



Universitat Autònoma de Barcelona

Facultat de Ciències.

Departament d'Informàtica

ESTUDI DEL COMPORTAMENT DE COMPONENTS
CROMÀTICS. APORTACIÓ A LA CORRECCIÓ DELS ERRORS
PROVOCATS PER LA SATURACIÓ DELS CCDs

Memòria presentada per Jordi Regincós i Isern
a la Universitat Autònoma de Barcelona amb la
finalitat d'obtenir el grau de **doctor en Ciències**
(Informàtica)

Director: Joan Batlle i Grabulosa

Bellaterra, novembre de 1996.



Facultat de Ciències.
Departament d'Informàtica

ESTUDI DEL COMPORTAMENT DE COMPONENTS
CROMÀTICS. APORTACIÓ A LA CORRECCIÓ DELS ERRORS
PROVOCATS PER LA SATURACIÓ DELS CCDs

Memòria presentada per Jordi Regincós i Isern
a la Universitat Autònoma de Barcelona amb la
finalitat d'obtenir el grau de **doctor en Ciències**
(Informàtica)

Director: Joan Batlle i Grabulosa

Bellaterra, novembre de 1996.

Agraïments

Durant aquests darrers anys he tingut la gran sort de treballar amb tota una sèrie de persones que sempre que ha calgut m'han ajudat, d'una manera o altra, a tirar aquest projecte endavant. No és senzill agrair només utilitzant paraules tot el suport rebut, però ho intentaré...

- D'entrada, no em queda més remei que agrair a en Joan Batlle, director d'aquesta tesi, tot el suport que m'ha donat durant tot el temps que ha durat aquest treball, i, especialment, per haver-me ofert la seva amistat des d'aquell gener de 1994 que vaig *aterrissar* al seu grup. Una menció especial mereix l'agost del 96, amb les hores compartides, el cinema DEIA, ...
- Al meu amic Enric Martí, pels anys que vàrem compartir despatx, pel seu suport tot aquest temps i per haver acceptat ser el meu tutor per aquesta tesi.
- Un agraïment especial per en Juanjo Villanueva, que va permetre'm introduir-me en el món universitari en acceptar-me al seu grup després d'acabar la carrera, i per totes les facilitats que m'ha donat tots aquests anys.
- A en Xavier Pueyo, per haver facilitat el meu trasllat a Girona i haver-me acollit ben cordialment al seu grup a la UdG... i per tots els tallats compartits a ca l'Enriqueta aquests dos darrers anys.
- A tota la gent del Grup de Visió de Girona per la seva amistat. Els dies de treball al laboratori costaran d'oblidar, doncs si un és un animal, l'altra encara ho és més ... (incloint-hi al *jefe*). Gràcies, Xevi Cufí, Toni Fernández, Pep Forest, Rafa García, Joan Martí, Lluís Pacheco, Quim Salvi, Josep Tomás i a tota la gent que ha passat per allà.
- A en Joan Surrell, primer company de despatx a Girona, per haver-me acceptat sense cap mena de problema i per haver-me posat al dia de la "realitat politècnica", i per totes les estones que hem anat compartint.

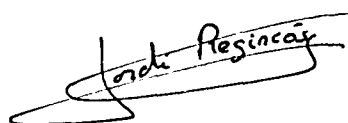
- He d'agrair també a tots els companys del departament d'Informàtica i Matemàtica Aplicada a la UdG l'ajuda que m'han ofert sempre que ha fet falta. En Fede Pérez i en Nacho Martín pels intercanvis d'informacions de L^AT_EX i altres programes de domini públic que hem mantingut, en Miguel Sainz per haver-me solucionat alguns dubtes matemàtics i, en general, a tota la gent del grup de gràfics que encara no he citat: Gonzalo Besuyeski, Imma Boada, Imma Busquet, Francesc Castro, Narcís Coll, David Fíguls, Carles Hernández, Pep Suy, Mateu Sbert i Antoni Sellarés. Sense oblidar a la resta de companys de la secció: Esteve Acebo, Pep Bosch, Miquel Feixas, Jaume Rigau, Pep Soler, L. Valdés i Josep Vidal i als companys de la resta de seccions del departament.
- Pel que fa al suport tècnic i administratiu aquests anys a Girona, em cal agrair a l'Albert Vergés i a la Mònica Arnau el seu esforç per tal que l'entorn informàtic del departament estigui sempre a punt, i a la Mercè Bautista per tota l'ajuda administrativa rebuda.
- Finalment, i pel que respecta a la meva etapa gironina, em resta agrair al rectorat de la Universitat de Girona per haver-me concedit un rebaix docent el curs 1995-96 dins del programa d'ajuda a la recerca d'aquesta universitat.

Mirant cap a la meva etapa a la UAB, on va començar tot, em venen al cap molts i molt bons records.

- Un record especial per en Xavier Binefa, amb el qual vaig compartir la meva etapa d'*exili* parisenc allà pel 1991, i d'on va sorgir una franca amistat.
- Un record per la gent que hi havia al grup quan jo vaig arribar-hi el 89 i pels consells i ajudes rebudes des del primer moment: Javier Sánchez (en especial pel seu desenvolupament del ViLi), en Joan Serrat i en Jordi Vitrià.
- Als companys més o menys de quinta per tots els patiments compartits (i les alegries també, eh?): en Ramon Baldrich, en Josep Lladós, en Felipe Lumbreras (gran jugador de bàdminton), en Narcís Parés, l'FX i en JA Pérez, en Quim Piera, en Xavi Roca, en Pep Massa, en Ricardo Toledo, i en Maria Vanrell.
- A la resta de companys que van anar arribant a la unitat: l'Antonio López, l'Esperanza López, l'Aleix Martínez, la M.Carmen Molina, la Petia Radeva, en Juan A. Rodríguez, la Gemma Sánchez, en Joan Ramon Serra i l'Ernest Valveny.

- En el terreny de suport tècnic i administratiu, cal mencionar el suport d'en Joan Masoliver i de l'Agustín Collado pel que fa a *les màquines* i el suport de la Carmen Ramírez, de la Pilar Villa, de la Carmen Alonso, de la Cristina Minguet i d'en Toni Ventura a la part administrativa.
- No puc oblidar-me de la meva etapa a la direcció de l'Escola d'Informàtica a Sabadell, i agrair des d'aquí tot el suport rebut per tots els companys, especialment per la Lola Rexachs (una gran directora/degana de centre), en Miquel Àngel Senar, en Javi Serrano, en Jordi Pons i la Montserrat Carné.
- I encara que l'hagi deixat pel final, per mi es el més important: agrair tot el suport rebut per part de tota la meva família en general (Anna Maria, Angeleta, Martí, Neus, Maria, Jaume, Àngela, Montse, Andreu, Teresa, Albert, Núria, Imma, Jordi C, Laia, Andreuet i Emma) i, en concret, l'assessorament ortogràfic/sintàctic/gramàtic de la Teresa, el suport tècnic i logístic proporcionat per en Jaume en el tema fotogràfic i, especialment, el suport en tots els aspectes de l'Anna Maria.

Espero no haver-me deixat a ningú. Si algú es troba a faltar en aquesta llista, li demano disculpes; que agafi alguna cosa que escrigui i que s'apunti en l'espai lliure que queda.



Jordi Regincós i Isern
Girona, novembre del 1996

Contingut

Agraïments	iii
1 Introducció	1
1.1 Descripció de les idees fonamentals d'aquesta tesi	1
1.2 Esquema de la tesi	2
2 Estudi del color	4
2.1 Aspectes fisiològics de la visió en color	4
2.1.1 Estructura de l'ull humà	4
2.1.2 La retina	5
2.1.3 Procés de percepció	6
2.2 Colorimetria	8
2.2.1 Espai tristimulus	8
2.2.2 Sistema Colorimètric de la CIE	12
2.3 Espais de color	14
2.3.1 Evolució històrica	14
2.3.2 Diagrama cromàtic de la CIE 1931	20
2.3.3 Diagrames cromàtics uniformes definits per la CIE	22
2.3.4 Espais de color uniformes definits per la CIE	23
2.3.5 Espais depenents dels dispositius	26
2.3.6 Models basats en el T_0 , la Saturació i la Intensitat	28
3 Treballs relacionats	34
3.1 Introducció	34
3.2 Il·luminació	35
3.2.1 Materials i llum	36
3.2.2 Model de la il·luminació global	37
3.3 Constància del color	39

3.3.1	Models basats en l'aberració cromàtica	40
3.4	Extracció de contorns i segmentació utilitzant el color	41
3.4.1	Extracció de contorns	41
3.4.2	Segmentació	46
3.5	Model de la reflexió dicromàtica	53
3.6	Indexació en Bases de Dades d'Imatges	57
3.7	Conclusions	60
4	Estudi del problema de la saturació dels CCDs	62
4.1	Introducció	62
4.2	Variació dels components cromàtics	64
4.2.1	La carta de colors <i>The Macbeth COLORCHECKER</i>	65
4.2.2	Gràfiques del comportament dels components cromàtics	66
4.3	Exemples pràctics	97
4.3.1	Colors en superfícies planes	97
4.3.2	Colors en objectes	113
5	Proposta de correcció del color	122
5.1	Introducció	122
5.2	Mètode proposat	124
5.2.1	Caracterització del color	124
5.2.2	Recuperació del color	130
5.3	Components de color a les que es pot aplicar el mètode proposat	137
5.3.1	Taules resum	150
5.4	Precissió del mètode	154
6	Validació del mètode de correcció del color i resultats experimentals	156
6.1	Implementació del sistema	156
6.1.1	El hardware utilitzat pel processat d'imatges	156
6.1.2	El sistema sencer	158
6.2	Resultats obtinguts	160
6.2.1	Conjunt d'imatges d'un <i>tren</i>	161
6.2.2	Conjunt d'imatges d'unes <i>pilotes</i>	182
6.2.3	Conjunt d'imatges d'uns <i>cables elèctrics</i>	188
6.2.4	Conjunt d'imatges d'uns <i>cotxes</i>	195

7 Conclusions	200
7.1 Conclusions	200
7.2 Futures línies de recerca	202
7.3 Publicacions fruit d'aquesta tesi	202
Referències bibliogràfiques	204

Capítol 1

Introducció

1.1 Descripció de les idees fonamentals d'aquesta tesi

El color està esdevenint, cada vegada més, una disciplina molt important en el camp de la visió per computador. És conegut que molts dels problemes d'anàlisi d'imatges per computador són resolubles sense utilitzar el color, però també és inqüestionable que aquest ens pot aportar informació addicional per tal de resoldre més eficientment aquests problemes. D'altra banda, hi ha aplicacions en les quals la utilització del color és imprescindible, especialment en l'anàlisi d'escenes naturals o en sistemes de control de qualitat.

