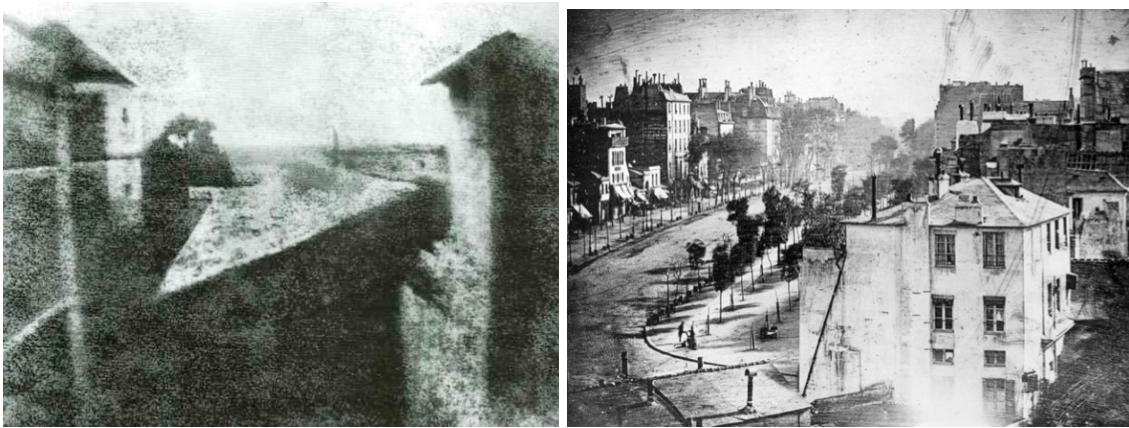


Capitolo 1

Introduzione

Storia della fotografia digitale

La **fotografia** nasce innanzitutto come passatempo e come divertimento per persone facoltose con molto tempo libero. All'inizio si trattava del tentativo di catturare e fissare un'immagine: sul catturare non c'erano problemi, i guai arrivavano quando si tentava di fissare l'immagine al supporto (fino all'800 era impossibile far durare una fotografia a lungo e presto l'immagine svaniva). Le foto si facevano impressionando uno strato di sali d'argento che, colpiti dalla luce, scurivano: così nasceva il **negativo**. La prima fotografia in quanto tale al mondo ha 150 anni ed è dovuta a Niepce, tuttavia le fotografie non durano in eterno (nemmeno quelle moderne, a prescindere dal formato cartaceo o digitale).



La prima foto al mondo. Il Getty Conservation Institute di Los Angeles ha annunciato di aver scoperto il metodo che avrebbe permesso a Joseph Nicéphore Niepce di realizzare la prima fotografia al mondo, comunemente identificata con quel «Point de vue du Gras» (qui sopra) datato 1826. Secondo quanto riferito dal «Washington Post», Niepce avrebbe utilizzato in particolare un procedimento denominato «eliografia», usando un tipo di catrame chiamato «bitume della Giudea», che è fotosensibile. Per Dusan Stulik (responsabile dell'istituto), Niepce avrebbe poi pennellato questo bitume e olio di lavanda sulla lastra di peltro, facendolo asciugare con una vernice trasparente. Quindi, avrebbe inserito la lastra in quella che sarà poi chiamata «camera oscura»; da un piccolo foro sul fondo di questa «camera», la luce avrebbe successivamente colpito per otto ore la lastra, lasciando infine la sua impronta. Il suo procedimento verrà perfezionato da un amico di Niepce, Louis Daguerre. Che, per fissare l'immagine, utilizzerà il vapore di mercurio riuscendo così a produrre fotografie in venti minuti invece che in otto ore. (archivistorico.corriere.it)

Ricopiare le fotografie è l'unico modo per poterle preservare in eterno. Con il passare del tempo per fissare le fotografie si è capito che bisognava “svilupparle” andando così a modificare la struttura dei sali d'argento.

La **camera oscura** è nota fin dal '700 tant'è che la usavano anche i pittori. Si tratta di una sorta di scatola su una parete della quale è effettuato in forellino: la luce entra dal foro e si proietta sulla parete opposta in maniera capovolta (oltre che leggermente depotenziata) → ricordiamo che la luce non è visibile, lo è solo un oggetto illuminato. La camera oscura è detta *pinhole camera* ed è facilmente riproducibile: ha la capacità di mettere a fuoco qualsiasi scena poiché non fa uso di lente alcuna.

Nel periodo tra il 1850 ed il 1920 la fotografia è un passatempo per facoltosi

disoccupati constando di un procedimento lunghissimo e costosissimo. Per produrre una fotografia era necessario rimanere in posa all'inizio per ore, poi solo per diversi minuti, e la fotografia finale veniva composta da una serie di lastre fotografiche. Il fotografo era anche un pittore: nel 1853 si faceva larghissimo uso di grandissime quantità di negativi per ottenere una singola fotografia che alla fine si imprimeva su una lastra di vetro → possedere una fotografia era come possedere un quadro, anche e soprattutto perché la fotografia era considerata un'arte alla stregua della pittura, con la stessa cura nella ricerca dei soggetti e nella composizione delle scene.

Nel 1920 **Mees** utilizza il primo negativo su supporto plastico abbandonando così quello vitreo. Assieme a **George Eastmann** fonda la *Kodak* facendo sì che la fotografia diventasse un passatempo alla portata di tutti. Siccome non era proponibile continuare a fare delle foto per le quali servissero ore intere di posa si inventò una pellicola che rese scattare una foto una questione di un istante (bastava spingere un pulsante). Le macchine fotografiche erano ancora costose poiché fatte artigianalmente in legno e pellame (Brownie Camera): la pellicola andava poi sviluppata in negozio.

Dal 1925 per la fotografia non ci furono passi in avanti significativi salvo adattamenti e raffinamenti. La produzione della pellicola era massiva ed una volta avviati gli impianti si producevano km e km di pellicola senza poter interrompere il ciclo produttivo (un po' come con gli altoforni), così era difficile calare la produzione per adeguarla ad una richiesta minore. Oggi il cinema utilizza ancora la pellicola (in grandi quantità) per girare film, ma non usa più per scattare foto, peccato che sia uno dei metodi più sicuri per conservare immagini.

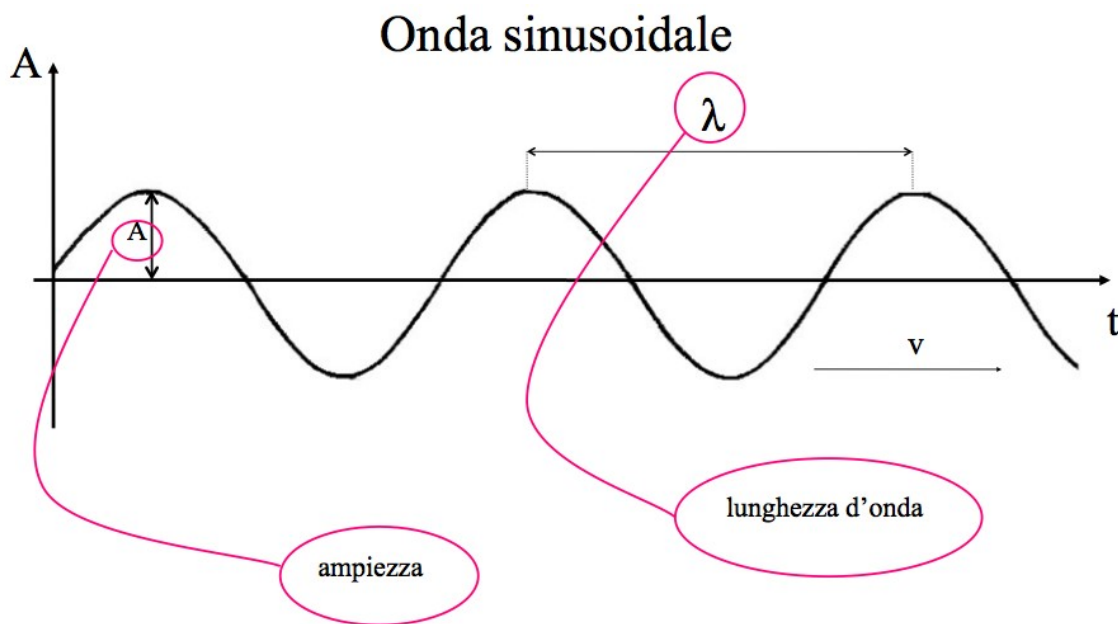


Dal 1970 avviene la rivoluzione: è di quell'anno la prima fotografia digitale (ed il primo apparato per la fotografia digitale che salvava su cassetta magnetica). La Kodak è stata ovviamente la prima a produrre macchine fotografiche digitali, tuttavia è fallita ugualmente: la Sony l'ha quindi soppiantata.

Colorimetria

La **colorimetria** è connessa al sistema visivo: dopotutto il colore non esiste e viene percepito solo come risposta del nostro sistema visivo ad un certo stimolo luminoso. Servono: una sorgente di illuminazione, un oggetto illuminato ed una persona che tramite il suo sistema visivo percepisce il colore.

La colorimetria è un modo per misurare il colore sotto certe condizioni. La **luce** è invisibile (infatti è energia propagata tramite un sistema di onde); ha una certa energia ma ha anche una natura corpuscolare resa possibile dai fotoni. Il nostro sistema visivo coglie al minimo 6 fotoni alla volta. L'onda luminosa è molto caotica e quando “sbatte” contro una superficie tende a propagarsi in maniera casuale disperdendosi in maniera ancora più caotica. Poichè la luce è un'onda di forma sinusoidale che viaggia ad una data velocità detta “velocità della luce” è possibile calcolare la sua lunghezza d'onda (data dalla distanza tra i picchi del relativo istogramma), mentre l'ampiezza è data dall'energia della luce.

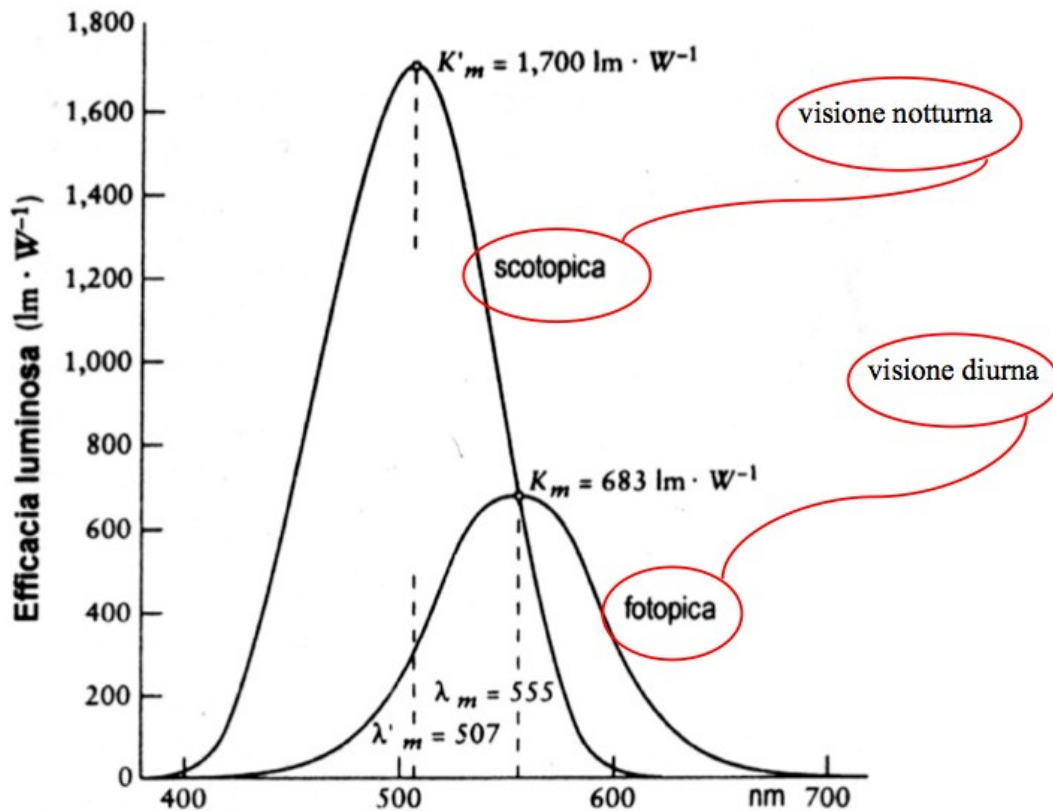


Lo **spettro del visibile** è stato notato, per la prima volta, proiettando su di un piano un fascio di luce attraverso un prisma: il prisma altro non fa che aprire a ventaglio lo spettro della luce visibile. Non è vero che ad ogni lunghezza d'onda corrisponde un colore sempre uguale perchè è il contesto a dare la sensazione di colore in ultima analisi (questo vuol dire che un dato colore può variare a seconda dell'illuminazione e del “contorno” nei quali questo è considerato). Un raggio di luce non è mono-spetturale tuttavia risulta essere una combinazione di onde: per distinguerle si utilizza lo **spettro di potenza** il quale, a seconda della frequenza, indica una diversa energia. Ogni onda viene rappresentata da un punto, per questo lo spettro di potenza è una specie di istogramma: ideato per descrivere l'emissione della luce (oppure la sua componente rifratta a partire da una superficie, detta *luminanza*, che viene poi catturata dal sistema visivo). La luce è una distribuzione di energia elettromagnetica nella parte dello spettro con lunghezza d'onda compresa tra 380 nm e 780 nm.

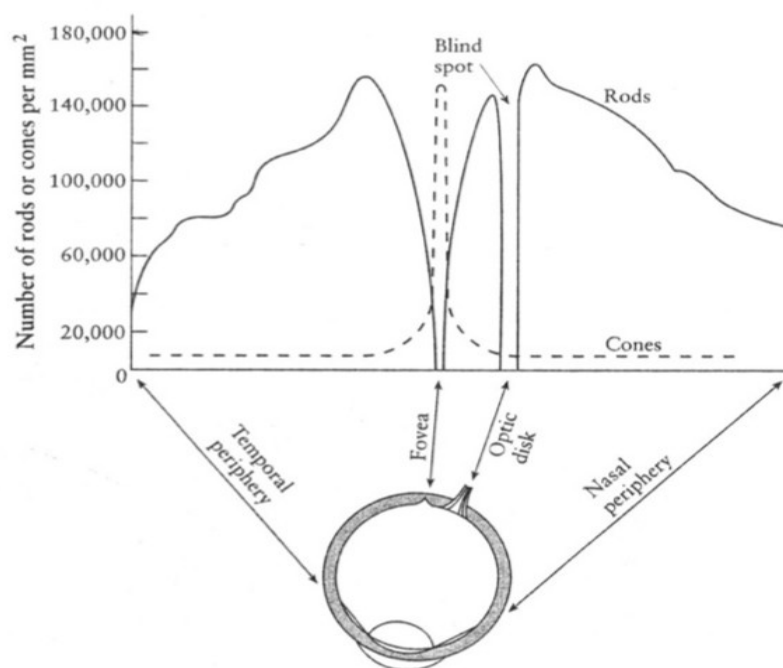
L'**occhio** incamera la luce attraverso una lente (il cristallino) che si contrae e si espande grazie ad appositi muscoli, i quali rendono possibile la messa a fuoco delle immagini. L'occhio è sferico ed è pieno di *umor vitreo* che permette alla luce di colpire la *retina* (oltre a mantenere la pressione dell'occhio costante). La *pupilla* si contrae per modificare la quantità di luce che entra nell'occhio: la sua continua trasformazione è legata alla sfera emozionale tant'è che il suo essere o meno dilatata comunica diversi stati emotivi. La retina è popolata da cellule di due tipi differenti: i *coni* ed i *bastoncelli*; è quindi collegata a *cellule gangliari*

che trasportano le informazioni visive al *nervo ottico* il quale le manda poi al *cervello*. La retina è “montata” al contrario e la luce non la colpisce direttamente evitando così che si bruci: la retina è, per certi versi, un'evoluzione del cervello estromessa dal cranio per poter meglio indagare il mondo. Tutti i tessuti sono abbondantemente irrorati di sangue.

Le *cellule orizzontali* mediano il segnale luminoso orizzontalmente mentre le *cellule amacrine* attuano una comunicazione tra i diversi segnali, è qui che si creano campi ricettivi e si amplifica il contrasto. I due segnali che arrivano a queste cellule si influenzano favorendo l'amplificazione del segnale alto a discapito del secondo segnale, alto anch'esso, che viene ignorato; se il secondo segnale è basso la loro combinazione è data dalla differenza tra i due. Fanno da filtro; infatti tutti i sensi umani sono tarati per accorgersi meglio delle differenze piuttosto che delle ampiezze d'onda. I *neuroni* si possono considerare come elettrodi in cascata: nel cervello gli impulsi sono modulati quindi i neuroni trasmettono tutti lo stesso tipo di impulso (da notare che se la carica è alta non ho un segnale ad impulsi lungo ma un lungo treno di impulsi). I **coni** sono cellule visibilmente a forma di cono con strutture a disco: ogni disco per osmosi scarica l'accumulo elettrico. I **bastoncelli** hanno, oltre ad una forma allungata, assai più dischi poiché sono molto più grandi dei coni: si usano per la visione notturna mentre i coni si sfruttano di giorno in visione fotopica. I coni hanno tre tipologie dovute alla rodopsina sensibili ad una certa porzione dello spettro (il colore è dato proprio dalla loro integrazione): i coni sensibili alle alte lambda (coni L), quelli sensibili alle medie (coni M) e quelli sensibili alle basse (coni S). La sensibilità dei coni è data da valori che si distribuiscono bene, ciascuno su una curva gaussiana, le cui curve tendono a sovrapporsi leggermente.



La **curva di efficacia luminosa** (detta curva V di lambda) riporta la quantità di energia che serve per rendere luminoso un colore come può accadere con una luce bianca colta dai coni e dai bastoncelli.



L'occhio non è esattamente come una macchina fotografica (e nemmeno una macchina fotografica funziona come l'occhio umano). L'occhio è una vera e propria estroflessione del cervello atto a campionare la realtà in modo che questa arrivi sino al cervello, elaborata sulla corteccia cerebrale occipitale che è la zona preposta alla visione. L'essere umano vede prima con il cervello che con l'occhio: casi di cecità totale in cui il cieco distingue i colori ma non le forme. Esiste un punto, sulla retina, dal quale escono come un fascio di cavi tutti i collegamenti neuronali che la connettono al cervello: quel punto è detto punto cieco. I bastoncelli sono molto numerosi nella periferia della retina mentre i coni si concentrano nella fovea: i bastoncelli sono molto sensibili al movimento e rendono possibile la visione periferica. Il cervello si impegna molto nella ricostruzione delle informazioni mancanti dalla visione ed in nessun istante esiste nel nostro sistema visivo una sorta di matrice immagine: il cervello campiona in continuazione la realtà mappandola nei punti in cui a lui sembra che manchino delle informazioni, così colleziona dettagli di volta in volta.

Circuiteria retinale. Formata da 10 layer differenti, consta di 12 tipi di cellule gangliari, 27 tipi di cellule amacrine e 12 rappresentazioni spazio-temporali. La retina compie delle distinzioni già prima di inoltrare il segnale visivo al cervello: si parla di preprocessing delle sequenze spazio-temporali, ecc. Due segnali che si assomigliano molto tendono ad inibirsi tra loro (sia in maniera verticale che in maniera orizzontale). Si ipotizza già da tempo che ogni layer abbia uno scopo ed è dedicato ad una specifica rilevazione → la retina quindi non è solo un banale sensore.

Codificare i colori dai coni. Il tutto a partire dai tre principali tipi di coni. Il colore blu ha una bassissima sensibilità perchè il numero dei coni ad esso dedicati è veramente molto basso, ma la quantità di coni per il rosso ed il verde

quasi coincide e la nostra retina è più sensibile a questi due colori.

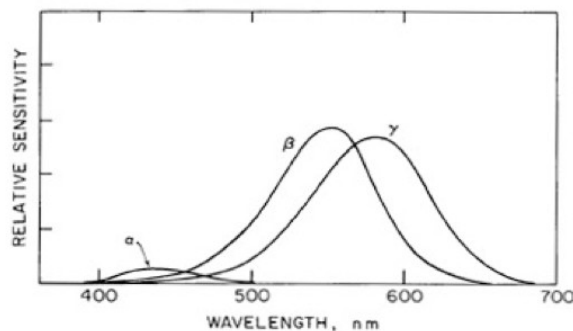


FIGURE 2.2-4. Typical spectral absorption curves of pigments of the retina (10).

A livello di segnale importa solo il rapporto massimo tra i differenti canali; è bassa la discriminazione tra i segnali. Il cervello processa i segnali per distinguere nettamente i colori → attraverso lenti distorte in funzione della fotografia (aggiunte via software) è stato possibile fotografare nel dettaglio e con accuratezza la retina di un gruppo di soggetti di sesso maschile (poiché gli uomini hanno tutti una distribuzione di coni diversa ma avvertono alla stessa maniera tutti i colori).

In un sensore digitale il segnale è proporzionale alla luce che lo colpisce, per gestire il colore devo filtrare la luce per ogni pixel creando così una differenza di comportamento spettrale. La retina fa la stessa cosa: dev'esserci una interpolazione spaziale per avvertire bene i colori. Con un laser a bassa intensità sparato nella retina si vuole così toccare un solo cono o al massimo un piccolissimo gruppo di coni (il laser ha ovviamente la dimensione di un cono), allora si chiede ai soggetti in esame di comunicare la sensazione di colore che il laser stimola loro di volta in volta: ne emerge che pensare alla retina come un dispositivo perfettamente elettronico è un'idea totalmente sbagliata. La sensazione di colore è data dalla differenza di stimolazione dei coni. La stimolazione delle cellule gangliari è tale per cui con stimolazioni semplici e complesse dimostrano diverse energie di stimolo (e con forme più complesse fanno meno fatica rispetto al riconoscimento di forme più elementari).

La luce

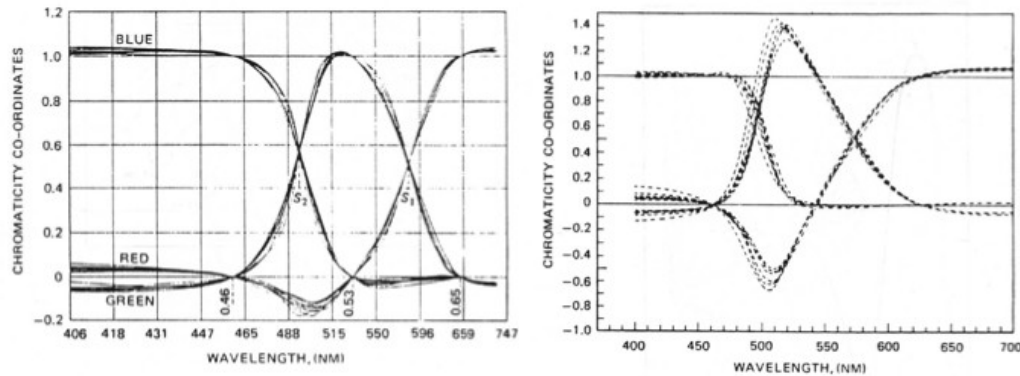
La luce è una distribuzione di energia elettromagnetica nella parte dello spettro con lunghezza d'onda compresa tra 380 nm e 780 nm. Un oggetto è di un determinato colore perchè assorbe l'energia di tutto lo spettro salvo quella di quel tal colore che viene tutta riflessa (es: una lampada alogena genera tantissimo rosso ma poco blu mentre una lampada al neon è da considerarsi come un illuminante moltiplicato per un oggetto colorato da un composto spettrale). Esiste una percentuale che riporta il colore di un oggetto e si ha una percentuale di energia riflessa da un'oggetto che determina il suo colore.

Hermann Von Helmholtz → eminente psicologo e studioso di ottica fisiologica della fine dell'800 che studiava la struttura dell'occhio (e tra le altre cose ha scoperto l'esistenza comprovata di coni e bastoncelli). Ha formulato la teoria del tristimolo.

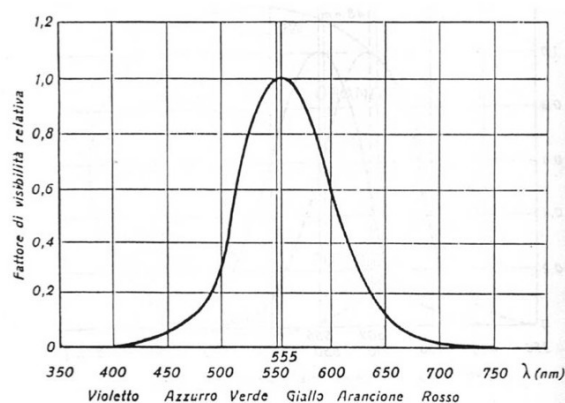
Wright (1928) e **Guild** (1931) → effettuano un esperimento basato sul proiettare tre luci monocromatiche colorate (una verde, una rossa ed una blu) in aggiunta ad una quarta luce precisamente colorata. L'utente regolava le tre luci per ottenere il colore esemplificativo della quarta luce. Durante questo

esperimento è stato disegnato un grafico con alcuni valori negativi perchè in certe configurazioni era necessario togliere un colore per ottenere quello dell'esempio (ovvero occorreva spostare uno dei colori dell'area campione, il che corrispondeva a sottrarlo dal test), tuttavia in generale bastano i tre colori primari per poter comporre tutti gli altri colori → allora per descrivere un colore ci vogliono tre variabili individuando così uno spazio tridimensionale che non si può contrarre e ridurre: lo **spazio colore**.

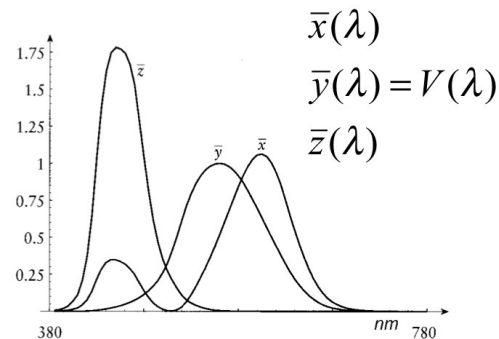
Tuttavia il test era psicologico, quindi diverse configurazioni di valori potevano dare luogo allo stesso colore (*metamerismo*).



La **curva V di lambda** (detta **curva di visibilità fotopica**) è stata definita nel 1931 da un organismo che regolava la colorazione industriale chiamato CIE, i quali hanno preso la decisione di identificare una curva senza valori negativi così hanno normalizzato le precedenti (da bravi ingegneri).



Osservatore Standard C.I.E. 1931

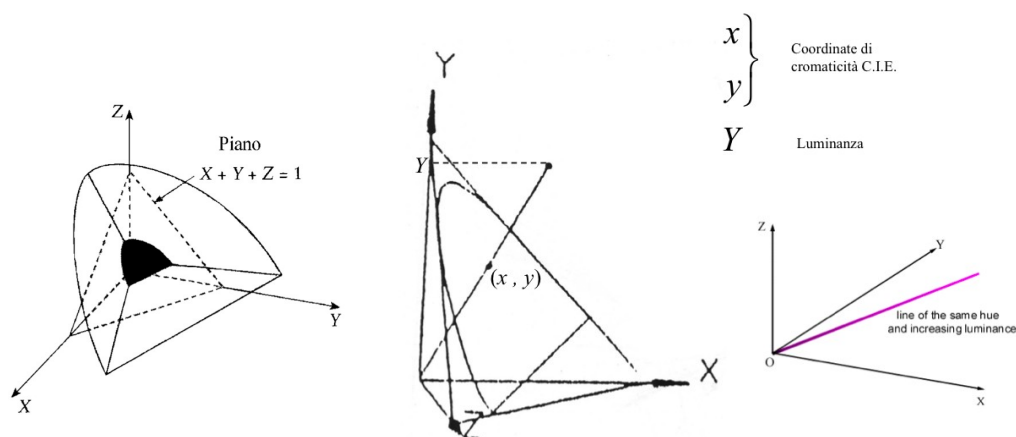


Tale curva identifica un osservatore standard e riporta non solo la sensibilità della retina ma la risposta del sistema visivo ad un campo bipartito in un campo di matching (traduzione: se osservo un colore attraverso un foro non ho percezione del contesto). È importante perchè se il colore non è in aperture mode non si può sfruttare la colorimetria (la colorimetria è spesso usata in modo improprio nonostante il lavoro egregio sul quale è stata edificata). Le curve dell'osservatore standard CIE si possono associare ai relativi colori blu (z), verde (y) e rosso (x) ma non li identificano a tutti gli effetti. Il verde sta vicino al valore di luminanza (il canale del verde è il più vicino e basta da solo per ottenere una foto in bianco e nero). Viene spesso usato come modello di visione ma NON è un vero modello di visione. È importante per trasformare un segnale colore in una tripletta (tristimolo) in un ambiente tridimensionale. Da una buona approssimazione per associare un valore numerico ad una sensazione del

colore. La colorimetria è lo strumento principe per la misurazione del colore.

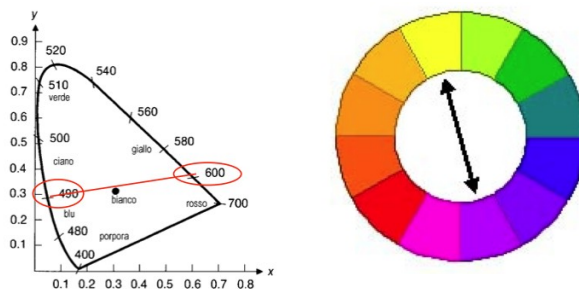
Metamerismo

Non esiste una vera corrispondenza biunivoca tra la distribuzione spettrale ed il tristimolo. Una stessa tripletta è ottenibile da infinite combinazioni di energia, questo definisce il metamerismo. Se non ci fosse il metamerismo si avrebbe enorme difficoltà a concepire triplette su diverse distribuzioni di energia. Se due distribuzioni di energia danno la stessa tripletta allora generano una situazione di metamerismo e conservano questo stato. Le curve CIE garantiscono il metamerismo la maggior parte delle volte.



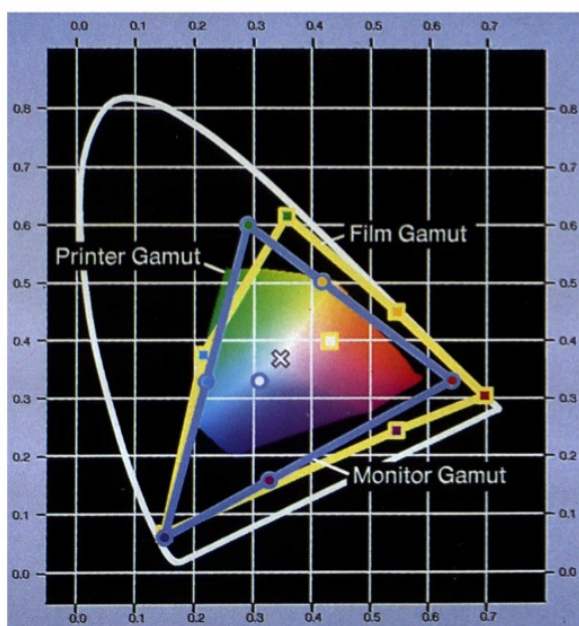
Le tre curve di cui sopra danno origine ad uno spazio tridimensionale detto **spazio colore** che determina la forma di tutti i colori possibili: i colori sono rappresentati da una terna di numeri XYZ detti valori di tristimolo ed ogni colore è rappresentato da un vettore nello spazio tridimensionale. Tale spazio prende il nome di **spazio XYZ**. Per semplificarne la rappresentazione lo spazio è stato secato da un piano in modo da ottenerne una certa sezione, detta **diagramma di cromaticità**. Si parla di cromaticità quando costringo lo spazio colore che è tridimensionale in uno spazio di due dimensioni. Si tratta di una semplificazione chiamata anche diagramma XY dove il "colore" che manca è costituito dalla luminanza.

Se genero un segnale mono-spetturale e da esso estrapolo una tripletta può capitare che cada sul bordo del diagramma di cromaticità che costituisce la linea dei colori saturi; la *linea delle porpore*, invece, non conosce nessuna corrispondenza nello spettro del visibile essendo inesistente nel diagramma spettrale ed ottenibile solamente con una mescolanza del rosso e del blu. Per descrivere un "colore" basta tracciare una linea dall'intersezione degli assi al punto del diagramma di cromaticità, ottenendo così una linea corrispondente sempre alla stessa tinta ma con luce ed illuminazione crescente. Posso passare da uno spazio all'altro senza perdere nessuna informazione.



Un **colore complementare** è il risultato di una mescola tra due luci di certe lunghezze d'onda all'interno della quale si elimina completamente il colore. Si usa un bianco come riferimento e per scegliere un colore complementare al colore dato basta tracciare una linea che congiunga il colore al bianco, la cui continuazione individua il colore complementare.

Dato uno spazio RGB ho sempre una certa energia di frequenza. Quando visualizzo una foto colorata su uno schermo RGB ottengo un grafico nel quale sono riportate le curve per ogni componente di colore. Per conservare la fedeltà cromatica della fotografia sullo schermo è necessario che i punti relativi ai colori della foto rientrino nel grafico del **gamut** relativo al dispositivo monitor. Il gamut costituisce il limite fisico del dispositivo ed è un problema fondamentale delle macchine. I gamut di macchine diverse non saranno mai uguali tra loro.



Attraverso una matrice è possibile passare dallo spazio XYZ a quello RGB (e viceversa) dove XYZ è uno spazio oggettivo e RGB costituisce una fluttuazione del colore. Lo spazio RGB è lo spazio del dispositivo ed è assoluto. I gamut spesso sono molto limitati e certi gamut (come quelli delle stampanti) non hanno forma triangolare per via della presenza di altri colori primari oltre i tradizionali. Attraverso il **gamut mapping** si passa per esempio dallo spazio colore del monitor a quello della stampante (che può essere locale ed allora si faranno aggiustamenti sul colore, oppure globale allora ogni nozione di colore verrà rivista e ridefinita).

Capitolo 2

Fotometria

Fotometria e sorgenti di luce

Bisogna conservare una certa distanza dalla sorgente luminosa per preservarne le proprietà particolari. La curva di sensibilità fotopica dimostra che lunghezze d'onda diverse comunicano diverse sensazioni di luminosità (maggiori o minori rispetto a ciò che ci si aspetta). La *curva fotopica* risulta essere più bassa mentre quella *scotopica* più alta per poter vedere meglio in assenza di luce (al buio, di notte).

