimentare nuove strade.

Dal complesso di queste brevi considerazioni conclusive si può pertanto affermare che la fotografia digitale è una vera rivoluzione anziché una semplice evoluzione.

## APPENDICE 1

### Color matching

Un osservatore, con un sistema visivo che rientra nella media, percepisce la stessa sensazione di colore (*color match*) guardando un colore di test oppure la somma di tre colori primari opportunamente dosati. Questa corrispondenza esiste anche per colori monocromatici caratterizzati da una specifica lunghezza d'onda.

In ogni caso i tre primari possono essere monocromatici, come è il caso della figura A1.1, o caratterizzati da un loro spettro, e vengono fissati dal CIE (*Commission Internationale de l'Eclairage*), come richiamato nella successiva **Appendice 2**.

Sperimentalmente, in un banco di prova con più osservatori, è stata valutata l'equivalenza di una luce di test monocromatica con la somma pesata di tre primari Rosso, Verde e Blu (Figura A1.1). Il valore medio delle sensazioni degli osservatori sulle pesature dà luogo, in funzione della lunghezza d'onda del colore di test, a tre curve definite "**Color Matching Function**".

Le funzioni  $r(\lambda)$ ,  $g(\lambda)$ ,  $b(\lambda)$  misurano i cosiddetti **valori** (o **componenti**) di **tristimolo** dei primari che realizzano l'equivalenza con il colore di test per ciascuna lunghezza d'onda esaminata. Valori negativi delle "**Color Matching Function**" equivalgono a dire che non esiste, per quel campo di valori di lunghezza d'onda, un'equivalenza tricromatica nel senso fisico della parola, ma una delle luci primarie va sommata alla luce di riferimento per ottenere l'equivalenza percettiva (cfr, eq. (1) e (2) del testo che riassume le leggi di Grassman). Conseguenza di ciò è che non tutti i possibili colori possono essere rappresentati a partire da un set di tre soli primari.

Le funzioni  $r(\lambda)$ ,  $g(\lambda)$ ,  $b(\lambda)$  misurano i cosiddetti **valori** (o **componenti**) di **tristimolo** dei primari che realizzano l'equivalenza con il colore di test per ciascuna lunghezza d'onda esaminata. Valori negativi delle "**Color Matching Function**" equivalgono a dire che non esiste, per quel campo di valori di lunghezza d'onda, un'equivalenza tricromatica nel senso fisico della parola, ma una delle luci primarie va sommata alla luce di riferimento per ottenere l'equivalenza percettiva (cfr, eq. (1) e (2) del testo che riassume le leggi di Grassman). Conseguenza di ciò è che non tutti i possibili colori possono essere rappresentati a partire da un set di tre soli primari.

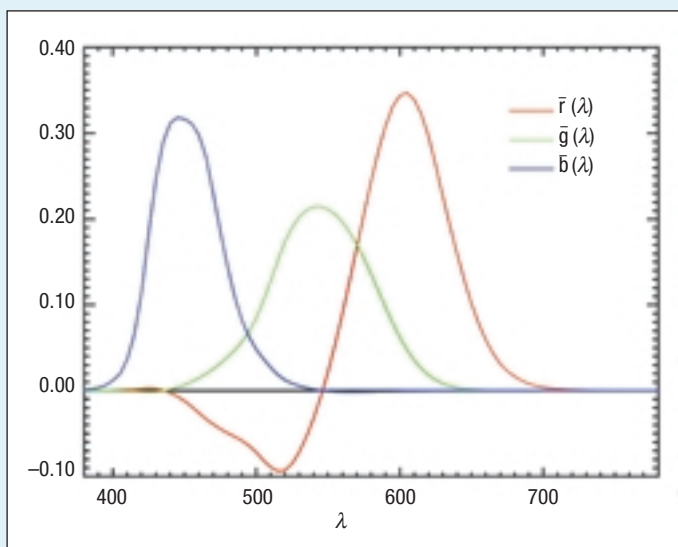


Figura A1.1 In questa figura le "Color Matching Functions" rappresentano, sull'asse delle ordinate, le entità ("valori di tristimolo") di tre luci primarie monocromatiche (Rosso, Verde e Blu con lunghezze d'onda standardizzate rispettivamente di 700, 546 e 436 nm) necessarie per l'equivalenza ad un colore di test monocromatico la cui lunghezza d'onda  $\lambda$  è quella indicata nella scala delle ascisse. Valori negativi di una delle curve di tristimolo indicano la necessità di aggiungere il relativo primario al colore di test per avere l'equivalenza. Le tre curve di questa figura (denominate "CIE-1931-RGB Color Matching Functions") sono normalizzate in modo da avere tutte e tre la stessa area.