Un dels grans problemes dels sistemes de visió és el de l'adquisició de les imatges a processar, degut al gran nombre de paràmetres que hi intervenen. Molts d'aquest paràmetres no sempre són controlables. Dos d'aquests paràmetres que afecten el procés d'adquisició són l'il·luminant i el dispositiu d'adquisició.

Un dels dispositius més extesos per a la captura d'imatges és la càmera CCD. La resposta dels CCDs és directament proporcional a la quantitat de llum incident per cada píxel. En les càmeres en color, la resposta ve donada pel que es coneix com un valor tristímulus, que correspon a un vector dins de l'espai RGB^* . És a dir, cada color s'expressa com una combinació de vermell, verd i blau.

La resposta de les càmeres no és infinita, com a dispositius físics que són, sinó que està acotada per un valor màxim. Anomenem rang dinàmic d'una càmera al rang de valors que pot donar aquesta. Quan per un determinat píxel la quantitat de llum incident supera el màxim permès, la resposta de la càmera queda retallada a aquest màxim i, per tant, el valor d'aquest píxel no es correspon amb la realitat. Direm que el sensor ha quedat

*Per ser coherents amb la bibliografia, en aquesta tesi utilitzarem les inicials RGB per referir-nos als colors Vermell (*Red*), Verd (*Green*) i Blau (*Blue*).

saturat en aquest píxel. En el cas de càmeres de grisos, aquests píxels seran de color blanc i no es pot fer res per intentar recuperar el valor correcte (sempre que no es tingui un coneixement molt precís de l'escena i de l'il·luminant)

En el cas de càmeres en color el problema és diferent: els tres components són independents i, per un determinat píxel podem tenir un, dos o tres components saturats. Aquest darrer cas seria el cas equivalent al de les càmeres de grisos: correspon al color blanc, i poca cosa podem fer per arreglar-ho. Fixem-nos, però, que quan hi ha dos components saturats, l'altre conté informació correcta i, quan només hi ha un component saturat, els altres dos contenen informació correcta. A partir d'aquesta observació és possible plantejar-se la possibilitat de recuperar la informació cromàtica dels píxels amb un o dos canals saturats. Quan algun dels tres components del color queda saturat direm que s'ha produït un retall en el color. Aquest retall modifica el color, de forma que els algoritmes que es basen en la informació cromàtica dels píxels no funcionaran correctament.

1.2 Esquema de la tesi

En aquesta tesi es presenta un estudi de com afecta la saturació dels CCDs a la informació cromàtica dels píxels en diferents espais de colors, i es proposa un sistema per corregir aquesta informació. Aquest sistema es pot utilitzar com a preprocessat per tota una família de treballs de visió artificial que utilitzen el color.

En el capítol 2 es donen unes pinzellades d'estudis sobre el color: es comença per la percepció del color, continuant per unes nocions de colorimetria i acabant amb l'evolució històrica dels espais de color, des dels primers treballs per intentar entendre i ordenar el color a finals del segle XVI fins als treballs normatius de la CIE i els models "computacionals" de color que s'estan utilitzant avui en dia.

El capítol 3 conté una relació de treballs relacionats amb aquesta tesi: models d'il·luminació, el model de la reflexió dicromàtica i un seguit de treballs de segmentació i detecció de contorns basats en el color que podrien utilitzar el nostre sistema en una etapa de preprocessat.

En el capítol 4 es presenta el problema que volem resoldre a partir d'alguns exemples. Es comença per una descripció teòrica de l'efecte de la saturació dels CCDs en els components cromàtics i s'acaba amb exemples pràctics de com pot afectar aquest efecte a l'hora de processar les imatges.

El capítol 5 conté el desenvolupament del mètode proposat per la correcció del color. Al final hi ha un estudi dels components cromàtics que han estat utilitzats per diferents

autors aquests últims anys per veure a quin d'aquests es podria aplicar el mètode proposat.

En el capítol 6 es presenta com ha estat implementat el sistema i es presenten i comenten els resultats obtinguts en diferents escenes i condicions.

Finalment, al capítol 7 es presenten les conclusions d'aquest treball, així com possibles vies que queden obertes i en les que es podria continuar treballant.

Capítol 2

Estudi del color

En aquest capítol introductori, oferim una visió de com s'ha anat tractant el color històricament, fins arribar a la seva utilització en la visió per computador. Es comença per donar alguns conceptes de la fisiologia de la visió en color, es continua amb unes nocions de colorimetria i s'acaba amb l'evolució dels espais de color que s'han anat utilitzant al llarg de la història fins arribar als espais utilitzats en la generació d'imatges per computador i en la visió artificial.

2.1 Aspectes fisiològics de la visió en color

Com a introducció a la visió per computador, donem unes pinzellades de com funciona un dels millors sensors: l'ull humà. Dues referències interessants sobre aquest tema són el llibre escrit per Fisher i Firshein [FF87], centrat en el tema de la intel·ligència artificial, i el llibre escrit per Hunt [Hun91] sobre la mesura del color.

2.1.1 Estructura de l'ull humà

La figura 2.1 mostra esquemàticament l'estructura d'un ull. L'ull disposa, com tot sensor, d'una òptica, protegida per la còrnea, i de fotoreceptors, situats a la retina. L'encarregat de l'obertura del sensor és l'iris, que disposa d'un forat circular al centre (la pupilla) per on entra la llum a l'ull, el diàmetre del qual varia entre 2mm i 8mm en funció de la intensitat de la llum incident. La llum que penetra per la pupilla arriba a la retina, que recobreix pràcticament tot l'interior de l'ull, oferint un gran camp de visió.

2.1.2 La retina

Tot i que la retina cobreix pràcticament tot l'interior de l'ull, només ofereix visió en color en un angle de 40° al voltant de l'eix visual, que no coincideix amb l'eix òptic, que està desviat 4° . La zona on la visió és més precisa correspon a una regió d'aproximadament un grau i mig al voltant de l'eix visual. Aquesta regió rep el nom de fovea (vegeu la figura 2.1). A la superfície de la retina s'hi troben dos tipus de cèl·lules recobertes de fotoreceptors anomenades cons i bastons*. La distribució d'aquests dos tipus de cèl·lules no és constant: a la foveola (part de la fovea que correspon a una amplitud d' 1°) només hi ha cons, i fora d'aquesta hi trobem cons i bastons, amb una proporció que varia uniformement fins arribar a regions més enllà dels 40° on només hi ha bastons. Com a curiositat, en un ull humà hi ha aproximadament 6 milions de cons i 120 milions de bastons.

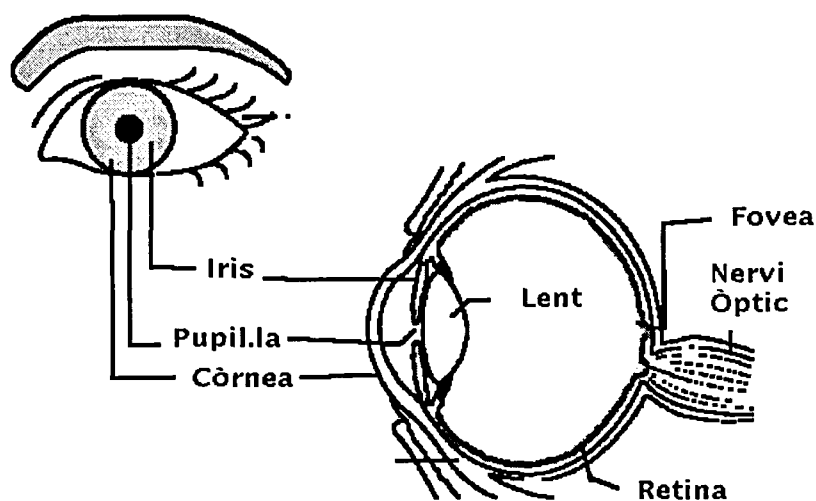


Figura 2.1: Esquema de l'estructura de l'ull.

Aquests receptors estan connectats al nervi òptic a través d'una xarxa de fibres nervioses, la densitat de la qual depèn de la regió de la retina on ens trobem. A la foveola hi ha pràcticament una fibra nerviosa per cada con i, a mesura que ens allunyem de l'eix

*L'origen dels noms *bastons* i *cons* està en les característiques morfològiques dels dos tipus de fotoreceptors.

visual, aquesta proporció decreix fins arribar a proporcions de centenars de cons i bastons per cada fibra nerviosa.