La fotometria e la radiometria trattano entrambe lo studio del comportamento della luce; la **fotometria** tratta di misure e grandezze fisiche della luce e le componenti spettrali vanno considerate come curve di un grafico, mentre la **radiometria** rende noto che le componenti spettrali pesano allo stesso modo. La grandezza fotometrica spettrale è uguale alla sensibilità spettrale (intesa anche come V di λ) per la grandezza radiometrica spettrale. Il rapporto tra fotometria e radiometria è dato dalla fotometria uguale alla radiometria moltiplicata per la V di λ .

Le *grandezze fotometriche* sono chiamate anche *grandezze radiofotometriche*. La sorgente luminosa possiede una energia radiante (radiometrica misurata in Joule) e luminosa (fotometrica misurata in Talbot) erogate in una certa unità di tempo che generano il flusso radiante/luminoso (all'interno di un area sferica che solitamente è intesa racchiudere la sorgente); preso un angolo solido che lega la superficie colpita dal flusso alla distanza in percentuale dalla sorgente di luce, il flusso colpisce la superficie determinando una certa misura di intensità radiante/luminosa. Il flusso però incide su un'area detta superficie di area individuando così una irradianza/illuminanza equivalente all'energia caduta sulla superficie in questione. Spesso l'energia rimbalza sulla superficie colpita (che può essere a sua volta *lambertiana* quando la diffusione dell'energia riflessa è omogenea all'interno di una ipotetica semisfera deposta sul piano e centrata sull'angolo di incidenza del flusso energetico; *speculare* quando la riflessione risulta essere opposta all'angolo di incidenza). Le riflessioni nella realtà sono comunque congiunzioni geometriche tra gli effetti della riflessione su superfici miste lambertiane e speculari. Quando al posto del piano si ha una lastra di vetro non si parla più di riflessione ma di trasparenza, comunque valgono ancora le regole di superficie lambertiana mista speculare (senza contare la rifrazione, tipica di un mezzo trasmissivo trasparente) dipendentemente comunque dalla sua natura liscia o smerigliata. Esistono superfici sia trasmissive che rifrattive dette traslucide. In qualsiasi caso, la luce allora torna indietro verso il sistema visivo (se la luce oltrepassa il mezzo si parla di radianza o trasmissione). Se la sorgente di luce non è puntiforme ma è distribuita l'area su cui si sparge la fonte luminosa aumenta e l'energia trasmessa cala: si parla allora di uscita radiante/luminosa.

Uscita radiante:

densità di flusso radiante uscente da una superficie

$$M_e = \frac{d\Phi_e}{dA}$$

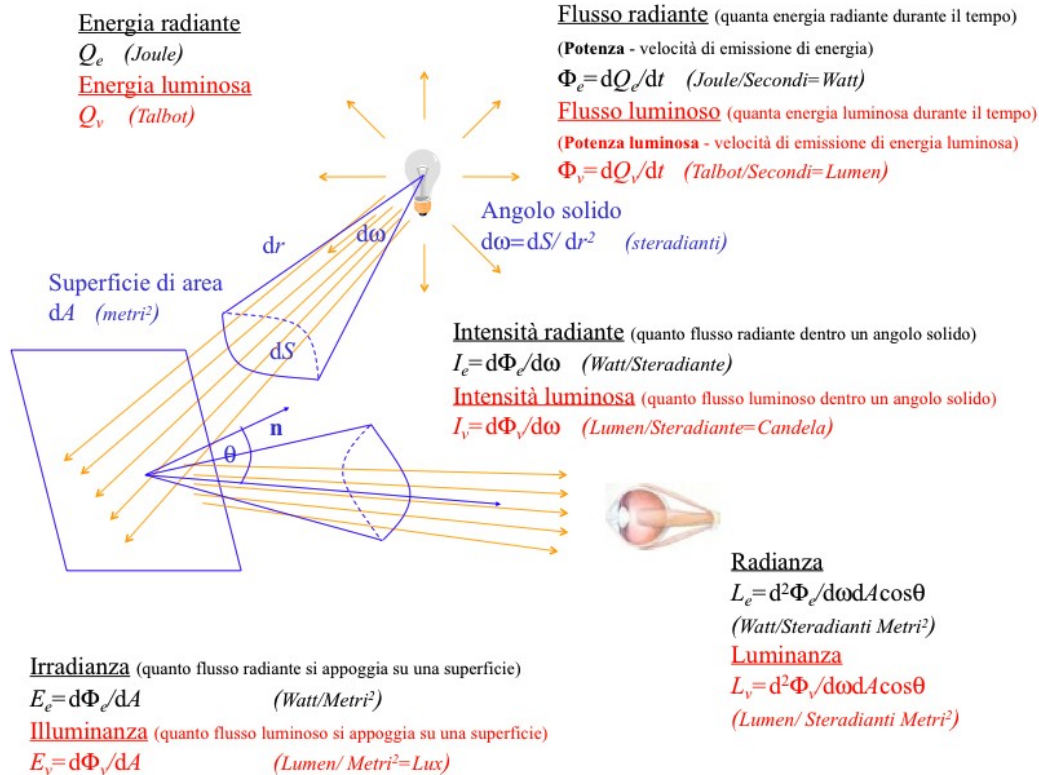
unità di misura: Watt/metri²

Uscita luminosa:

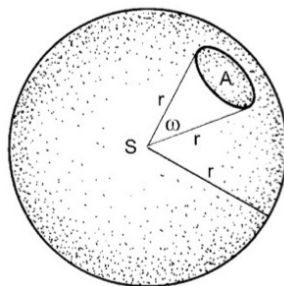
densità di flusso luminoso uscente da una superficie

$$M_v = \frac{d\Phi_v}{dA}$$

unità di misura: Lumen/metri²



Legge fondamentale della radiometria. Detta anche *legge dell'inverso del quadrato della distanza*, dice che l'energia della luce si disperde a seconda della distanza dalla sorgente in maniera quadratica (oppure la stessa area tende a proiettarsi con un angolo minore).



$$d\Phi_e = I_e d\omega \quad \text{definizione di intensità}$$

$$dA = r^2 d\omega \quad \text{definizione di angolo solido}$$

$$E_e = \frac{d\Phi_e}{dA} = \frac{I_e d\omega}{r^2 d\omega} = \frac{I_e}{r^2}$$

Non si è ancora parlato di componente spettrale ma tutte le grandezze viste sino ad ora possiedono ugualmente una componente spettrale: è noto infatti che tutte le grandezze radiofotometriche dipendono in generale dalla lunghezza d'onda λ .

$$\text{Grandezze totali} = \int \text{Grandezze spettrali} \, d\lambda$$

Tutte le grandezze radiofotometriche hanno una componente λ e tutti gli integrali di questo tipo sono da considerarsi compresi tra 380 nm e 780 nm all'interno dello spettro del visibile.

Solitamente una **sorgente luminosa** è caratterizzata da aspetti *geometrici-direzionali* e da aspetti *spettrali-energetici*. Una fonte di luce λ si rappresenta con un solido goniometrico misurando in ogni punto la percentuale di energia irradiata all'interno di una stanza completamente nera.

Grandezze radiofotometriche spettrali

Energia radiante

$$Q_e(\lambda) \quad (\text{Joule/Metro})$$

Energia luminosa

$$Q_v(\lambda) \quad (\text{Talbot/Metro})$$

Flusso radiante (quanta energia radiante per ogni λ durante il tempo)

$$\Phi_e(\lambda) = dQ_e(\lambda)/dt \quad (\text{Joule/Secondi Metro} = \text{Watt})$$

Flusso luminoso (quanta energia luminosa durante il tempo)

$$\Phi_v(\lambda) = dQ_v(\lambda)/dt \quad (\text{Talbot/Secondi Metro} = \text{Lumen})$$

Intensità radiante (quanto flusso radiante per ogni λ dentro un angolo solido)

$$I_e(\lambda) = d\Phi_e(\lambda)/d\omega \quad (\text{Watt/Steradiano Metro})$$

Intensità luminosa (quanto flusso luminoso dentro un angolo solido)

$$I_v(\lambda) = d\Phi_v(\lambda)/d\omega \quad (\text{Lumen/Steradiano Metro} = \text{Candela})$$

Irradianza (quanto flusso radiante per ogni λ si appoggia su una superficie)

$$E_e(\lambda) = d\Phi_e(\lambda)/dA \quad (\text{Watt/Metri}^2)$$

Illuminanza (quanto flusso luminoso si appoggia su una superficie)

$$E_v(\lambda) = d\Phi_v(\lambda)/dA \quad (\text{Lumen/ Metri}^2 = \text{Lux})$$

Radianza

$$L_e(\lambda) = d^2\Phi_e(\lambda)/d\omega dA \cos\theta \quad (\text{Watt/Steradiani Metri}^3)$$

Luminanza

$$L_v(\lambda) = d^2\Phi_v(\lambda)/d\omega dA \cos\theta \quad (\text{Lumen/ Steradiani Metri}^3)$$

Valori tipici di Luminanza

luna	0,03 cd/m ²
cielo nuvoloso	300 cd/m ²
cielo limpido	30.00 cd/m ²
sole coperto	30.000 cd/m ²
sole	2.000.000.000 cd/m ²

Valori tipici di Illuminanza

luna piena	1 lux
illuminazione stradale	10 lux
illuminazione interna	100 lux
illuminazione interna diurna	1000 lux
sale operatorie	10.000 lux
luce diurna pieno sole	100.000 lux

Grandezze radiofotometriche totali

$$Q_e = \int Q_e(\lambda) d\lambda \quad \text{Energia} \quad Q_v = \int Q_v(\lambda) d\lambda$$

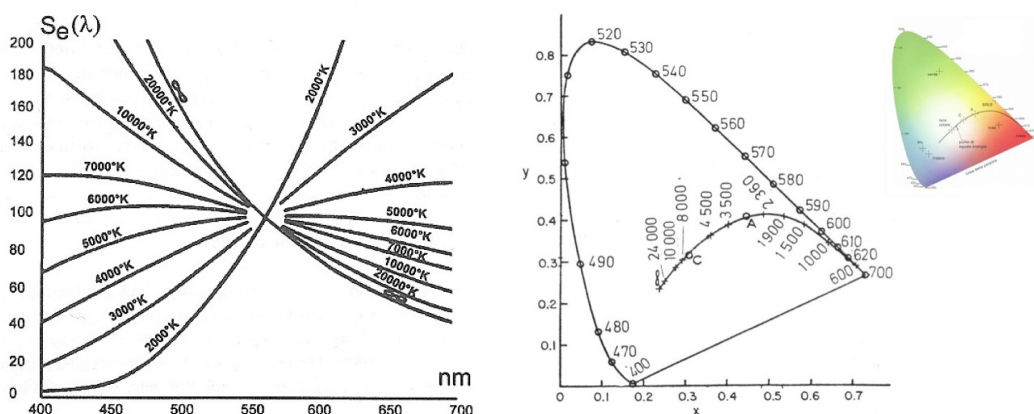
$$\Phi_e = \int \Phi_e(\lambda) d\lambda \quad \text{Flusso} \quad \Phi_v = \int \Phi_v(\lambda) d\lambda$$

$$I_e = \int I_e(\lambda) d\lambda \quad \text{Intensità} \quad I_v = \int I_v(\lambda) d\lambda$$

$$E_e = \int E_e(\lambda) d\lambda \quad \begin{matrix} \text{Irradianza} \\ \text{Illuminanza} \end{matrix} \quad E_v = \int E_v(\lambda) d\lambda$$

$$L_e = \int L_e(\lambda) d\lambda \quad \begin{matrix} \text{Radianza} \\ \text{Luminanza} \end{matrix} \quad L_v = \int L_v(\lambda) d\lambda$$

Gli aspetti energetici e spettrali più importanti sono relativi a quanta energia viene emessa per ogni λ . Si definisce **corpo nero** un oggetto teorico (che non esiste) il quale assorbe tutta l'energia proiettata su di esso senza emetterla. Tale oggetto emette anche e soprattutto in funzione della temperatura quindi se scaldato ad alte temperature emette energia in tutto lo spettro del λ : le percentuali allora vanno confrontate in modo assoluto; normalizzando le temperature rispetto ad un certo λ (555) le curve colorimetriche tendono a convergere, tuttavia man mano che le temperature si alzano aumentano le percentuali relative al colore blu (la lampada al tungsteno, ora in disuso, era una sorgente luminosa perfetta perchè possedeva una grande componente blu). Esistono sorgenti reali (ma artificiali) comprese tra i 2000 ed i 3000°K mentre tutte le altre sono sorgenti naturali, come la luce del sole. Le lampade ad incandescenza possiedono uno spettro continuo con energia diffusa in tutto lo spettro.



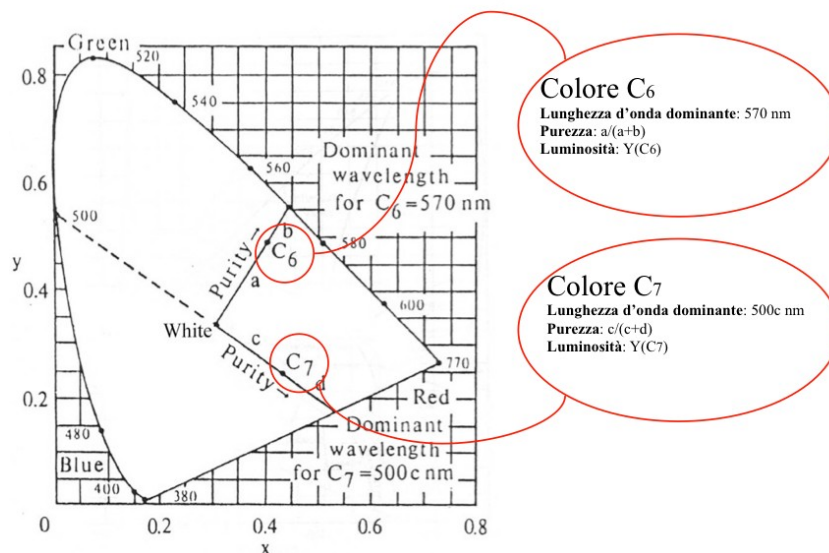
È detta **linea degli illuminanti** quella linea che riassume e colloca qualsiasi illuminante a spettro pieno e la si può definire su di una specifica curva grazie alla nozione di metamerismo: da un punto di vista percettivo e su una superficie bianca la luminanza è trasformata in una tripletta simile a quella della luce solare, ma su una superficie blu/verdastra ci sono delle differenze dovute alla *distorsione cromatica*. Le sorgenti individuate in questa maniera sono standard tanto che la sorgente D65 è più blu che rossa ed è quella che assomiglia di più alla luce del sole in una giornata serena, misurata in parti differenti del mondo in condizioni differenti (è una luce che nella pratica però non esiste perchè è il risultato di una media). A è la luce del tungsteno, mentre B e C sono lampade al tungsteno con filtro blu con un ottima qualità di spettro che tende a raddrizzare la curva della lampada A (usate spesso sui tecnigrafi, non è certamente a risparmio energetico).

Tornando al processo di visione manca all'appello la riflettanza spettrale che in pratica è una percentuale di energia riflessa con luminanza ed illuminamento diretti verso il nostro sistema visivo; la luminanza spettrale è pesata sulla riflettanza (che è sempre e solo percentuale). La riflettanza torna il colore degli oggetti.

I colori neutri sono il modo migliore per controllare e testare il vero colore degli illuminanti. Il passaggio dal riferimento XYZ a xyz è puramente matematico (e banale) poiché il piano è unitario e si pone il vincolo ad 1, allora si può ignorare z perchè è sempre ricavabile. Quando si passa da uno spazio all'altro si commettono comunque errori numerici e di approssimazione. Con il tristimolo e le tre relative curve è sempre possibile codificare il comportamento spettrale.

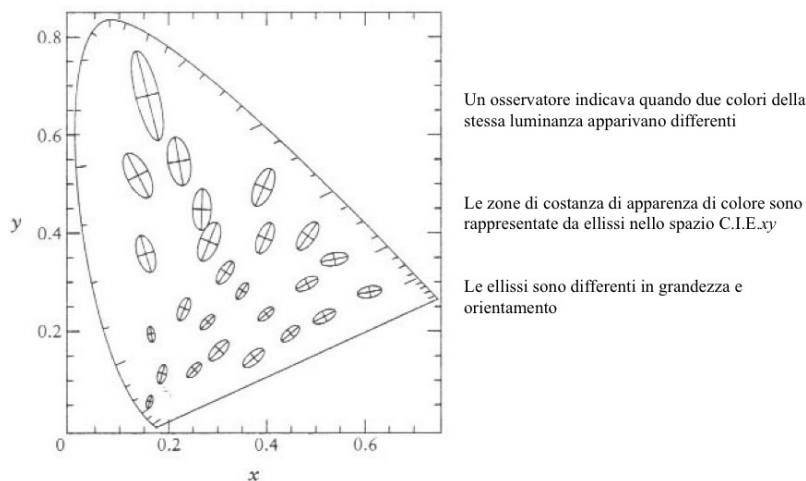
Calcolo sul diagramma di cromaticità. Si prende un colore (per esempio C6) ed a partire da questo colore si tira una linea che passi per il colore e per il bianco di riferimento (tale bianco costituisce l'illuminante grazie al quale è possibile la visione di quel dato colore, la sua importanza è fondamentale per la percezione del colore): la continuazione di quella linea incrocerà il bordo del diagramma, allora la lunghezza d'onda dominante corrisponde al punto in cui la linea tocca il bordo dei colori monospettrali: assoceremo quindi C6 a quel colore.

Caratteristiche intuitive del diagramma di cromaticità



La **purezza** del colore è data dalla varianza del bianco oppure dalla tinta, detta **saturazione**. Se la linea di cui sopra tocca la linea delle porpore (le quali hanno una corrispondenza percettiva ma non a livello di spettro del visibile) la si prolunga dall'altra parte. La **tinta** è la lunghezza d'onda dominante. Si può utilizzare lo *spazio Munsell* come diagramma della cromaticità (ne parleremo meglio più avanti). La **luminosità** è data quindi dalla luminanza.

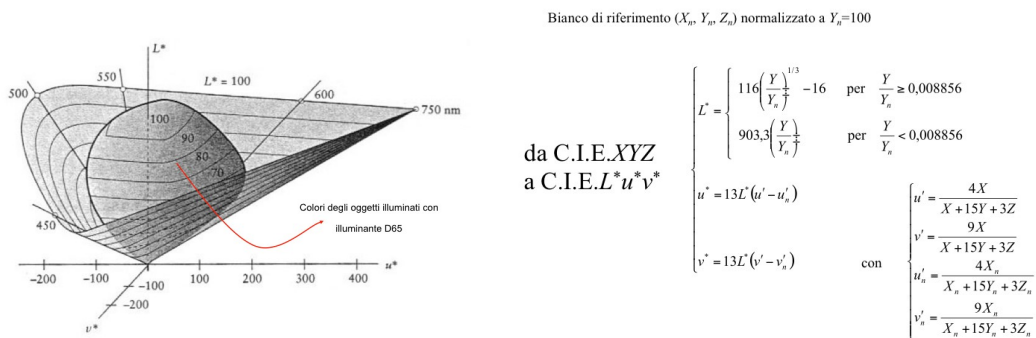
La **legge di Grassm** sostiene che l'additività di due colori è lineare. Questa è la ragione per cui è possibile parlare di **gamut**: qualsiasi combinazione lineare è compresa all'interno del triangolo del gamut. Negli anni '70 ci si accorse che gli spazi CIE XYZ e xyY non erano percettivamente lineari: **Mc Adam** scopre che la distanza tra due punti (corrispondenti a relativi colori sul diagramma di cromaticità) non misura effettivamente la differenza della loro apparenza:



ovvero, se si scelgono due colori con una certa distanza tra loro sullo schema di cromaticità non è assolutamente detto che abbiano la stessa distanza anche a livello di percezione degli stessi colori → non si possono associare tinte e percettibilità, gli spazi in questione non hanno alcuna metrica percettiva.

Mc Adam prese una serie di punti sul diagramma di cromaticità (almeno 10 volte più piccoli di come appaiono nella rappresentazione) ed ha creato un campo bipartito da una parte con tutti i colori correttamente scelti e dall'altra con delle variazioni sugli stessi colori; ha quindi eseguito alcuni test (sempre sulla stessa persona) ed il soggetto appena aveva percezioni differenti rispetto al colore di riferimento doveva segnalarlo. Si evince dal test che nell'area dei verdi la percezione di verdi differenti è inibita e diverse tinte di verde risultano diverse solo dopo una lunga osservazione (l'ellisse allora corrispondente alle tinte verdi può essere ampliato). Per uniformare le ellissi Mc Adam ha allora deformato lo spazio di cromaticità dando vita così agli spazi **CIE $L^*u^*v^*$** per i colori emessi ed additivi e **CIE $L^*a^*b^*$** per i colori sottrattivi (che in fin dei conti è lo spazio di riferimento più sfruttato).

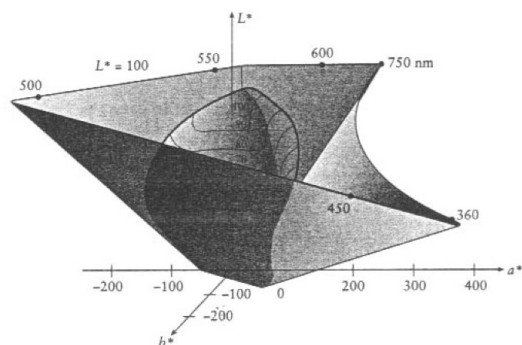
Lo spazio CIE $L^*u^*v^*$ rappresenta la trasformazione lineare dello spazio CIE XYZ in modo da ottenere uno spazio percettivamente uniforme → tale spazio è prodotto da un tot di formule matematiche ma il bianco di riferimento rimane fondamentale, perchè finchè ci si trova in XYZ si è in uno spazio oggettivo senza alcun problema di percezione e si preserva il metamerismo, tuttavia non viene detto nulla sulla relazione tra gli stimoli; quando si passa ad una mappatura percettiva è fondamentale sapere in che ordine sono distribuiti i colori (quindi si devono definire delle *scalature*) e, poiché i colori si vedono solo quando sono illuminati, si ha bisogno di individuare un bianco di riferimento che influisce quindi sui colori (il colore viene mescolato con l'illuminante) → si cerca un trucco per poter contare il contesto senza però voler davvero tener conto di tale contesto.



Gli assi u e v vogliono modificare i verdi ed i blu, la radice cubica corrisponde alla luminosità legata ad un fattore percettivo della luminosità stessa che si comporta logaritmicamente, tanto da assomigliare, per l'appunto, ad una radice cubica. Il metodo fallisce per tutti i colori molto scuri (è la luminanza a favorire la corretta percezione dei colori). Si tratta di uno spazio percettivo fluttuante, la sua forma assomiglia a quella di una sfera imperfetta (anche se su Internet sostengono invece che assomiglia ad una sfera perfetta, è una imprecisione!!); con una serie di illuminanti la quasi-sfera si può espandere e spostare sino a ricoprire tutto lo spazio esterno.

Lo spazio CIE $L^*a^*b^*$ è praticamente la stessa cosa salvo per la scalatura della cromaticità che cambia drasticamente: si tratta di una trasformazione lineare dello spazio CIE XYZ in modo da ottenere uno spazio percettivamente uniforme.

Da un punto di vista matematico lo spazio è decisamente peggiore, tuttavia si tratta dello spazio di riferimento più utile ed esauriente (molto in uso nell'industria tessile, anche Photoshop lo utilizza ma sbaglia a non indicare il bianco di riferimento).



Bianco di riferimento (X_w, Y_w, Z_w) normalizzato a $Y_w=100$

da C.I.E.XYZ
a C.I.E. $L^*a^*b^*$

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_w} \right)^{1/3} - 16 & \text{per } \frac{Y}{Y_w} \geq 0,008856 \\ 903,3 \left(\frac{Y}{Y_w} \right) & \text{per } \frac{Y}{Y_w} < 0,008856 \end{cases}$$

$$a^* = 500L^* \left[f\left(\frac{X}{X_w}\right) - f\left(\frac{Y}{Y_w}\right) \right]$$

$$b^* = 200L^* \left[f\left(\frac{X}{X_w}\right) - f\left(\frac{Z}{Z_w}\right) \right]$$

$$\text{con } f(r) = \begin{cases} r^{1/3} & \text{per } r \geq 0,008856 \\ 7,787r + \frac{16}{116} & \text{per } r < 0,008856 \end{cases}$$

Una volta fissato il bianco di riferimento basta calcolarne la distanza euclidea per calcolare la distanza dei punti dello spazio \rightarrow le distanze allora suppergiù coincidono, idem per i colori (e se così non è vince in rilevanza il colore percettivo e non quello calcolato) e due coppie di colori con la stessa distanza euclidea sono percettivamente simili. Lo spazio non è percettivamente uniforme allora si tende a variare adeguatamente la misura della differenza perchè il colore non si presta MAI ad una metrica perfetta: non sono misure precise quindi non andrebbero prese troppo sul serio (se la discordanza è minore di 10 i due colori sono lontanissimi, se è maggiore di 1 allora i colori sono praticamente identici ma se è maggiore di 10 allora i colori sono quasi simili, ma non vale in assoluto).

Lo spazio Yuv è detto anche EBU (European Broadcasting Union) ed è di fatto lo standard adottato per la codifica delle informazioni sul colore negli schermi TV Pal/Secam. Y è chiamato "luma" e per passare da RGB a Yuv si usa una determinata matrice. Il sistema NTSC usa Y'i'q' con una seconda matrice di trasformazione. Si tratta di una codifica oggettiva che dipende dall'illuminante di riferimento e coincide con lo spazio del dispositivo (come nel caso di RGB) senza considerare una compressione della Y. Si prende il valore di RGB correttamente codificato e lo si traduce nel nuovo spazio.

Lo spazio Ac1c2 è lo spazio di massima separazione e varianza, orienta il sistema di riferimento secondo l'asse di massima varianza approssimando lo spazio della sensibilità retinica. In ogni spazio c'è una correlazione tra canali x, y e z ma questo sponde al massimo lo spazio per permettere di de-correlare i canali (cosa che di solito fa il sistema visivo).

Teorie del colore

Quanto detto sino ad ora avviene fuori contesto ed è matematicamente robusto ma assai distante dalla percezione reale. Le teorie del colore sono fortemente legate al contesto.

Nel 1789 **Gaspard Monge** si chiede come è possibile percepire un colore mediante l'osservazione dello stesso attraverso un filtro del medesimo colore, accorgendosi che il colore studiato appare quindi bianco. Allora il filtro colorato tende a sbiancare tutti gli oggetti del medesimo colore.

Disputa Goethe-Newton \rightarrow Newton sosteneva che il colore venisse prodotto da luce di determinate lunghezze d'onda, allora associò le lunghezze d'onda alla

percezione del colore. Goethe, invece, oppostamente a Newton (ovviamente moltissimi anni più tardi) si innamorò del ritratto di una fanciulla e passò lungo tempo a fissarlo senza mai distogliere lo sguardo, nel momento in cui lo distolse scoprì l'effetto dell'*afterimage* e scrisse “Die Farbe” insultando Newton perchè, secondo Goethe, il colore era un effetto prettamente personale e soggettivo prodotto dal pensiero dell'osservatore. Quello che Goethe non sapeva era che avevano ragione entrambi.

È dovuto a Monge l'esperienza delle “ombre colorate”, durante la quale si illumina uno schermo bianco con due proiettori, uno che emette luce bianca e l'altro che emette luce rosa. Se si pone un oggetto lungo il fascio di luce rosa tale oggetto proietta un'ombra sullo schermo: l'ombra è di colore blue-verdastro (complementare del rosa) ma la regione in ombra riflette solo la luce bianca. Il colore qui è analizzato nel contesto.

Anni dopo, **Hering** studia il passaggio dal giallo al blu e si rende conto che è impossibile farlo in maniera diretta ma è possibile solo attraverso una gradazione intermedia con altri colori; scopre così che esistono contrapposizioni tra i colori individuando i colori complementari (codificando così $L^*u^*v^*$ e $L^*a^*b^*$), individuando 6 stimoli cromatici (rosso, verde, blu) e tre coppie antagoniste (bianco/nero, giallo/blue, rosso/verde), appoggiato da Wittgenstein che sostiene che è impossibile concepire un “rosso verdastro” o un “giallo bluastro”. Cos'è quindi il colore se non un misto tra tinta ed illuminante? Il colore è una rappresentazione interna che la mente umana associa ad uno stimolo visivo in relazione ad altri stimoli sensoriali (una macchina fotografica coglie solo lo stimolo ignorando totalmente il contesto). Gli esseri umani percepiscono meglio per una questione di sopravvivenza che permette di essere indifferenti ai cambiamenti di illuminazione (\rightarrow *color constancy*) aumentando il contrasto in maniera locale su tutta l'area.

Esperimento di Edwin Land \rightarrow si hanno un pannello ricoperto di patch colorate ed un trio di proiettori colorati (rosso, verde e blu) i quali vengono proiettati su una specifica regione del pannello corrispondente ad un certo colore: si capisce che nonostante la luce proiettata il colore avvertito rimane sempre lo stesso. Se però la patch viene isolata dal contesto, cambia anche la percezione del colore. Grazie alla **color constancy** il sistema visivo umano tende ad attenuare i colori molto accesi ed attribuisce approssimativamente il medesimo colore ad un oggetto in maniera svincolata ed indipendente dal particolare illuminante che lo investe; mentre la **lightness constancy** agevola l'adattamento alla luminosità ed attribuisce approssimativamente la medesima luminosità ad un oggetto indipendentemente dal particolare livello di intensità dell'illuminante.

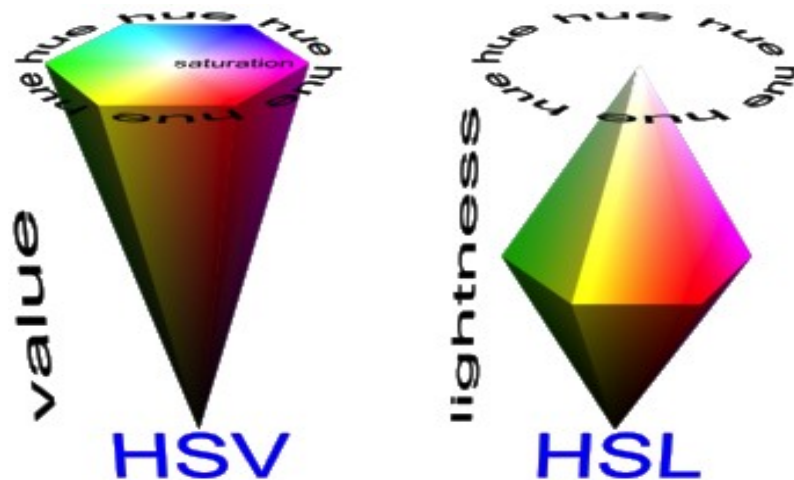
Capitolo 3

Il colore

Formazione del colore

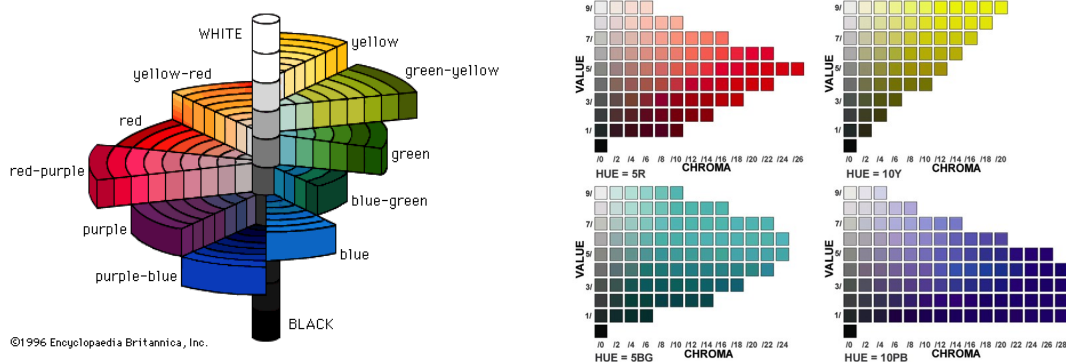
Il colore gode di una categorizzazione particolare; si pensi che l'azienda HP sta studiando un sistema accurato per comunicare la sensazione di un colore semplicemente chiamandolo per nome. Il **colore**, per come lo pensiamo, si può descrivere attraverso degli attributi come tinta, saturazione e luminosità. La **tinta** costituisce un termine autoreferenziale abbastanza difficile da descrivere in maniera assoluta, tant'è che lo si fa per negazione (*un colore senza tinta è acromatico* → è un ossimoro) tuttavia se si pensa al “cerchio dei colori” la tinta corrisponde alla direzione di quel dato colore all'interno di quel dato spazio; la **saturazione** invece descrive la purezza di una tinta, quel comportamento tipico di quando la tinta risalta a tal punto che sembra brillare (grigio, bianco e nero tendono, se aggiunti alla tinta, a diminuire la brillantezza, quindi anche la saturazione); la **cromaticità** costituisce quel modo di descrivere il colore senza la luminosità, e si può intendere la caratteristica della cromaticità come la somma composita di tinta e saturazione e di qualsiasi altro attributo che prescinde dalla luminosità (solitamente si sfrutta la cromaticità per fare delle normalizzazioni nello spazio RGB); la **chiarezza** invece è quanto la luminosità fa apparire chiara una certa tinta, si tratta di una percezione assolutamente non misurabile.