## APPENDICE 2

### Spazi cromatici

Le curve indicate nella figura A1.1 dell'Appendice 1, caratterizzate da primari monocromatici alle lunghezze d'onda di 700 nm (rosso), 546.1 nm (verde) e 435.8 nm (blu), sono quelle storiche derivate dagli esperimenti di D. Wright e J. Guild alla fine degli anni '20 e recepiti in ambito CIE (*Commission Internationale de l'Eclairage*) nel 1931 con il nome di "Color Matching Function CIE-1931-RGB". Questo spazio, anche se non è in grado di rappresentare tutti i colori, ha un **gamut sufficientemente esteso**.

Esistono molteplici spazi cromatici in cui poter rappresentare le informazioni di cromaticità, ciascuno dei quali possiede proprie peculiarità.

Oltre lo spazio CIE-1931-RGB, è significativo ricordare lo **spazio sRGB**, leggera modifica dello spazio precedente, utilizzato attualmente per rappresentare i colori associati alle immagini acquisite dalle moderne **fotocamere digitali** in cui l'illuminante di riferimento è il D65 (luce bianca di riferimento definita dal CIE per rappresentare la luce solare indiretta) e che include una correzione gamma delle componenti di colore per favorire un *rendering* sui monitor che sono dispositivi di riproduzione fortemente non-lineari. Il gamut è sufficiente ampio e questo spazio è diventato uno standard de-facto (la s di sRGB sta per standard) per le fotocamere digitali con presentazione delle immagini su monitor da computer.

La rappresentazione di tutti i colori è invece raggiunta nello spazio C.I.E. XYZ derivato matematicamente da quello visto in precedenza e che, tuttavia, possiede dei primari non fisicamente realizzabili. Poiché lo scopo era quello di un metodo di rappresentazione matematica efficace per tutti i colori, questo limite non deve preoccupare più di tanto.

In generale il passaggio fra uno spazio di colore ed un altro avviene attraverso una trasformazione lineare (moltiplicazione del vettore a tre dimensioni che rappresenta un colore per una matrice  $3 \times 3$  opportunamente definita).

Quando si vogliono rappresentare i colori in uno spazio che sia indipendente dal particolare dispositivo di acquisizione o restituzione si fa in genere riferimento al cosiddetto spazio XYZ. Al contrario di quanto accaduto per la definizione delle *Color Matching Function* (CMF) nel sistema RGB, le CMF dello spazio XYZ (riportate nella Figura A2.1) non sono state individuate tramite esperimenti psicovisivi bensì attraverso una trasformazione delle corrispondenti curve del sistema RGB misurate nel 1931. La trasformazione utilizzata è stata individuata, andando alla ricerca di primari che potessero soddisfare alle seguenti condizioni:

a) fare in modo che le *Color Matching Function* XYZ non avessero valori negativi, caratteristica necessaria per soddisfare il requisito di riproducibilità di tutta la gamma dei colori nella fascia del visibile;

b) valore di tristimolo Y scelto in modo tale da essere equivalente alla luminanza del colore rispetto cui ottenere equivalenza visiva;

c) coordinate del bianco equienergetico nel **triangolo dei colori** (vedi dopo Figura A2.2) pari ad  $1/3$  per le tre coordinate.

È possibile rappresentare le informazioni contenute nei valori di tristimolo XYZ in maniera non assoluta, ma normalizzate in un campo di valori che va da 0 a 1, utilizzando le coordinate cromatiche (o coefficienti di tristimolo)  $x$ ,  $y$ ,  $z$ . Queste ultime sono legate ai valori assoluti delle componenti di tristimolo tramite le seguenti relazioni:

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

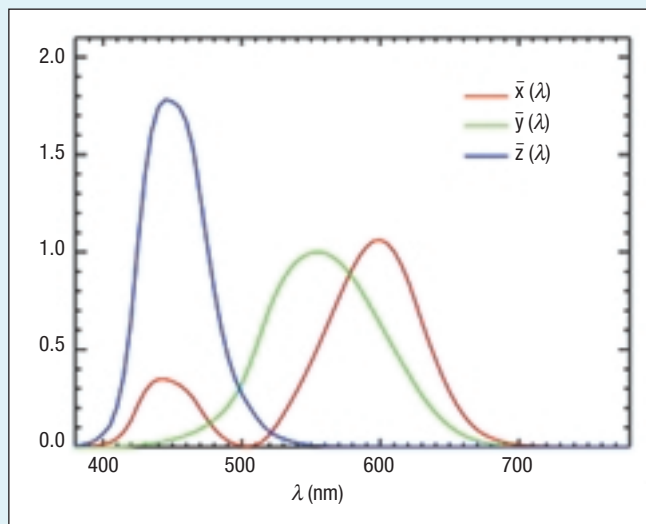


Figura A2.1 "Color Matching Functions" per i primari CIE-XYZ (non fisicamente realizzabili)



Dalle precedenti relazioni segue che:

$$x + y + z = 1$$

Si noti che poiché la somma di questi coefficienti vale 1, sono sufficienti due di essi, per esempio,  $x$  e  $y$ , per definire una luce monocromatica, in quanto il terzo si ricava per complemento a 1.

Poiché le coordinate di cromaticità risultano linearmente dipendenti, dallo spazio  $(x, y, z)$  non si è in grado di risalire alle coordinate  $(X, Y, Z)$ . Per questa ragione si considera equivalente al sistema di riferimento XYZ quello costituito dalla luminanza  $Y$  e da due coordinate di cromaticità, in particolare  $x$  e  $y$  (CIE- $xyY$ ). In questo modo si tende a separare l'informazione di intensità (luminanza) rispetto alle informazioni di colore (cromaticità). Nella figura A2.2 è rappresentato il diagramma di cromaticità, denominato triangolo dei colori, ottenuto considerando le sole coordinate  $(x, y)$ .

La linea curva superiore è nota con il nome di *spectrum locus* e rappresenta il luogo dei punti di  $x$  e  $y$  corrispondenti a **luci monocromatiche pure** (righe spettrali). La linea retta tracciata alla base del "ferro di cavallo" rappresenta invece i valori  $x, y$  corrispondenti a qualunque combinazione tra 0 e 100% del rosso e del blu. Questa linea viene definita la purple line dal momento che lo stimolo ottenuto è di varie gradazioni del colore purpureo.

**Tutti i colori sono rappresentabili come punti interni allo "spectrum locus"** e la zona centrale rappresenta il bianco. La posizione di ogni punto (colore) sulla retta che lo congiunge con il bianco puro rispetto all'intersezione con la curva dà la **saturazione** del colore considerato, mentre le variazioni angolari del fascio di rette che partono dal bianco dà la variazione di **tinta**. Per la rappresentazione dell'intensità del colore occorrerebbe sfruttare anche la terza dimensione.

Tutti i colori visibili hanno nello spazio  $xy$  una rappresentazione che cade all'interno dell'area definita dallo *spectrum locus* e dalla *purple line*. Nello spazio  $xy$  sono rappresentate le informazioni cromatiche associate ad un particolare colore, mentre le informazioni di intensità sono associate a  $Y$ .

Per le fotocamere digitali, nel caso di utilizzo del sistema di primari RGB, già menzionato poco sopra, la rappresentazione del **"gamut" del sistema** nello spazio  $xy$  è descrivibile unendo i valori di lunghezza d'onda di tali primari che danno luogo al **triangolo con linea continua** della figura A2.2.

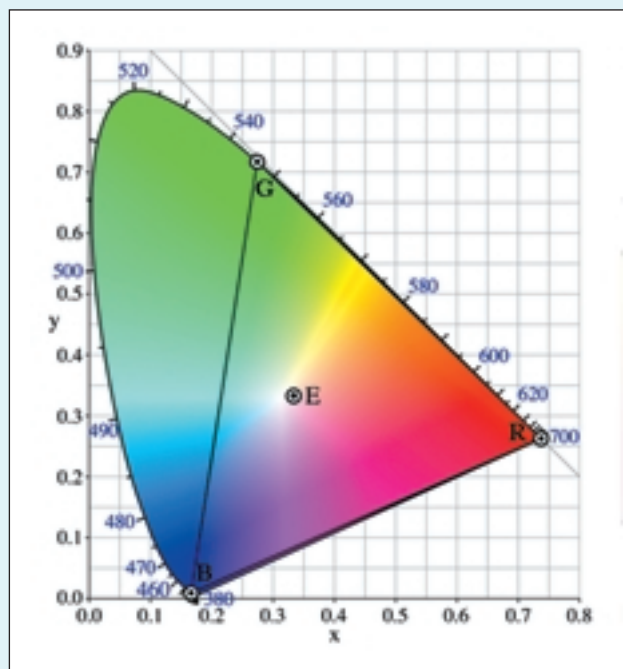


Figura A2.2 Triangolo dei colori nello spazio  $x, y$  (spectrum locus) con E punto del bianco.