2.1.3 Procés de percepció

Els bastons són els encarregats de la visió monocromàtica en baixos nivells d'il·luminació (visió *escotòpica*), i els cons s'encarreguen de la visió en color amb nivells d'il·luminació normals (visió *fotòpica*). Els bastons són 500 vegades més sensibles a la llum que els cons. En la transició entre la visió escotòpica i fotòpica, hi ha la visió *mesòpica*, en la que hi intervenen cons i bastons.

Els bastons són més sensibles a les longituds d'ona de la part blava i verda de l'espectre. Pel que fa als cons, no tots tenen la mateixa resposta i s'ha establert que n'hi ha de tres tipus, en funció de la zona de l'espectre on tenen una millor sensibilitat. Habitualment se'ls identifica com cons sensibles al blau, al verd i al vermell, tot i que aquests últims la màxima sensibilitat la tenen al taronja. La figura 2.2 mostra les corbes que representen la sensibilitat dels quatre tipus de receptors; la corba amb línia discontinua representa els bastons, l'etiquetada β representa els cons amb més sensibilitat al blau, la γ correspon als cons amb més sensibilitat al verd i la ρ als que tenen una resposta més alta al vermell. El fet d'escollir ρ, γ, β és per distingir-ho de les funcions R, G, B usades més endavant per representar els components vermells, verds i blaus d'estímuls lluminosos. La quantitat de cons de cada tipus no és la mateixa; s'estima que la proporció $\rho : \gamma : \beta$ és de 40:20:1.

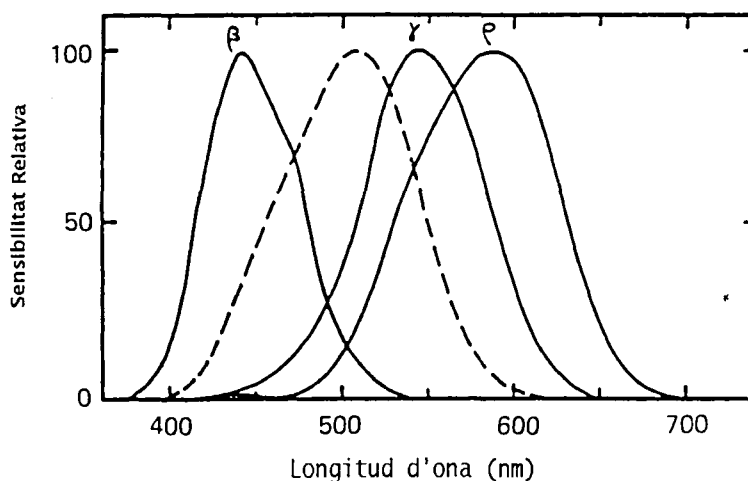


Figura 2.2: Sensibilitat dels bastons (línia discontinua) i dels tres tipus de cons ρ, γ, β .

Així doncs, a l'ull s'hi troben quatre tipus de receptors: els bastons i els tres tipus de cons. Tot i que no es coneix exactament com es transmet la informació recollida en els quatre senyals generats pels receptors, se'n té una idea aproximada. L'esquema de la figura 2.3 ofereix una aproximació de com es transmet la informació. Tal i com es pot veure, es creu que els quatre senyals s'integren en un senyal acromàtic i dos de cromàtics. El senyal acromàtic A s'obté a partir de la suma dels senyals dels quatre receptors:

$$A = 2\rho + \gamma + \beta/20 + S \quad (2.1)$$

En l'expressió anterior S representa el senyal emès pels bastons.

Els senyals cromàtics s'obtenen a partir de la diferència entre tipus de cons:

$$\begin{aligned} C_1 &= \rho - \gamma \\ C_2 &= \gamma - \beta \\ C_3 &= \beta - \rho \end{aligned} \quad (2.2)$$

Es pot veure que aquests tres senyals C_1, C_2, C_3 són redundants i que poden ser expressats amb dos senyals:

$$C_1 = \rho - \gamma \quad (2.3)$$

$$\begin{aligned} C_2 - C_3 &= \gamma - \beta - (\beta - \rho) = \\ &= \rho + \gamma - 2\beta \end{aligned} \quad (2.4)$$

2.1.3.1 Atributs del color

Des del punt de vista de la percepció humana, es poden definir dues famílies d'atributs del color:

bàsics: brillantor (més o menys quantitat de llum), t_o (vermell, verd, groc, ...) i cromaticitat (sensació de que un color exhibeix més o menys del seu t_o).

relatius: lluminància (brillantor respecte al blanc), cromàtic (diferència amb un estímul acromàtic de la mateixa brillantor) o saturació (diferència amb un estímul acromàtic independentment de la seva brillantor)

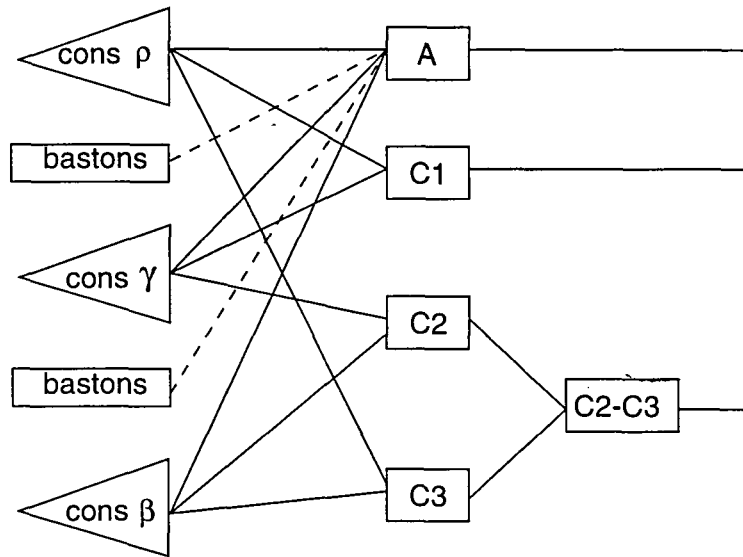


Figura 2.3: Esquema de les connexions entre els quatre tipus de receptors i les fibres nervioses del nervi òptic.

2.2 Colorimetria

La colorimetria és la ciència que s'encarrega de l'estudi de com especificar numèricament el color corresponent a un estímul físic i de com mesurar la diferència entre colors. Un llibre fonamental en aquesta disciplina és el de Wyszecki i Stiles *Color Science* [WS82]. Dins la colorimetria l'enfoc més extès és el de la generalització tricromàtica, en la que tot color s'expressa numèricament a partir d'una combinació de tres colors primaris.

2.2.1 Espai tristimulus

Tot estímul \mathbf{Q} pot expressar-se, segons la generalització tricromàtica, com un vector en un espai tridimensional, on cada dimensió correspon a un dels estímuls primaris \mathbf{R} , \mathbf{G} i \mathbf{B} :

$$\mathbf{Q} = R_Q \mathbf{R} + G_Q \mathbf{G} + B_Q \mathbf{B} \quad (2.5)$$

R_Q, G_Q, B_Q s'anomenen els valors tristimulus de \mathbf{Q} i corresponen a les quantitats necessàries d' \mathbf{R} , \mathbf{G} i \mathbf{B} per posar en correspondència \mathbf{Q} i els tres primaris.

La informació cromàtica de l'estímul \mathbf{Q} la dona la direcció del vector $\mathbf{R}_Q, \mathbf{G}_Q, \mathbf{B}_Q$, i el seu mòdul aporta informació sobre la intensitat de l'estímul. A la pràctica, la informació cromàtica sol expressar-se sobre la projecció de l'espai \mathbf{RGB} al pla $R+G+B = 1$. Aquesta

projecció rep el nom de diagrama cromàtic (r, g, b) . La figura 2.4 mostra gràficament els dos conceptes. Com es pot veure, un punt del diagrama cromàtic descriu de forma única la direcció del vector \mathbf{Q} , però no conté informació sobre la intensitat de l'estímul.

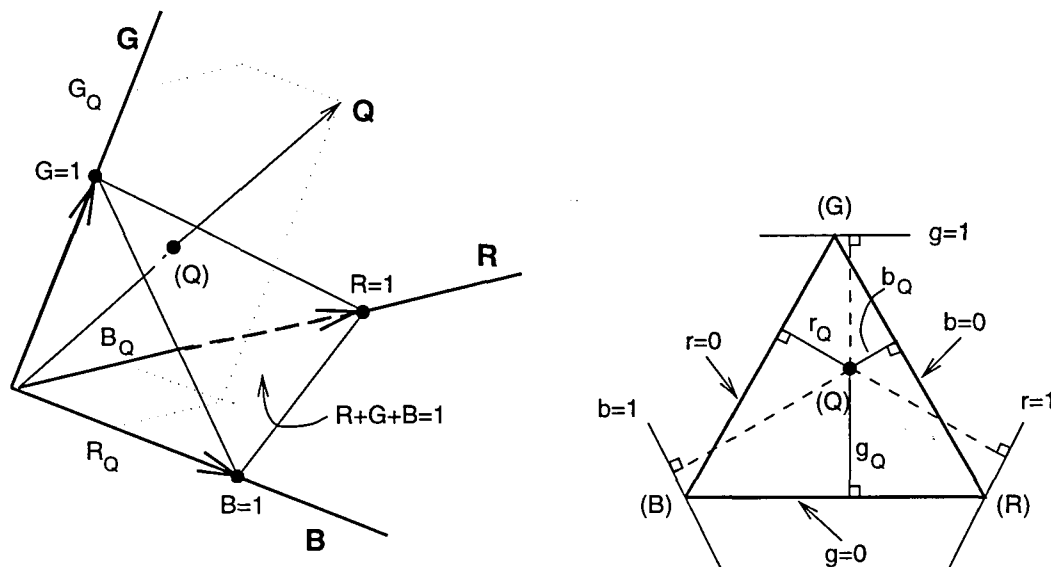


Figura 2.4: Representació de l'espai tristimulus (esquerra) i del diagrama cromàtic (r, g, b) .