1. La **tinta** si stabilisce sulla ruota dei colori che simula la sequenza di successione dei colori dello spettro; di queste ruote ne esistono di diversi tipi, ciascuna con una funzionalità ben precisa a seconda dello scopo. La tinta si può dire corrisponda in tutto e per tutto al colore;
2. Per definire la **saturazione** si sfrutta ancora la ruota dei colori dando ad essa una nuova dimensione all'interno della quale gli stessi colori possono scurire oppure schiarire muovendosi in direzione diverse ed opposte: in questa maniera si nota la variazione di “carica” ed intensità del colore (*saturazione* → quantità di tinta pura che dev'essere mescolata al bianco per produrre il colore percepito, è così che si estende la ruota con una seconda dimensione);



3. La **luminosità** costituisce la terza dimensione del colore: lo schema **HSL** sostiene che tutte le tinte si trovano a metà di una distribuzione di luminosità (quindi se si disegnano i colori saturi li si piazza al centro e si forza lo spazio in modo che le tinte sature in realtà si trovino sulla fascia del grigio medio, tuttavia non è una rappresentazione corretta perchè se ci sono tinte cariche gli spazi colore poi non corrispondono a come si descrive il colore per come lo concepiamo noi). **HSV** invece è il risultato della compressione della punta chiara di HSL; **HSB** invece costituisce un cilindro (rappresentazione che utilizza Photoshop) e coincide con HSL salvo per il fatto che il bianco ed il nero non costituiscono dei punti ma dei cerchi, così capita che alcuni colori vengano moltiplicati.

Il sistema di Munsell. Ideato da un pittore, *Munsell* appunto, il quale ha descritto i colori nella maniera più corretta possibile. Il suo schema presenta una linea degli acromatici al centro (dove trovano posto il bianco ed il nero); un valore stabilito di *chroma* come numero di colori per tassello o “irraggiamento” dal centro mentre la saturazione misura la distanza di ogni sfumatura di colore dall'asse centrale. È stato quantizzato in modo che un salto da un colore adiacente all'altro sia fondamentalmente equiperceptivo (quindi salti tra tinte, saturazioni ecc a patto che siano in “cellette” che comunicano tra loro). Si nota che all'interno dello schema non sono presenti tutte le sfumature per ciascun colore, perchè non è detto che esistano davvero.



È possibile passare dallo spazio di Munsell a quello XYZ per definizione, tuttavia la percezione del colore comunque cambierà anche se di poco a seconda dell'illuminante → **M L*a*v*** (con M di Munsell) è uno spazio utilissimo ma raramente sfruttato che è come CIE L*a*b* ma si vede applicato allo spazio di Munsell.

Le tinte si possono misurare in uno spazio a piacere, solitamente si prendono come riferimento spazi che vanno da 0 a 360 (come i gradi della ruota dei colori) oppure da 0 a 255 (come nello spazio RGB) od ancora da 0 a 100 per un calcolo unitario.

Percezione

La percezione richiede partecipazione. Nella **percezione** esiste una grande percentuale di ambiguità per definizione: un oggetto può risultare colorato per sua natura oppure tale colore può essergli stato conferito dall'illuminazione, poiché il colore possiede dei meccanismi geometrici e spaziali.

Il colore nonostante tutto è svincolato dalla distribuzione spettrale e costituisce

una realtà fondamentale per analizzare il colore nel suo contesto. Il nostro cervello è più sensibile ai bordi degli oggetti e delle immagini assegnando a ciascuno di essi un colore di riempimento, questo costruisce la nostra personale sensazione del colore (quando infatti vengono a mancare i bordi il cervello si sente autorizzato a tralasciare il colore, ha a che vedere con la sensazione del colore all'interno del proprio contesto). L'attivazione corticale e retinale tornano valori differenti a seconda della visione di una scena semplice anziché complessa; il nostro cervello vede naturale il movimento “strisciato” dei bordi all'interno di una scena dinamica.

Il metamerismo individua una tripletta per ogni distribuzione di energia e costituisce una cosa fondamentale per lo studio dell'imaging digitale, perché se i colori percepiti sono tra loro metamERICI allora nonostante non condividano la stessa tripletta questi sono tra loro equivalenti (come se la condividessero davvero) ed appaiono fondamentalmente uguali. Se il metamerismo è una proprietà della riflettanza dipende allora anche dall'illuminante della scena ed il suo cambiamento può rischiare di invalidare il metamerismo. Il metamerismo è importante non solo nella visualizzazione di immagini attraverso monitor ma anche nel processo di stampa (se non esistesse non si potrebbe stampare a colori); il metamerismo dice che si può avere la stessa percezione anche se i segnali dello spettro sono differenti → le **color matching functions** si vorrebbe collegassero alla stessa tripletta i colori che appaiono simili e per colori che invece appaiono dissimili identificassero triplette differenti (il tutto basandosi però sull'apparenza del colore). Poiché non esiste un correlato fisico del colore (o per lo meno diciamo che non esiste), il colore dovrebbe essere misurato solamente fuori contesto.

Dal colore all'apparenza. Oggetti che hanno la stessa riflettanza cambiano colore a seconda della lucidatura della loro superficie e del contesto in cui si vedono immersi. L'aumentare del contrasto modifica la percezione del colore. L'effetto **Cornsweet** dimostra che un'immagine costituita da righe dello stesso colore grigio sembrano più chiare oppure più scure in corrispondenza dei bordi confinanti con zone più chiare o più scure, quando ovviamente non lo sono: è l'effetto dato proprio dalla presenza del bordo.

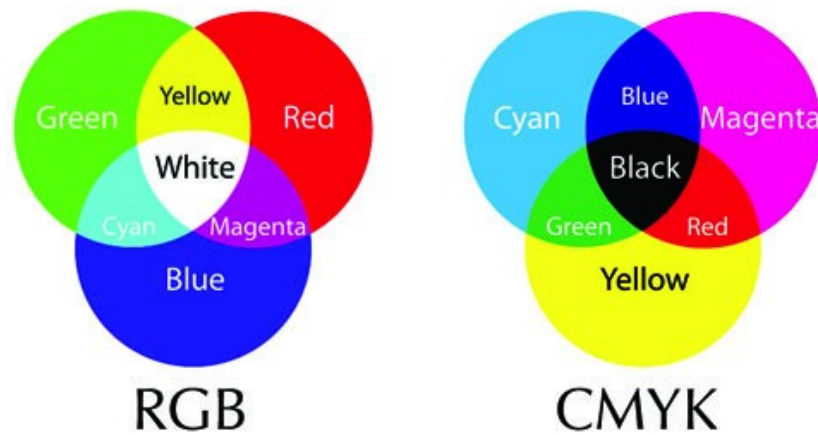
L'assimilazione capita in certi contesti quando è la conformazione spaziale a definire la presenza di bordi. L'effetto “bordo” non solo cambia la percezione del colore di riempimento ma cambia anche la polarità del bordo stesso e si costringe a percepire lo stesso colore in maniera totalmente differente.

Il Gamut

Il **gamut** incarna un problema fisico tant'è che se non si riesce a rappresentare un colore, non si riesce e basta: per gamut s'intende lo spazio colore di un dato dispositivo (dispositivi diversi possiedono gamut differenti). I colori desaturi non hanno nessun problema infatti si preferisce utilizzare colori poco saturi. Un'immagine raster è costituita da una matrice di informazioni dette pixel; un'immagine vettoriale invece è creata tramite e rappresentata da curve matematicamente definite (quindi può essere modificata, ingrandita o diminuita in dimensione in qualsiasi momento senza perdere dettagli ed informazioni).

Il computer *rasterizza* quando trasforma un'immagine in un “qualcosa” che possa essere stampato. Il rasterizzatore è quell'unità di calcolo che implementa

il controllo del gamut, a metà tra la stampante ed il suo driver. Infatti, anche se la tripletta RGB è assoluta il colore tuttavia cambia su display differenti (si pensi che monitor diversi possiedono colori primari diversi, alcuni in numero maggiore o minore di altri). Quando si stampa si passa da RGB a CMY; RGB è uno spazio cubico compreso tra 0 ed 1 (oppure tra 0 e 255) dove a 0 si ha il nero, seguendo la sola direttiva di un colore al massimo della sua potenza si ha il colore stesso mentre se si potenziano tutti i colori, al valore 1 si ottiene il bianco.



Si parla di **sintesi sottrattiva** quando si mescolano i pigmenti tra loro, tipica dello spazio CMY (che è lo spazio opposto a RGB); RGB invece fa **sintesi additiva** mescolando luci anziché pigmenti; lo spazio CMY non è uno spazio colore reale poiché sommando tutti i colori non ottengo il nero puro (bensì un marroncino sporco molto costoso a livello di inchiostro) → si usa allora lo spazio CMYK con UGR il quale prende una tripletta di ciano-magenta-giallo che nella realtà è, per esempio, azzurro-verde ma che viene poi resa grigiastra, così la sostituisco con una tripletta simile a quella di CMY più un valore K, dove K è un misto di giallo, magenta e ciano nella stessa intensità (per dire che i valori di CMY che sono uguali vengono riassunti nella variabile K lasciando indietro i valori discordi che andranno quindi rappresentati in altre maniere). Passare da CMY a CMYK non è univoco e richiede una strategia, il tutto per risparmiare denaro. K costituisce una variabile in più all'interno dell'equazione ma non causa perdite e disagi a livello di stampa (anzi, ne risolve parecchi).



Periferiche. Se cambiano le macchine attraverso le quali si visualizza un colore cambiano anche gli effetti finali. I vecchi monitor a tubo catodico possono utilizzare un contrasto più spinto e secco rispetto i monitor moderni rendendo la visione di immagini molto più piacevole per l'occhio umano (sono consigliati a coloro che lavorano col colore). Qualità e fedeltà non sono spesso sinonimi.

Se si fotografa un *color checker* ogni colore viene poi riportato dai sensori della macchina fotografica e dipende fortemente dai filtri rosso/verde/blu applicati dal costruttore del dispositivo su ogni pixel catturato. Dai numeri colorimetrici della fotocamera alla colorimetria tradizionale esiste un passaggio che descrive le caratteristiche proprie della macchina fotografica; le cose si complicano di molto nel momento in cui si fotografa un'immagine che è già digitale.

Per caratterizzare un monitor come nella fotografia è necessario convertire la luce di un pixel colorato in tristimolo (per caratterizzare il dispositivo devo convertire il colore in un valore associato corrispondente ai suoi valori interni decisi dal costruttore). Utilizzare il punto di bianco aiuta ma non è detto che aiuti la corrispondenza in maniera perfetta perchè è innanzitutto da misurare e ci sarà sempre qualcuno disposto a cambiarlo. La gamma dei grigi allora non è mai lineare anche se ci si aspetterebbe il contrario: alla metà della suddetta gamma la luminanza d'ingresso dovrebbe coincidere con quella d'uscita, invece cambia a seconda del dispositivo (lo stesso problema si verifica nella stampa di ogni canale) → tutto ciò non coincide col gamut, la **gamma** rappresenta la non linearità del trasferimento dei toni. Il gamut possiede una rappresentazione tridimensionale e la gamma costituisce quel valore che rappresenta la distribuzione della densità del gamut. Una cosa importante che influenza la stampa è ovviamente il colore della carta.

Quando si rappresenta un gamut all'interno del diagramma di cromaticità lo si fa per comodità ma invece il gamut è sempre tridimensionale ed è difficilissimo caratterizzarne i contorni; il gamut è spaziale ma si accetta comunque la rappresentazione bidimensionale così da trasportare il gamut tridimensionale in spazi di natura diversa. La sovrapposizione dei gamut bidimensionali costituisce un'informazione incompleta.

I colori che si trovano nei pressi nel bordo del diagramma di cromaticità non sono mai stampabili mentre altri non si possono mai rappresentare all'interno di uno schermo. Per passare da una fotocamera ad un monitor c'è la necessità che le triplette siano correttamente traslate allora avrò la sicurezza che sia lo stesso pixel ad essere rappresentato, tuttavia nella colorimetria non è detto che esista per forza la stessa tripletta quando non c'è un'intersezione tra i gamut. Allora bisogna provare di cambiare la tripletta per ottenere lo stesso colore: tuttavia per i colori fuori gamut si ha bisogno di un gamut mapping eseguito a livello software, dipende da ACC (→ unione di aziende che hanno creato una serie di standard) che con i suoi 4 standard definisce altrettanti tipi di gamut mapping. La calibrazione di un solo dispositivo non basta a risolvere il problema e la calibrazione si estende a tutti i dispositivi integrati in modo tale che sia fatta con la massima precisione. La gestione del colore allora migliora ma non si risolve il problema in maniera decisa poiché i colori, per loro natura, non coincideranno mai del tutto.

XYZ è uno spazio assoluto mentre RGB è uno spazio relativo al dispositivo e si passa dall'uno all'altro attraverso una trasformazione lineare (resa anche non lineare per via della gamma). Si tratta di una semplificazione perchè nella realtà fisica i colori sono il risultato della combinazione di tre curve e non di tre soli valori numerici. Esistono moltissime matrici di trasformazione, tra cui:

1. **SRGB** che nasce dall'accordo tra software ed hardware per il trattamento del gamut, in questa maniera il gamut è “gammato” moltissimo rendendo la distribuzione non esattamente lineare per avvicinarsi il più possibile all'aspetto percettivo del colore. Più il gamut è piccolo più è probabile che sia compreso in altri più ampi (SRGB possiede un gamut molto piccolo);
2. **Adobe RGB** possiede invece un gamut enorme con la scusa di non aver così alcun problema a comprendere qualsiasi altro gamut esistente essendo esente così da limitazioni, tuttavia costringe a fare moltissime interpolazioni per tentare di riprodurre correttamente certi valori senza mettersi al sicuro da eventuali problemi di conversione (si passa dal dispositivo al profilo del dispositivo prima di passare direttamente ad un altro dispositivo → PCS)

Comprimere un gamut non rappresenta un problema, lo è invece espanderlo; comunque sia, si guadagna e si perde un po' di informazioni in entrambi i casi.

Gamut Mapping. Si divide in **globale** (quando si prende tutto lo spazio e lo trasforma interamente; tutti questi metodi tendono a scalare i valori RGB) e **locale** (quando i valori comuni ai due gamut rimangono inalterati a discapito di quelli che “sfiorano” che verranno quindi appiattiti sul corpo comune).

I metodi locali si dividono in:

1. **colorimetrico assoluto:** si è in uno spazio e si vogliono conservare più punti possibili di un gamut (è un metodo che si basa sul *clipping* perchè tutti i valori fuori dal gamut di riferimento verranno schiacciati sul bordo perdendo informazione). Si tratta di una mappatura geodesica con un punto oppure un asse centrale di riferimento: ahimè non è il metodo migliore (ma dopotutto nessuno lo è) perchè così è possibile perdere la possibilità poi di distinguere colori poco prima ancora distinti tra loro. Tutti i punti in comune non vengono interessati dal cambiamento. Il problema sta nel comportamento imprevedibile tra i punti all'interno ed all'esterno del gamut (e si spera che quelli interni siano di più rispetto a quelli esterni);
2. **colorimetrico relativo:** come il colorimetrico assoluto salvo che prima di appiattare le differenze tra i gamut si trasla il gamut da adattare in modo tale che coincidano i punti di bianco, dopodiché la procedura è sempre la stessa. Dipende sempre da quello che si vuole ottenere: preservare il bianco risulta comodo tuttavia si riduce l'area effettiva di sovrapposizione, tuttavia un colore unico di riferimento fa comodo → CAT (Color Adaptation Transport).

I metodi globali si distinguono tra:

1. **percettivo:** nome inadeguato perchè per via di un fraintendimento non è possibile definirlo percettivo nel senso della percezione visiva del colore. Fa un morphing tra uno spazio e l'altro adattando uno dei due gamut sull'altro (espandendo e contraendo a seconda delle zone) mantenendo sempre l'asse dei grigi invariato così da preservare le tinte. Costituisce tra tutte la soluzione migliore. Non essendo image dependent non è realmente percettivo;
2. **saturazione:** questo metodo preserva la saturazione perchè è considerata l'attributo colore più importante, quindi mappa ogni colore sul punto saturo più vicino il che comporta una variazione di tinta estrema. Torna utile quando si mappano i colori di un grafico nel quale non importa che i colori siano fedeli ma che si distinguano gli uni dagli

altri. Anche qui l'asse dei grigi rimane centrale e preservato.

Adattamento cromatico. (secondo il professore non esiste specificatamente, più che altro è sbagliato chiamarlo *adattamento* perchè non esiste un sensore del colore, sarebbe stato meglio *aggiustamento*) Definito come un meccanismo che mappa i toni dipendentemente dal contesto. Nel bianco, facendo così si ottiene una trasformazione globale; per il mapping percettivo il gamut andrà sempre ricalcolato.

Modello di Von Kriss → massimizzando ogni colore del modello si ottiene il bianco più adeguato.

Tutta la colorimetria ed alcuni tipi di gamut mapping tendono a salvare i colori di alcuni pixel dei quali non si conosce il vero colore; non si tratta degli strumenti migliori e si scelgono di volta in volta a seconda della necessità.

Profilare un dispositivo e fare gamut mapping sono cose differenti. La somma di gamma e gamut torna uno standard per la profilazione: avendo un gamut si può scegliere uno spazio XYZ e per ogni dispositivo si sceglierà il modo opportuno per passare dallo spazio del dispositivo a quello del profilo del dispositivo (si impone allora un criterio di gamut mapping). Il profilo del dispositivo è da intendersi come una grande tabella e la *gammatura* avviene tra spazi e non tra immagini.

Capitolo 4

Fotografia digitale

Introduzione alla Fotografia Digitale

Nella fase di acquisizione delle immagini all'interno del dispositivo avvengono alcuni procedimenti particolari; il risultato tuttavia non è mai una riproduzione fedele della realtà poiché, se così fosse, non risulterebbe gradevole al nostro sistema visivo.

Prima di scattare una fotografia è necessario considerare la **scena** che costituisce il fulcro ed il contesto del soggetto che si desidera fotografare. Si possono effettuare diverse rilevazioni sulla scena in modo da poterla così riprodurre, lo scatto allora verrà caratterizzato anche da valori di luminosità e qualità del colore.

La **fotocamera** che cattura la scena è formata da diversi dispositivi, primi tra tutti le **ottiche** che costituiscono sistemi di lenti le quali, filtrando la luce della scena, tendono a distorcere le immagini (si parla quindi di aberrazione cromatica quando attraversando la lente la luce si disperde in maniera non uniforme ed a seconda della componente colorata come se attraversasse un prisma; si tende al solito a correggere questo comportamento indesiderato con un'accurata progettazione della lente tramite computer); le *lenti* costituiscono solo la prima piccola stazione del viaggio della luce all'interno del dispositivo, la quale incontra presto alcuni **sensori**, oggi molto evoluti: alcuni sono dedicati al controllo dei principi fisici di base che risultano essere di buona qualità a seconda della risposta lineare che danno al bombardamento energetico solo ed esclusivamente in un certo range. Esistono sensori per il rumore del segnale tuttavia la qualità finale della foto è per l'80% merito del software di computazione del dispositivo, anche se la sensibilità è data dal meccanismo di amplificazione del segnale (che amplifica sì il segnale, ma non il rumore che spesso si porta appresso).

Il **camera control** si occupa di controllare aspetti importanti come l'apertura del diaframma ed il tempo di esposizione e di scatto. La messa a fuoco si dedica a misurare la luminosità sul piano del fuoco. Il **camera image processing** costituisce il valore aggiunto di qualsiasi macchina fotografica digitale moderna: tra le altre cose ricostruisce il segnale colore e modifica la risposta al colore in base all'esposizione; davanti ad ogni sensore, poiché è posizionato un filtro passa basso smerigliato, l'immagine subisce una sfocatura che asseconda la predilezione del nostro cervello per i bordi non troppo netti (senza questo sensore le forme sarebbero troppo nitide e darebbero troppo fastidio; esistono macchine nate senza questo specifico filtro che fanno spesso foto solo in bianco e nero eliminando così tutti i problemi connessi alla gestione del colore); nonostante ciò si tenta di recuperare parte dell'informazione eliminata dal filtro tentando così di acquistare nuovamente qualità e nitidezza attraverso l'utilizzo di un software.

Solo alla fine l'immagine viene **compressa** e **salvata**. Esistono diversi formati digitali a seconda della macchina e del costruttore, anche se solitamente si prediligono il *jpeg* ed il *raw* (quest'ultimo presenta l'assenza del colore). Se si tratta di una videocamera i formati di compressione sono ancora più variegati, molti dei quali sono basati sul *jpeg*, e solitamente si memorizzano in memorie

interne al dispositivo (fisse o mobili a seconda della tecnologia).

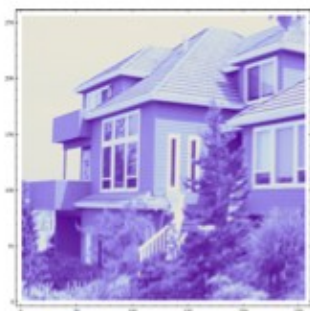
Trasformazione delle immagini

La matematica utile per lavorare e trattare con le immagini digitali è quella relativa alla trasformazione dello spazio.

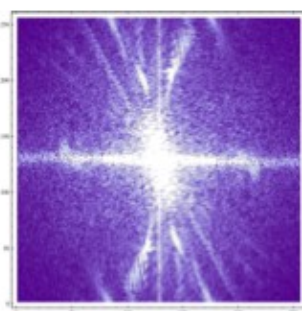
1. **Fourier** → utilizzo di procedure che descrivono uno spazio attraverso una serie di funzioni seno e coseno le quali, sommate tra loro, portano al segnale originale; da questo presupposto tutto è descrivibile mediante onde frequenziali;
2. **Cosine** → come sopra ma si utilizzano solo le funzioni coseno. Si rende il tutto più semplice e facile da calcolare;
3. **Wavelets** → dette anche funzioni *ondine*, si basano sullo steso principio di Fourier salvo per il fatto che in Fourier andava inevitabilmente persa la componente frequenziale del segnale così da modificare la parte irrazionale della fase che andava perduta.

Per leggere il risultato grafico di una trasformata di Fourier si prende in considerazione il quadrato in alto a destra (poiché il resto del grafico costituisce una ripetizione dello spazio positivo): alla base del grafico le frequenze sono basse ma salendo verso l'alto aumentano d'intensità. I punti bianchi indicano un'alta energia nelle basse frequenze che sono orizzontali, verticali e diagonali contemporaneamente. I punti scuri indicano invece degli abbassamenti di energia su quelle date frequenze. Con Fourier si analizza solo l'aspetto della luminanza perchè non si può applicare Fourier sui colori di un'immagine per via del fatto che come funzione è stata pensata per essere applicata ad un solo canale alla volta (ed il colore è costituito contemporaneamente tra 3 canali).

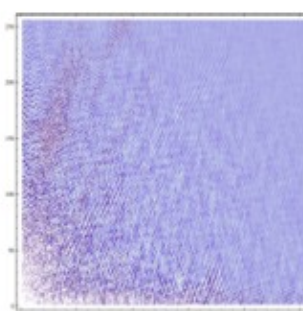
Per un'immagine il cambiamento di frequenza corrisponde a come essa si rappresenta nello spazio: sezioni scure e chiare possiedono frequenze differenti per via della loro natura.



Natural Image



Fourier Transform



Cosine Transform

La trasformata bidimensionale presenta doppie sommatorie discrete (ortogonali tra loro) rispetto a quella monodimensionale, tuttavia non cambia il principio matematico sottostante.

Per costruire un onda quadra si fa una sommatoria di onde spazzolando a partire dalla frequenza zero. Quando accade il contrario si ottiene il duale di quest'onda, ed è tale che si compie una **convoluzione**.

$$x_1(t) \otimes x_2(t) = \int x_1(\tau)x_2(t-\tau)d\tau \Leftrightarrow X_1(f)X_2(f)$$

Quando si sfoca un'immagine si fa una convoluzione della funzione gaussiana associata alla sfocatura pesando ogni pixel dell'immagine relativamente a tutti i valori della gaussiana stessa. La convoluzione esegue una sommatoria dell'onda in ogni suo punto per ogni suo punto. Nel campo delle frequenze una convoluzione diventa una moltiplicazione nel campo dello spazio. Tutti i filtri passa basso sono fatti in modo tale da simulare una convoluzione la quale sta alla base del processo di campionamento (si veda il *teorema di Shannon*).

Trasformata monodimensionale Trasformata bidimensionale

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df$$

$$F(u,v) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m,n) e^{-i \frac{2\pi um}{M}} e^{-i \frac{2\pi vn}{N}}$$

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt$$

$$I(m,n) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) e^{i \frac{2\pi um}{M}} e^{i \frac{2\pi vn}{N}}$$

$$e^{-j\alpha} = \sin \alpha - j \cos \alpha$$

$$e^{-i\alpha} = \sin \alpha - i \cos \alpha$$

Un segnale che non ha componenti in frequenza superiori ad una frequenza massima si dice limitato in banda, allora è ragionevole pensare che anche una immagine digitale sia per definizione limitata in banda; il segnale ricostruito deriva dalla frequenza di campionamento e può capitare di sotto-interpolare il segnale originale. Il teorema del campionamento stabilisce che un segnale limitato in banda può essere ricostruito senza errori da una serie di campioni rilevati ad una frequenza almeno doppia della frequenza massima, chiamata *frequenza di Nyquist*. Se il segnale è campionato ad una frequenza inferiore si dice che è sotto-campionato.

Lo spettro di un segnale corrisponde alla trasformata di Fourier tagliata, la quale presenta un collasso dei valori verticali ed orizzontali. Il limite in banda è dato dalla limitazione del campionamento del segnale e le riduzioni in frequenza corrispondono a delle relative riduzioni delle dimensioni delle immagini. Nella pratica abbiamo a che fare con segnali che non sono limitati in banda ma sono limitati nel tempo. Essi infatti vengono acquisiti solo in un intervallo di tempo (o di spazio) finito. Lo spettro di un segnale limitato nel tempo per poter essere correttamente ricostruito richiede una limitazione in banda. Infatti, nei segnali limitati nel tempo, lo spettro si ripete periodicamente ad una frequenza pari alla frequenza di campionamento. Queste diverse repliche dello spettro di sovrappongono alle altre cose se la frequenza di campionamento è inferiore alla frequenza di Nyquist, così lo spettro in questo modo non si annulla mai. In tal modo le alte frequenze del segnale si sommano alle basse frequenze della replica producendo il fenomeno dell'**aliasing**, chiamato anche equivocazione spettrale. La soluzione a questo problema consiste nel limitare in banda il segnale prima di procedere al campionamento, in modo da separare le repliche dello spettro. L'aliasing, nella concretezza della foto, tende a sfumare i bordi delle figure.

La **trasformazione wavelet** costituisce un sistema di funzioni applicabili non all'immagine ma ad un sotto-campionamento strutturato della stessa. Ciascuna

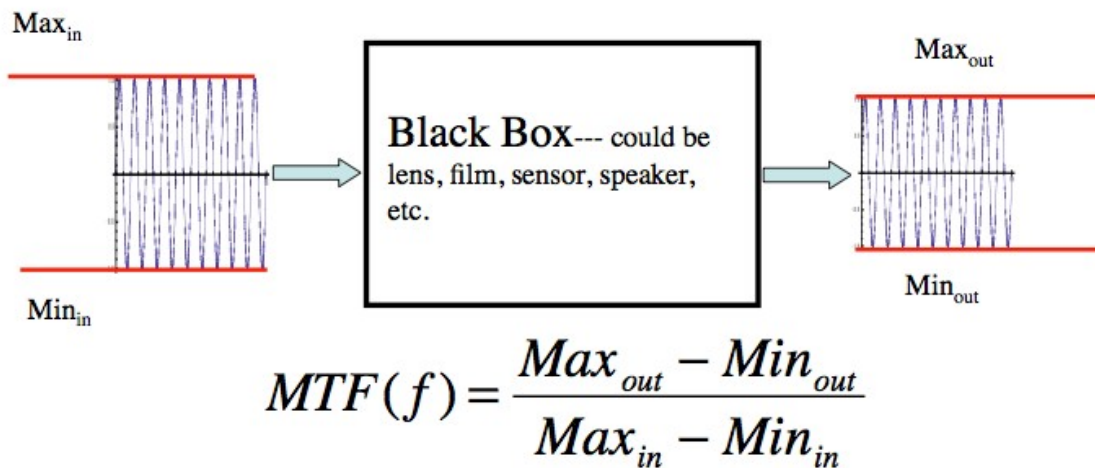
delle sotto-immagini possiede componenti orizzontali e verticali che descrivono in maniera iterativa lo spettro del segnale (le funzioni in questione sono strutturate in maniera multi-livello su una funzione di base). Si tratta di funzioni spaziali con repliche frequenziali che riducono mano a mano sempre più l'immagine di partenza.

Il **filtraggio gaussiano** costituisce un filtro passa basso con convoluzione dalla forma di una funzione gaussiana: si tratta di riportare la gaussiana su un supporto finito e di forma quadrata il quale schematizza i valori e la forma della gaussiana in maniera regolare. Ogni pixel dell'immagine risultante dal filtraggio corrisponde a tutte le possibili combinazioni di pixel dell'immagine originale, il tutto diviso la sommatoria dei pesi di ogni valore della gaussiana in ogni punto; la luminosità della foto risultante sarà la stessa di quella di partenza. In questa maniera il segnale può solo peggiorare; la procedura opposta mira ad ottenere una maggiore nitidezza (e la gaussiana di riferimento non poggia sull'asse delle x ma lo oltrepassa andando sotto lo zero): si amplifica allora la differenza tra pixel posizionati vicini aumentando il contrasto dei bordi.

Poiché i dispositivi digitali sono molto sensibili alla luce infrarossa sui sensori sono spesso posizionati dei filtri che ne smorzano un po' l'intensità.

Si definisce *risposta all'impulso* la risposta di un sistema al quale è dato un impulso ad una certa frequenza. Agire nel campo del frequenziale è più semplice che farlo nel campo dello spazio (consentendo manipolazioni più rapide).

La **funzione di trasferimento di modulazione** (MTF) definisce come un sistema lineare risponde ad input frequenziali. La pipeline è ovviamente limitata in frequenza e più pixel si hanno più componenti frequenziali si possono rappresentare, tuttavia non è detto che il sistema copra tutti i pixel necessari (allora la lente non *risolve* abbastanza). La lente risponde alla frequenza in ingresso come se si trattasse di un sistema lineare.



La funzione MTF è il risultato dell'ampiezza del segnale in ingresso sull'ampiezza del segnale in uscita in termini di frequenza, la cui frequenza quando sale fa decrescere il segnale. Se si ha una frequenza e la si vuole tagliare non si ottiene mai uno stacco netto (fisicamente impossibile) ma si presenterà un decrescimento graduale; non si eliminano i contenuti ad alta frequenza ma solo quelli a bassa frequenza. La MTF costituisce un filtro passa banda che filtra segnale per segnale; la frequenza spaziale è data dal numero di linee per centimetro ma quella ottica è data dal numero di linee per grado, allora la distanza in termini visivi non ha alcun peso. Qualsiasi supporto di stampa si

serve di una funzione MTF come qualsiasi sistema ottico tende ad agevolare il degrado della frequenza e del suo andamento. Una MTF serve sostanzialmente per ottenere uno spettro compatto che va a zero ad un certo punto nemmeno troppo distante dall'origine.

La lente

Una **lente** è un dispositivo che, presa in entrata una serie di raggi luminosi (solitamente appartenenti ad un'immagine), va a mettere a fuoco nella stessa identica maniera tutti i suddetti raggi su di un solo punto del piano. Spesso il punto, nella realtà, coincide con un dischetto un po' sfocato e decisamente più largo di un solo semplice punto. Un **diaframma** è un cerchio che regola la luce che entra nell'obiettivo fotografico ed è spesso descritto da un valore numerico in rapporto rispetto alla focale. Il diametro della lente coincide con la massima apertura del diaframma, mentre la **focale** costituisce la misura della lunghezza dell'obiettivo. Più la misura della focale è piccola maggiore sarà la grandezza del diaframma il quale sarà contenuto sempre meno volte nella lunghezza della focale. Il circolo di confusione della lente si riduce con il diaframma chiuso e la MTF lavora più agevolmente.

Più l'immagine è sfocata più risulta essere spazialmente autocorrelata con punti sovrapposti. Il sistema ottico serve per la creazione di immagini digitali: un'*immagine continua* è quella tipica delle pellicole la quale riporta il segnale in maniera, appunto, continua. Qualsiasi sensore da un segnale elettrico analogico e continuo, allora la risposta del sensore è lineare e dopo aver letto il segnale questo viene quantizzato dalla macchina all'interno di un range prestabilito in una fase detta fase di compressione.

La quantizzazione fa la differenza a parità spaziale; le HDR sono utili perchè, dato uno spazio può essere che si abbia a disposizione uno spazio più ampio per la registrazione delle differenze locali. Se la funzione MTF tende a distruggere le informazioni allora non è più così importante il numero di pixel di cui è capace il dispositivo.

Per passare dal continuo al digitale la quantizzazione è inevitabile perchè riporta il numero di bit di cui si ha bisogno per riportare fedelmente le informazioni dell'immagine. Il procedimento può tranquillamente essere non lineare scalando in maniera logaritmica così da dare più bit alle informazioni che si trovano sotto il valore medio (quindi le informazioni sulle ombre scure); il cinema sfrutta molto questa tecnica scalando logaritmicamente poiché fa uso di moltissime scene continue. I bit layer si aggiungono per ultimi e sono scelti a seconda del formato del file. La funzione MTF sarà allora legata alla risoluzione dell'immagine finale.

Il **sensore** è un fotodiodo bianco e nero che restituisce un segnale a seconda dell'intensità della luce che lo colpisce. Per poter cogliere il colore dell'immagine che si sta catturando è necessario apporre sopra il sensore un filtro chiamato **Bayer Pattern** il quale contiene molto più verde rispetto al rosso ed al blu (per motivi di trasmissione di senso di luminosità che il verde è più propenso a favorire nel nostro sistema visivo rispetto agli altri colori). Anche se il nostro sistema visivo risulta più affine al colore verde è importante dedicarsi per prima cosa alla restituzione della luminosità. Questo tipo di filtro (praticamente come qualsiasi altro filtro in quanto tale) tende a frenare molto la luminosità, così il verde tenta di ristabilire i valori naturali. Il campionamento è dato dalla misura della dimensione del sensore; la quantizzazione influisce molto sull'aspetto

finale della fotografia (determinandone la risoluzione spaziale).

Il campionamento della risoluzione è tutta un'altra cosa.

Il **bit planes** consiste nella rappresentazione delle immagini mediante 8 matrici tutte quante da un solo bit. La quantizzazione è legata quindi al dettaglio. Combinando le 8 matrici si aggiungono frequenze spaziali e dettagli un più. Il **bit depth** assomiglia al bit planes ma non è esattamente la stessa cosa poiché utilizza un numero differente di matrici.