La linea curva, a ferro di cavallo, soprastante il segmento B-R rappresenta i colori monocromatici puri, mentre il segmento B-R (detto "purple line") rappresenta tutte le combinazioni del Rosso e del Blu. I vertici del triangolo RGB rappresentano i primari CIE-1931. L'area interna al triangolo rappresenta il "gamut" dei colori riproducibili con tali primari. Il punto E è la rappresentazione del bianco puro. Il diagramma mostra come colori monocromatici nell'intervallo 440-550 nm non sono riproducibili in tricromia in quanto esterni "al gamut" del sistema di primari RGB scelti (come risulta anche da Figura A1.1)

## Bibliografia

### Sulla colorimetria:

- [1] Hunt R.W.G.: *Measuring Color*. Fountain Pr Ltd; 3-rd edition, 2001.
- [2] Wyszecki G., Stiles W.S.: *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley-Interscience - 2-nd edition, 2000.
- [3] IEEE Signal Processing Magazine: *Special Issue on Color Image Processing*. January 2005.

### Sui sensori:

- [5] Litwiller Dave: *CCD vs. CMOS: facts and fiction*. Photonics Spectra, January 2001.

### Sul processamento dei colori nelle fotocamere digitali:

- [6] Gunturk B.K., Glotzback J., Altunbasak Y., Schaffer R.W., Mercerau R.M.: *Demosaicing: Color Filter Array Interpolation*. IEEE Signal Processing Magazine, January 2005.

### Sui vari tipi di processamento nelle fotocamere digitali:

- [7] Kawamura S.: *Capturing Images with Digital Still Cameras*. IEEE Micro, December 1998.
- [8] AAVV: *Image Processing for Digital Still Camera*. ST Journal of System Research, Dec. 2001.

MASSIMO MANCUSO, laureato con lode in Ingegneria Elettronica all'Università di Palermo nel 1990. Inizia la carriera professionale al COMILITER (Consorzio per la Ricerca Microelettronica nel Mezzogiorno) occupandosi di elaborazione di immagini. Dal 1994 è in STMicroelectronics con responsabilità sulla definizione di metodologie ed architetture per i dispositivi di cattura d'immagini. Dal 2004, operando nella sede di Agrate di STM, è Direttore del Team di Ricerca su "Imaging Rendering and Human Machine Interaction" nel Gruppo "Advanced System Technology".

È autore di diversi brevetti ed articoli nel campo dell'Image -Video Processing.

E-mail: massimo.mancuso@st.com

STEFANO TUBARO è Professore Straordinario di Telecomunicazioni presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano (DEI-PoliMi). Svolge attività di ricerca nell'ambito dell'elaborazione numerica dei segnali, in particolare mono e bidimensionali (audio ed immagini).

Coordina le attività del gruppo ISPG (*Image and Sound Processing Group*) presso il DEI-PoliMi e partecipa (anche in veste di coordinatore) a numerosi progetti di ricerca finanziati dal Governo Italiano, dalla Commissione Europea e da enti industriali.

Autore di più di 200 contributi a congressi e riviste internazionali e di numerosi brevetti.

E-mail: stefano.tubaro@elet.polimi.it

GUIDO VANNUCCHI, laurea in Ingegneria Industriale all'Università di Bologna nel 1958, "Master Science EE" alla Stanford University nel 1963, Libera Docenza in Comunicazioni Elettriche nel 1971.

Direttore Generale Telettra dal 1983 al 1990, "Senior Consultant" di Olivetti Telemidia, Vice Direttore Generale della RAI dal 1996 al 1998. Attualmente Docente al Politecnico di Milano di "Architetture per reti e sistemi multiservizio" e Presidente Associazione AEIT.

Laurea "ad honorem" in Ingegneria delle Telecomunicazioni, conferita dall'Università di Padova nel 1998, per i contributi scientifici e manageriali nel campo della trasmissione dei segnali.

E-mail: guido.vannucchi@fastwebnet.it