Les coordenades d'un punt al diagrama cromàtic vénen expressades per les equacions:

$$\begin{aligned}
 r &= \frac{R}{R+G+B} \\
 g &= \frac{G}{R+G+B} \\
 b &= \frac{B}{R+G+B}
 \end{aligned}
 \tag{2.6}$$

De l'equació 2.6 pot deduir-se que $r + g + b = 1$ i, per tant, el diagrama cromàtic queda unívocament definit per dues coordenades. Habitualment el diagrama cromàtic es representa com un triangle rectangle, on un dels catets representa la coordenada r i l'altre representa la g .

Tot estímul \mathbf{Q} té associada una distribució espectral $\{P_\lambda d\lambda\}_Q$, i \mathbf{Q} es pot veure com la combinació additiva de n estímuls acotats cada un d'ells a un interval de l'espectre d'amplada $\Delta\lambda_i$, de tal forma que els n intervals corresponen a una partició del conjunt de longituds d'ona de la part visible de l'espectre. D'acord amb això, \mathbf{Q} es pot reescriure com:

$$\mathbf{Q} = \sum_{i=1}^n \mathbf{Q}_i = \left(\sum_{i=1}^n R_{Q_i} \right) \mathbf{R} + \left(\sum_{i=1}^n G_{Q_i} \right) \mathbf{G} + \left(\sum_{i=1}^n B_{Q_i} \right) \mathbf{B} \quad (2.7)$$

on els valors tristimulus vénen donats per:

$$R_Q = \sum_{i=1}^n R_{Q_i}, \quad G_Q = \sum_{i=1}^n G_{Q_i}, \quad B_Q = \sum_{i=1}^n B_{Q_i} \quad (2.8)$$

Portant aquesta propietat de la combinació additiva d'estímuls al límit, i suposant que els n intervals tenen tots la mateixa amplada, tenim que:

$$\int_{\lambda_a}^{\lambda_b} P_\lambda d\lambda = \lim_{\Delta\lambda \rightarrow 0} \sum_{i=1}^n P_{\lambda_i}(\Delta\lambda) \quad (2.9)$$

La quantitat $P_\lambda d\lambda$ representa el poder radiant a la longitud d'ona λ i s'anomena un estímul monocromàtic de longitud d'ona λ . Per cada estímul monocromàtic \mathbf{Q}_λ es compleix que:

$$\mathbf{Q}_\lambda = R_\lambda \mathbf{R} + G_\lambda \mathbf{G} + B_\lambda \mathbf{B} \quad (2.10)$$

Un cas particular d'estímul és aquell que té un poder radiant igual a 1 per cada longitud d'ona, i que es coneix com un estímul d'igual energia \mathbf{E} . Si es posa en correspondència un estímul d'igual energia respecte tres primaris monocromàtics s'obtenen unes funcions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ com les que es mostren a la figura 2.5 que es coneixen com *colour-matching functions*. L'equació que posa en correspondència els primaris i l'estímul d'igual energia \mathbf{E} és:

$$\mathbf{E}_\lambda = \bar{r}(\lambda) \mathbf{R} + \bar{g}(\lambda) \mathbf{G} + \bar{b}(\lambda) \mathbf{B} \quad (2.11)$$

Observant la figura 2.5 es veu que per certes logituds d'ona, \bar{r} , \bar{g} o \bar{b} prenen valors negatius. Per exemple, per $\lambda = 475$,

$$\mathbf{E}_{475} = -0.045 \mathbf{R} + 0.032 \mathbf{G} + 0.186 \mathbf{B}. \quad (2.12)$$

La interpretació física d'aquesta quantitat negativa de vermell és que no és possible expressar l'estímul \mathbf{E}_{475} com una combinació additiva dels tres primaris, però que sí que és possible "sumar-li vermell" de forma que la combinació d' \mathbf{E}_{475} i \mathbf{R} equivalgui a una combinació de \mathbf{G} i \mathbf{B} tal i com es mostra a continuació:

$$\mathbf{E}_{475} + 0.045 \mathbf{R} = 0.032 \mathbf{G} + 0.186 \mathbf{B}. \quad (2.13)$$

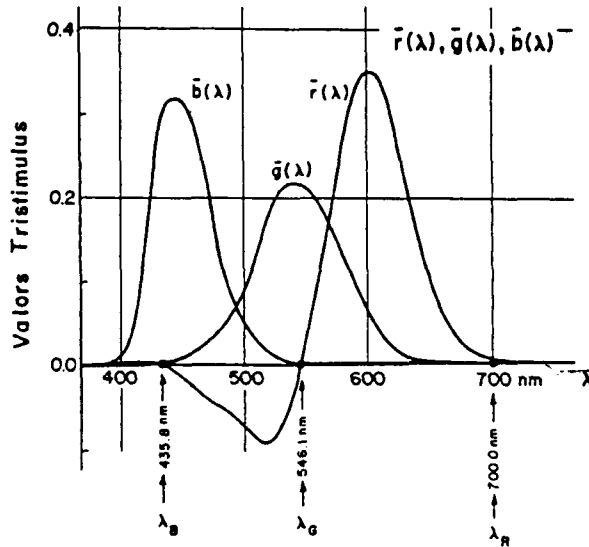


Figura 2.5: Distribució dels valors tristimulus de E_λ per tres primaris monocromàtics ($\lambda_R = 700$, $\lambda_G = 546.1$, $\lambda_B = 435.8$) en aquest cas.

Aplicant l'equació 2.6 als valors tristimulus espectrals, obtenim

$$\begin{aligned} r(\lambda) &= \frac{\bar{r}(\lambda)}{\bar{r}(\lambda) + \bar{g}(\lambda) + \bar{b}(\lambda)} \\ g(\lambda) &= \frac{\bar{g}(\lambda)}{\bar{r}(\lambda) + \bar{g}(\lambda) + \bar{b}(\lambda)} \\ b(\lambda) &= \frac{\bar{b}(\lambda)}{\bar{r}(\lambda) + \bar{g}(\lambda) + \bar{b}(\lambda)} \end{aligned} \quad (2.14)$$

on $(r(\lambda), g(\lambda), b(\lambda))$ corresponen al lloc geomètric de tots els punts tals que $r(\lambda) + g(\lambda) + b(\lambda) = 1$, i que es coneix com el lloc geomètric de l'espectre al diagrama cromàtic- (r, g) , representat a la figura 2.6.

Finalment, a partir de l'equació 2.11 podem expressar qualsevol estímul \mathbf{Q} en funció de $(\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda))$:

$$\mathbf{Q}_\lambda \equiv (P_\lambda d\lambda) \mathbf{E}_\lambda = (P_\lambda d\lambda) \bar{r}(\lambda) \mathbf{R} + (P_\lambda d\lambda) \bar{g}(\lambda) \mathbf{G} + (P_\lambda d\lambda) \bar{b}(\lambda) \mathbf{B} \quad (2.15)$$

i, en aquest cas, els valors tristimulus (R_Q, G_Q, B_Q) serien:

$$R = \int_{\lambda_a}^{\lambda_b} P_\lambda \bar{r}(\lambda) d\lambda, \quad G = \int_{\lambda_a}^{\lambda_b} P_\lambda \bar{g}(\lambda) d\lambda, \quad B = \int_{\lambda_a}^{\lambda_b} P_\lambda \bar{b}(\lambda) d\lambda, \quad (2.16)$$

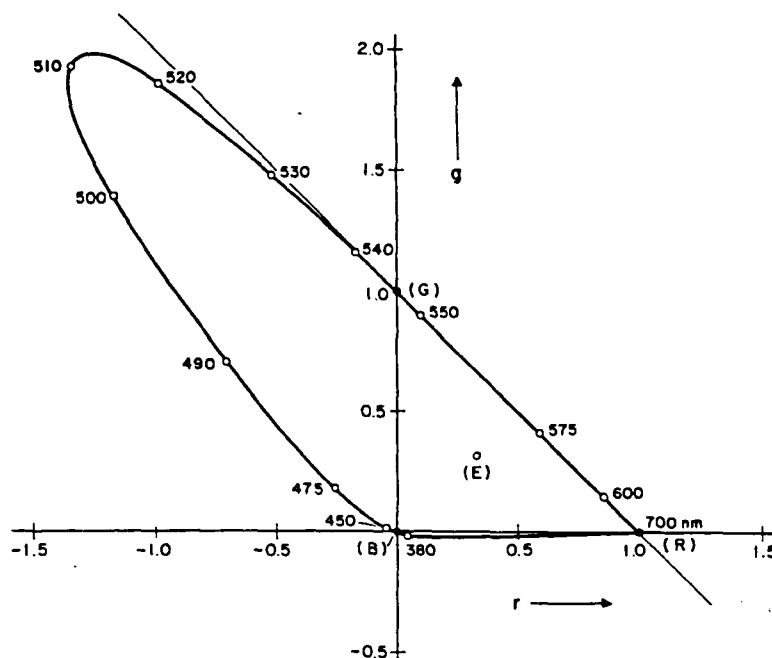


Figura 2.6: Lloc geomètric de l'espectre al diagrama cromàtic- (r, g) .