Esistono formati d'immagine con solo 8 bit associati per pixel per comunicare le informazioni di colore di quel determinato pixel, forzando così l'utilizzo delle cosiddette *palette* nelle quale si riportano tutti i colori utilizzati dalla fotografia scegliendo tra una quantità precisa di colori (tipico delle immagini GIF). La *look up table* genera triplette RGB a partire da delle entry di colore. La differenza tra look up table e palette sta anche nella risoluzione.

Di formati d'immagine se ce sono moltissimi, eccone alcuni:

1. **Layer/Plannar Format** memorizza tutti i canali colore uno per volta;
2. **Pixel Interleav Format** salva la tripletta RGB di ciascun pixel un pixel alla volta;
3. **Linear Intrleav Format** fa la stesa cosa di cui sopra, ma una linea di immagine alla volta (anzichè un pixel alla volta).

Si definisce **halftone** la tecnica di stampa che realizza l'immagine colorata tramite macchie di colori primari che si combinano e sovrappongono per ottenere la giusta tonalità. Esistono tantissime versioni di halftone tuttavia sono tutti modi alternativi per distribuire colore sulla carta quando non è possibile creare il punto colore grande a piacere. La diffusione dell'errore per quanto riguarda le immagini mono-colore è terribile, mentre funziona molto meglio per quelle immagini che di colori ne contengono una certa quantità. La tecnica di Floyd e Steinberg è molto sfruttata per determinare la distribuzione dell'errore in modo da calcolarlo facilmente.

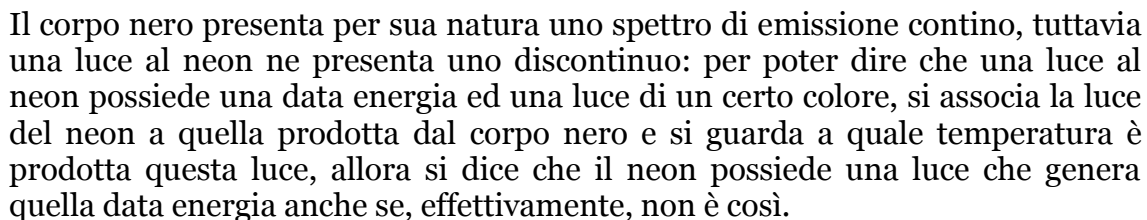
Il colore della luce ed il bilanciamento del bianco

Per riuscire a scattare una foto serve sfruttare una **sorgente di luce** che dovrebbe scaldare venendo assorbita dall'oggetto o dalla scena (soggetto della fotografia) e che in parte verrà riflessa colpendo così il sensore della fotocamera. Per **luce** s'intende quella specifica porzione dello spettro elettromagnetico che risulta visibile all'occhio umano; infatti l'occhio umano non è sensibile alla **luce infrarossa**, tuttavia lo sono i sensori delle fotocamere (anche se in esse gli IR sono in parte filtrati). Si può fare in modo che una macchina fotografica possa cogliere solo la luce infrarossa rimuovendo il filtro apposto dalla fabbrica sul sensore, tuttavia ne invalida la garanzia (e risulta essere alquanto laborioso).

La luce infrarossa aiuta a vedere attraverso la foschia in situazioni meteo che impedirebbero il filtrare della luce dello spettro visibile (ma con la nebbia non funziona perchè composta per la maggior parte da gocce d'acqua che per loro natura assorbono la luce, anche infrarossa, e non la lasciano filtrare).

Di solito la luce viene classificata in base alla propria direzione, intensità, colore e "durezza" (con il cui termine s'intende la dimensione della sorgente luminosa che se è piccola proietta ombre più nette mentre se è più ampia e distribuita nello spazio ne proietta di più sfocate). Il colore della luce non è mai il bianco puro perchè, se si pensa a quella solare, dipende da molti fattori come l'ora del giorno e la latitudine caratteristiche dell'osservazione. Il colore della luce si indica con la **temperatura colore** misurata in gradi kelvin (°K). Per far ciò si

Il **luogo plankiano** all'interno dello schema di cromaticità (disegnato come una curva che lo taglia a metà sfiorandone il centro) indica la cromaticità di un corpo nero esposto a diverse temperature.



31

valori ottenuti la quale restituisce un *valore di resa cromatica* compreso tra 1 e 100 (dove 100 è scelto come un valore a piacere per indicare la massima qualità ottenibile). La CAT è una funzione di *trasformazione di adattamento cromatico* la quale, in un certo senso, aiuta a bilanciare i colori di due scene diversamente illuminate. Se i campioni di colore osservati sono molto simili si avrà un valore di resa cromatica molto alto → allora anche il CRI sarà molto vicino al 100.

Il CRI è stato inventato negli anni '60 quando tutte le lampadine in commercio presentavano spettri continui, oggi purtroppo non è più così poiché le lampade moderne presentano picchi molto alti tanto da ridurre il CRI a valori tra il 70 e 80. Il valore del CRI, specialmente negli ultimi anni, non rispecchia più la percezione reale della qualità della luce (per lo meno non in tutti i casi), così si è tentato di studiare qualche alternativa ma nessuna è ancora diventata standard: i procedimenti moderni per il calcolo del CRI mirano a penalizzare gli estremi (luci troppo gialle oppure troppo azzurre) così da dire che una lampada con luce giallastra che magari percettivamente risulta ottima invece sia di scarsa qualità. Non trovando un'accordo sulla stesura di un nuovo metodo per il calcolo del CRI si mira a modificare le lampade in fase di costruzione affinché agevolino i calcoli matematici e “facciano tornare i conti”, tuttavia anche in questo caso ad un ottimo valore di CRI potrebbe non corrispondere una buona qualità percettiva della luce.

La **costanza cromatica** è data dalla visione di una scena illuminata in una certa maniera la quale diventa “normale” agli occhi del nostro sistema visivo così da rendere tale anche qualsiasi cosa esista in essa (così la luce verrà bilanciata automaticamente dal sistema visivo affinché risulti essere una luce normale). La macchina fotografica non riesce a fare la stessa cosa, allora viene in aiuto il **bilanciamento del bianco**. Quando si scattava ancora in pellicola esistevano delle pellicole apposite le quali venivano spesso combinate con lenti dedicate allo scopo: in questa maniera il bilanciamento del bianco serviva per donare un tono più naturale alle fotografie con una predominanza giallastra. Le macchine fotografiche digitali invece se sono impostate in automatico fanno quello che possono, in alternativa si può comunicare loro di utilizzare una scena di riferimento (solitamente pre-impostate nel dispositivo) oppure manualmente comunicando la temperatura in kelvin dell'illuminazione o fotografando un oggetto (come un muro oppure un foglio) bianco dicendo alla macchina che quello dev'essere il bianco di riferimento (quest'ultimo metodo funziona solo se la sorgente luminosa è singola oppure se sono più di una ma tutte dello stesso tipo e stessa temperatura).

Quando si scatta una foto la luce catturata sulla scena passa attraverso il sensore il quale trasforma la luce in un segnale che poi viene salvato in **file raw**. Il file raw possiede svariati bit per canale il quale viene poi convertito e compresso in file jpeg che possiede solo 8 bit per canale. Il bilanciamento del bianco si può fare all'interno della macchina fotografica (metodo preferibile) oppure in post-produzione (anche se pure il procedimento all'interno della macchina lo si può considerare in una qualche maniera post-produzione poiché si lavora comunque sul file raw come si lavorerebbe sul computer con software appositi sempre sui medesimi file raw). Il bilanciamento del bianco in post-produzione (quindi sull'immagine jpeg e non su quella raw) prevede diverse tecniche, tra cui:

1. ricerca nella fotografia del pixel più bianco possibile (ovvero la cui tripletta RGB sia la più vicina possibile a quella del bianco puro); una volta trovato si divide il valore del bianco puro per il valore del bianco della fotografia (canale per canale) ottenendo così un certo fattore

moltiplicativo per ciascun canale, dopodiché si bilanciano tutti gli altri pixel della fotografia moltiplicando i loro valori RGB con il risultato della suddetta divisione. Se il valore trovato non è un bianco allora siamo nei guai: si parte infatti dal presupposto che nella foto ci sia almeno un valore bianco (sennò è preferibile utilizzare un altro metodo);

2. si può scegliere il punto di bianco indicandolo;
3. *grey world*: si parte dalla fotografia e si calcola la media per ogni canale dei valori di tutti i pixel ottenendo così un valore che assomiglia molto ad un colore grigio (si dice allora che *la scena è mediamente grigia*); si divide allora 128 per il valore di ciascun canale ricavato dalla media; i canali di ogni pixel della foto originale allora verranno moltiplicati per il valore risultante dalla divisione per ciascun canale così da ottenere il bilanciamento finale.

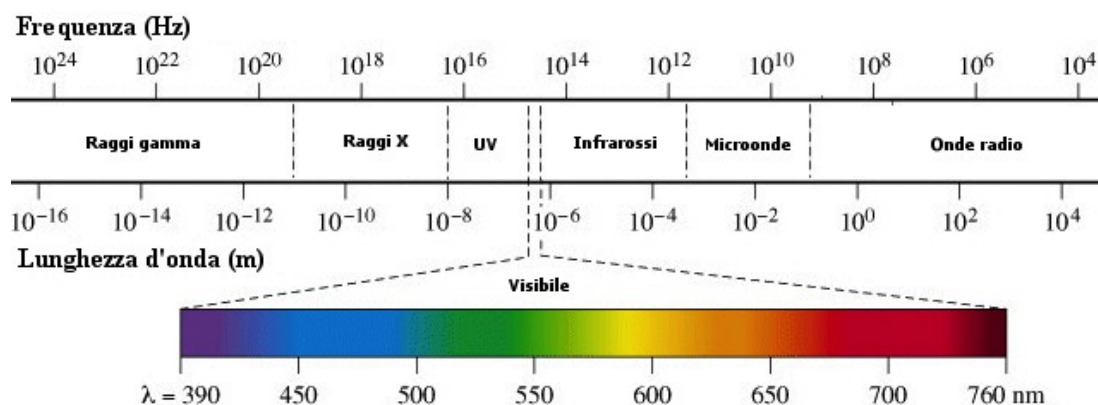
Ogni metodo risulta appropriato a seconda del caso in esame: non è detto che siano sempre tutti validi come non è detto che non lo siano.

La riflettografia infrarossa

Esistono foto che non sono solo descrittive ma che servono anche per studiare e misurare fenomeni fisici (come le foto al microscopio oppure le radiografie). La prima radiografia risale all'800 quando ci si rese conto che una pellicola fotografica lasciata accanto ad una sorgente di radiazioni rimaneva impressionata (da qui anche i primi studi sulla radioattività e le onde gamma). La fotografia astronomica invece serviva (e serve tutt'ora) per catturare lo spettro della luce proveniente dalle stelle così da poterne studiare i colori che lo compongono e poterne conseguentemente conoscere la composizione chimica. La fotografia è d'aiuto anche nello studio e nella conservazione dei beni culturali, poiché consente studi e rilevazioni anche se non ci si trova sul sito.

La *riflettografia* è praticamente una *fotografia* anche se per “disegnare” non usa la luce visibile, ma quella infrarossa; costituisce una tecnica ottica di analisi la quale produce un'immagine nell'infrarosso sfruttando radiazioni luminose invisibili all'occhio umano (NIR → Near InfraRed). Questa tecnica si sfrutta nello studio della natura e della storia dei dipinti e degli affreschi (ma anche dei papiri) per poter vedere così ciò che altrimenti non sarebbe visibile.

Lo **spettro dell'infrarosso** è decisamente molto ampio (più dello spettro del visibile) all'interno del quale si possono verificare molteplici fenomeni (come quelli termici che però non ci interessano).



L'infrarosso costituisce una radiazione riflessa da una superficie, tuttavia capita che a temperature più alte un oggetto riscaldato emetta a sua volta radiazioni

infrarosse (fenomeno legato alla natura dell'oggetto ed alla sua temperatura, che al momento non ci interessa). Per rilevare le radiazioni infrarosse serve una fotocamera sensibile alla luce infrarossa ed un corpo che emetta o rifletta questo tipo di luce, emessa da una sorgente.

Sfruttiamo le onde infrarosse per una questione di diffusione della luce. È noto che il colore del cielo e quello del sole cambino durante la giornata, per effetto della quantità variabile di atmosfera che si frappone tra il sole ed il nostro occhio: l'atmosfera infatti tende a filtrare la luce poiché composta non solo da gas ma anche da corpuscoli particellari che possono riflettere e deviare le onde luminose. Poiché la radiazione luminosa non è monocromatica, le lunghezze d'onda più corte (e più blu) tendono a rimbalzare in tutte le direzioni sulle particelle dell'atmosfera; al crescere della lunghezza d'onda il comportamento varia sino al punto in cui le lunghezze più ampie (di colore rosso) tendono a viaggiare nella direzione incidente rispetto a quella di propagazione, smettendo di essere isotrope (questo significa che la radiazione, man mano che entra nell'atmosfera, tende a viaggiare in linea retta ed in maniera quasi indisturbata → vicino al sole la luce tende a disperdersi, ma a debita distanza invece tende a viaggiare in linea retta e diretta). Questo comportamento è stato formalizzato da **Rayleigh**: *all'aumentare della lunghezza d'onda lo scattering (ovvero la dispersione luminosa dovuta ai corpuscoli dell'atmosfera) diminuisce*; è vero per lunghezze d'onda molto più ampie in dimensione rispetto alle particelle. Gli infrarossi fanno al caso nostro perché, trovandosi esattamente dopo le onde del rosso visibile, sono le radiazioni che viaggiano in linea più retta in assoluto.

Nel caso dei dipinti, si ha un oggetto colorato da pigmenti tenuti assieme da leganti ed addensanti (cosa che non avviene nell'inchiostro dove i pigmenti sono sciolti nell'acqua in maniera disomogenea). Si può dire che in un dipinto il pigmento è il corrispettivo della particella nell'atmosfera. Per creare un colore per la pittura i componenti venivano macinati e ridotti in polvere (il pigmento, appunto); la granulometria della polvere dava effetti di colore differenti, per questo motivo la macina dei componenti era una questione molto importante. Su una tavola di legno (oppure su una tela) andava quindi stesa una mano di patina bianca sulla quale si andava a disegnare il soggetto del quadro che poi passava al colore e per ultima cosa veniva ricoperto con una vernice trasparente atta a fissare i pigmenti sulla tela (vernice che nei secoli tende ad ingiallire perché filtra gran parte delle onde luminose). Il disegno era fatto solitamente con una matita (a mina oppure a punta metallica), oppure un carboncino, una penna, un pennello oppure una sanguigna; si utilizzava anche la tecnica dello spolvero per copiare su larga scala un disegno (la cui tecnica intendeva disegnare il soggetto su un foglio di carta sottile, forellare la sagoma del disegno ed appoggiare il foglio alla tela prima di spolverarlo di polvere di carbone: così la polvere passava attraverso i forellini tracciando i contorni del disegno).

Si sfrutta la radiazione infrarossa perché si spera che i pigmenti lascino passare qualche informazione invisibile ad occhio nudo. Si utilizzano le onde NIR (ovvero quelle dalla frequenza più bassa) perché oltre ad una certa lunghezza d'onda si è bloccati nell'acquisizione dall'assorbimento dei leganti che tengono assieme i pigmenti i quali assorbono la luce nonostante la diffusione in avanti della radiazione. Capita a volte di trovare delle macchie sui dipinti più scure o più chiare, le quali indicano interventi di restauro in tempi andati durante i quali si è scelto un colore molto simile a quello del dipinto, ma non del tutto uguale. Siccome capita spesso che i papiri siano talmente ingialliti nello strato superiore da non riuscire più a leggerne il contenuto, gli infrarossi aiutano a

distinguere i caratteri dipinti su di essi.

La pellicola per gli infrarossi è sensibile fino agli 800 nm. Al giorno d'oggi per la fotografia ad infrarossi si utilizzano macchine digitali con rilevatori al silicio, oppure con rilevatori all'arsenico, gallio ed indio oppure ancora con rilevatori al cadmio e mercurio (perchè sono tutti elementi chimici sensibili agli IR), spesso apparecchiatura militare che passa nelle mani degli studiosi solo quando viene dismessa. Per lavorare con i quadri si utilizzano macchinari molto ingombranti i quali vengono posizionati davanti all'opera d'arte perchè venga scandita punto per punto. Utilizzando un banco ottico normale non si riuscirebbe ad avere un rilevatore grande quanto l'immagine da acquisire, così si sono costruiti macchinari che tengono fermi ottica e dipinto ma muovono solo il rilevatore (così da non avere la distorsione tipica dell'ottica ripetuta per ciascuno scatto): si tratta di macchine di grandi dimensioni poiché serve una grandissima precisione. Oggi si utilizzano anche telecamere a movimento sferico le quali riprendono una superficie sferica → il tutto poi viene passato direttamente al software.

Per fotografare in infrarosso è necessario apporre diversi filtri davanti al sensore della macchina fotografica i cui filtri passa basso spesso filtrano anche gli IR (c'è anche un filtro che vibra per pulire l'immagine dalla sfocatura del filtro passa basso). Senza il filtro del colore passano solo le lunghezze d'onda del visibile (senza ultravioletti ed infrarossi); se si toglie il filtro IR la curva del rosso aumenta perchè ora il sensore è sensibile anche alla luce infrarossa. Se si vuole catturare solo l'infrarosso si deve apporre un nuovo filtro che blocchi tutte le lunghezze d'onda del visibile.

Nel caso degli affreschi è impossibile utilizzare la riflettografia poiché per loro natura costituiscono dei casi di pittura fatta su intonaco fresco così da incorporare i pigmenti nella parete; si salvano i casi di ripresa a secco (dove gli affreschi sono stati staccati dalla parete) o di tecniche di non affresco.

Color Appearance Model (CAM)

Il *colore* è un attributo fisico che si può misurare perchè è fondamentalmente un'onda elettromagnetica che si propaga nello spazio e nel tempo ma è al tempo stesso un attributo percettibile della percezione visiva: a dimostrazione di ciò si pensi che i fari di una macchina di notte risultano molto più fastidiosi che di giorno, che le stelle si vedono solo di notte e che non esiste una luce colorata di grigio o di marrone.

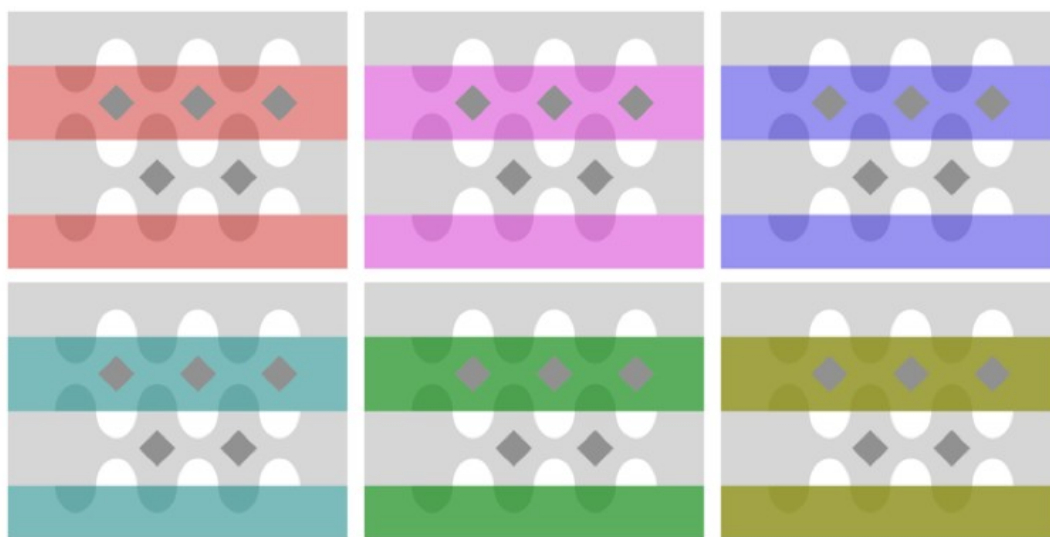
Le sensazioni percettibili non sono misurabili dalla colorimetria base, la quale torna però utile per dimostrare da parte di un osservatore standard che due stimoli distinti visti sotto le stesse condizioni di osservazione, risultano uguali. La colorimetria avanzata però mira ad estendere la colorimetria di base per poter specificare l'apparenza del colore e trattarne le dinamiche dell'effettiva percezione. Sensazione e percezione sono due termini simili che però indicano realtà differenti: la **sensazione** è quella che emerge dall'apparenza tuttavia la **percezione** incorpora elementi e fattori prettamente cognitivi (per esempio, se si avesse un disegno da colorare, la sensazione spingerebbe a riconoscere tutte le sfumature di colore che danno dimensione e rendono autentico il disegno mentre la percezione punterebbe a dare uno specifico colore a ciascun pezzo della composizione indifferentemente dal contesto come se fossero gli oggetti stessi ad essere “pitturati” e non una loro rappresentazione). La sensazione è comunque connessa all'apparenza e nella percezione entrano in gioco per forza

dei fattori cognitivi (la linea che separa queste due realtà è molto sottile). Lo scopo della funzione CAM è quello di studiare e regolarizzare l'apparenza.

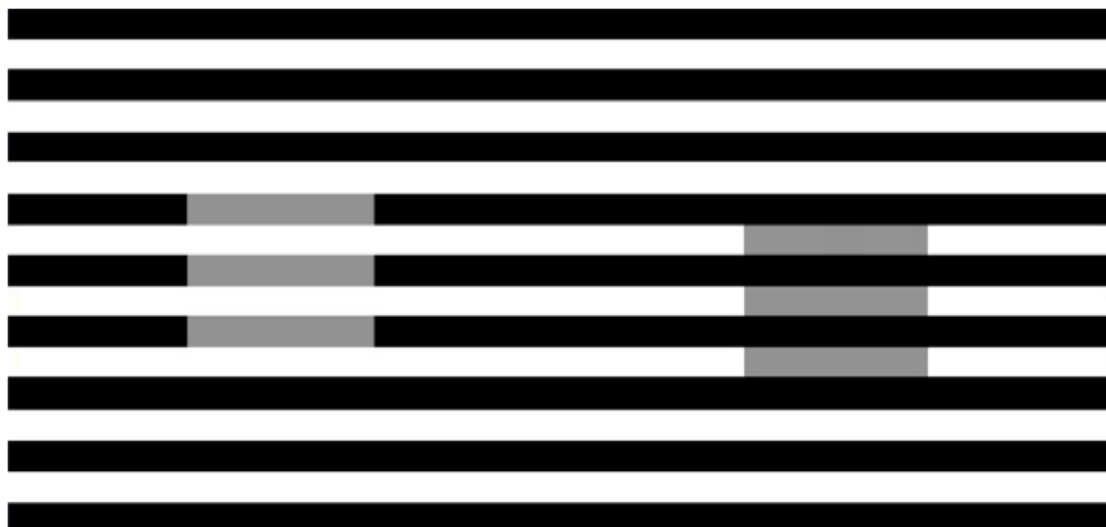
L'**afterimage** viene "scoperto" e descritto da Goethe nel suo trattato sui colori pubblicato già nel 1810. Si tratta della sensazione dovuta al fissare a lungo una certa immagine che si ha quando, distogliendo lo sguardo, si ricrea la stessa immagine, ma con colori opposti, se si fissa nuovamente lo sguardo su una parete monocromatica. La stessa cosa era già stata notata da Leonardo da Vinci nel 1570, il quale dedicava i suoi studi al contrasto tra colori opposti.

Il sistema visivo umano è composto da occhio, nucleo genicolato laterale (LGN) e corteccia visiva, il quale si dedica all'elaborazione delle informazioni visive. Nell'occhio si trova la retina la quale è ricoperta in superficie da 3 tipi differenti di coni (i quali si dedicano alla visione fotopica diurna in presenza di luce) ed un solo tipo di bastoncelli (i quali invece si occupano della visione scotopica in assenza d'illuminazione); i tre tipi di coni assorbono diverse lunghezze d'onda così da dare i natali alla **teoria tricromatica** (*Young 1802 ~ Helmholtz 1866*) la quale afferma che esistono tre diversi sensori nell'occhio umano, ciascuno dedito alla cattura rispettivamente delle lunghezze d'onda rosse, verdi e blu. Il principio del **tristimolo** sta alla base della colorimetria moderna, si pensi ai monitor con pixel di tre tipi illuminati da triplette RGB. La **teoria dei colori oppponenti** (*Hering 1872*), invece, afferma che esistono tre colori fondamentali ai quali si aggiunge il giallo e ciascuno di questi colori possiede un colore opposto: questa teoria fornisce la spiegazione per alcuni fenomeni visivi, come l'afterimage del quale si era già accorto Goethe. Le due teorie sono in leggero contrasto solo apparentemente perchè rappresentano due fasi diverse dello stesso procedimento visivo: la teoria del tristimolo è valida a livello retinico mentre il segnale che viene poi codificato in segnali oppponenti viene poi trasmesso al cervello.

Il nostro occhio, nonostante tutto, è diversissimo da una fotocamera. A riprova del fatto che sono strumenti differenti, ci vengono in soccorso le **illusioni ottiche** che sono comportamenti inattesi del nostro sistema visivo. L'immagine è elaborata nel cervello; non percepiamo il segnale in ciascun punto della scena ma facciamo attenzione alla composizione spaziale ed all'interazione degli elementi al suo interno. L'*esperimento di Daw* dimostra che i contorni degli oggetti sono importanti a livello cerebrale: sfruttando l'effetto dell'afterimage si facilita la ricostruzione dell'informazione se si utilizzano delle forme vuote (che poi l'afterimage va a colmare col colore impresso sulla retina).



La **costanza cromatica** dice che il segnale che giunge ai nostri occhi è una combinazione dell'interazione del colore della luce che funge da illuminazione con il colore degli oggetti che costituisce la riflettanza. Da un punto di vista fisico il colore degli oggetti varia a seconda dell'illuminazione di contorno nonostante noi percepiamo in maniera piuttosto stabile il mondo attorno a noi malgrado i cambiamenti spesso repentini di illuminazione. Il **contrasto di simultaneità** di Adelson dimostra che forme dalla stessa colorazione poste su sfondi collimanti di natura cromatica differente si percepiscono colorati in maniera diversa. Kitaoka aggiunge che non solo cambiano percettivamente colore anche nel caso cromatico. L'**assimilazione** si nota di più sulla base di pattern irregolari, funziona anche nel caso cromatico. È l'opposto del concetto di contrasto di simultaneità.



È il cervello che inganna perchè l'occhio il colore lo vede correttamente.

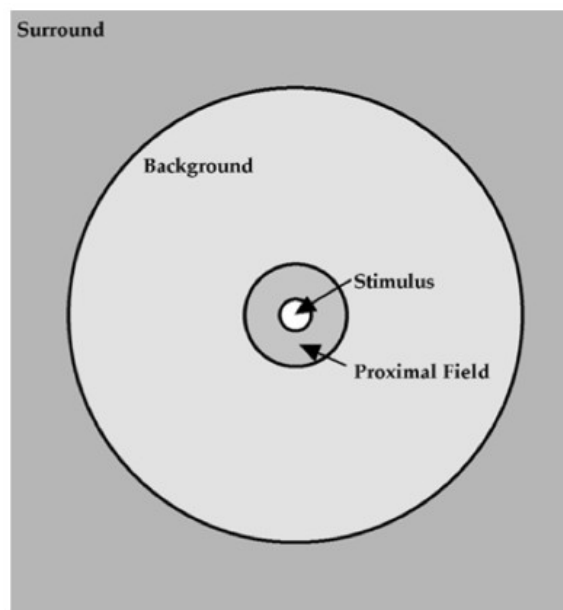
La colorimetria classica non è in grado di misurare e riportare questi fenomeni di illusione ottica, perchè per lei i colori che sono percettivamente differenti in realtà corrispondono alla stessa tripletta di tristimolo. La colorimetria avanzata invece tiene conto del contesto in cui lo stimolo viene visto, così da riportare che i colori percettivamente diversi sono diversi davvero (e per questo vengono istituiti dei termini specifici per descrivere il colore, come la tinta, il contrasto, la brillantezza, ecc).

Nel 1973 Wyszecki definisce la colorimetria di base come “in senso stretto, un tool usato per fare una predizione su luci qualsiasi (da intendersi come stimoli visivi) con distribuzione di energia spettrale differente le quali corrispondono in colore sotto date condizioni di osservazione. La predizione è fatta determinando i valori di tristimolo dei due stimoli visuali e se i valori di tristimolo degli stimoli sono identici a quelli dell'altro stimolo, un colore corrispondente sarà osservato da un diverso osservatore nella normale visione a colori”. La colorimetria include alcuni metodi per specificare il colore percepito in un contesto molto complicato (cosa molto difficile da fare).

Esistono moltissimi CAM (Color Appearance Model) costituendo un insieme di modelli che hanno tra l'altro molte proprietà in comune; lo spazio XYZ e tutti gli attributi del tristimolo non riescono più a descrivere in maniera esauriente il colore apparente, allora si inizia a parlare di **tinta**, **brightness** (o *brillanza*), **lightness** (o *chiarezza*) corrispondente alla *brillanza relativa*, **pienezza** (*colorfulness*) che indica l'intensità della tinta, **pienezza relativa** (*chroma*) e **saturazione**.

Brightness è un attributo per indicare una sensazione visiva in accordo con ciò che in un'area appare più o meno luminoso; la **lightness** costituisce la brightness di un area giudicata relativa alla brightness di un area similmente illuminata che sembra essere bianca o altamente riflettente (misura della brillantezza di un oggetto la quale rimane sempre la stessa nonostante l'osservazione in ambienti differenti). Per esempio, la pagina di un libro vista in un aula o vista in esterno al sole avrà una brillantezza diversa (al sole ci sarà maggiore brillantezza) ma stessa chiarezza (probabilmente in entrambi i casi sarà la cosa più chiara nel nostro campo visivo). La **tinta** invece è misurata come un angolo tra 0° e 360° all'interno del cerchio dei colori.

Le **condizioni di vista** determinano le condizioni dello stimolo: l'apparenza di uno stimolo dipende dallo stimolo in sé e dal contesto. Solitamente si ha uno stimolo ed un proximal field nelle immediate vicinanze dello stimolo (che viene considerato solo nel modello di Hunt) seguiti poi da una grossa porzione di background e quindi di surround (che costituisce l'ambiente nel quale avviene l'osservazione che può essere solo poco o molto illuminato). Il background viene inteso acromatico per comodità, tuttavia nella realtà non lo è mai (il che fa in modo che questo modello generico di condizioni visive non si possa attuare mai a nessuna immagine del mondo reale). Lo stimolo è dato dai valori di tristimolo e risalta a causa della natura acromatica (in scala di grigio) del background. Il surround indica la luminosità relativa rispetto al valore di tristimolo.



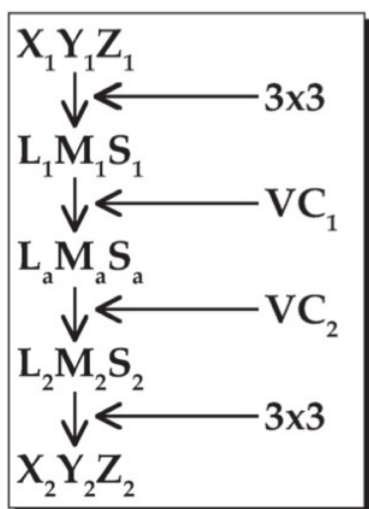
L'**adattamento cromatico** fornisce una spiegazione plausibile per il fenomeno di quei piccoli cambiamenti dell'apparenza dei colori, anche se varia di molto il colore dell'illuminante → possibile spiegazione della costanza cromatica come capacità del sistema visivo di adattarsi all'illuminante. L'adattamento alla luce induce un calo di sensibilità visiva, mentre l'adattamento al buio è il suo esatto contrario con conseguente aumento della sensibilità visiva nel caso di diminuzione di luminosità nella scena. Sono entrambi dei meccanismi fisiologici.

Nel **corresponding color data** due stimoli visivi posti sotto condizioni di illuminazione diversa appaiono comunque dello stesso colore (così due valori di tristimolo diversi danno invece lo stesso risultato cromatico). Per schematizzare,

uno stimolo specificato dai valori di tristimolo $X_1Y_1Z_1$ visto sotto certe condizioni VC_1 può apparire lo stesso di un secondo valore $X_2Y_2Z_2$ visto sotto altre condizioni di vista VC_2 . $\{X_1Y_1Z_1, VC_1, X_2Y_2Z_2, VC_2\}$ costituiscono coppie di colori corrispondenti CC; servono per testare la funzione CAM, così che dato $X_1Y_1Z_1$, VC_1 e VC_2 si riesca a predire $X_2Y_2Z_2$ e si paragona il risultato ottenuto con la CC. Le coppie di colori del CC si possono indicare sul grafico di cromaticità in modo da indicare i colori corrispondenti ed i relativi valori di tristimolo sotto le due diverse illuminazioni. I colori corrispondenti CC si ottengono in maniera sperimentale:

1. **Haloscopic matching:** si aspetta che gli occhi si adattino a due condizioni di luce differenti. Il sistema vale solo se l'adattamento avviene in maniera indipendente; non vale se intervengono meccanismi cognitivi.
2. **Memory matching:** si fanno corrispondere due stimoli visivi raccolti però in momenti differenti.
3. **Magnitude estimation:** in seguito alla valutazione dello stimolo luminoso, si assegna un valore arbitrario agli attributi di apparenza quali tinta, chiarezza e saturazione.

La CAT (conosciuta anche come Trasformazione di Adattamento Cromatico) esegue una trasformazione a partire dai valori di tristimolo raccolti in certe condizioni di vista in altri valori di tristimolo corrispondenti ad un'altra condizione di vista, senza includere però gli attributi di apparenza come la tinta (che sono di pertinenza della CAM). Serve per modellare i meccanismi fisiologici di adattamento cromatico, tant'è che si converte l'informazione di tristimolo dallo spazio XYZ a quello LMS che è quello corrispondente all'eccitazione dei coni (cosa che si può approssimare con una trasformazione lineare mediante una matrice 3×3).



S'incomincia con i valori di tristimolo $X_1Y_1Z_1$ rappresentati per convenzione nello spazio CIE i quali riportano i valori dello stimolo alla prima condizione visiva $VC_1 \rightarrow$ li si trasforma in una tripletta di valori $L_1M_1S_1$ i quali corrispondono al valore di eccitazione dei coni \rightarrow vengono quindi incorporate le informazioni sulla prima condizione visiva VC_1 servendosi del modello di adattamento cromatico per predire il relativo adattamento del segnale dei coni $L_aM_aS_a \rightarrow$ si esegue il processo inverso con il secondo dato sulle condizioni visive VC_2 per determinare il colore corrispondente in termini di eccitazione dei coni $L_2M_2S_2 \rightarrow$ quindi lo si trasforma in un valore di tristimolo $X_2Y_2Z_2$ nello spazio CIE.

La CAT non è la CAM (sono due funzioni differenti); la CAM fornisce in più le indicazioni sugli attributi percettivi del colore, come tinta, chiarezza, ecc.

Von Kries definisce a parole (quindi non si serve di formule matematiche) il procedimento di adattamento cromatico: dice che le componenti ricettive dell'occhio sono tra loro indipendenti e si adattano alla situazione in maniera indipendente. Modernamente si è voluto formalizzare tale assunto indicando che i coni adattatisi alla situazione luminosa $L_aM_aS_a$ sono matematicamente individuati come i coni di partenza LMS moltiplicati per certi coefficienti detti **coefficienti di scala** (*gain control*). Tali coefficienti si possono ricavare dall'inverso del valore dello stimolo massimo presente in scena per ciascun cono

oppure dall'inverso del valore per lo stimolo bianco. L'operazione può essere fatta anche in forma matriciale.

$$\begin{aligned}
 L_a &= k_L L & k_L &= 1/L_{\max} \quad \text{or} \quad k_L = 1/L_{\text{white}} \\
 M_a &= k_M M & k_M &= 1/M_{\max} \quad \text{or} \quad k_M = 1/M_{\text{white}} \\
 S_a &= k_S S & k_S &= 1/S_{\max} \quad \text{or} \quad k_S = 1/S_{\text{white}}
 \end{aligned}$$

$$\begin{vmatrix} L_a \\ M_a \\ S_a \end{vmatrix} = \begin{vmatrix} 1/L_{\max} & 0.0 & 0.0 \\ 0.0 & 1/M_{\max} & 0.0 \\ 0.0 & 0.0 & 1/S_{\max} \end{vmatrix} \begin{vmatrix} L \\ M \\ S \end{vmatrix}$$

La funzione di **Corresponding Color (CC)** serve per calcolare i segnali di adattamento nelle condizioni di vista iniziali, le quali devono essere uguali alle condizioni di adattamento delle condizioni di vista finali.

$$\begin{aligned}
 L_a &= L_2 / L_{\max 2} \\
 L_a &= L_1 / L_{\max 1}
 \end{aligned}$$

$$\begin{aligned}
 L_2 / L_{\max 2} &= L_1 / L_{\max 1} & L_2 &= (L_1 / L_{\max 1}) L_{\max 2} \\
 M_2 / M_{\max 2} &= M_1 / M_{\max 1} & \longrightarrow & M_2 = (M_1 / M_{\max 1}) M_{\max 2} \\
 S_2 / S_{\max 2} &= S_1 / S_{\max 1} & S_2 &= (S_1 / S_{\max 1}) S_{\max 2}
 \end{aligned}$$

La **CAM** (conosciuta anche come Color Appearance Model) è un modello scelto dalla CIE per includere quegli attributi percettivi che non considera la funzione di adattamento cromatico. Definito come un modello per l'apparenza del colore il quale non è incluso o previsto da nessun altro modello. Le specifiche del CAM prevedono in input uno stimolo e le condizioni di vista iniziali espresse in termini di valori di tristimolo XYZ (che sono quindi facili da misurare); la trasformazione in uno spazio dei coni in modo da modellare i processi fisiologici ed altri dati riguardanti l'ambiente entro il quale avviene l'osservazione (come il relativo background e surround). Si procede utilizzando la CAT con l'inclusione di un fattore D di adattamento, quindi si trasformano i segnali in forma di colori opponenti ed infine si opera una combinazione dei segnali per predire gli attributi della percezione visiva come tinta, chiarezza, ecc.

Il CAM standard ora è indicato come CIECAM02; qualsiasi CAM, comunque, stima l'apparenza dei singoli stimoli osservati su un background uniforme e spesso acromatico: in realtà è rarissimo trovare stimoli su sfondi grigi, così il CAM non risulta idoneo per calcolare l'apparenza nelle immagini fotografiche. Per ovviare al problema nasce **iCAM**, un modello ideato nel 2002 da Fairchild in modo tale che si possa applicare sulle immagini (nato inizialmente come *tone mapper* per immagini HDR, dove tali immagini, che non possono essere visualizzate a monitor salvo mediante un'approssimazione, risultano essere il prodotto di sovrapposizioni di immagini catturate a diversi toni di saturazione

per ottenere un'immagine ad alta dinamicità con valori di luminanza reali) estendendo così il modello CIECAM02. L'immagine HDR viene separata in due immagini una contenente tutti i dettagli dello scatto in alte frequenze (immagine dei dettagli) e l'altra contenente le basse frequenze detta immagine di base. Successivamente si lavora solo sull'immagine di base, la quale subisce un *white layer* che ne fornisce una versione sfocata (considerata il punto di riferimento per il punto di bianco dell'immagine HDR). In seguito si attua un adattamento cromatico all'immagine sfocata la quale verrà poi compressa e ricongiunta con l'immagine in alte frequenze contenente i dettagli. Si ottiene poi un'immagine finale RGB a bassa dinamicità che passa da 32 bit a soli 8 bit di colore. L'adattamento cromatico di iCAM sottintende la divisione di ogni punto dell'immagine per il corrispettivo punto di bianco dato dal white layer (allora il punto di bianco sarà diverso per ogni layer).

iCAM mira a separare quei processi che nel nostro cervello sono una cosa sola; il problema di iCAM è il tentativo di simulare esattamente il processo visivo quando ancora non è chiaro come il sistema visivo funzioni esattamente; per questo si fa largo uso di tantissime equazioni empiriche, nate da esperimenti di difficile valutazione.

Termografia infrarossa

La termografia sfrutta le onde dell'infrarosso più lontane dal visibile; queste onde non generano un riflesso quando colpiscono la superficie dell'oggetto poiché l'oggetto tende ad assorbirle ed a scaldarsi: in questa maniera è più facile studiare il comportamento della radiazione emessa, appunto, nell'infrarosso.

Nel 1800 Herschel scopre l'infrarosso attraverso un banale esperimento durante il quale con un prisma scisse un fascio di luce nelle sue componenti, quindi misurò la temperatura di ciascuna componente colorata rendendosi conto che la temperatura aumentava anche oltre lo spettro visibile, là dove si trovavano gli infrarossi; si tratta di una radiazione che parte da un fenomeno globale il quale testimonia che ogni corpo riscaldato ad una temperatura maggiore dello zero assoluto emette energia ad almeno una lunghezza d'onda nello spettro elettromagnetico. L'energia emessa in questione è proporzionale alla temperatura del corpo, piccole differenze di temperatura tuttavia scatenano grandi differenze di energia emessa. Se la temperatura aumenta la lunghezza d'onda diminuisce.

Un'immagine termica mostra la distribuzione della temperatura nella scena. Tutto ciò che viene rilevato è la temperatura superficiale dell'oggetto visionato poiché non si può vedere oltre la sua superficie (questo significa che magari la temperatura interna sarà maggiore o minore, ma non ci è dato di saperlo). Si tratta di informazioni associate all'irraggiamento dell'energia a partire dal corpo (se lo toccassimo sarebbero informazioni dedotte per conduzione, se fosse l'aria o qualsiasi altro mezzo in movimento a portarle sarebbe per induzione) ottenute senza entrare in contatto diretto con lui.

Quasi nessun corpo emette energia come se fosse un corpo nero; attraverso un **fattore di emissività** (scelto da una tabella o tramite un esperimento in laboratorio) si indica quanto veramente un corpo assomiglia ad un corpo nero.

Un **termogramma** è una fotografia termica, inizialmente resa in scala di grigi, associata ad una scala colorata la quale riporta la temperatura dei corpi fotografati. Materiali alla stessa temperatura possono possedere una differente emissività.

La **termografia** misura l'irraggiamento dal quale, con opportune conversioni, si può ricavare la temperatura superficiale di un corpo. Per far ciò si utilizzano delle **termocamere**.

L'atmosfera scherma diverse lunghezze d'onda. Sulle lunghezze d'onda del sole (ovvero quelle della luce visibile) non assorbe nessuna radiazione, idem in certi punti dell'infrarosso (è per questo che lo possiamo sfruttare) le cui onde possono essere viste con strumenti per la rilevazione delle onde corte (short wave) oppure per le onde lunghe (long wave) a seconda delle onde che si vogliono vedere. Il vetro non è trasparente all'infrarosso, anche se lo è alla luce visibile, così come certe lunghezze d'onda dell'atmosfera.

La termografia non è affatto invasiva o distruttiva e permette di individuare certi fenomeni che non sono visibili ad occhio nudo; lo si può fare a comoda distanza con dispositivi con ottiche non di vetro (quindi abbastanza costosi). In questo modo si osserva la temperatura superficiale di diversi materiali nascosti magari dall'intonaco (se si parla di costruzioni in muratura) che è abbastanza sottile da essere praticamente ignorato. In modalità passiva la rilevazione in termografia è fatta senza applicare alcuno stimolo luminoso o termico a ciò che viene osservato sfruttando così eventuali stimoli naturali come la luce del sole (si può ottenere lo stesso risultato non solo riscaldando ma anche raffreddando, l'importante è spezzare l'equilibrio termico). In modalità attiva invece si punta ad alterare l'equilibrio termico in maniera attiva sparando sull'oggetto osservato aria calda (o fredda) per ottenere una certa temperatura. Nei beni architettonici si sfruttano queste tecniche per:

1. vedere se nelle pareti ci sono delle **infiltrazioni d'acqua** che provocano umidità (l'acqua che evapora porta via calore ed induce un raffreddamento evidenziando di conseguenza l'eventuale *risalita capillare dell'acqua*/umidità nelle pareti). Attraverso il *taglio chimico* s'intende combattere l'effetto della risalita capillare poiché l'acqua che sale e poi evapora lascia all'interno delle pareti degli accumuli di sali e detriti che alla lunga, come cunei, spaccano i laterizi squarciando e destabilizzando le pareti. Il taglio chimico serve a combattere questo fenomeno intervenendo fisicamente e chimicamente alla base dei muri.
2. diagnosticare dei **distacchi** dell'intonaco che può formare delle bolle d'aria per poi distaccarsi definitivamente dalla parete. Il distacco è visto con la termografia come una zona più calda dovuta al fatto che la bolla d'aria si riscalda più facilmente rispetto al resto del muro se irraggiata (a questo punto però bisogna assicurarsi che l'effetto riscontrato sia dato proprio da un fenomeno di distacco e non semplicemente da un'illuminazione localizzata).
3. vedere come **certe murature** sono state edificate e **si caratterizzano**, così si può vedere come i diversi materiali emettono radiazioni infrarosse (distinguendo così materiali più nuovi oppure pattern di muratura in disaccordo con il resto della costruzione).

La termografia è utilissima per studiare eventuali ponti termici (ovvero zone nelle quali c'è uno scambio termico attivo tra l'interno e l'esterno di un edificio). Oggetti scuri tendono a riscaldarsi più velocemente rispetto ad oggetti chiari, per evitare disuguaglianze quando si riscalda un'oggetto qualsiasi per studiarlo termograficamente si preferisce l'irraggiamento, tuttavia l'aria calda tende comunque a salire ed a scaldare senza volerlo oggetti più alti (come per esempio soffitti). Lo studio termografico cambia a seconda delle fasi di riscaldamento o raffreddamento nelle quali si possono notare elementi differenti su uno stesso

oggetto. Per studiare i distacchi è meglio sfruttare un termogramma a colori, ma per studiare la tessitura muraria è più facile guardare un termogramma in scala di grigio.

La termografia può essere impiegata in altri modi, come nello studio medico della risposta muscolare in soggetti atleti e non.

L'esposizione

Si parte sempre dal presupposto che una foto venga scattata correttamente, allora il *sensore*, considerato come un insieme di piccole vaschette una per ogni pixel, viene riempito di luce come se si riempisse d'acqua allo scattare di una qualsiasi fotografia. Può capitare tuttavia che i contenitori di luce vengano riempiti troppo oltre la loro portata: la luce comincia ad esondare (effetto di blooming) ed a riempire nuovamente le vaschette adiacenti (innescando un effetto a catena poiché influenzerà la ricezione luminosa dei pixel vicini), in questa maniera la luminosità non viene più correttamente quantificata. Se il sensore si riempie troppo poco, invece, si crea del rumore di fondo.

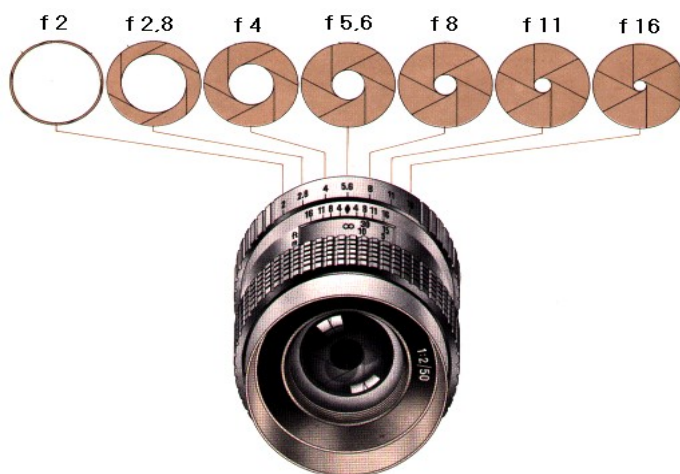
Analizzare il valore di un solo pixel all'esterno del contesto ed isolato dai suoi vicini non ha nessun significato: si preferisce che l'intera foto/matrice abbia un senso complessivo. Se si fotografano due ambienti adiacenti bisogna scegliere l'esposizione dell'uno o dell'altro e la scelta viene influenzata dall'obiettivo finale della foto: interessa maggiormente massimizzare l'acquisizione di luce in maniera automatica (altrimenti significa che viene fatta un'esposizione detta a *forcella manuale*).

Il **diaframma** è un dispositivo meccanico a lamelle il quale crea una forma regolare (circolare o rettangolare) ma variabile (ovvero che si allarga e si restringe a seconda dell'occasione). Il diaframma regola l'apertura ed il flusso di luce in entrata mentre l'**otturatore** regola il tempo d'esposizione. Nella fotografia digitale non è necessario utilizzare l'otturatore perchè virtualmente sarebbe possibile leggere la luce che colpisce il sensore per una durata di tempo stabilita, tuttavia sarebbe opportuno poter “svuotare” il sensore dalla luce che lo colpisce in modo tale da farla fluire per evitare l'accumulo ed il trabocco, così si preferisce (per evitare complicazioni) comunque montare un piccolo otturatore anche nelle macchine digitali.

L'otturatore è catalogabile in:

1. centrale: in questo caso l'apertura avviene verso i bordi del dispositivo e la chiusura verso il suo centro esatto. Questo significa che il centro sarà maggiormente irrorato di luce rispetto alla periferia che ne riceverà per meno tempo. Si tende a posizionare l'otturatore al centro della focale direttamente sull'obiettivo e non all'interno della macchina fotografica nei pressi del sensore, in questa maniera risolve temporaneamente il problema ma aumentano i costi di produzione;
2. a lamelle o a saracinesca: per risolvere il problema dell'irrorazione di luce più abbondante al centro si è creato un otturatore di forma rettangolare grande quanto il sensore (e posizionato su di esso) il quale consta di due tendine a lamelle che scendono una dopo l'altra (con un certo scarto temporale precedentemente settato) in maniera tale che ogni porzione dell'immagine venga illuminata alla stessa maniera per la stessa quantità di tempo. Purtroppo questo tipo di otturatore non va molto d'accordo con l'utilizzo del flash la cui velocità non permette tempi di sincro troppo lunghi.

Il diaframma è teoricamente chiudibile a qualsiasi delle sue altezze, nella realtà per una questione di comodità sono state comunemente istituite delle tacche dette **stop** (misura per la quale la quantità di luce incamerata raddoppia o si dimezza tra una tacca e l'altra). Larghezza del foro e numero indicatore sono inversamente proporzionali → F16 vuol dire che il diametro del diaframma è ripetuto 16 volte all'interno della lunghezza della focale. Il valore del diaframma raddoppia ogni due stop in maniera vicendevole e tra un valore e l'altro c'è una differenza di radice di 2. I tempi dell'otturatore sono espressi in secondi ed anche tra di loro c'è una differenza esprimibile in stop (si ricorda che in questo caso il tempo è da considerarsi ancora lineare).



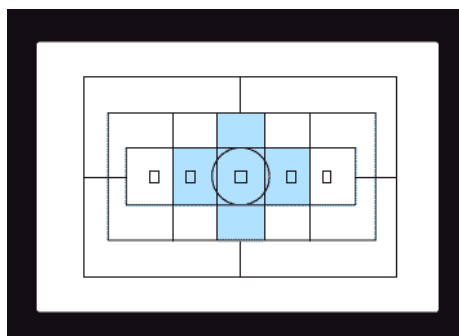
Si può giocare con tempo e diaframmi (attraverso conformazioni e coppie diverse ma con lo stesso effetto) in modo da forzare sempre la stessa quantità di luce ad entrare nell'obiettivo. Un diaframma molto aperto con una minima profondità di campo può essere associato ad un tempo molto veloce (ideale per la cattura di immagini sportive). Tali combinazioni sono riassunte nella tabella di valori EV.

Su una macchina fotografica è possibile esporre in modalità manuale (ovvero scelgo autonomamente la giusta ed opportuna combinazione di otturatore e diaframma con un occhio di riguardo ai tempi che vanno scelti per primi, allora il valore di diaframma verrà scelto di conseguenza anche se è preferibile scegliere il valore del diaframma per una questione di comodità) sennò si lavora in automatico/program (ed è la macchina a scegliere le coppie di valori già preimpostati dal costruttore all'interno di una serie di modalità → tutti “falsi” perchè sono solo degli escamotage per la scelta della coppia tempo/diaframma quando invece le modalità sono appena 3 con lo *sportivo* con tempi molto corti, il *panorama* con la chiusura più stretta del diaframma e lo *standard* che è una via di mezzo; tutti gli altri sono specchi per le allodole per far vendere di più l'oggetto macchina fotografica).

La sensibilità della pellicola/sensore dev'essere di un certo tipo per fare in modo di ottenere un buon risultato, inoltre c'è il problema del grigio medio.

Misurare la luce incidente (illuminamento) e riflessa (luminanza) di una scena è fondamentale. Utilizzando un **esposimetro** si misurano entrambe le forme di illuminazione: l'oggetto possiede una calottina sferica e bianca per raccogliere meglio la luce diffusa in maniera il più possibile omogenea (per la misurazione della luce riflessa basta togliere la calottina). Anche la macchina fotografica possiede al suo interno un piccolo esposimetro che però coglie solo la luce

riflessa (dipende però da quanto è chiaro o scuro il soggetto). L'esposimetro basa la sua quantificazione sul calcolo del grigio medio dando per scontato un valore medio di illuminazione: con luci molto intense il bilanciamento fa chiudere il diaframma (il contrario per luci meno intense) anche se in verità non è la scena ad essere luminosa ma è il soggetto (es: una sposa in bianco). In presenza di oggetti chiari è necessario sovraesporre per bilanciare l'errore commesso dal sensore di luminosità della macchina fotografica. In alternativa si misura il grigio medio basandosi su un grigio di riferimento (che è la stessa cosa che si fa mediante la misurazione tramite la calottina bianca). Può capitare che sia lo sfondo ad essere molto chiaro oppure molto scuro: se si appone una calottina sull'obiettivo la misurazione può essere eseguita direttamente con la macchina fotografica. L'esposimetro insiste su diverse geometrie a seconda dei casi, in prevalenza sul fuoco centrale oppure su un solo punto al centro (spot): l'importante è conoscere bene i settaggi della propria macchina fotografica.



All'esposizione corretta concorre anche il valore della **sensibilità** della pellicola (o sensore) detta **ISO**. Una pellicola molto sensibile ha bisogno di poca luce (diaframma chiuso) adducendo però molto più rumore. I valori di ISO non hanno una numerazione che raddoppia. Una volta c'era la grana che oggi però è sostituita dal rumore che raccoglie il sensore.

La tabella degli EV può essere sfruttata per conoscere i giusti valori a seconda della sensibilità della pellicola: se la sensibilità aumenta i valori vanno traslati da destra verso sinistra di un passo per ogni salto di sensibilità.

☆ Tabella della relazione fra valore dell'apertura e valore della velocità dell'otturatore, valore EV

$$EV = AV + TV$$

AV \ TV		0	1	2	3	4	5	6	7	8	9	10
		1.0	1.4	2.0	2.8	4.0	5.6	8.0	11	16	22	32
0	1s	0	1	2	3	4	5	6	7	8	9	10
1	1/2	1	2	3	4	5	6	7	8	9	10	11
2	1/4	2	3	4	5	6	7	8	9	10	11	12
3	1/8	3	4	5	6	7	8	9	10	11	12	13
4	1/15	4	5	6	7	8	9	10	11	12	13	14
5	1/30	5	6	7	8	9	10	11	12	13	14	15
6	1/60	6	7	8	9	10	11	12	13	14	15	16
7	1/125	7	8	9	10	11	12	13	14	15	16	17
8	1/250	8	9	10	11	12	13	14	15	16	17	18
9	1/500	9	10	11	12	13	14	15	16	17	18	19
10	1/1000	10	11	12	13	14	15	16	17	18	19	20

* L'asse orizzontale è per i valori dell'apertura e l'asse verticale è per i valori della velocità dell'otturatore. I numeri in grigio indicano l'EV.

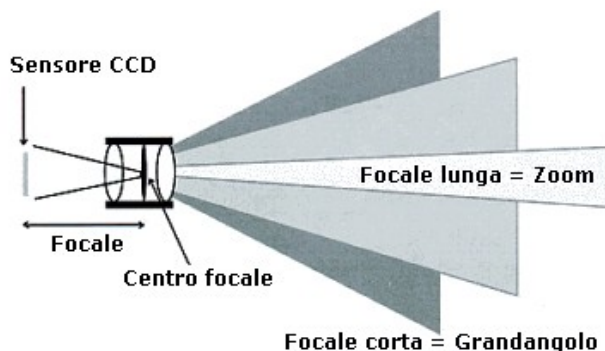
Quantificare la quantità di luce è importante, ma è altrettanto importante capire da dove questa luce proviene. EV può essere sfruttato come misurazione della quantità di luce solo se si lavora con una sensibilità di 100 ISO.

La **messa a fuoco** rende possibile focalizzare alcune porzioni di una scena anche se non necessariamente centrali. Tale tecnica sfrutta la trasformata di Fourier massimizzata: quando si sfoca una scena si applica la trasformata la quale schiaccia il contenuto frequenziale. La messa a fuoco applica alla stessa maniera la trasformata di Fourier tenendo in considerazione l'immagine ottenuta che la permette in maniera più ampia ed esauriente.

La **linea di fuoco** è una linea infinitesimale, in essa un solo piano della realtà viene mappato correttamente. Il piano di messa a fuoco può essere spostato, in questa maniera non permette la mappatura di un solo pixel ma al suo posto di una serie di pixel ad esso circostanti istituendo una sorta di cerchietto (detto *cerchio di confusione*, accettabile quando la sua dimensione è minore di quella del pixel che lo genera e lo contiene). La profondità di fuoco sta per 1/3 prima del soggetto a fuoco (linea focale) e per 2/3 dietro tale soggetto. Un diaframma stretto significa un cerchio di confusione altrettanto stretto. Più il soggetto è vicino più è bassa la profondità di campo; più la focale è corta più aumenta la profondità di campo.

Si parla di **iperfocale** se voglio fotografare un soggetto enorme e lontanissimo (come una montagna) e non mi interessa che sussistano quei 2/3 di scena a fuoco alle sue spalle. Il soggetto allora viene posto al limite della distanza focale (in maniera arbitraria senza per forza mettere la montagna sulla linea focale). In pratica l'iperfocale costituisce il fuoco oltre la focale normale. Focali maggiori danno tuttavia profondità di campo.

La **focale** costituisce la distanza tra l'incrocio dei raggi luminosi che entrano nell'obiettivo ed il piano della pellicola/sensore, corrisponde per certi versi alla lunghezza dell'obiettivo. Al cambiare dell'angolo di campo (focale alta → angolo stretto che però non ingrandisce le foto), più l'escursione è alta più la qualità si abbassa (per questo lo zoom è sconsigliato).



La dimensione della focale cambia anche al variare della dimensione del sensore tant'è che con sensori più piccoli si costruiscono focali più corte (sennò si otterrebbero dei teleobiettivi se si montassero delle focali di dimensioni medie o normali). Il sensore deve avere però pochi pixel per essere molto piccolo. Con focali e sensori ridotti si avranno anche diaframmi ed obiettivi più piccoli.

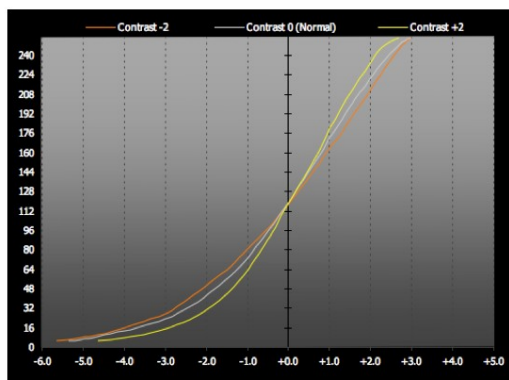
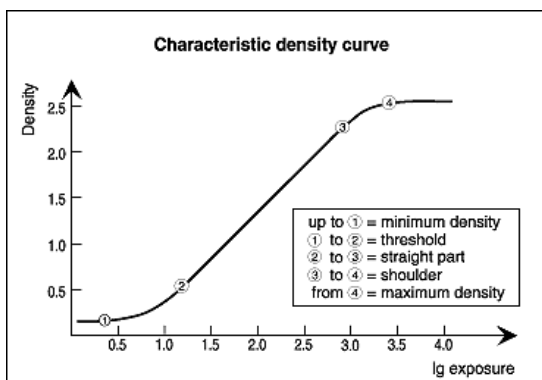
Una *focale equivalente* costituisce uno schema che illustra il rapporto tra i vari angoli focali e la lunghezza della focale: una focale lunga è ottimale per fotografare un oggetto molto lontano, tuttavia più si spinge la focale più i piani dell'immagine si appiattiscono favorendo spiacevoli parallelismi.

Le coppie tempo/diaframma sono disgiunte dalla macchina fotografica ed

aiutano la compensazione della quantità di luce rispetto alla focale, solo in questa maniera la quantità di luce che colpisce il sensore può rimanere sempre la stessa al variare della distanza focale.

Livelli di densità. La densità minima è data dal massimo punto di bianco e la massima dal massimo punto di scuro, si parla anche di intensità del segnale. La trasmissione è data dalla luce trasmessa sulla luce totale, il tutto messo a logaritmo rispetto la densità.

La **curva caratteristica** è data dalla misura del logaritmo della luminanza per determinare la coppia tempo di esposizione/diaframma ($\log E$) ovvero la luce che cade sul sensore. Si voleva un rapporto proporzionale perchè tutto ciò che è lineare è più facile e bello, doveva essere una retta invece è saltata fuori una curva a forma di S. Tale curva è di importanza fondamentale: riporta la quantità di luce che colpisce il sensore e la quantità di luce restituita. Allora regolare il contrasto significa cambiare la pendenza di tale curva: più l'andamento è rigido e verticale più il contrasto è alto. Ci si interessa molto alla cosa perchè al nostro occhio l'elevato contrasto nelle zone medie ed il poco contrasto nelle zone alte e basse piace molto, il tutto perchè nelle zone medie trovano posto i colori della pelle (sono tutti grigi medi e vanno mappati bene a livello cerebrale anche per una questione di sopravvivenza). Il tutto aiuta ad ottenere foto più belle: la foto è una mappatura lineare del contrasto.



La **legge di reciprocità** per la quale e grazie alla quale scegliamo qualsiasi coppia diaframma/tempo tende a fallire sugli estremi della curva.

High Dynamic Range Image (HDR)

Le **HDR** innanzi tutto sono immagini. All'interno di una macchina fotografica la scena è riassunta attraverso gli oggetti e gli elementi che la compongono in maniera disgiunta dalla loro effettiva luminosità. Il **range dinamico** costituisce, in una scena, il rapporto tra la sua zona più chiara e quella più scura. In natura la riflettanza non supera mai valori di 50:1 (costituendo quindi range abbastanza bassi); un range dinamico alto si ha invece quando la quantità di luce e quella di ombra sono tra loro sproporzionate.

Il sensore della macchina fotografica, dopo l'acquisizione della scena, non rende le informazioni subito disponibili (bisogna attendere che il segnale venga gestito e salvato in un formato file). Il sistema visivo invece è composto di due lenti sovrapposte (purtroppo non ottimali) le quali filtrano la scena che sarà poi il cervello ad elaborare (ed a migliorare).

L'HDR costituisce un settore dell'imaging digitale il quale ha subito un successo

assai maggiore rispetto alla comprensione del fenomeno stesso. Si può definire HDR un'immagine che sussiste nel caso in cui devo riprodurre un'immagine più grande del dispositivo di riproduzione. Tutti i quadri dipinti dai pittori sono immagini in HDR: ciò che si percepisce con il sistema visivo è molto simile a ciò che si vede in un quadro, tuttavia con la fotografia non è possibile riprodurre la percezione come è possibile fare attraverso un dipinto (i pittori riproducono la percezione, le macchine fotografiche no); i problemi sembrano risiedere nelle ombre e nelle zone controluce che un sensore digitale rende tali ma che il nostro cervello in qualche modo gestisce in maniera da renderle visionabili in maniera piacevole. Ciò che appare non è fotografabile.

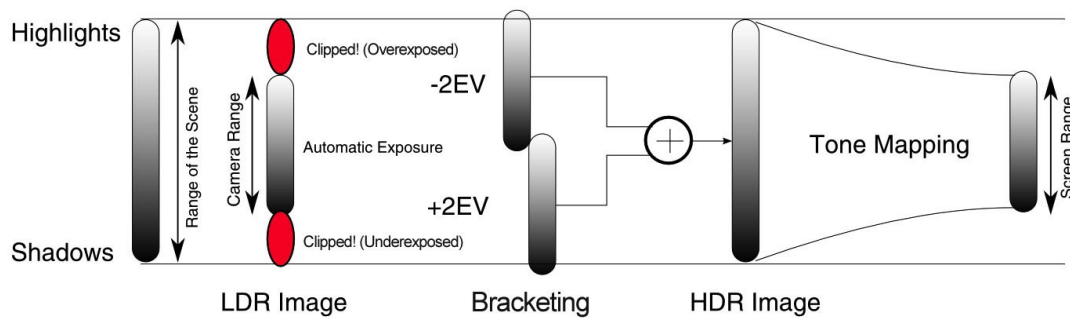
Il range dinamico è misurabile in qualsiasi scena si stia contemplando, allora qualcuno ha voluto quantificarlo per il mondo intero (combinando qualche pasticcio). Il valore del range dinamico a livello mondiale non ha alcun senso poiché non sempre e non ovunque il range dinamico è così quantificabile ma dipende fortemente dalla scena e dalle sue caratteristiche (ora del giorno, luogo del pianeta, conformazione, ecc). La maggior parte delle volte il range dinamico è abbastanza contenuto. Il sistema visivo, in completa autonomia, riduce i gradienti di luminosità.

La metafora del salame. Alcune volte il range dinamico può essere espresso in bit. Bisogna chiedersi però che dimensioni abbiano questi bit per poter dare un senso a questa affermazione. C'è chi sostiene che un range molto ampio ottenuto con stimoli isolati non sia ottenibile in una immagine (non è corretto). Si pensi al range dinamico come ad un salame: la sua lunghezza (qualunque essa sia) corrisponde al range effettivo. Il salame può essere affettato in maniera più o meno fina, producendo un numero variabile di fette che dipendono anche dal loro spessore: si riconosce in questo la quantizzazione del range. Più bit (quindi più fette di salame, ma più fini) non significano un range più alto, come meno fette ma più spesse non stanno ad indicare un range più basso.

Un sensore può tollerare un massimo di un tot di bit, allora ci si deve riferire alla curva caratteristica che lega la luminanza al digit e che quindi ci suggerisce con che finezza la macchina fotografica “taglia” il salame: fette più fini rendono una quantizzazione più fine facendo sì che si possano cogliere dalla scena più caratteristiche e dettagli in maniera bilanciata (che è ciò che fa localmente il sistema visivo). Per rendere al meglio una zona con molto dettaglio è necessario servirsi di molti bit, tuttavia una macchina fotografica non può cambiare il modo in cui “affetta” i bit.

Una dinamica di scena molto grande (o molto piccola) necessita di un sensore con un range molto ampio (o molto ridotto), la quantizzazione del range però è un'altra cosa. Quando il dispositivo non possiede un range molto alto tende a bruciare quella parte dell'immagine che non riesce a quantizzare a causa dei suoi limiti strutturali. Si vorrebbero allora dei sensori enormi per dinamiche enormi: nel 2013 la dinamica di un sensore digitale si avvicina sempre più a quella della pellicola fotografica (è auspicabile che nell'immediato futuro la raggiunga e la sorpassi), tuttavia dotarsi di un sensore/pellicola con un range ampissimo non è la priorità: bisogna adattare anche tutta la pipeline fotografica dell'acquisizione dell'immagine.

Detto questo comunque si è affrontata una parte del problema, così si arriva ad applicare l'idea fondamentale della fotografia HDR: con un sensore ovviamente a dinamicità minore rispetto alla scena posso effettuare uno scatto sovraesposto ed uno sottoesposto che poi andranno combinati assieme.



L'idea è buona ma non funziona, non è così che si deve risolvere il problema. Innanzitutto per sommare tra loro due immagini devono essere identiche, questo implica un'acquisizione perfetta e simultanea di soggetti che nel frattempo non si sono mossi; se le immagini sono identiche si possono sovrapporre ed all'inizio si era pensato di calcolarne la media, tuttavia sovra e sottoesponendo non si rilevano i corretti valori della luminanza (rendendo la media impossibile da calcolare; sarebbe calcolabile solo se la curva caratteristica fosse una retta, invece con una curva i valori non combacerebbero, salvo un'approssimazione che però non è preventivata). Si può provare a ricostruire la curva caratteristica dello scatto.