2.2.2 Sistema Colorimètric de la CIE

Un dels objectius de la colorimetria és que els resultats obtinguts siguin aplicables a qualsevol observador tricromàtic normal. Per assolir aquest objectiu es defineix el comportament d'un observador estàndard. La majoria de treballs sobre colorimetria s'han basat en la definició d'observador estàndard que va fer la CIE (*Comission Internationale de l'Éclairage*) l'any 1931 i en la definició de l'observador estàndard suplementari de l'any 1964, que s'exposen tot seguit.

2.2.2.1 Observador estàndard de la CIE

La definició d'aquest observador estàndard es basa en els treballs de Guild (1931) i Wright (1928-1929), que analitzaren la resposta de 7 i 10 observadors respectivament. Les mesures es van fer considerant camps de visió de 2° i consistien en presentar a l'observador un conjunt d'estímuls monocromàtics. La persona en qüestió havia d'anar construint, a partir de tres primaris, colors que fossin equivalents als estímuls presentats. Per cada estimul monocromàtic s'anotaven les quantitats utilitzades de cada primari. La figura 2.7 mostra esquemàticament aquest procés. Els primaris utilitzats en els

dos experiments no foren els mateixos, però per comparar els resultats es varen posar les mesures en funció de tres nous primaris monocromàtics ($\lambda_R = 700, \lambda_G = 546.1, \lambda_B = 435.8$) obtenint-se així les funcions de correspondència ($\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda)$) il·lustrades a la figura 2.5.

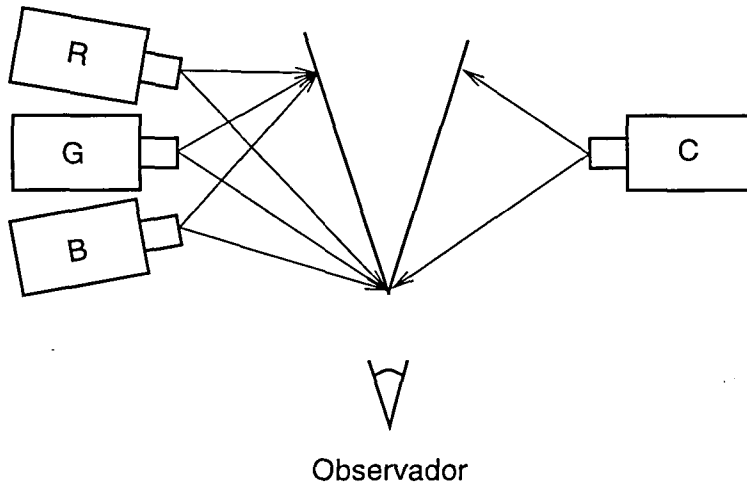


Figura 2.7: Exemple de la posada en correspondència d'un color C amb els tres primaris R, G i B.

El problema d'aquest espai **RGB** és el de l'existència de quantitats negatives en algun dels primaris: $\bar{r}(\lambda)$ és negativa en un interval dels blaus i en un dels verds (per $\lambda \in [440, 545]$), $\bar{g}(\lambda)$ és negativa en una zona de blaus (per $\lambda \in [380, 435]$), i $\bar{b}(\lambda)$ ho és en una zona de verds, grocs i taronges (per $\lambda \in [550, 650]$).

Per solucionar aquest problema, la CIE proposa un nou espai, de primaris imaginaris **X, Y i Z** que tinguin la propietat que tot color es pugui expressar com una barreja additiva de quantitats positives de cada un dels tres nous primaris. La transformació de l'espai **RGB** al nou espai **XYZ** es fa aplicant la següent transformació lineal:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (2.17)$$

Els elements de la matriu que expressa la transformació lineal han estat escollits de forma que X, Y, Z sempre siguin valors positius, i que la component Y es correspongui amb la lluminositat. Les noves funcions de correspondència així obtingudes es coneixen com a ($\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$). Aquestes funcions estan dibuixades a la figura 2.8

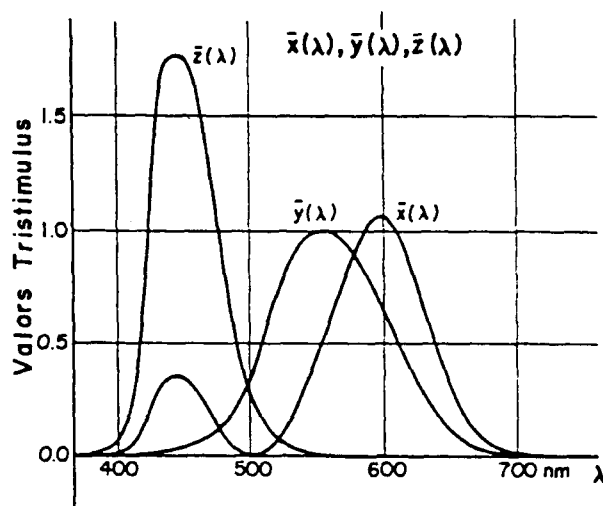


Figura 2.8: Distribució dels valors tristimulus de E_λ per als primaris imaginaris X , Y i Z .

El 1931 $(\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda))$ es definiren entre 380 i 780 nanòmetres, a intervals de 5 en 5, expressant els valors en 4 decimals. El 1971 la CIE reescriu les funcions de 1931, donant valors entre 360 i 830 nanòmetres, per intervals d'un nanòmetre.

2.2.2.2 Observador estàndard suplementari de la CIE

És una redefinició del model del 1931 basada en els treballs de Stiles i Burch (1959) i Speranskaya (1959), per camps de visió de 4° a 10° . S'obtenen també dos conjunts de funcions de correspondència equivalents a les $(\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda))$ i $(\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda))$, que es coneixen amb el nom de $(\bar{r}_{10}(\lambda), \bar{g}_{10}(\lambda), \bar{b}_{10}(\lambda))$ i $(\bar{x}_{10}(\lambda), \bar{y}_{10}(\lambda), \bar{z}_{10}(\lambda))$. Les diferències entre aquestes noves funcions i les anteriors són els següents:

- Són aplicables per camps de visió superiors a 4° .
- Les funcions es donen amb 6 decimals.

2.3 Espais de color

2.3.1 Evolució històrica

El primer model experimental amb base científica reconeguda és fruit dels treballs de Isaac Newton al Trinity College pels volts de 1660 tal i com és comentat per Norman

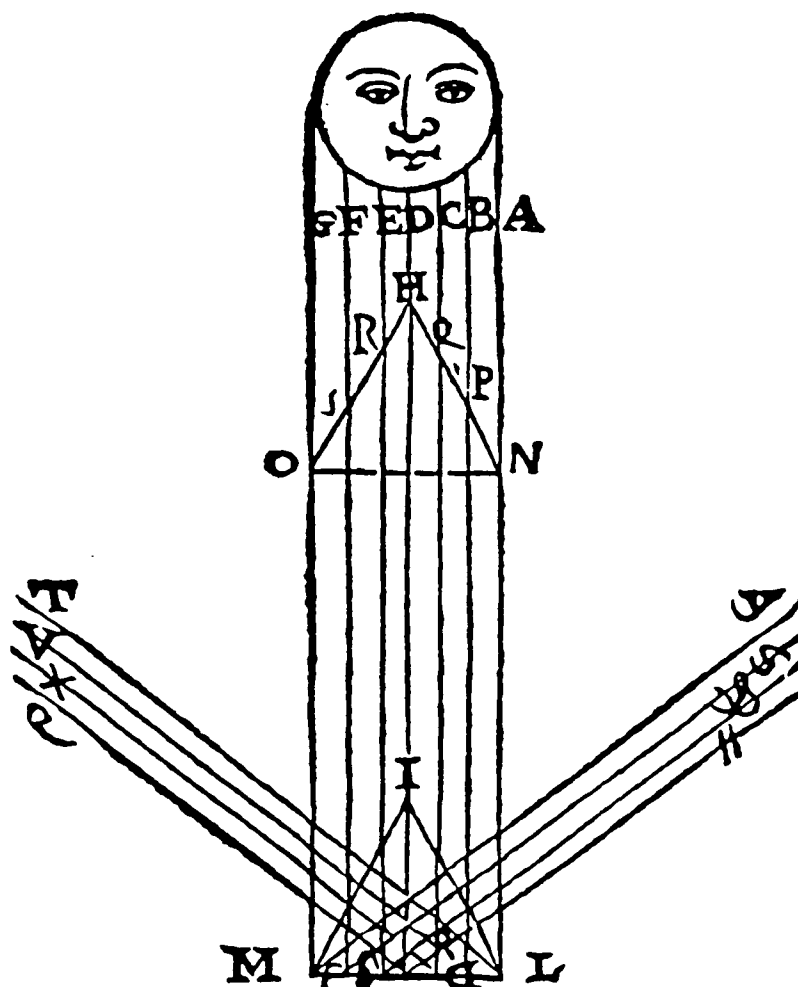


Figura 2.9: Model de color de Della Porta (1593).

[Nor90]. Abans però, tal i com explica Gerritsen al seu llibre [Ger76], a finals del segle XVI i principis del segle XVII, trobem dues representacions prèvies dels colors: la de Della Porta, basada en la refracció de la llum en un prisma i la d'Aguilonius, un monjo jesuïta belga al que també es deuen uns dels primers treballs sobre la visió estereoscòpica, basada en aspectes més *espirituals*: l'espectre variava continuament del blanc (*albus*, matinada) fins el negre (*niger*, nit), deixant el vermell al centre de l'espectre. Les figures 2.9 i 2.10 mostren aquests dos espais.