Con che criterio è possibile mappare nella stampa tutti i dati raccolti in questa maniera? Si può optare per l'approccio dell'artista, in questa maniera “vado a sentimento” e riporto i dettagli come meglio credo; sennò tento di imitare oggettivamente il sistema visivo in maniera da ottenere dei risultati il più possibile apprezzabili. Questo viene spesso trattato come un problema di *signal processing*: meglio lavorare con dati oggettivi, quindi con registrazioni e visualizzazioni della luce fisica nella distribuzione colorimetrica, oppure con dati soggettivi ottenuti dalla rappresentazione dell'apparenza o da un'intento di rendering differente? La filosofia che segue la fotografia HDR intende capire l'immagine e la scena non solo a livello di dispositivo sensore, ma anche a livello retinale, quindi s'interroga riguardo al range dinamico (se è uguale o diverso da quello della scena) per finire col capire le dipendenze dalla scena ed il risultato finale al quale il processo mira.

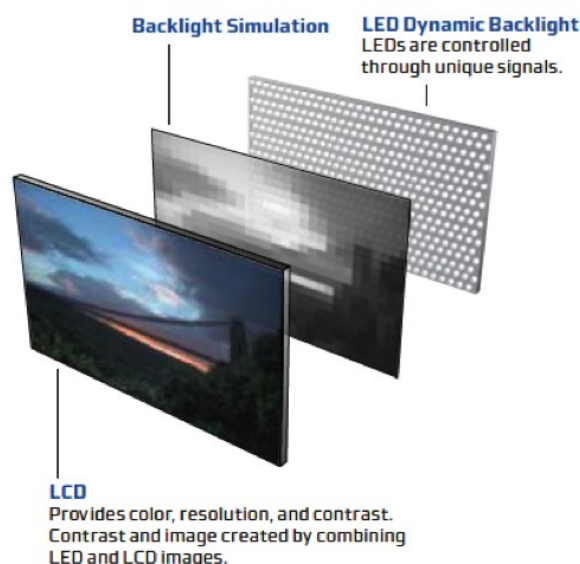
La **curva caratteristica** tra le altre cose ci aiuta a capire come e quanto la luce sia mappata in ogni *digit* (\rightarrow pixel), tuttavia non costituisce un rapporto lineare perchè, se così fosse, al nostro sistema visivo non piacerebbe. Si è pensato di lavorare di più sull'esposizione perchè è attraverso questo tecnicismo che si possono fissare maggiori dettagli della scena (in luce come in ombra).

Nella storia (della fotografia e non) si sono trovate diverse soluzioni al problema della fotografia in HDR (tanto per dire che non si tratta solo di una scoperta recente, tantomeno di una preoccupazione solo moderna): già nel 1858 ad H.P. Robinson per fissare una fotografia in HDR servivano 5 scatti (per rendere il range dinamico più esteso) di una scena perfettamente immobile i quali, una volta fissati su lastra di vetro, venivano sezionati in maniera tale da isolare le porzioni di scena con dinamica maggiore, così da comporre una fotografia definitiva (ritoccata a china nei punti necessari) la quale riproducesse il più fedelmente possibile la scena reale (*“Fading Away”*). Nel 1920 Mees (che poi prestò servizio presso la Kodak per la cui azienda sviluppò una pellicola particolarmente sensibile al range dinamico) disegnò la prima autentica curva caratteristica semplicemente testando la chimica delle pellicole presenti ai suoi

tempi per studiarne una migliore di tutte. Ansel Adams (fotografo di paesaggi prettamente in bianco e nero) sviluppò il suo sistema zonale grazie al quale dava luogo ad una mappatura molto precisa della scena e delle sue caratteristiche grazie ad un controllo serrato della pellicola il tutto per ottenere il range dinamico più fedele in assoluto: voleva riprodurre il più possibile l'apparenza. Nel 1941 Jones e Condit effettuano un lungo viaggio negli Stati Uniti durante il quale scattano diverse fotografie e misurano la differenza di range dinamico di una quantità notevole di scene differenti producendo, di conseguenza, un grafico del *flare* (rumore ottico).

Il **Retinex** è un algoritmo facente parte di una nutrita famiglia di procedimenti che mirano a riprodurre l'apparenza. Presa una scena in parte in ombra ed in parte luminosa e posizionati due campioni di colore identici da entrambe le parti, si nota che il bianco posto in ombra ha lo stesso digit del nero illuminato: si deduce che la percezione viene scalata sul rapporto tra la luce e l'ombra, allora la luminanza si riduce ad essere la medesima anche se l'apparenza è totalmente diversa (l'apparenza è data dal rapporto con il contesto).

Come detto poco prima, l'HDR sembra un concetto moderno ma nei fatti è da più di un secolo che ci si lavora. Come è possibile fare scatti multipli della stessa scena? Utilizzando magari dei dispositivi con specchi traslucidi che splittano l'immagine acquisita, altrimenti lo si fa a mano armati di cavalletto e di tanta pazienza. Per sovrapporre i due scatti, tuttavia, si è già visto che calcolare la media non basta, così si stima la curva caratteristica delle due acquisizioni: si ottiene il CameraDigit a partire dalla moltiplicazione di radianza e tempo, tuttavia interessa sapere solo qual'è la luminanza per poterla associare al digit. La radianza moltiplicata per il tempo di esposizione torna il valore del digit associato, si può quindi prendere una serie di scatti fatti ad uno stesso soggetto ma ad esposizioni differenti (e magari crescenti ma con il diaframma tenuto sempre fisso). In questa maniera si è in grado di stimare la luminanza che verrà data dal rapporto con il digit su di una determinata curva caratteristica. Il meccanismo tuttavia non funziona salvo nel caso in cui ci si ritrovi con una dinamica limitata.



Nel display di un'immagine si presenta lo stesso problema: solitamente gli schermi sono retroilluminati, alcuni da un unico grande led per garantire uniformità e stabilità al massimo delle possibilità. Se lo schermo è LCD avviene

anche una modulazione ed il range dinamico non va oltre il 100:1 → ovviamente il nero non può essere inteso come assenza di luce così è reso da un grigio molto scuro (ma ovviamente non nero). Uno schermo LCD tra i più tecnologici e moderni però non si limita a fare una semplice modulazione ma si compone di un pannello a serie di led retrostante che si possono accendere o spegnere a seconda dell'esigenza. L'idea è di poter creare led abbastanza piccoli da poter essere contenuti in un pixel, ma oggi ci si accontenta di associare un gruppetto di 20 pixel ca. a ciascuno di quei led, in questa maniera il range sale davvero a 10.000:1 tuttavia bisogna fornire al pannello retroilluminato un'immagine gemella a quella visualizzata ma sotto campionata (si tratta di tecnologia molto costosa).

Anche nell'acquisizione si riscontrano un certo numero di problemi. Il **glare** è un difetto dovuto alle lenti che si verifica quando un raggio di luce viene rifratto dalla lente mentre l'attraversa ed una certa percentuale dell'energia che lo compone rimbalza in maniera caotica all'interno della lente (il fenomeno risulta visibile solo quando la sorgente luminosa è molto intensa, ad es: il sole). Il glare ha effetti diffusi su tutto lo scatto comportando cali localizzati di contrasto. Si tratta di un rumore ottico ineliminabile. Si può simulare la consistenza di una lente costruendo una sottile camera dalle pareti di vetro e riempita d'acqua: se nell'acqua si sciolgono poche gocce di latte ecco che si riproduce l'effetto del glare. Qualsiasi punto di un'immagine può generare glare, con predilezione per i punti più chiari e luminosi. Se si mette in relazione il glare proveniente da zone diverse si nota che nelle porzioni di foto più scure il glare fa danni più consistenti spesso causati anche (e soprattutto) dal glare proveniente dalle zone più chiare.

Un sistema ottico come quello della macchina fotografica possiede un certo numero di immagini parassita dovuto al numero dei gruppi di lenti presenti nell'obiettivo; si tratta di immagini esterne all'asse ottico che si presentano in trasparenza come se fossero dei fantasmi, si tratta di un glare strutturato che segue le immagini all'interno della macchina fotografica (il diaframma ed il non essere in perfetto asse dell'obiettivo creano ed aggiungono, oltre alle immagini parassita del glare, diverso rumore ottico).

La curva caratteristica della pellicola è più sinuosa (più a forma di S) rispetto a quella del sensore della macchina fotografica digitale anche e soprattutto perché la pellicola è capace di un range dinamico molto più alto.

L'HDR riesce ad aumentare sul sensore il range dell'informazione da acquisire ma spesso quello della scena reale risulta essere molto più basso, così non si può sfruttare un set di fotografie HDR per stimare correttamente la luminanza di una scena.

Il glare è un elemento importante per il sistema visivo umano tant'è che spesso in computer grafica viene aggiunto in maniera fittizia. Bisogna saper distinguere anche dal range dinamico reale (quello della scena) e da quello apparente (ovvero quello che vede l'occhio umano e che reputa essere quello reale). Ansel Adams esponeva a metà una foto su di un muro uniforme stendendo un velo di luce sulla pellicola, in questa maniera le ombre si alzavano riproducendo un range dinamico apparente più alto.

Attraverso il test dell'HDR si prendono 4 cerchietti a spicchi che si scuriscono in maniera graduale passando dal più chiaro al più scuro (*Kodak Projection Print Scale*) e li si montano prima su una maschera completamente nera, poi si una completamente trasparente, quindi si retro illuminano le due situazioni e si osserva l'effetto del glare sui 4 cerchietti: si capisce che non c'è una correlazione

oggettiva tra luminanza reale e stimata: se ottengo una stima di luminanza corretta non posso attribuire alla stima una curva corretta; se prendo in considerazione il glare mi rendo conto che nella situazione con lo sfondo scuro certi cerchietti (quelli più chiari) non li vedo nemmeno. Più il glare aumenta più la zona visibile diminuisce e si restringe, va da sé che il glare dà un limite molto rigido al range dinamico di una scena.

A livello retinale nel contrasto di simultaneità si percepisce più chiara qualsiasi forma immersa nel colore bianco anche se nella realtà accade il contrario. Il cervello si attiva per contrastare il glare (chi ha gli occhi chiari subisce di più l'effetto del glare; dove c'è molta luce la gente è fenotipicamente più scura; nel '700 i pirati si scurivano la zona oculare per diminuire il glare dovuto alla luce diffusa). Per questo motivo il glare non è assolutamente trascurabile. Sulla retina avviene infatti una grande caduta del segnale in corrispondenza dei contrasti (da non dimenticare che il segnale sulla retina è abbastanza scadente allora il cervello lavora moltissimo per recuperare il contrasto combattendo il glare). Nella scena e nella retina accadono cose drasticamente differenti con dinamiche distantissime tra loro: sulla retina avviene una riproduzione della scena reale schiacciata tuttavia il cervello fa in modo di recuperare tutte le informazioni che la retina non è capace di segnalare.

Capitolo 5

Visione Stereoscopica

Introduzione alle tecniche e alle tecnologie di produzione e visualizzazione stereoscopica

S'intende per **stereoscopia** il processo di realizzazione e visualizzazione di tutte quelle immagini, disegni, fotografie e filmati finalizzata a riprodurre il più fedelmente l'aspetto della tridimensionalità; analogamente alla visione generata dall'apparato visivo binoculare del sistema ottico umano (gli occhi possiedono due disparità retiniche che il cervello elabora e fonde per creare la visione tridimensionale alla quale siamo abituati; il 12% della popolazione mondiale non è capace di vedere in tre dimensioni, tuttavia questo non costituisce un problema vitale).

L'occhio destro e l'occhio sinistro vedono una differente prospettiva della scena, in questa maniera è più facile ottenere informazioni sulle profondità e le distanze. La visione binoculare da sola non basta e la percezione della profondità è rinforzata da altri indizi come l'accomodazione, le occlusioni, le ombre e le riflessioni, le dimensioni relative, le dimensioni familiari, ecc. Molti degli indizi di cui sopra sono monoculari, quindi è su queste caratteristiche che le persone con un occhio solo vedono il mondo.

Nella stereoscopia si generano/acquisiscono due viste diverse:

1. sistemi per il cinema stereoscopico;
2. software/API per la stereoscopia in computer grafica;

Entrambi costituenti visioni generate con videocamere o con software di computer grafica dedicati; attraverso degli appositi occhialini i segnali sono allineati ed il cervello crea l'effetto finale di 3D.

Ogni tipo di vista deve raggiungere solo l'occhio corrispondente (e per il quale è stata studiata); viene simulata la visione binoculare ed il sistema visivo procede da sé alla fusione dei due stimoli (diversi) che riceve.

La **parallasse** è il concetto chiave della stereoscopia; si tratta della differenza nella posizione apparente di un oggetto osservato da due punti di vista diversi. Determina se un oggetto apparirà “davanti” o “dietro” lo schermo. La parallasse zero si ottiene quando gli oggetti sono percepiti sullo schermo; la parallasse positiva si ha quando gli oggetti sono percepiti dietro lo schermo; la parallasse divergente si verifica quando il valore della parallasse è molto ampio, si tratta di un effetto da evitare perché molto fastidioso per la vista (poiché la sforza ad allargare la vista in maniera innaturale), tipico di quando si avvertono degli oggetti posizionati davanti allo schermo, che se vengono così percepiti si può parlare anche di parallasse negativa.

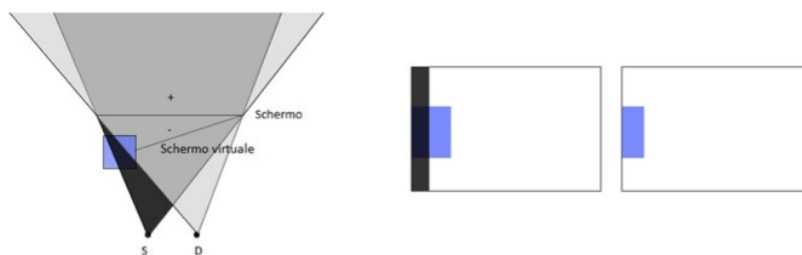
La *massima parallasse positiva* è data dalla distanza che si misura tra gli occhi (o comunque tra i dispositivi di registrazione, come le videocamere, all'interno di un sistema di visione stereoscopica); si parla di assi di visione posizionati in parallelo e non dipende solo dalla distanza interoculare ma anche dalle dimensioni del device di visualizzazione (che può essere uno schermo da proiezione oppure un semplice display), dalla distanza dell'osservatore dal dispositivo di visualizzazione e dalla focale della camera (ed anche da tanto altro). Qualsiasi contenuto stereo va prima testato in condizioni equiparabili alla situazione di fruizione finale prima di essere distribuito: ovvero il setup al quale

si rifà il consumo del contenuto dev'essere conosciuto a priori e studiato in anticipo.

Esiste un certo conflitto tra accomodazione e vergenza, dove l'accomodazione costituisce il meccanismo di messa a fuoco dell'oggetto da parte del sistema visivo e viene attuato attraverso la contrazione del cristallino mentre la vergenza è la rotazione dei bulbi oculari grazie alla quale l'oggetto osservato è portato all'interno della zona foreale (comporta uno sforzo muscolare notevole se l'oggetto è troppo vicino e da qualche indizio sulla profondità che è sempre molto forte per oggetti vicini). Il conflitto va ben gestito per ottenere un buon setup. Durante la visualizzazione stereoscopica l'accomodazione avviene sullo schermo e la vergenza cambia a seconda del valore della parallasse. L'effetto finale stereoscopico risulta essere fastidioso se non è studiato bene soprattutto per effetti di parallasse negativa (è inoltre più fastidioso per le persone anziane ed è facile da riconoscere per tutti coloro che sono esperti/sviluppatori, ecc).

Le zone di comfort indicano come e dove la vista può affaticarsi o meno durante la visualizzazione di qualsiasi soggetto. Una sensazione di comfort ottimale avviene quando l'oggetto visualizzato si trova intorno al punto di parallasse zero e costituisce un basso impatto visivo. Il comfort più basso è dato da alti valori di parallasse accomunati ad una forte attività muscolare: è tollerata solamente per brevi periodi di visione, oltre i quali il fastidio diventa esagerato. Le rivalità retiniche positive costituiscono delle rivalità naturali che creano fastidio quando i valori di parallasse sono troppo alti: risultano accettabili solo se ci si trova nei pressi della parallasse zero. Le rivalità retiniche negative invece sono estremamente fastidiose perché consentono ad un oggetto un'entrata/uscita nella scena visibile molto (e spesso troppo) veloce; si intende anche come *violazione di finestra*.

Solitamente quando si guarda un'immagine 2D si ha a che fare con una mera rappresentazione piatta (appunto a due dimensioni) delimitata da una cornice che è il contorno dello schermo come se si trattasse di una specie di quadro; non è la stessa cosa per le immagini 3D perché si guardano oggetti posti in maniera da risultare davvero tridimensionali e lo schermo non è più un confine ma una finestra su di uno spazio completamente aperto: nel 3D c'è una profondità variabile che non esiste nelle due dimensioni. Tutti i tradizionali effetti della cinematografia 2D come il flare, gli effetti speculari e lo sfruttamento della polarizzazione sono sconsigliati in quella stereoscopica, la stessa cosa vale per il framing: semplicemente perché alcune convenzioni non vanno più bene (come per esempio la ripresa "over the shoulder") incorrendo così nella *violazione della finestra*. La **violazione di finestra** implica che tutti gli oggetti posti sullo sfondo dell'immagine vengano tagliati (e questo va bene) ma che lo siano anche quelli in primo piano (e nella stereoscopia questo è inammissibile): in questo caso il valore di parallasse indica quindi che un oggetto è davanti lo schermo e l'occlusione comunica invece che l'oggetto dovrebbe trovarsi dietro (l'occlusione è un indizio di parallasse molto forte).



L'unica soluzione è rendere fluttuante la finestra nelle scene nelle quali è presente la violazione: sfruttando il principio della **floating window** si fa in modo di ruotare la finestra tanto da nascondere le parti che la violano, quindi in pratica si applicano delle maschere nere ai lati delle immagini che, così facendo, nascondono i dettagli che non devono essere visibili ad entrambi gli occhi. La floating window può interessare qualsiasi (e tutti) i lati della finestra, può essere dinamica ed è sicuramente una soluzione molto efficace perchè tra le altre cose non viene mai percepita. Viene applicata nel campo cinematografico nel momento della post produzione e permette diverse soluzioni creative.

Il setup di un film in 3D è molto difficile da fare ed ottimizzare, per questo motivo sono nate nuove figure professionali che si occupano dell'aspetto tecnico di tutti i settaggi del caso (3D supervisor, stereografo, tecnico del settaggio stereo, tecnico dei rig, ingegnere 3D ecc). Si vive un ripensamento ed un adeguamento di molte convenzioni e tecniche vedendo anche l'introduzione di nuova terminologia:

1. **depth budget:** termine che indica la quantità di “profondità” che si ha a disposizione; dipende dalla quantità di parallasse complessiva a disposizione (quindi utilizzabile) e dipende anche dalla dimensione dello schermo di visualizzazione della risorsa stereo;
2. **depth bracket:** termine che indica la profondità totale della scena (che è una sottomisura del depth budget; si noti che la profondità totale della scena non uguaglia mai quella che si ha a disposizione, ma va tarata per essere sicuri di non oltrepassarla). Deve sempre stare entro i limiti del depth budget;
3. **depth script:** indica la modulazione della profondità scena per scena con tanto di indicazioni su quale tipo di parallasse si è scelto per ciascuna scena e di come si passa da una scena all'altra (e da una profondità all'altra);
4. **depth continuity:** istituita per evitare di saltare in maniera brusca da scene caratterizzate da forte parallasse a scene che invece hanno una parallasse negativa; deve sovrapporsi con il depth bracket.

Per la generazione e l'acquisizione dei contenuti stereoscopici esistono diversi metodi: nella cinematografia “live” si lavora con foto/videocamere configurate in maniere differenti a seconda delle necessità; in computer grafica ci si avvale di rig camere stereo accoppiate a software dedicati alla computer grafica ed ad API con supporto per la visualizzazione di immagini stereografiche. È ovvio che la composizione tra immagini live e computer grafica è possibile ed incentivata.

La *distanza interoculare di una camera stereoscopica* collima con la distanza tra la camera sinistra e la camera destra (dando per scontato che un sistema di questo tipo sia composto sempre da una coppia di videocamere gemelle); costituisce il parametro cruciale della ripresa stereoscopica, setta il depth bracket della scena e controlla le dimensioni apparenti degli oggetti ripresi.

La fotografia e quindi l'acquisizione di contenuti stereoscopici nasce all'inizio del 900 con la comparsa della fotografia. Negli anni '50 scoppia la mania per il cinema 3D ma la febbre passa in fretta perchè la visione di contenuti di questo tipo stressa molto la vista e provoca forti mal di testa (un po' come succede anche oggi). Con un sistema di acquisizione a camere affiancate si lavora con due camere gemelle (quindi con la stessa lente) controllandone gli assi ottici. Poiché hanno gli stessi parametri di acquisizione bisogna gestire in maniera sincronizzata i tempi. Importante da settare è la distanza interoculare dei dispositivi che è comunque limitata dalla dimensione delle videocamere (si

possono avere setup diversi a seconda dei budget di cui si dispone). Con sistemi half-mirror si ha una coppia di macchine posizionate ad un angolo di 90° frammezzate da uno specchio inclinato: permettono distanze interoculari basse tuttavia con sistemi molto professionali. Lo specchio è semitrasparente ed è di importanza fondamentale tenerlo sempre in ottima manutenzione (qualsiasi imperfezione potrebbe danneggiare l'acquisizione). Ha molti fattori in comune con il sistema a camere parallele tuttavia vede una notevole perdita di luminosità dovuta allo specchio. Utilizzando i sistemi a doppia ottica/sensore integrati si dispone di sistemi all-in-one con una buona praticità (un solo set di controlli, non serve sincronizzare i due dispositivi in maniera manuale) tuttavia hanno pochissima flessibilità con una lunghezza focale ed anche una distanza interoculare fissata e non modificabile. C'è anche poco controllo sugli effetti finali tuttavia è ottima per tutte quelle situazioni in cui basta filmare (run and gun). Per la scelta del sistema più opportuno bisogna stimare diversi parametri tra cui il budget a disposizione, il tipo di scena/contenuto da riprendere, il peso e l'ingombro, ecc.

La **regola dell'1/30** è una regola approssimata per effettuare una ripresa stereo accettabile: dice che la distanza interoculare dev'essere $1/30$ della distanza complessiva tra la videocamera e l'oggetto più vicino. È una regola che però non è sempre valida ed andrebbe aggiustata mettendola in relazione con la lunghezza focale delle videocamere e la grandezza (dimensione) dello schermo di visualizzazione. Esistono diversi studi assai più complessi per il calcolo della relazione che intercorre tra la distanza interoculare, la parallasse sullo schermo ed il setup di qualsiasi videocamera stereo (il tutto anche per evitare problemi di serializzazione).

Anche le videocamere possono convergere (non solo i bulbi oculari): attraverso un setup a camere parallele si ottengono solo valori di parallasse negativi; convergendo tuttavia spostato la scena lungo l'asse di profondità quindi non intervengo sul depth bracket ma devo comunque controllare di essere ancora compreso del depth budget. Esistono tuttavia dei problemi di convergenza come la distorsione di Keystone la quale introduce la parallasse verticale ed il difficile controllo meccanico del rig delle videocamere. La soluzione sta nell'usare un setup valido per le camere parallele cambiando quindi la sovrapposizione delle due immagini per ottenere l'effetto desiderato → Horizontal Image Translation (HIT) richiede l'overshooting: l'immagine viene croppata ai bordi e dipende dalle dimensioni di visualizzazione (quindi dal massimo della parallasse).

Anche nella computer grafica esistono varie conformazioni per le videocamere stereo con un estremo controllo su tutti i parametri. Si evitano così i soliti problemi tecnici/meccanici caratteristici della ripresa cinematografica con la possibilità di replicare in scala le dimensioni dell'ambiente di visualizzazione finale. Il *multirigging* è di facile utilizzo in questo contesto. Se gli assi delle camere sono paralleli non sfruttano tutto lo schermo e necessitano dello HIT (dando diversi problemi in realtime) mentre se gli assi sono convergenti si introduce la parallasse verticale. Il *frustum asimmetrico detto di "off-axis"* permette di sfruttare tutto lo schermo di proiezione e non necessita di HIT come nel caso degli assi paralleli (pur avendo comunque gli assi a suo modo paralleli). Si tratta di un effetto stereo modificato cambiando la distanza interoculare delle camere sul piano di proiezione: bisogna definire un frustum asimmetrico accedendo alle primitive di basso livello. Il multirigging solleva un certo problema: c'è una differenza rilevante tra l'oggetto ripreso in primo piano e quello che si trova sullo sfondo. Con un unico setup di videocamere si otterrebbe

una cattiva resa dello sfondo causando troppo parallasse. La soluzione si intravede dividendo la scena in due parti (soggetto e sfondo) quindi si renderizza le sue parti con due setup di camera differenti e le si unisce poi tramite compositing, allora si diminuisce il gap di profondità tra le due parti. Sia nella cinematografia live che nella computer grafica esistono diverse modalità per ottenere l'effetto del multirigging.

Per la visualizzazione di immagini e contenuti stereoscopici bisogna che l'informazione destra arrivi all'occhio destro (idem per la sinistra), esistono quindi diverse soluzioni (tutte comunque da fruire da soli):

1. **free-viewing;**
2. **anaglifo:** si tratta di una tecnica famosa anche ai più inesperti che si basa sull'utilizzo di filtri colorati (solitamente rosso-ciano) applicati alle due viste in fase di produzione; le due viste sono poi sovrapposte e salvate in una unica immagine. Gli stessi filtri sono presenti non solo sull'immagine ma anche sugli occhiali da indossare per poter visualizzare il contenuto, e permettono il passaggio solo alla vista trattata con lo stesso filtro (perchè si tratta di colori complementari). Si possono sfruttare in vari contesti come sul monitor, in una proiezione oppure sulla stampa. La produzione digitale è facilitata da software appositi per l'elaborazione delle immagini o di compositing. Il tutto procura un discreto effetto di profondità specialmente al monitor tuttavia la resa dei colori è alquanto scarsa. Oggi si usa molto questa tecnica in fase di test;
3. **sistemi attivi:** si basano sulla proiezione alternata delle viste destra e sinistra. Per farlo basta un proiettore oppure un monitor e degli occhiali con otturatore IR a cristalli liquidi, parecchio costosi (che per via della loro natura fanno perdere un po' della luminosità originale del contenuto visualizzato). Le lenti destra e sinistra si aprono in modo alternativo e sincronizzato con la proiezione (richiede comunque un'alta frequenza di visualizzazione).

Nel caso di un sistema basato sulla proiezione la luminosità si vedrà abbassata (per via dello spessore degli occhiali e del fatto che si vede praticamente una sola immagine alla volta); ci saranno poi problemi di sickness oltre ad un alto costo per il reperimento delle attrezzature;

4. **sistemi passivi:** sono i sistemi "classici" che si basano sulla proiezione contemporanea di entrambe le viste. Sono necessari almeno 2 proiettori dotati a loro volta di filtri specifici, quindi servono anche occhiali dotati essi stessi di filtri polarizzati che fanno passare esclusivamente la vista corrispondente alla lente (ed il loro costo è assai più ridotto); in questo caso la luminosità sarà più alta per via degli occhiali meno spessi e perchè entrambe le immagini vengono proiettate contemporaneamente. I filtri possono essere *anaglifici* (ma al giorno d'oggi non sono proprio più utilizzati), *polarizzati* (che sono lineari o circolari e necessitano di uno schermo molto costoso che mantenga la polarizzazione detto *silver screen*);
5. **sistemi ibridi:** costituiscono sostanzialmente una soluzione ibrida tra sistemi attivi e passivi con proiezione alternata tra vista destra e sinistra; "z-screen" posizionato davanti al proiettore; si alternano i due filtri e si utilizzano gli stessi occhiali usati in un sistema passivo classico. Si tratta del sistema maggiormente utilizzato nelle sale cinematografiche (altre utilizzano direttamente dei sistemi attivi), il che permette loro di utilizzare un solo proiettore per facilitarne la manutenzione ed abbassare

i costi. La stessa cosa è concretizzata da qualche sistema di televisione 3D;

6. **sistemi basati sulla polarizzazione:** sono sistemi che utilizzano filtri polarizzati in maniera diversa per la vista destra e sinistra. I filtri bloccano la radiazione elettromagnetica a seconda della polarizzazione e sfruttano la direzione di oscillazione del vettore campo elettrico durante la propagazione dell'onda nello spazio.

Nella polarizzazione lineare la direzione è costante nel tempo perchè l'onda si propaga su un solo piano ed i due filtri hanno polarizzazioni ortogonali, così l'osservatore è costretto a non ruotare mai la testa.

Nella polarizzazione circolare il vettore in questione ruota attorno all'asse di propagazione ed i due filtri hanno senso di rotazione opposto: il senso orario significa polarizzazione circolare destra mentre quello antiorario coincide con la polarizzazione circolare sinistra. L'osservatore in questo caso può muovere la testa senza perdersi nulla durante la visione.

Esistono purtroppo dei problemi come ad esempio il “ghosting” ed il “crosstalk” dovuti ad un filtraggio non perfetto: parte dello spettro raggiunge lo stesso l'occhio sbagliato (ghosting) ed il tutto dipende molto dalla posizione dell'osservatore;

7. **sistemi INFITEC™:** si tratta di una evoluzione dell'anaglifo (detto per questo “super anaglifo”), in essi lo spettro della luce emessa è diviso in 2 sottoinsiemi complementari: si hanno allora due filtri passabanda uno più sul magenta ed uno più sul verde, effetti meno forti rispetto l'anaglifo, si eliminano i problemi di ghosting e si ha una buona resa cromatica (con conseguente adattamento visivo anche se c'è qualche problema con i colori molto saturi come il rosso e la rivalità retinica). Si tratta di una tecnologia usata negli USA nelle sale Dolby (3D).

Nel caso di *schermi autostereoscopici* è possibile la visualizzazione stereo senza occhiali perchè si visualizzano le due viste allo stesso momento su tutto il pannello tramite colonne di pixel alternate. Con l'utilizzo di sistemi selettivi cerca di indirizzare la vista sinistra solo all'occhio corrispondente e viceversa. Esistono due diversi approcci come la *barriera di parallasse* e le *reti lenticolari*. È una tecnologia non ancora matura per l'uso su televisori e monitor per via dei pochi angoli di visione disponibili: si tratta dei primi dispositivi che utilizzano display basati su barriera di parallasse.

Frame Compatible Formats (FCF)

Le immagini stereo possono essere create (quindi combinare le due viste in un unico frame 2D) solo in una certa maniera. Permette di riutilizzare tecnologia di fruizione già esistente e si usano spesso player esterni per allineare le immagini. Risulta piuttosto facile la gestione e l'estrazione dei singoli frame. Comporta una certa perdita di risoluzione e necessita una scalatura dei due frame per rimanere in formati standard (e si rischia l'anamorfismo).

L'anaglifo è a tutti gli effetti un FCF, è però contemporaneamente una codifica “lossy”: le due viste sono salvate in maniera sovrapposta e si tende a perdere le informazioni cromatiche viste nell'originale. Un'immagine stereo va salvata in un formato differente ed ogni player di video stereoscopici permette di scegliere la visualizzazione anaglifica a partire da un altro FCF selezionato:

1. top-and-bottom (TAB): a partire da una coppia di immagini, l'immagine alta va a destra e quella bassa va a sinistra;

2. side-by-side (SBS): date due immagini affiancate vanno fuse nello stile dell'anaglifo (quella a sinistra verrà polarizzata circolare sinistra e quella destra circolare destra);
3. line interleaved: sfrutta le linee interlacciate, data un'immagine singola le righe di pixel di numero pari si vedono solo con un occhio, le altre solo con l'altro (ovviamente i due occhi devono vedere informazioni diverse per poterle interlacciare);
4. quincunx: permette una qualità maggiore prendendo i pixel delle immagini destre e sinistra tramite uno schema a scacchiera: la ricostruzione di ogni pixel mancante si baserà sui 4 vicini e non solo su 2 come dei metodi precedenti.

2D + “qualcosa”. Si tratta di tecniche che si basano sulla trasmissione di una sola vista 2D accompagnata da informazioni supplementari utilizzate per calcolare sul momento la seconda vista. L'immagine possiede diversi fattori perchè venga resa tridimensionale, come per esempio una mappa di profondità (che indica la profondità di ogni pixel), una mappa dei livelli di grigio e sono immagini che richiedono poca larghezza di banda aggiuntiva (quindi sono adeguate alle infrastrutture già esistenti, anche 2D). Capita però che molte delle informazioni aggiuntive possano essere scartate così rimane solo la vista a 2D. Le informazioni aggiuntive sono calcolate a partire da due immagini affiancate (destra e sinistra).

Si studia la conversione di un video standard 2D in 3D ma sostanzialmente è una questione difficile ed onerosa (oltre che costosa) con risultati spesso molto scadenti. Nella computer grafica, tuttavia, è più facile produrre da zero filmato direttamente in 3D.

Capitolo 6

TMO

Tone Mapping Operator (TMO)

I **Tone Mapping Operator** sono dei software che convertono una immagine ad alta dinamicità (HDR) in una immagine visualizzabile e riproducibile da una stampante o da un display (dove qualsiasi cosa accada non ho mai la dinamica necessaria per rendere una qualunque immagine HDR); in soldoni è ciò di cui si occupa il software di una stampante, così si può dire che la stampante stessa è un TMO.