Newton, a partir de les seves observacions sobre el comportament de la llum, va suggerir una representació del color en forma de cercle, amb els set colors de l'arc de Sant Martí: vermell, taronja, groc, verd, blau, indi i violat (sembla que el fet d'utilitzar set colors



Figura 2.10: Model de color d'Aguilonius (1613).

prové de voler establir un cert paral·lisme amb les set notes musicals). L'ordenació de colors proposada per Newton està representada a la figura 2.11. La disposició dels colors en forma de cercle dóna una continuïtat perceptual a l'espectre, passant del violeta al vermell (tal i com es proposa en molts dels models de color utilitzats avui en dia).

El següent treball interessant el trobem al 1810: Goethe proposa una ordenació dels colors basada en les sensacions que evocuen. No pretén descriure físicament el fenomen del color, sinó que es centra en un estudi psicològic de la percepció dels colors. Defineix una estructura circular amb sis divisions, tot i que sovint s'expressa com un triangle equilàter (com el que es mostra a la figura 2.12): els vèrtexs del triangle corresponen als tres colors primaris vermell (imaginació), groc (raonament) i blau (enteniment). A partir de barreges d'aquests, podem obtenir la resta de colors (per exemple, entre el blau i el groc hi trobem el verd, que correspon a l'esperança).

Una mica més tard, el 1872, un físic escocès, J.C. Maxwell, identificà el vermell, el verd i el blau com a colors primaris, i defineix un triangle amb els tres primaris als vèrtexs, situant la resta de colors als costats del triangle o al seu interior. Recorrent els costats del triangle es van obtenint els colors en l'ordre proposat per Newton. Al centre del triangle hi ha el color blanc. Inicialment es proposa un rectangle equilàter (figura 2.13.a) en el qual qualsevol color es pot representar amb tres coordenades. Si el triangle es converteix a triangle rectangle (figura 2.13.b), llavors amb només dues coordenades podem especificar un color.

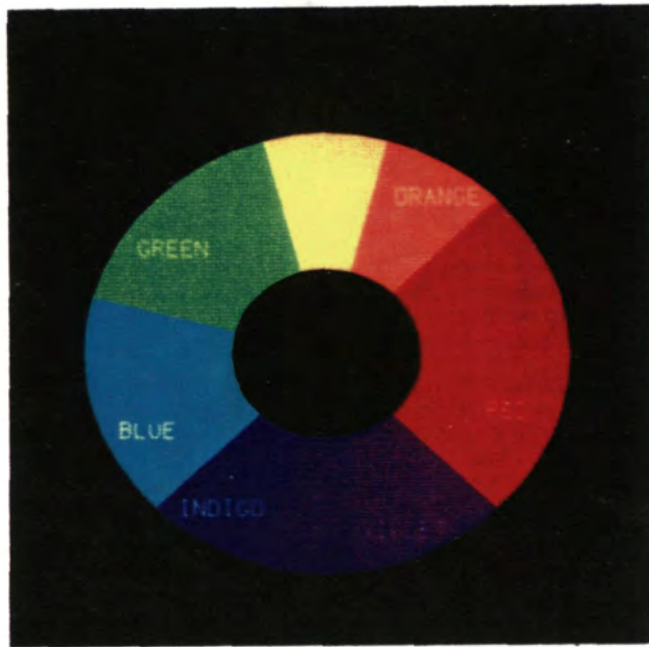


Figura 2.11: Ordenació dels colors segons les observacions de Newton.



Figura 2.12: Ordenació dels colors segons Goethe. Criteris perceptuals.

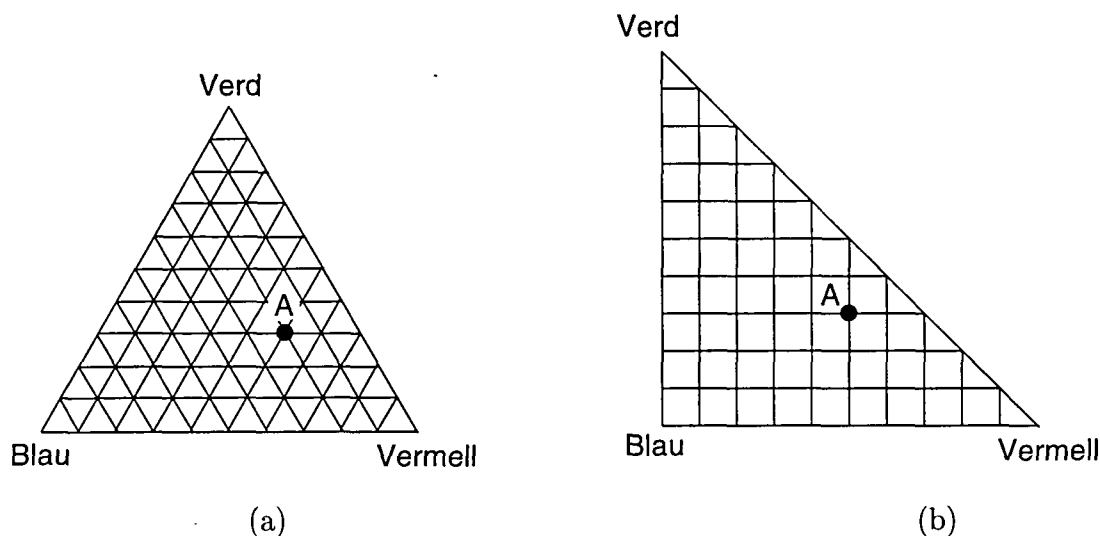


Figura 2.13: Triangle de colors de Maxwell. Versió original (a) i versió transformada (b).

Fins a principis del nostre segle no apareix un sistema tridimensional de color. Es el sistema proposat pel professor d'art Albert Munsell (1905), que organitza els colors tenint en compte tres paràmetres: to (*Hue*), cromacitat (*Chroma*) i lluminositat (*Value*), estructurats en un sistema de coordenades cilíndriques (la lluminositat correspon a l'eix, la cromacitat a la distància a l'eix i el to a la posició angular).

El sistema de Munsell defineix cinc tons bàsics: vermell, groc, verd, blau i porpra etiquetats 5R, 5Y, 5G, 5B i 5P respectivament. Entre cada un d'aquests 5 tons bàsics i ha un to intermig. Aquests cinc tons estan etiquetats com 5YR, 5GY, 5BG, 5PB i 5RP. Finalment, entre cada to bàsic i cada to intermig hi ha 10 divisions etiquetades, per exemple per l'interval 5R i 5YR, com 6R, 7R, 8R, 9R, 10R, 1YR, 2YR, 3YR i 4YR. En total, el sistema contempla 100 tons diferents.

Pel que fa a la lluminositat, considera l'eix partit en 10 regions, amb el blanc etiquetat com a 10 i el negre com a 0. La cromacitat es dóna en una escala entre 0 i 12. El valor màxim de la cromacitat depèn del valor del to i de la lluminositat. La figura 2.14 mostra dues seccions de l'espai de colors de Munsell. A l'esquerra es mostra una secció horitzontal corresponent a tots els colors representables amb una mateixa lluminositat. A la dreta es mostra una secció vertical del sòlid de colors corresponent als colors existents per dos tons constants: els rectangles de la columna central corresponent a l'eix acromàtic

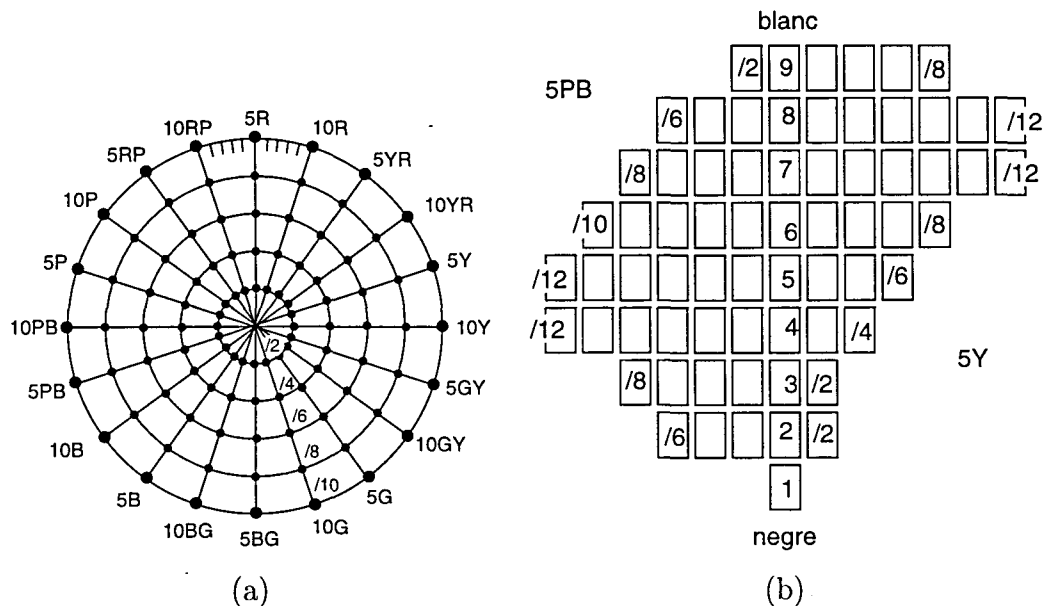


Figura 2.14: Esquema de l'espai de color de Munsell. Organització per lluminositat constant (a) i organització per dos tons constants (5Y i 5PB) (b).

(cromacitat = 0), els que hi ha a la banda esquerra de l'eix són les variacions existents del to 5PB (entre porpra i blau) i les de la dreta són les del to 5Y (groc). Es pot veure que el rang de cromacitats possibles depèn del to i de la lluminositat.