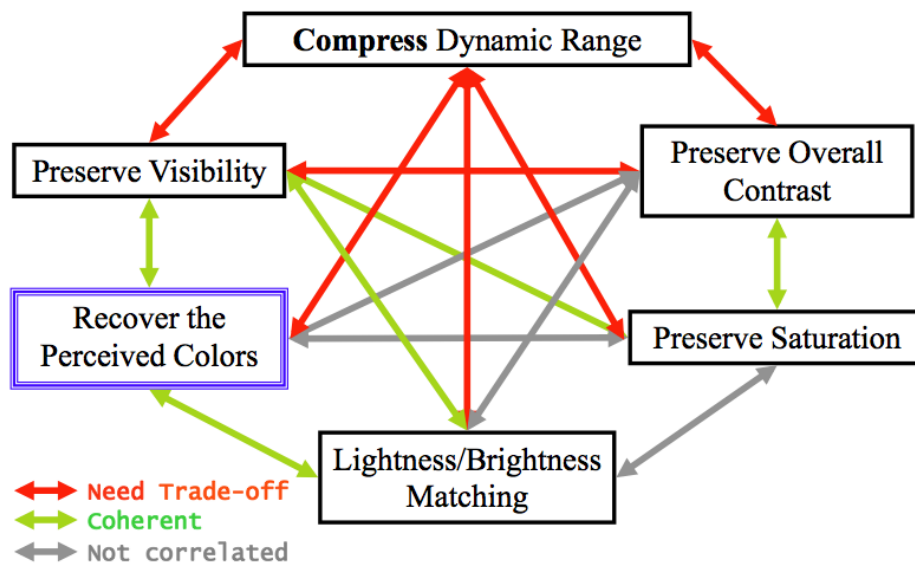
Si parte sempre da una scena reale oppure da un'immagine in 3D (sfruttata nel cinema per rendere gli effetti speciali perchè è più facile creare un'immagine tridimensionale in computer grafica che dal vivo attraverso modelli in scala come si usava decenni fa: costituiscono una grossa fetta dell'HDR moderno con il vantaggio, rispetto alla scena reale, di non soffrire l'effetto alquanto fastidioso ed inevitabile del glare, tuttavia proprio per questo motivo e per il fatto che il glare è fondamentale per la vista umana una scena 3D appare sempre poco verosimile). Lo scopo del TMO è quello di riprodurre la sensazione visuale della scena originale; si vorrebbe che la mappatura della scena HDR desse la stessa sensazione visiva di una scena guardata con i nostri occhi dal vivo (visto che non è possibile mappare in maniera fedele, ci si accontenta di ottenere il giusto effetto). Esistono metodi molto efficienti che riescono nell'intento ed alcuni che proprio non sono opportuni.

I problemi che TMO si propone di risolvere nel mappare un'immagine HDR in un'immagine a minore dinamicità sono:

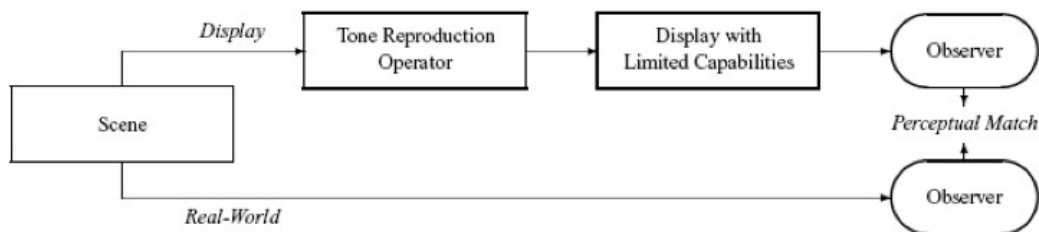
1. la compressione del range dinamico preservando comunque la visibilità dei dettagli (cosa che però dipende anche dalla fotografia e dall'intento con cui la foto è stata scattata);
2. preservare l'impressione generale di contrasto (se c'è una luce molto forte nella scena il range dinamico esplode ed il contrasto nelle zone scure invece tende a perdersi mentre in quelle chiare nulla apparentemente cambia);
3. conservare la saturazione che è il prodotto della differenza tra gli stimoli dei canali RGB, se si comprime la fotografia si schiaccia anche questa differenza comportando perdita d'informazione e di bellezza/risalto dei colori;
4. mantenere il contrasto di simultaneità (il nostro sistema visivo cambia il grigio di riferimento a seconda della casistica ma in configurazioni con apparenze diverse tali dettagli possono non essere conservati falsando così l'effetto globale finale);
5. conservare l'apparenza del colore reale;
6. ...

Tutti gli obiettivi appena elencati devono scendere a compromessi (oppure collaborare per raggiungere un obiettivo comune per cui rispettato uno si ottiene anche l'altro che è diretta conseguenza del primo e viceversa, oppure semplicemente ignorarsi) se vogliono essere rispettati tutti con proficuità (la lista è lunga e ci si affida allo schema a seguire per visualizzare i contrasti e le somiglianze tra gli obiettivi). Si ricorda che per *brightness* s'intende la chiarezza

percepita mentre per *lightness* la chiarezza misurata oggettivamente (tuttavia in certe altre rappresentazioni si può leggere il contrario). Per visibilità s'intende anche il contrasto locale che è diverso dal contrasto globale (non è detto che al variare del primo cambi il secondo e viceversa). Conservare la visibilità significa rinunciare al contrasto globale.



Per testare la validità di quanto è appena stato assunto si fa un'ipotesi di test: si osserva una scena direttamente e si osserva la stessa scena attraverso un display (quindi attraverso una fotografia HDR proiettata su un display), allora le due immagini dovrebbero essere uguali almeno dal punto di vista percettivo; non è possibile riprodurre perfettamente e fedelmente la scena vista nella realtà quindi si prova a forzare e distorcere l'immagine di rappresentazione della realtà in maniera che ci assomigli il più possibile sul piano della percezione (non si può proprio fare altrimenti).



La dinamica della retina è limitata tuttavia se ci si trova nelle giuste condizioni tale dinamica può essere espansa consentendo di godere di un range più ampio. È auspicabile effettuare il test sempre immersi nello stesso contesto, anche se non è fisicamente possibile: si deduce che a questo punto è meglio mappare la sensazione che suggerisce quella scena in quel contesto e tentare di riprodurla. Purtroppo non esiste un metodo standard per effettuare questa procedura così tutti i risultati ottenuti di volta in volta ed ottenuti ad ogni approccio vanno valutati come in un *beauty context*, tuttavia quest'approccio non risulta essere molto scientifico. Di **tone mapper** ne esistono di molti tipi e vanno valutati uno per uno per le loro caratteristiche intrinseche. Ne esistono di globali che mappano il valore di ogni pixel dell'immagine a prescindere da dove si trova (una curva caratteristica è facilmente identificabile come un autentico tone

mapper globale) facendo quindi una trasposizione lineare; solitamente una mappatura globale è facile da eseguire e da computare ma non si hanno molte possibilità di trovare una *soluzione generale* (quindi che valga per qualsiasi fotografia in qualsiasi contesto e non solo per casi specifici oculatamente prescelti) che punti in direzione della percezione visiva. Al contrario, un tone mapper locale lo potrebbe fare tranquillamente (dipende comunque dai casi): un metodo di questo tipo ha una complessità computazionale molto più alta e dev'essere snello da processare. Tra questi, ce ne sono alcuni realtime.

Ward (1997). Greg Ward, professionista che ha pubblicato molti articoli e molte ricerche con il nome di Greg Larson (dove Larson era il nome di sua moglie preso dopo le loro nozze ed abbandonato dopo la separazione), ha lavorato allo sviluppo di un software gratuito chiamato *PhotosphereU* per Mac OS facendo un ottimo lavoro dal punto di vista dell'analisi dei segnali ma senza approfondire nessuna componente percettiva; ha poi redatto un ottimo lavoro di studio sulle HDR. Il suo metodo si basa sulla mappatura della brightness (quindi dell'adattamento all'illuminazione locale) di un'immagine dalla quale poi si ricava un istogramma per l'aggiustamento della tecnica. Le informazioni che si ricavano provengono dalla luce presente nella scena e che definisce anche il range dinamico, tuttavia l'informazione più imponente ed importante proviene dalle zone d'ombra. Data l'immagine caso di studio, si applicano delle curve di mappatura per ogni dettaglio dell'immagine che risulta interessante (nel caso di una foto d'interni ad una grande finestra ci sarà una curva per l'interno della stanza, una per ciò che si vede dalla finestra ed una che mitiga le due prima e le riassume): assomiglia, solo come concetto, a quello della curva caratteristica unita alla compressione del contrasto globale e all'estensione di quello locale. Il tutto è ricavabile dall'analisi dell'istogramma: dove risulta esserci dell'informazione (ovvero nei punti più scuri dell'immagine) si allarga il range dinamico, dove risulta esserci meno informazione lo si restringe; il principio è quello di stiracchiarne i pixel.

Il glare aumenta la dinamica apparente del nostro sistema visivo (è un difetto derivante dalla luminanza ma il sistema visivo umano ha imparato a convivere ed a smorzarlo). La mappatura globale così fatta non è perfetta e tramite diverse soluzioni si tenta di arginare la sua incompiutezza.

Jobson (1997). L'algoritmo in questione è molto simile al *Retinex* (che in breve è stato sviluppato da colui che ha inventato i filtri polarizzati e le foto istantanee, un pilastro della Polaroid; nel 1982 ha cercato di migliorare il suo algoritmo che ad oggi rimane incompiuto ed imperfetto). Jobson, dipendente della NASA, escogitò un algoritmo locale che è in parte un filtro di sequenza (*mexican hat*) basato sul center/surround ed ampiamente ritoccato per avere risultati decenti, si compone di due fasi: MultiScaleRetinex e Color Restoration. Siccome una funzione gaussiana non è mai negativa, non lo è nemmeno una funzione di tipo mexican hat che, contrariamente ad una gaussiana tradizionale, enfatizza i bordi. Lo fa con un kernel la cui forma è data da una certa quantità di calcoli e dalla larghezza della funzione mexican hat (la cui forma è fissa e non va sempre bene per tutte le immagini). Il MultiScaleRetinex ha un buon effetto a livello di fedeltà al dettaglio e di mappatura della dinamica dal punto di vista percettivo: le zone scure vengono sovracomprese per far emergere a forza il contrasto così il glare viene compensato come accade nel nostro sistema visivo (consecutivamente c'è anche il bilanciamento del bianco). La saturazione però

tende a calare: con il Color Restoration il colore torna vibrante come dovrebbe essere (e come ci si aspetta, aumentando la luminosità degli scuri, si tratta di un effetto visivo). È un metodo locale tuttavia non costituisce una soluzione ottima.

Pattanaik (1998). Crea un modello il più vicino al sistema visivo così che tale modello sia d'ispirazione per la creazione dell'iCAM. È detto “modello di tipo Frankenstein” perchè risulta essere una somma di sottoparti disgiunte risultato di un'applicazione spartana di svariati dati ottenuti in condizioni controllate e con particolari stimoli, per questo motivo poco maneggevole e funzionale.

Utilizza due curve: la Threshold Versus Intensity (TVI, legge di Weber-Fechner) e la Contrast Sensitivity Function (CSF) che torna la risposta del sistema visivo ad una situazione di contrasto (per far ciò si mostra ad un osservatore una immagine bianca e nera a righe fitte sotto la quale si nasconde un segnale con i picchi corrispondenti alle zone nere; l'osservatore può e deve interagire con una speciale manopola che cambia la luminosità dell'immagine fino alla scomparsa delle righe, e quindi anche del segnale). In questo caso lo stimolo è semplice, isolato e foreale. La TVI è una funzione simile, si tratta in entrambi i casi di misure fisiologiche. Quello di Pattanaik è un metodo assoluto che non può in nessun modo essere relativo: si deduce che quando c'è poca luce il sistema visivo non è capace di mettere a fuoco la scena e più luce c'è più i colori saranno saturati. Si tratta di un sistema complesso che sfrutta i dati del sistema visivo in maniera molto arbitraria.

Reinhard (2002). Si tratta di un metodo relativo e globale ispirato ad uno speciale metodo fotografico di Ansel Adams (il metodo zonale): compie una mappatura globale massimizzando il bianco con l'intenzione di sollevare al massimo il contrasto locale (aggiunta delle componenti locali di *dodging* e *burning*). Si compone di due variabili: la Key Value e la Fi-sharpening. Conta come l'immagine viene acquisita e va bene per molte immagini (ma non per tutte) a patto di fare un corretto settaggio dei parametri.

Durand (2002). Costituisce un metodo segnalistico (non imita il sistema visivo umano e non è un metodo di tone mapping secondo la definizione). Si vuole ridurre il contrasto e contemporaneamente preservare comunque il maggior numero di dettagli. Lavora solo sulla luminosità. Il *bilateral filtering* analizza l'immagine pixel per pixel (come un filtro unilaterale) ed in più si basa anche sul valore dei pixel che filtra (se il filtro cambia forma a seconda della frequenza sottostante si può scegliere quale rumore è possibile eliminare).

Fattal (2002). In questo caso si tratta di un metodo concettualmente semplice, computazionalmente efficiente, robusto e facile da utilizzare; l'idea è quella di manipolare il gradiente comprimendo direttamente la dinamica; utilizzando il valore del gradiente è possibile comprimere/espandere localmente in maniera drastica il range dinamico preservando comunque i dettagli più fini. Si tratta di un metodo utile e funzionante.

Il livello di compressione di un pixel viene determinato in modo “morbido” considerando i livelli di compressione ottenuti su varie e differenti scale di sotto campionamento: l'obiettivo è evitare gli artefatti.

Hitah (2003). Metodo che costituisce un'applicazione diretta dell'algoritmo ACE all'alta dinamicità, si ispira a HVS ed ha un comportamento sia locale che

globale. Non è supervisionato, è robusto ma di contro ha un altissimo costo computazionale. Come effetto ha la correzione del colore.

Non è difficile ottenere un'immagine HDR corretta e soddisfacente, a patto che sia percettivamente accettabile. Bisogna sempre distinguere tra un metodo di tone mapping globale ed uno locale, detto anche tone render: nel tone rendering si considera tutta l'immagine per mappare un singolo pixel (il ch  può sembrare un controsenso). La corretta riproduzione della realt  non ci soddisfa a pieno. Mees ricav  ai suoi tempi la curva caratteristica e con essa cre  un buon compromesso per il tone mapping. Da solo (e da sempre) il sistema visivo compensa il gradiente di luminosit : quello che conta non   il valore in un solo punto ma si vogliono preservare gli edge per ottenere una buona percezione. Se due colori risultano uguali le curve che li descrivono si sovrappongono, senn  le curve sono tra loro molto distanti: la visione umana tende ad incrementare i colori chiari quando si trovano su di uno sfondo pi  scuro. Si vorrebbe quindi ottenere un tone render che mappi i pixel in quella quantit  che piace al nostro sistema visivo. La dinamica apparente dev'essere uguale arrangiandosi a livello spaziale.

Il sistema visivo vuole combattere il glare ed il tone scale della vista umana varia a seconda del glare percepito. Lavorare sui toni scuri induce un apparente ampliamento della dinamica (come avviene nell'effetto STRESS di GIMP).   difficile mescolare l'andamento globale e locale dei vari metodi come lo fa il nostro sistema visivo, specialmente con componenti locali molto forti.

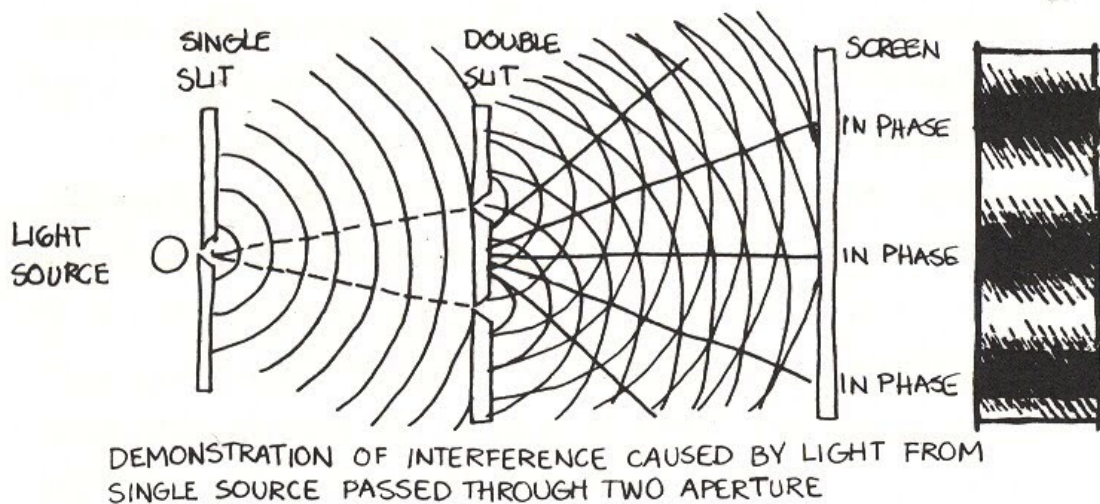
Capitolo 7

Sensori Fotosensibili

Sensori Fotosensibili per Imaging Digitale

La **luce** è un onda elettromagnetica caratterizzata, tra le diverse cose, da una velocità di propagazione, da una lunghezza d'onda propria ed una frequenza d'oscillazione, tuttavia è anche una particella (fotone) e come qualunque altra particella può avere un comportamento ondulatorio (proprio delle onde) come le onde possono avere comportamenti tipici delle particelle. L'energia di una particella deriva dalla sua frequenza d'onda.

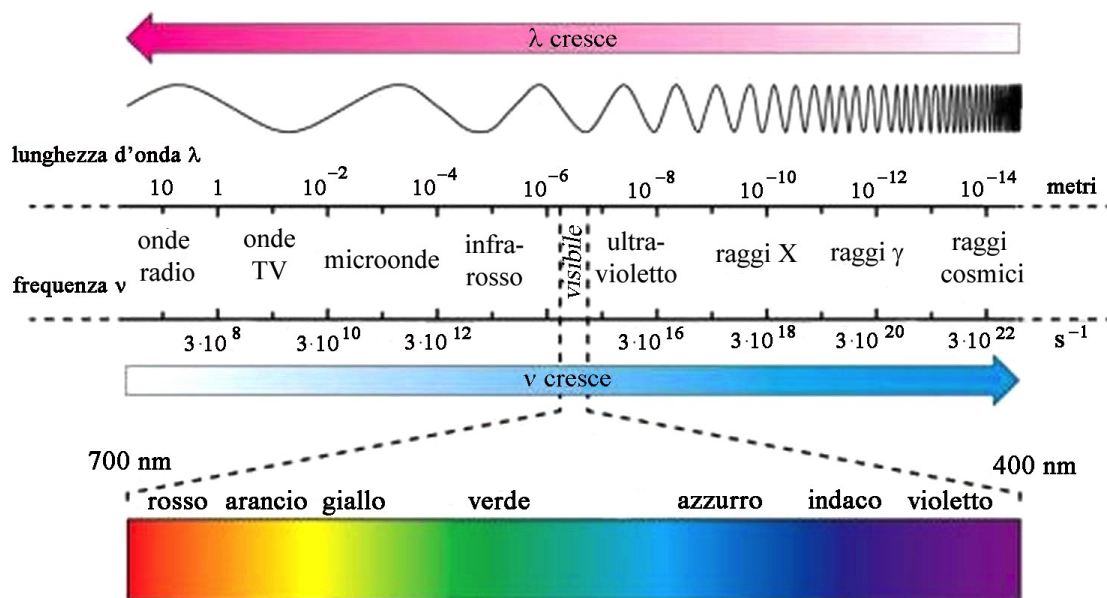
L'esperimento di Thomas Young (1803) constava di un cartoncino nel quale era stato effettuato un taglio verticale ed attraverso questo taglio andava proiettato un fascio di luce: sulla parete opposta la luce si andava proiettando in una certa conformazione (vedi figura). Se nel cartoncino si facevano due tagli paralleli la luce proiettata sul muro cambiava forma: si assisteva ad un fenomeno autentico d'integrazione tra i due fasci di luce (onde) come avviene con le increspature sul pelo dell'acqua con punti in cui le onde si sommano ed altri distinti dove si annullavano. Si è dimostrato così il comportamento (e la natura) ondulatorio della luce.



Einstein scopre l'**effetto fotoelettrico**: nel momento in cui si irraggia (o semplicemente si illumina) un materiale con la luce solare i fotoni/onde della luce che possiedono una certa energia fanno *saltare* gli elettroni esterni appartenenti agli atomi del materiale in esame, tuttavia non dipende dalla quantità di fotoni ma dall'energia/calore della luce. In questa maniera si è dimostrata la quantizzazione della luce. Il fotone è una particella con una certa quantità di energia al suo interno.

A livello di dubbio su quale sia la natura della luce non si è ancora arrivati ad una risposta unanime. Definire la luce come un'onda e contemporaneamente come una particella è corretto, tuttavia si tratta di descrizioni riduttive per un fenomeno che è assai più ampio e complicato di quanto si riesca a descrivere a parole. Non solo la luce si comporta in questo duplice modo ma anche qualsiasi altro materiale possiede una natura ondulatoria oltre che particellare (che nella realtà di tutti i giorni non si nota perchè risulta visibile solo per masse molto

piccole mentre per quelle più grandi le onde tendono ad essere molto più lunghe quindi invisibili ad occhio nudo).



Lo spettro elettromagnetico. Risultano interessanti tra tutte soprattutto le frequenze dello spettro del visibile. La frequenza come unità di misura d'onda aumenta quanto la lunghezza dell'onda stessa diminuisce. La maggior parte degli esseri viventi vede solo la parte di spettro della luce visibile (con leggere variazioni a seconda dell'essere vivente, alcuni vedono più negli ultravioletti altri più nell'infrarosso). Lo spettro riporta i vari tipi di onda elettromagnetica anche se dal punto di vista fisico tutte queste lunghezze d'onda rappresentano variazioni poco significative dello stesso fenomeno.

Le onde elettromagnetiche generano campi elettrici e magnetici e si producono decelerando od accelerando una carica elettrica. Come tutte le onde, anche quelle elettromagnetiche hanno una direzione di propagazione. Si tratta di increspature nell'equilibrio tra campo elettrico e campo magnetico (che si solito sono in quiete) le quali si propagano alla velocità della luce.

Come interagiscono con la materia? Dipende dalla frequenza della radiazione elettromagnetica, dalla chimica della materia irraggiata e dalle caratteristiche fisiche delle particelle che la compongono. Si prende lo spettro e, passando di frequenza in frequenza con maggior attenzione alle principali, si creano delle categorie a seconda dei comportamenti del materiale irraggiato: al variare dell'energia delle onde si controlla cosa succede.

1. **Microonde.** Sono onde che hanno una tale energia da far ruotare su loro stesse le molecole dei materiali irraggiati non conduttori, sprigionando così calore ed energia. Certe molecole sono agitate solo da certe specifiche lunghezze d'onda, il che costituisce il concetto di base del *forno a microonde* che mira ad agitare solamente le molecole d'acqua. Di solito la stragrande maggioranza degli oggetti è trasparente alle microonde;
2. **Infrarossi.** Un po' come le microonde agitano stiracchiando i legami tra gli atomi e le molecole che costituiscono la materia facendoli vibrare (si da per assodato che in un oggetto caldo le molecole si muovono più velocemente di quelle di un oggetto più freddo → lo *zero assoluto* riporta

la temperatura oltre la quale non avviene più nessun movimento a livello molecolare tanto da non poter misurare un freddo più acuto di quello). I materiali tendono ad assorbire di più gli infrarossi rispetto alla luce visibile ed alcuni oggetti sono trasparenti all'infrarosso. Le microonde sono decisamente più pericolose per gli esseri umani ed è per questo che, contrariamente agli infrarossi, non sono utilizzate per il riscaldamento degli ambienti. Il passaggio tra infrarossi e microonde è molto sfumato, così per il riscaldamento si utilizzano le onde infrarosse più vicine allo spettro visibile e più lontane dalle microonde;

3. **Luce visibile.** Viene tutta assorbita dai materiali e veramente pochi ne sono trasparenti (tipo il vetro). Ha una potenza tale da far saltare completamente i legami tra elettroni ed atomi/molecole così da generare calore (fenomeno molto importante);
4. **Ultravioletto.** Vari materiali/atomi irraggiati con queste onde tendono a raggiungere l'energia di ionizzazione strappando gli elettroni dalle molecole. Hanno banda molto larga e sono onde molto assorbite dai materiali. Le onde più alte tendono ad alterare le proprietà chimiche (quindi i legami molecolari) tra atomi: ecco perchè si sente la necessità di filtri appositi per schermare queste onde, anche se l'atmosfera fa già da sola gran parte del lavoro;
5. **Raggi X.** Vengono solo in parte assorbiti dagli oggetti e possiedono una quantità di energia tale che dopo aver causato il salto di un elettrone in un atomo l'onda genera un fotone che continua a viaggiare anche se drasticamente depotenziato e con lunghezza d'onda più bassa (fonte di quel bagliore tipico degli oggetti radioattivi). La materia è generalmente trasparente ai raggi X, alcuna meglio di altra: infatti si utilizzano per fare le lastre mediche perchè la pelle li lascia passare ma le ossa sono più resistenti. Queste onde sono ionizzanti quindi si tende ad utilizzarle poco perchè leggermente pericolose. Dopo di loro, nello spettro, ci sono i *raggi gamma*.

Perchè noi vediamo solo la luce visibile? Perchè tra tutte le lunghezze d'onda dello spettro solo la luce visibile viene riflessa dagli oggetti e se così non fosse non potremmo vedere quali oggetti ed ostacoli si trovano sul nostro cammino rendendoci incapaci di schivarli (ma anche di inseguirli o cercarli). Inoltre il sole emette innanzitutto luce visibile (a dire il vero emette praticamente tutto lo spettro ma l'atmosfera filtra le altre lunghezze d'onda lasciando passare solo la luce visibile). Esistono vari tipi di animali che vedono più nell'infrarosso o nell'ultravioletto, come per esempio le api per le quali le corolle dei fiori sono molto più variopinte di come le vediamo noi. Il cielo è blu per via dell'effetto *daylight* dovuto all'interazione tra le onde luminose e le particelle/molecole che compongono l'atmosfera: detta anche interazione di scattering, si parla della relazione tra la lunghezza d'onda che attraversa l'interno della molecola atmosferica e l'angolo d'uscita il quale cambia a seconda dell'energia dell'onda in entrata.

Semiconduttori. Ogni atomo è composto da un nucleo di protoni e neutroni attorno al quale orbitano degli elettroni in quantità (sia per atomo che per ogni orbita) commisurata alla propria natura; se a partire dal numero per il quale l'atomo risulta stabile si sottraggono degli elettroni dall'orbita più esterna questo farà di tutto per recuperarli, di rimando se se ne aggiungono in più l'atomo farà di tutto per disfarsene per recuperare l'equilibrio perduto. Esistono

però atomi che nell'ultimo orbitale possiedono la metà del numero concesso per quell'orbita e si trovano in un stato intermedio tra il volerne cercare per completarla od il volerli cedere: sono gli atomi dei materiali semiconduttori. In un conduttore gli elettroni dell'orbita esterna sono liberi di andare a venire da altri atomi simili creando così una corrente elettrica (in presenza di una differenza di tensione che stimoli lo scambio di particelle); se questi elettroni invece sono vincolati all'atomo il materiale è isolante. Nei materiali isolanti gli elettroni rimangono fermi nella banda di valenza e non si muovono salvo interventi straordinari di raggi gamma grazie ai quali vengono spostati nella banda di conduzione. Un materiale isolante vede molta distanza tra le due bande; un conduttore invece le ha molto vicine (si sovrappongono) tanto da rendere il salto degli elettroni naturale; un materiale semiconduttore presenta una via intermedia dove lo spazio tra le due bande (detto spazio inerte) è abbastanza da non stimolare il passaggio naturale degli elettroni ma abbastanza piccolo da facilitarlo quando necessario.

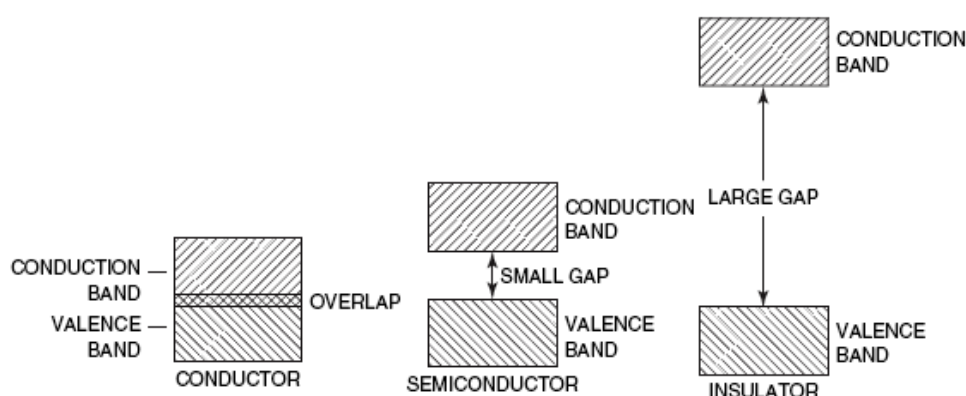


Figure 9-1. Valence and conduction bands in conductor, semiconductor, and insulator.

Quando gli elettroni fanno il salto all'interno dei semiconduttori rendono il materiale conduttore a tutti gli effetti. Se un fotone colpisce un atomo può costringere uno dei suoi elettroni a compiere il salto tra bande: per creare un rivelatore di fotoni si sfrutta appunto la natura dei semiconduttori.

Drogaggio. I semiconduttori esistono già in natura tuttavia si vuole forzare il loro comportamento tipico mescolandoli ad altri materiali (creando quindi delle impurità): si ottengono allora degli *elementi drogati*. Il comportamento di un semiconduttore naturale è così esaltato in modo tale che quando arriva un fotone che colpisce un atomo del materiale si genera un flusso di corrente, altrimenti rimane inerte. Se si avvicinano materiali semiconduttori drogati in maniera diversa (e magari opposta) nel punto di giunzione gli atomi privati di elettroni tenderanno a strapparne a quelli che invece ne hanno in eccesso perpetuando un scambio di cariche: in questa maniera chi cede cariche si polarizza positivamente mentre chi le acquista lo fa negativamente. Si viene a creare un punto di perfetto equilibrio tuttavia lo scambio di elettroni e la creazione di cariche positive e negative crea un campo elettrico che inibisce gli scambi tra gli atomi più distanti dal punto di giunzione: è così che avviene la creazione di un diodo utilizzato per la rilevazione di fotoni. Quando un fotone colpisce un atomo del diodo fa saltare un elettrone il quale bypassa il campo elettrico creato prima e raggiunge la zona di atomi che lo richiedono (così per

ogni fotone, ovviamente non se ne presenta uno solo alla volta) generando così una corrente elettrica. In questa maniera si avrà una misura di quanti fotoni sono passati a seconda della quantità di elettroni spostati.

I sensori. In genere i sensori delle macchine fotografiche digitali sono costituiti da array/matrici regolari di fotodiodi. Di solito i fotodiodi sono sensibili a qualsiasi fotone di luce sia nello spettro del visibile che no (e per questi fotoni spesso si utilizzano dei filtri per limitarne l'accesso) e vedono in bianco e nero: per cogliere i colori si appongono sopra i sensori dei filtri appositamente colorati che lasceranno passare solo il corrispondente colore (quindi solo certe lunghezze d'onda). Il problema sta nel come si faccia a leggere la quantità di fotoni e di cariche elettriche collezionate dai fotodiodi in seguito alla loro impressione.

1. **Full Frame CCD (Charge Coupled Device).** Si tratta di uno dei primi sensori digitali tra i più famosi, costituito da un array/matrice di fotodiodi il quale copre tutta l'area del sensore (da qui il nome *Full Frame*). Quando la luce colpisce un diodo come spiegato prima si istigano degli elettroni a spostarsi generando un flusso elettrico. Si divide l'acquisizione della luce in due fasi: la prima composta dall'apertura e chiusura dell'otturatore il quale fa entrare una certa quantità di luce e la seconda composta dall'impressione della piastra metallica posta sotto i fotodiodi la quale, precedentemente polarizzata in modo da avere una tensione che possa attirare le cariche elettriche, cambia polarizzazione in corrispondenza di ogni fotodiodo riportando così un certo valore per ciascuno di essi. A partire dalla placca le cariche vengono convertite ad un certo livello di tensione la quale dà un valore numerico alla quantità di fotoni catturati e cariche spostate da ogni diodo. Questi tipi di sensori hanno un'alta sensibilità e sono semplici da costruire (ancora non necessitano di microlenti), tuttavia c'è la necessità di montare un otturatore meccanico grazie al quale discriminare la quantità di luce immagazzinata dai diodi e non è possibile leggere i loro valori se l'otturatore è aperto (ed i diodi sono impressionati), inoltre la lettura dei valori riga per riga è molto lenta ed è difficile sincronizzare tutti questi procedimenti;
2. **Frame Transfer CCD.** Si tratta di un sensore costruito alla stessa maniera di quello precedente salvo per il fatto che la sua area d'estensione è larga il doppio: il secondo array/matrice gemello del primo in realtà è coperto e non viene impressionato direttamente dalla luce, in esso verranno immagazzinate le cariche generate per poi essere lette più in fretta magari durante una seguente acquisizione. Produrre sensori di questi tipo ha un costo molto elevato (per via della produzione moltiplicata dei diodi) e la lettura dev'essere veloce abbastanza anche se c'è sempre bisogno di un otturatore meccanico per non rischiare di sovraccaricare i diodi (effetto di *charge smearing*);
3. **Interline Transfer CCD.** Come nel primo caso si ha un sensore completamente popolato da un array/matrice di fotodiodi i quali sono in maniera alternata riga per riga un po' esposti ed un po' nascosti (quelli nascosti sono ovviamente utilizzati per il readout). In seguito all'apertura dell'otturatore ed alla esposizione dei diodi del primo tipo le cariche vengono spostate immediatamente in quelli nascosti. Finalmente si può implementare un otturatore elettronico il quale semplicemente decide

quando è ora di spostare le cariche accumulate dai diodi impressionati a quelli nascosti, tuttavia permane il problema del charge smearing. Si ha però una ridotta sensibilità e si utilizzano delle microlenti per convogliare la luce sui fotodiodi dedicati alla cattura dei fotoni (cercando di escludere quelli nascosti). L'acquisizione e l'integrazione dei dati sono parallele;

4. **CMOS APS.** Si tratta di una tecnologia alternativa che però non ha conquistato il favore del pubblico: si integra nel fotodiodo la traduzione in segnale della tensione che prima veniva fatta in un secondo momento così da ottenere direttamente il valore del pixel. Grazie alla tecnologia CMOS per la componentistica elettronica si sono creati dei piccolissimi transistor da poter inserire direttamente sotto al fotodiodo. Però se si vuole integrare tutto all'interno del diodo si riduce drasticamente la dimensione reale del fotodiodo così servono per forza delle microlenti per intensificare la luce sulla regione ricettiva del diodo. Questo sistema permetteva la lettura continua del valore del pixel con l'ausilio di un otturatore elettronico. Il processo produttivo era molto economico poiché si sfruttavano macchinari utilizzati anche per altre produzioni; anche l'elettronica era molto semplice con un conseguente risparmio energetico tuttavia con ridotte sensibilità e range dinamico (il quale dipende da quante cariche possono essere immagazzinate dal diodo il quale, se ridotto nelle dimensioni, riduce anche la sua capacità).

Prima dei fotodiodi si utilizzavano le valvole come fotosensori, poi nella produzione di sensori low cost è diminuita la presenza di CCD in vantaggio al CMOS il quale rimane comunque preferibile per certi tipi di impieghi (come per esempio la fotografia astronomica) e va lentamente a sostituire il CCD anche se non in tutti i casi.

Foveon Technology. Tecnologia che mira a risolvere la sensibilità del colore dei sensori senza utilizzare nessun filtro apposto. Nei sensori tradizionali ogni pixel è reso da 4 fotodiodi i quali corrispondono 2 al verde, 1 al rosso ed 1 al blu e si vuole fare in modo che ogni pixel sia indipendente e discerna da solo il colore da attribuire al pixel senza avvalersi di 4 diodi (ma solo di sé stesso). Si fanno 3 layer sovrapposti di diodi (uno per ogni canale cromatico) e si sfrutta la capacità di penetrazione del fotone per determinare quali fotoni arrivano a quali layer.

Questo sistema non ha molto spopolato perché costoso e la discriminazione del colore risulta essere probabilistica e non deterministica e poco precisa. I sensori di questo tipo sono molto piccoli per via del fatto che per costituire un pixel non ne servono più altri 4 (con conseguente eliminazione del color aliasing). Si ha però una risoluzione più alta ed aumenta di conseguenza il rumore di fondo legato alla probabilità che gli elettroni saltino autonomamente sfalsando le rilevazioni in presenza di calore (in questo caso per ovviare il problema basta raffreddare il sensore).

Capitolo 8

SCA

Spatial Color Algorithm (SCA)

Anzichè sintetizzare il colore direttamente dallo stimolo lo si fa a partire da tutto il contenuto della scena (quindi dallo stimolo e dal suo contesto). Si passa dalla colorimetria tradizionale (che vede sempre costante il rapporto tra stimolo e descrizione) all'apparenza cromatica. Esistono due modi per farlo: o ci si affida allo CIECAM e si “stiracchia” il colore a seconda delle informazioni di contesto oppure ci si avvale di uno **Spatial Color Algorithm (SCA)** il quale calcola l'apparenza utilizzando le relazioni spaziali proprie dei pixel come del resto fa anche l'apparato visivo. Il range dinamico si mappa quindi secondo principi percettivi e questi algoritmi (dei quali esistono diversi tipi) dovrebbero fornire un mappaggio in linea con l'apparenza.

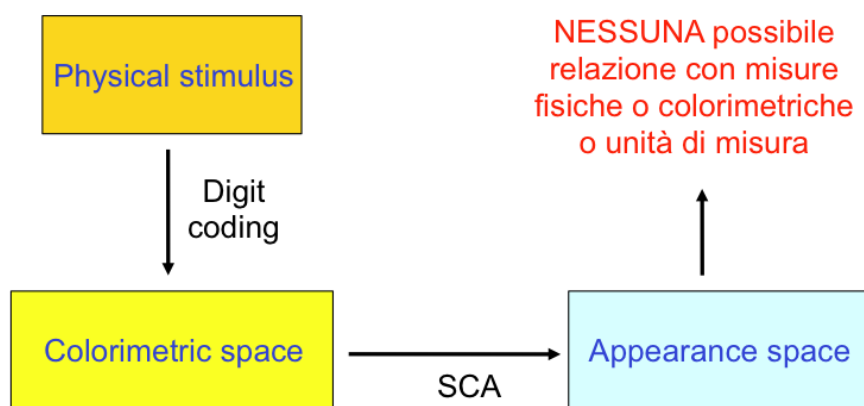
Effetti spaziali. Siamo abituati a chiamare gli effetti spaziali “illusioni ottiche” tuttavia non si tratta di vere e proprie illusioni, sarebbero da intendere tali solo in relazione ad un modello che, comunque sbagliato, noi riteniamo corretto. Il modello che si dà per scontato, comunque, risulta essere il modello corretto anche se, quando lo si smentisce, il modello diventa improvvisamente superato e l’“illusione” smette di essere tale. *La nostra percezione è una manifestazione della realtà che a noi sembra corretta perchè è la nostra realtà*; misurare la realtà è facile ma non lo è misurare la percezione. Realtà e percezione non coincidono e questo non coincidere è oggetto di studio. Chi crea gli algoritmi SCA vorrebbe computare la realtà per ottenere la percezione che deriva da quella stessa realtà: ci si riesce solo per il 70% delle scene reali computate. Per l'80% degli effetti di contrasto di simultaneità si basa sul fatto che è difficile ottenere l'apparenza a partire dalla realtà creando così delle specie di “illusioni” ottiche. Dal punto di vista di CIE CAM si prendono i valori numerici del colore di sfondo della scena (ed anche qui bisogna definire che cosa è sfondo e che cosa è in primo piano, quindi soggetto: non è una cosa scontata specialmente in immagini complicate e meno geometriche di quelle dei casi di studio) e lì si sfalsano dal colore chiaro a quello scuro (e viceversa a seconda del colore di partenza) aumentando così la differenza ed il contrasto con il colore in primo piano per simulare la percezione; questo discorso vale quasi per tutti i casi salvo per quelli dell'assimilazione. L'assimilazione costituisce un problema nel quale statisticamente è coinvolto il concetto della *ripetizione* (patternizzazione). Si individua così un discorso di contrasto che spiega la maggior parte degli effetti “illusionistici” con una specie di onda/segnale che simula a modo suo essa stessa il contrasto. Nelle configurazioni 3D l'effetto illusionistico è ricondotto ad una segmentazione della scena effettuata in maniera strategica dove i colori vengono appositamente scambiati dando un significato di profondità alla scena. Avere un alto range dinamico aiuterebbe a rendere meglio l'apparenza ma non sarebbe comunque possibile ottenerlo (a causa della solita limitatezza della tecnologia). SCA garantisce l'avvicinamento del proprio risultato al risultato autentico della percezione.

A partire da una scena reale della quale, per un qualsiasi motivo, non posso ricostruirne e riprodurne lo stimolo, utilizzando un algoritmo SCA l'apparenza

finale risulterà molto simile alla reale apparenza.

Per fare tutto ciò bisogna essere locali, ma quanto? La spazialità della località è l'aspetto centrale della questione. Dopotutto non è più interessante ricalcolare il valore del colore del pixel perchè l'apparenza, in fondo, localmente coincide con la realtà (il che deriva dai rapporti reciproci ed assai meno dalla componente del segnale in sé). Allora la comparazione dev'essere fatta sì localmente ma anche globalmente.

Per creare uno SCA si parte da uno stimolo fisico facilmente caratterizzabile il quale corrisponde ad un digit che rappresenta una buona riproduzione della scena reale (poi da caratterizzare all'interno di uno spazio colorimetrico). Poiché lo stimolo non corrisponde all'apparenza si deve riprodurre l'apparenza attraverso l'impiego dello SCA tenendo sempre in conto il contesto all'interno del quale lo stimolo si verifica. Si passa così allo spazio dell'apparenza che è ancora più complicato dello spazio reale. In tutto ciò non è mai possibile determinare un'unità di misura che quantifichi queste operazioni.



Gli spazi di colore fisici (anche se non sono propriamente “fisici” ma sono ciò che di più si avvicina ad una misura fisica) sono XYZ, RGB ecc; gli spazi invece psicofisici sono il CIEL*a*b* che è una deformazione di XYZ, il Munsell, ecc; gli spazi dell'apparenza legati al modello/algoritmo colorimetrico spaziale sono il CAM che considera anche la luce all'interno del contesto, il Retinex, l'ACE, ecc che ricavano l'apparenza ricalcolando il colore del pixel come esso stesso una stima dell'apparenza.

L'**apparenza**, a seconda dello scopo, un po' cambia. Si pensi alla fotografia di un oggetto diversamente illuminato, anche se si sa che si tratta di un oggetto con tutti i lati dipinti dello stesso colore apparentemente sembra che ne abbia alcuni colorati in maniera differente (allora la radianza varia). Un pittore che dipinge la scena vorrà allora utilizzare due colori fisicamente differenti per riportare l'apparenza corretta. In verità l'oggetto è tutto dello stesso colore e lo si può capire grazie ad un ragionamento mentale. L'apparenza è la rappresentazione che il cervello si costruisce riguardo la scena svincolandola dal contesto; questo s'intende per sensazione visiva. Tutti gli SCA riportano la sensazione visiva. Non ci si occupa in questa sede della costruzione della scena a livello mentale: vedere un oggetto significa ragionare su quell'oggetto. Come esseri umani noi siamo abituati ad effettuare una normalizzazione del contesto.

Si tenta in questo modo di costruire un modo per confrontare stessi pixel in contesti diversi. Il sistema visivo umano fa, autonomamente e con la complicità del cervello, diversi aggiustamenti. La costanza di chiarezza è quel concetto per

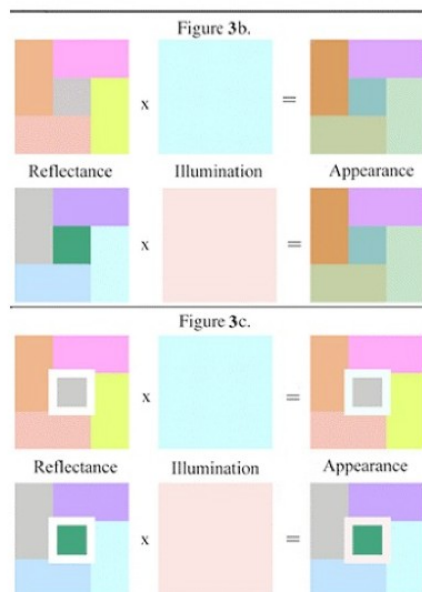
il quale si percepisce stabile una scena anche se il cambio di illuminazione avviene in maniera repentina, a livello di contrasti: una macchina fotografica, tuttavia, non lo fa come il nostro cervello che si aiuta mediando il procedimento attraverso la retina. C'è sempre ambiguità tra la riflettanza e l'illuminante ed otteniamo per prima cosa, sempre, informazioni sulla riflettanza (per una questione relativa alla sopravvivenza) tuttavia non riusciamo mai a dirimere l'ambiguità con la luminanza, cosa che siamo sempre in grado di stabilire.

Vincolo computazionale → serve qualcosa che si adegui ad un range dinamico alto. Il **gray world** consiste nel calcolo della sommatoria dei pixel di una intera immagine la quale dovrebbe dare un valore di grigio medio complessivo per quell'immagine (la cosa però non è del tutto veritiera e questo metodo viene utilizzato raramente nella determinazione di modelli).

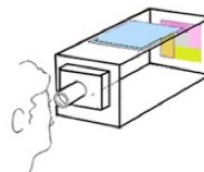
Costanza cromatica → al variare della luce non si smette mai di riconoscere gli oggetti facenti parte della scena; il sistema visivo, da solo, tende ad eliminare un po' della luce presente. A riguardo di ciò esistono moltissimi lavori, la maggior parte sbagliati perchè partiti da presupposti erronei.

White Patch → regolazione cromatica con un bianco di riferimento che funge da punto centrale per il sistema visivo (in ogni scena ed in ogni situazione è sempre possibile determinare un bianco di riferimento). Avviene indipendentemente su ogni canale.

Linea di Ipparco (di Micea) → Ipparco era un filosofo/scienziato greco il quale, partendo dalle osservazioni astronomiche sulla luminosità delle stelle, descrisse con una linea retta che al variare della radianza si ha una sensazione di cambiamento della luminosità. Il bianco è sempre bianco tuttavia la sensazione cambia a seconda dell'intensità dello stimolo; se il bianco fosse assoluto non si riuscirebbe a capire che cosa è più o meno luminoso. Tramite una white patch si prendono i pixel con i valori massimi per ogni canale (all'interno di una scena) e li si porta ai massimi valori consentiti adeguando di conseguenza tutti gli altri pixel della scena. Se il sistema visivo è composto da 3 canali distinti ed uno di questi spicca sugli altri vuol dire che nella scena c'è maggioranza di luce in quel canale e si tende ad alzare gli altri valori così da operare un *bilanciamento del bianco*. Tuttavia non è esattamente quello che fa il sistema visivo.



Maximov shoeboxes



color constancy is controlled by channel maxima

Scatola da scarpe di Maximov. Si prende una scatola da scarpe e ci si fa un

foro in cima, sul coperchio, da coprire a seconda dei casi con una serie di filtri colorati di riflettanza. Nella scatola, su una delle pareti corte, va apposto uno schema colorato e sulla parete opposta va effettuato un secondo foro. L'idea è quella di combinare filtri di riflettanza e schemi colorati in maniera da ottenere sempre la stessa impressione (apparenza) di colore, e la cosa funziona finché non si inserisce una cornice bianca; questo accade perché il bianco è il punto di riferimento del sistema visivo.

Quando lo fondo di un'immagine è più scuro dello stimolo in oggetto, lo stimolo decade e lo sfondo prende il sopravvento. Se il bianco si trova vicino al grigio, il grigio si scurisce per contrasto; quando invece si allontana anche il grigio si schiarisce e se proprio il bianco manca allora è il grigio a sembrare bianco (il tutto su un sfondo scuro). Si possono leggere le stesse configurazioni anche nel modello dell'apparenza.

In generale il nostro sistema visivo amplifica i bordi e sopprime i gradienti.

Riassunto. Quando si caratterizza il colore come informazione nel punto preciso in cui lo si studia ci si serve della **colorimetria classica** la quale risulta inadatta se utilizzata all'interno di un contesto. Il **contesto** modifica l'apparenza e spesso lo si caratterizza mediante alcuni aspetti: lo **sfondo** è un insieme di pixel colorati del quale non importa la struttura (allora sarà un *contributo spaziale*) → su questa concezione dello spazio e dello sfondo si basano i modelli SCA. L'assimilazione ed il contrasto di simultaneità sono due effetti ottici opposti ed è difficile che uno SCA qualsiasi riesca ad elaborarli ed a riprodurli fedelmente. Il sistema visivo pretende che la percezione sia sempre il più ricca possibile così, nel caso dell'effetto di simultaneità, si tende sempre ad aumentare il contrasto delle scene. Il **sistema visivo** è fondamentalmente un soppressore di gradienti ed un esaltatore di bordi: tuttavia uno SCA potrebbe non essere per forza un modello fedele del sistema visivo (può capitare che la colorazione prenda una direzione opposta a quella che prenderebbe all'interno dell'apparenza). Gli SCA si dividono in due tipologie: *modelli quantitativi* e *modelli qualitativi* dove quelli *quantitativi* tendono a seguire l'andamento del sistema visivo in maniera numericamente/matematicamente corretta e coerente (tramite la calibrazione dei dati in ingresso ed in uscita).

Il *meccanismo* detto del *fill'in* corrisponde a quello dell'*effetto edge* e pochi modelli SCA tendono a tenere conto di questi procedimenti di segmentazione dell'attività visiva. Un modello del sistema visivo comunque dev'essere sempre **locale**: il tipo di variazione della computazione dell'immagine dev'essere quindi locale.

L'**apparenza** non possiede né uno spazio di riferimento tantomeno un'unità di misura; l'apparenza si compone della distinzione tra sensazione e percezione e cambia a seconda della risposta del sistema visivo ad un certo stimolo (con una risposta alta oppure bassa, tuttavia è impossibile separare una percezione d'alto livello da quella di basso livello, poiché si compongono indissolubilmente ed è compito del cervello fare questo tipo di elaborazioni in maniera totalmente automatica). La modellizzazione della percezione è molto complessa e va tenuta in conto per livelli medi di caratterizzazione (la vista, come tutto del nostro organismo, va educata ed allenata). In generale con SCA ci si occupa della percezione ai bassi livelli preferendo un approccio dal basso verso l'alto (e non viceversa, perché comunque si tratta di modellizzazioni diverse).

Gli algoritmi SCA fanno riferimento alla sensazione; il CIECAM è invece un

approccio top down perchè deve incominciare riconoscendo per prima cosa lo sfondo e la figura in primo piano.

Lavorando con gli SCA sorgono moltissime complicazioni, tuttavia non è detto che vadano risolte tutte in un colpo solo: ci saranno algoritmi che come scopo si dedicheranno maggiormente a certi dettagli rispetto ad altri.

I meccanismi dell'**esposizione** sono molto simili a quelli messi in atto dal nostro sistema visivo il quale tende ad aggiustare la vista a seconda dello stimolo esterno → *gray world*: per ogni canale (in maniera indipendente) si opera un aggiustamento cromatico mirato partendo da un colore grigio di riferimento che corrisponde ad una sommatoria centrata su un certo valore. Il valore finale dei colori così trattati, tuttavia, potrebbe non essere del tutto corretto.

Il **colore** dopotutto è un misto di riflettanza e del colore dell'illuminante, tuttavia la vista tende a togliere un po' del colore dell'illuminante purificando l'impressione dei colori indipendentemente dalla luce della scena, il tutto per ottenere una percezione il più stabile possibile.

La **linea di Ipparco** aiuta a capire che uno stimolo più luminoso appare comunque più chiaro rispetto ad altri meno luminosi (allora il bianco sembra più bianco). Lo *spazio di Munsell* va calibrato sotto un certo illuminante e costituisce un modo per quantificare l'apparenza. Il vincolo computazionale si trova nella zona più chiara che diventa automaticamente il colore bianco di riferimento in un certo intorno della scena, quindi localmente → **white patch**. Si vogliono normalizzare tutti i canali cromatici rispetto a questo bianco ma non a livello globale che sarebbe comunque più semplice da fare ma avrebbe poco significato, così si cerca la località opportuna: il filtro varia a seconda della qualità della scena che si osserva, solo così si può rendere l'idea del sistema visivo a livello computazionale. Il bianco spazialmente comanda su tutti gli altri stimoli, specialmente se si prende in considerazione lo spazio di Munsell. Più il bianco si trova distante dal colore preso in esame più questo sembrerà chiaro (quindi ha un minore effetto locale), viceversa più il bianco è vicino più comanda e per contrasto il colore sembrerà scuro → *peso della distanza*; più il bianco si dirada meno tende a contrastare → allora conta anche la *densità della distribuzione dei pixel di bianco* attorno allo stimolo in esame. Forma VS Colore → conta la distribuzione dei pixel ed il fatto che nella loro conformazione si creino o meno dei bordi netti.

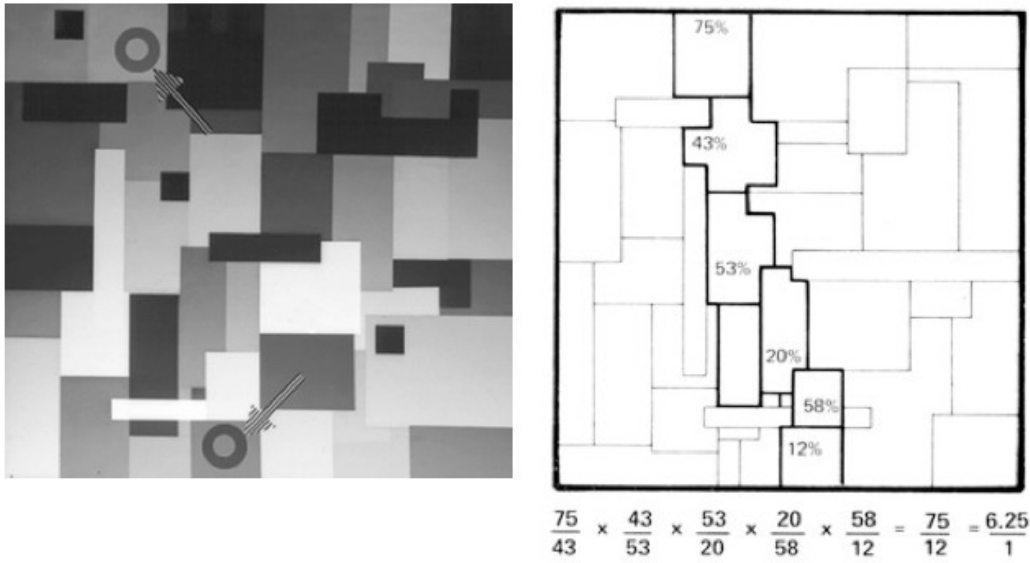
Aggiustamenti dell'SVU e la loro modellizzazione

Si parla di modelli qualitativi e quantitativi tra i quali ci si concentra su quelli qualitativi (parametri, calibrazioni I/O, dinamica, ecc).

La struttura comune a tutti gli algoritmi di SCA è solitamente costruita attorno ad alcune operazioni fondamentali:

1. **prima fase**: dato un punto/pixel si ricomputa il suo valore colorimetrico che non dipende più da sè stesso in quanto tale ma dal legame che ha con i pixel che gli stanno nelle immediate vicinanze (*pixel recomputation*);
2. **seconda fase**: dato il valore ricomputato lo si scala (solitamente a partire dalla condizione in cui si trova lo si riporta in una scala del tipo 0-255, ma dipende di caso in caso) in maniera indipendente sui 3 canali principali; comodo da fare per l'assenza di vincoli colorimetrici e, straordinariamente, crea pochissimo rumore di fondo (*rescaling*).

Retinex. Il termine RETINEX è composto dall'unione delle parole *retina* e *cortex* (corteccia) dal momento che non è ben chiaro dove avvenga il passaggio dal flusso radiante alla luminosità. Creato negli anni '60 da McCann e Land al fine di simulare e capire al meglio il funzionamento del sistema visivo. Si parte dal presupposto che esistono 3 computazioni separate del segnale visivo relative ai 3 principali canali percettivi. Quando si deve ricomputare il valore del pixel preso in esame rispetto agli altri che lo circondano ci si serve della *lightness* intesa come *chiarezza relativa di canale*, perchè non si ha interesse nel ricomputare lo stimolo d'ingresso ma solo di rendere ciò che avviene all'interno del sistema visivo, ovvero come un punto/pixel rischia di apparire più chiaro o più scuro a seconda della luminosità relativa ai vari canali di colore. Non si ottiene esattamente la riflettanza ma qualche cosa di correlato; Retinex esattamente come il sistema visivo umano non ha intenzione di separare l'illuminazione della riflettanza.



Si parte da un modello in bianco e nero il quale è stato illuminato in maniera sfumata (quindi presenta un lato più illuminato del lato opposto) sul quale si individuano due regioni apparentemente differenti ma che nella realtà presentano lo stesso valore di luminanza. Dopo esser stato fotografato si prova ad individuare un cammino casuale che colleghi le due aree in modo da calcolare il rapporto tra i bordi delle due aree; anche il sistema visivo esplora una scena attraverso dei cammini casuali. Il valore di un pixel dell'immagine allora è dato da una catena di rapporti (anche se la forma originale era individuata da una produttoria sequenziale di rapporti) nella quale si sfrutta il *concetto di soglia* e si calcola la produttoria solo se effettivamente c'è differenza tra pixel vicini: si tratta di un procedimento che non è fondamentale ma che risulta essere efficace per eliminare i gradienti. I cammini sono individuati in maniera randomica: si cercano i vicini locali/globali grazie ai quali rendere possibile la computazione dei rapporti. La *lightness* relativa lungo un cammino è data da:

$$r_{l,m,s}^{i,j} = \sum_{x \in path} \delta \cdot \log \frac{I_{x+1}}{I_x} \quad \delta = \begin{cases} 1, & \text{if } \left| \log \frac{I_{x+1}}{I_x} \right| > threshold \\ 0 & \text{else} \end{cases}$$

mentre quella relativa nel punto/pixel, separatamente per ciascun canale, è individuata da:

$$R_{l,m,s}^i = \frac{\sum_{k=1}^N r_{l,m,s}^{i,j_k}}{N}$$

Si tratta di un algoritmo pesantissimo perchè la computazione avviene per ogni pixel, tuttavia l'idea di fondo ed il procedimento risultano essere abbastanza semplici. Non è possibile individuare un solo percorso e calcolare tutto in relazione di un solo cammino perchè la sommatoria avviene tra una selezione di cammini alternativi; i cammini sono ideati perchè percorrano l'immagine alla ricerca di un bianco di riferimento a livello locale il quale più è visibile e maggiori saranno le probabilità di incrociarlo attraverso diversi cammini. Il meccanismo di ricerca del bianco di riferimento elegge a “bianco” la zona più chiara che incontra in maniera iterativa all'interno di un certo range; sfrutta un meccanismo statistico poco utilizzato e studiato ma molto efficace.

Ma come ottengo la *località*? Poiché la natura dei cammini individuati non è dritta, si sfruttano i *cammini browniani* (quindi casuali) i quali si adattano perfettamente all'esigenza di campionare moltissimi dati; se la campionatura è abbastanza fitta si riuscirà a raggiungere tutti i punti del sistema raggiungendo una copertura globale a livello di algoritmo (il che rende tutto più efficiente, anche se non collima affatto con ciò che accade davvero nel sistema visivo, tuttavia se si campiona troppo poco si crea del rumore). Poiché la percezione è multilivello si prova ad effettuare una campionatura multilivello: si dispone di una quantità di versioni della stessa immagine ma a qualità differenti, dopo aver individuato e creato il percorso casuale per l'esplorazione dell'immagine si prende in considerazione l'esemplare della fotografia in esame alla qualità più opportuna a seconda dei casi e della direzione del cammino (con l'idea di simulare il lavoro della fovea oculare); si tratta di un algoritmo che funziona bene e risparmia energie di calcolo. Se serve trovare il pixel più bianco di un certo intorno si campiona una quantità di pixel vicino a quello in esame così da individuare il più chiaro: questo è il sistema più efficiente. Fissare un intorno fisso nel quale cercare il bianco di riferimento significa fissare il comportamento delle frequenze. L'obiettivo di Retinex è quello di campionare l'intera scena: si tratta della realizzazione di un modello qualitativo. Si fa finta che l'input sia la realtà e si calcola l'apparenza del pixel rispetto al bianco di riferimento più vicino a quel pixel.

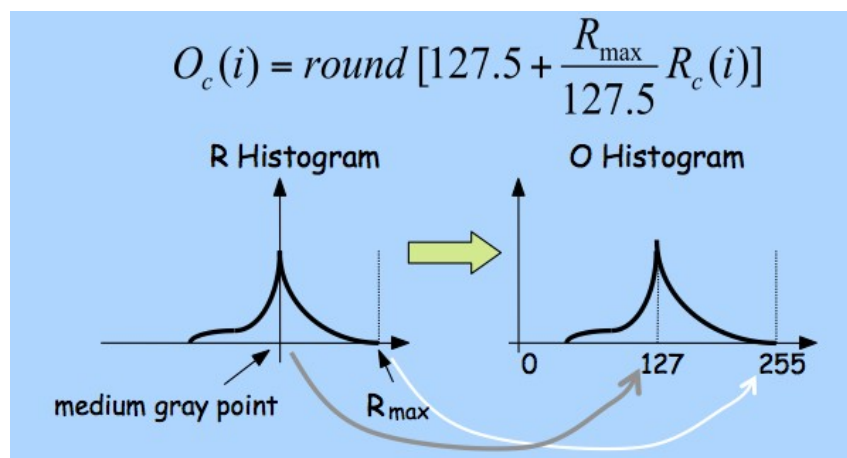
Automatic Color Equalization (ACE). Se in Retinex si operava un rapporto tra lunghezze e distanze, in ACE si compie una differenza così che conti più la distanza tra pixel che il cammino per raggiungerne uno.

ACE prende un'immagine e compie una computazione in base a sottrazioni precedenti: si prende il pixel di una matrice immagine e si sottrae il valore del pixel al valore di un pixel lì vicino, ciò che si ottiene va sommato per tutte le differenze possibili; il tutto (perchè solo la differenza non basta) va amplificato mediante un certo procedimento così che un pixel molto diverso da un suo vicino vada inteso come più importante (= più pesante): contano le distanze meglio se lineari e non quadratiche.

Il valore finale di un pixel è dato dalla differenza tra gli altri pixel pesata rispetto la distanza: si amplifica così il range d'azione. Il valore d'ingresso va da -255 a 255 (da intendere comunque da 0 a 1); più il grafico della funzione si alza più si stringono le differenze per le quali questi valori hanno significato (si avrà una funzione molto alta per pixel vicini e decisamente più bassa per pixel lontani, idem per il chiaro e lo scuro).

$$R_c(i) = \sum_{j \in \text{subset}, j \neq i} \frac{s(I_c(i) - I_c(j))}{d(i, j)}$$

In Retinex la località è data dai cammini mentre in ACE dalla distanza tra pixel chiari e scuri. Retinex godeva di un meccanismo di reset che lo rendeva non lineare mentre il comportamento di ACE è definito dalla funzione S. Con tutti i numeri sommati, se computati bene, si ottiene una matrice immagine finale con valori decimali per ogni pixel (dati distorti frutto di diversi passaggi che possono essere sia positivi che negativi, dipende se più chiari o più scuri), allora ci si aspetta un istogramma con estremi difficilmente determinabili facilmente centrato sullo zero: il bilanciamento sarà dato dalla distribuzione locale (il tutto va preso in considerazione nell'intervallo 0-255). Lo si fa scalando linearmente oppure ci si disinteressa del valore relativo al dato più scuro guardando dove va a collocarsi il dato più chiaro che, anche se non corrisponde al valore d'ingresso, è sicuramente quello più attorniato da valori scuri.



STRESS. Si prende un punto/pixel in esame e si cercano a caso tutti i pixel ad esso più vicini; quando si trova il più chiaro diventa il “bianco” di riferimento ed il più scuro diventa automaticamente “nero”, allora dopo tutti gli altri colori verranno adattati di conseguenza. Si fa una interpolazione lineare tra i due valori di riferimento. I punti cardine si calcolano per ogni pixel ma possono coincidere, in questa maniera espando il valore locale.

Si tratta di un algoritmo presente anche nel software GIMP 2.8 (C2G → Color to Gray), quando lo si usa, poiché un po' pesante, è meglio utilizzarlo senza usare l'anteprima. I valori di default per STRESS sono molto bassi i quali danno risultati molto rumorosi; sono tutti procedimenti locali, se fossero globali l'intera idea non funzionerebbe molto bene.

L'adattamento al colore del sistema visivo umano non è completo: in una

immagine normalizzata i punti neutri vengono tirati fuori; idem per quanto riguarda l'aggiustamento della dinamica. Se l'immagine in esame possiede una dominante cromatica questi metodi non la conservano. Il contrasto è molto legato alla percezione e non c'è sostanzialmente una misura oggettiva: locale e globale si aggiustano rispetto alle modifiche di volta in volta.

In Retinex si ricalcola il pixel rispetto ad un nuovo bianco locale; il valore di un pixel è visto sempre in relazione ad un bianco teorico, rifacendo tutti i conti d'accapo se il nuovo bianco è più chiaro; un'immagine trattata con SCA si schiarisce (tutte intraprendono un comportamento tipo white patch ma non tutti possono comportarsi in maniera gray world). Se si filtra un'immagine con pochi colori (tipo .gif) il numero di colori finale non rimane lo stesso (dovuto alla scalatura, così il numero di colori aumenta).

Si compie una dequantizzazione spaziale dovuta alla posizione spaziale del bit (una cosa simile accade anche nel nostro cervello). Se si filtrano le illusioni ottiche ci si aspetta una reazione qualitativa simile a quella della percezione. Se si filtra un risultato ciò che si ottiene non si discosta di molto dal risultato precedentemente ottenuto.

Sfruttando SCA si possono fare molte cose, come per esempio la ricerca di una precisa immagine all'interno di un database di immagini (allora deve minimizzare la differenza, solitamente si fa col colore) e con SCA, che è un normalizzatore cromatico molto potente, si prefiltra l'immagine.

Una color matching function serve a calcolare quanto due valori siano simili: il colore spettrale al quale si associa una tripletta preservando il metamerismo (allora le due triplette sono uguali se i colori in esame sembrano uguali); se la informazione spettrale è diversa ma appare uguale anche la tripletta appare uguale. Se si prende un'immagine multispettrale, 3 CMF diverse e 3 immagini ad alta dinamica → tone mapper puntuale (lavoro globale) e mappato in modo spaziale; calcolo la differenza tra le immagini mappate puntualmente con le 3 CMF e quanto le CMF pesano se la mappatura è spaziale. La mappatura puntuale offre delle differenze tra CMF risolvendo delle differenze tra colori. La mappatura spaziale normalizza e rende le CMF meno importanti → coni diversi ma contano le differenze nella scena.

Disclaimer. Si tratta delle trascrizioni di appunti presi durante le lezioni del corso di **Fondamenti di Imaging Digitale** tenuto dal prof. Rizzi presso l'*Università degli Studi di Milano* (laurea magistrale in Informatica per la Comunicazione) ad opera della *dott.ssa Farinelli Agnese*. Poiché si tratta di appunti trascritti potrebbero esserci diversi errori concettuali, passaggi fraintesi ed errori di battitura. L'autore non si accolla l'onere di garantire la veridicità dei suddetti appunti né è perseguibile in alcuna maniera qualora studiarli non sortisca l'effetto desiderato.

Puoi scaricare questi appunti presso l'indirizzo (www.thalionwen.altervista.org) e diffonderli come meglio credi, stampandoli o inviandoli a tutti coloro che ne desiderano una copia. Puoi utilizzare questo lavoro come punto di partenza per un'opera derivata, aggiungendone nuove parti oppure togliendone a piacimento, a patto che l'opera derivata venga condivisa alla stessa maniera dell'opera originale. Non attribuirli la paternità di quest'opera spacciandola come tua né è permesso scambiarla con lo scopo di ottenere denaro.



<http://creativecommons.org/licenses/by-nc-sa/3.0/it/deed.it>
<http://creativecommons.org/licenses/by-nc-sa/3.0/it/legalcode>