A la figura 2.15 es mostra una representació del sòlid de colors associat a l'espai de Munsell. Tot color en aquest espai ve especificat per un codi format per tres valors: HV/C (To Luminositat/Cromacitat). Per exemple, 2Y4/8 indica que és un color amb To = 2Y, amb Luminositat = 4 i amb Cromacitat=8. És a dir, que el to està entre 10YR i 5Y (vegeu figura 2.14 esquerra), que té una lluminositat una mica per sota de la meitat (la meitat seria 5) i que és un color bastant allunyat de l'eix acromàtic (la lluminositat està entre 0 i 12).

L'any 1931 es produeix un fet molt important en l'avenç dels treballs sobre el color: es defineix el sistema colorimètric de la CIE i s'adopta com a estàndard mundial de referència per tota la gent que treballa amb temes relacionats amb el color. Degut a la seva importància, aquest model serà tractat més endavant en aquest mateix capítol).

D'una forma general podríem dir que en tota aquesta evolució històrica els models semblen orientar-se cap a dues vessants diferents: d'una banda apareixen models orientats a pintors, és a dir, models basats en la barreja de pigments, i d'altra banda apareixen models orientats més al fenomen de la llum. Com a exemples dels primers trobem els

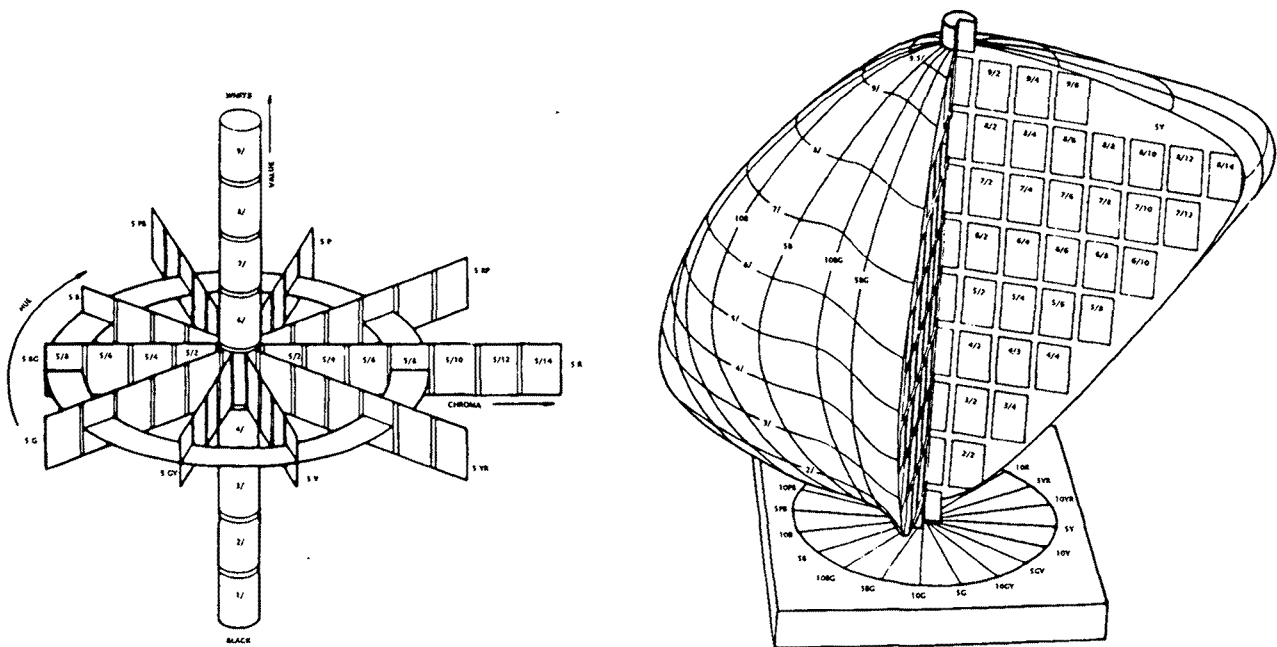


Figura 2.15: Representació del sòlid de colors obtingut a partir de l'espai de Munsell.

models de Goethe o Munsell, i com a exemples dels segons trobariem els de Newton, Maxwell, o els de la CIE.

2.3.2 Diagrama cromàtic de la CIE 1931

Com ja s'ha comentat, la informació cromàtica d'un estímul pot expressar-se amb dos paràmetres, utilitzant el que es coneix com a diagrama cromàtic. En el cas de la CIE, el diagrama cromàtic x, y es defineix com:

$$x = \frac{X}{X+Y+Z} \quad (2.18)$$

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

Fixem-nos que la z pot expressar-se com $1 - x - y$ i que, per tant, no aporta informació addicional. Expressant tots els estímuls cromàtics en funció de x i y obtenim el diagrama que es mostra a la figura 2.16. Aquest diagrama defineix totalment la informació cromàtica, però no té en compte la intensitat de l'estímul. Per això, per caracteritzar completament un color es donen les seves coordenades xyY (les coordenades sobre el diagrama cromàtic xy més el valor de la lluminància Y).

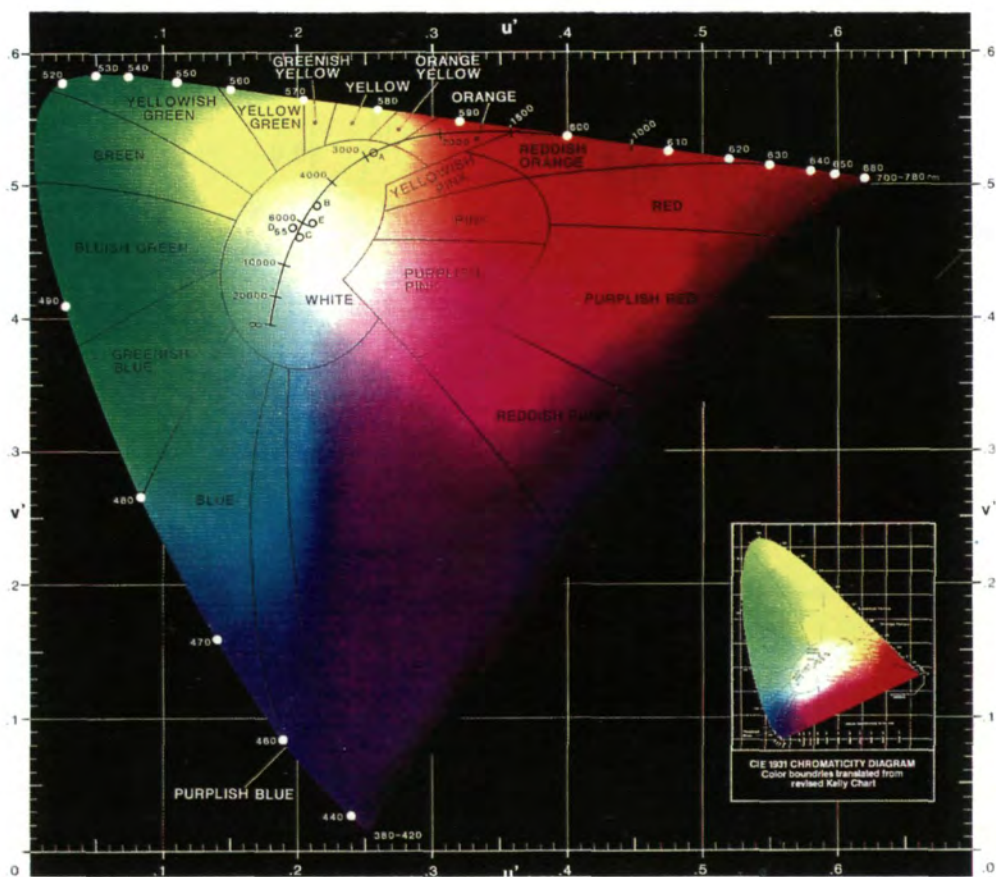


Figura 2.16: Diagrama cromàtic de la CIE u', v' (1976) i diagrama cromàtic de la CIE x, y (1931).

2.3.3 Diagrames cromàtics uniformes definits per la CIE

El diagrama cromàtic definit per la CIE el 1931 presenta el següent problema: la distància entre parelles de colors que tinguin perceptualment la mateixa diferència depèn del lloc del diagrama en el que es trobin. Així, per exemple, les distàncies són més grans a la zona dels verds i molt més curtes a la zona dels violats tal i com s'il·lustra a la figura 2.17.

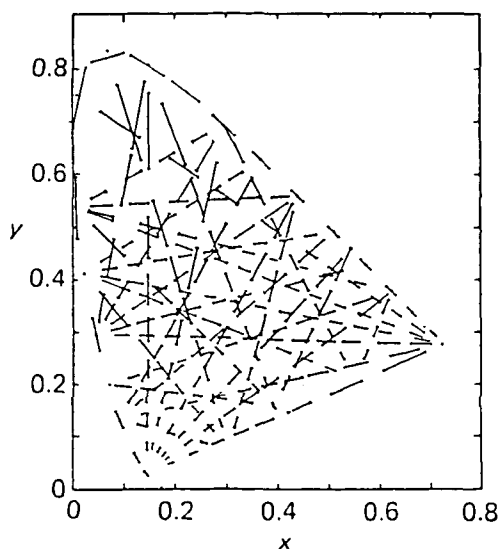


Figura 2.17: Parelles de colors amb la mateixa distància perceptual, representats sobre el diagrama cromàtic x, y .

Per reduir aquest problema, la CIE defineix el 1960 el diagrama cromàtic u, v , també conegut com diagrama cromàtic CIE 1960 UCS *Uniform Chromaticity Diagram*. Més tard, el 1976, la CIE el retoca, establint el diagrama cromàtic u', v' . La definició d' u, v en funció dels primaris X, Y, Z ve donada per:

$$u = \frac{4X}{X + 15Y + 3Z} \quad (2.19)$$

$$v = \frac{6Y}{X + 15Y + 3Z}$$

Si s'utilitzen les coordenades del diagrama cromàtic CIE-1931 obtenim una definició per u, v equivalent a l'anterior:

$$u = \frac{4x}{-2x + 12y + 3} \quad (2.20)$$

$$v = \frac{6y}{-2x + 12y + 3}$$

Aquesta transformació encara tenia el problema de concentrar els colors a la regió dels grocs, taronges i vermells. Per corregir això últim, es defineix 16 anys més tard el nou diagrama cromàtic u', v' , també conegut com a CIE 1976 UCS :

$$\begin{aligned} u' &= u \\ v' &= 3v/2 \end{aligned} \quad (2.21)$$

Aquests dos diagrames estan representats a la figura 2.18

2.3.4 Espais de color uniformes definits per la CIE

Com a pas posterior a la definició de diagrames cromàtics uniformes, la CIE fa front a la problemàtica de definir espais de color uniformes, afegint-hi la component de lluminositat. El 1964 es defineix el model $(U^*V^*W^*)$, precursor del $(L^*u^*v^*)$ (1976); any que també s'adopta el $(L^*a^*b^*)$.

2.3.4.1 CIE 1964 $(U^*V^*W^*)$

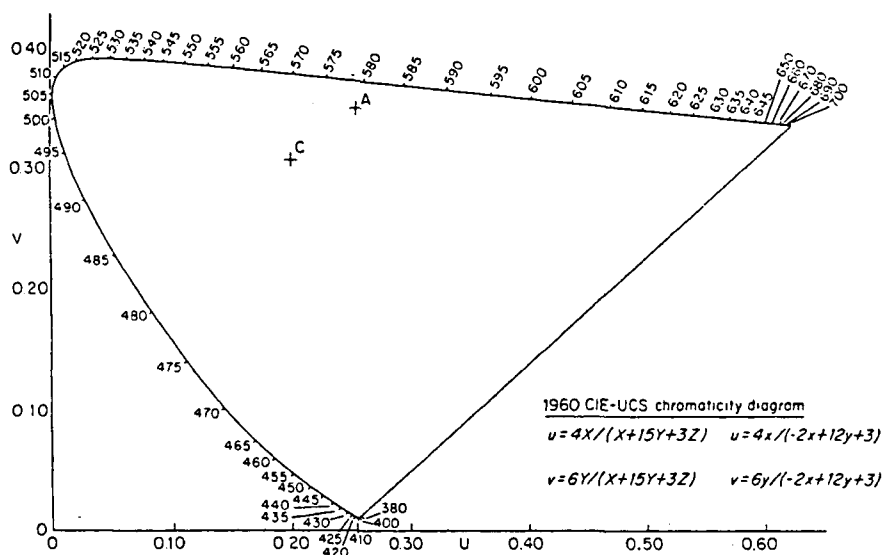
És un sòlid de colors definit a partir del diagrama cromàtic CIE 1960 UCS. La transformació que passa de l'espai XYZ a l'espai $(U^*V^*W^*)$ és:

$$\begin{aligned} U^* &= 13W(u - u_n) \\ V^* &= 13W(v - v_n) \\ W^* &= 25\sqrt[3]{Y} - 17 \text{ per } 1 \leq Y \leq 100 \end{aligned} \quad (2.22)$$

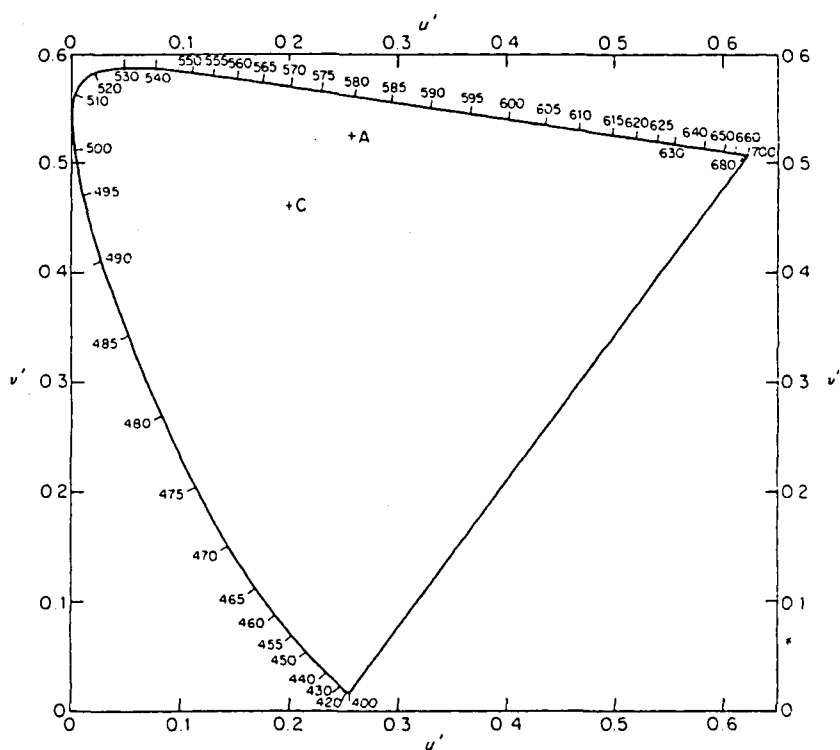
on u i v són els definits a l'equació 2.19 i u_n, v_n són les coordenades cromàtiques de l'estímul acromàtic de referència.

2.3.4.2 CIE 1976 $(L^*u^*v^*)$

El 1976, la CIE defineix la lluminositat L^* . A partir d'aquesta i del diagrama cromàtic CIE 1976 UCS, es redefineix l'espai de colors $(U^*V^*W^*)$, essent rebatejat com espai de colors CIE 1976 $(L^*u^*v^*)$, o CIELUV. Les coordenades d'aquest nou espai s'obtenen com:



(a)



(b)

Figura 2.18: Diagrama cromàtic u, v (a). Diagrama cromàtic u', v' (b). (Segons Chamberlin i Chamberlin [CC80]).

$$\begin{aligned}
L^* &= \begin{cases} 116\sqrt[3]{Y/Y_n} - 16 & \text{si } Y/Y_n > 0.008856 \\ 903\sqrt[3]{Y/Y_n} & \text{si } Y/Y_n \leq 0.008856 \end{cases} \\
u^* &= 13L^*(u' - u'_n) \\
v^* &= 13L^*(v' - v'_n)
\end{aligned} \tag{2.23}$$

on els valors de u', v' són els definits a l'equació 2.21, i u'_n, v'_n corresponen a les coordenades d'un blanc de referència.

Sobre aquestes coordenades es poden definir els atributs perceptuals de to[†] h_{uv} , cromà C_{uv}^* i saturació s_{uv}^* de la següent forma:

$$\begin{aligned}
h_{uv} &= \arctan(v^*/u^*) \\
C_{uv}^* &= \sqrt{u^{*2} + v^{*2}} \\
s_{uv}^* &= C_{uv}^*/L^*
\end{aligned} \tag{2.24}$$

2.3.4.3 CIE 1976 ($L^*a^*b^*$)

El mateix any en que defineix l'espai ($L^*u^*v^*$), la CIE defineix un altre espai de colors: el CIE 1976 ($L^*a^*b^*$). Aquest espai és més orientat als fabricants de colorants. Les definicions dels components d'aquest espai són:

$$\begin{aligned}
L^* &= \begin{cases} 116\sqrt[3]{Y/Y_n} - 16 & \text{si } Y/Y_n > 0.008856 \\ 903\sqrt[3]{Y/Y_n} & \text{si } Y/Y_n \leq 0.008856 \end{cases} \\
a^* &= 500(\sqrt[3]{X/X_n} - \sqrt[3]{Y/Y_n}) \\
b^* &= 200(\sqrt[3]{Y/Y_n} - \sqrt[3]{Z/Z_n})
\end{aligned} \tag{2.25}$$

La definició del to h_{ab} i del cromà C_{ab}^* sobre aquest espai és:

$$\begin{aligned}
h_{ab} &= \arctan(a^*/b^*) \\
C_{ab}^* &= \sqrt{a^{*2} + b^{*2}}
\end{aligned} \tag{2.26}$$

El problema de tots aquests espais de color és que cal tenir un sistema perfectament calibrat per poder expressar els valors adquirits amb els dispositius en funció d'aquests

[†]Si no es diu el contrari, en el que resta de tesi suposem que l'arctangent utilitzada per l'obtenció del to retorna un valor entre 0 i 2π (per això es té en compte la relació entre el numerador i el denominador).

sistemes de coordenades (habitualment els sensors donen tres components (r, g, b) per cada píxel) . I aquesta calibració no sempre és possible.

2.3.5 Espais depenents dels dispositius

En aquest apartat hi inclouríem tots els espais de color en els quals la posició que ocupa un color es calcula directament en funció dels valors tricromàtics rebuts des del sensor.

2.3.5.1 Espais de color equivalents als dels dispositius

Són espais que tenen el seu origen en els dispositius físics relacionats amb l'adquisició o reproducció d'imatges en color. L'espai RGB s'utilitza en sensors (càmeres, escàners) i monitors, mentre que el CMY s'utilitza per imprimir en color. L'espai RGB és un espai additiu (els colors es formen sumant al color negre els valors dels tres components) i l'espai CMY és un espai substractiu (els colors es formen restant del blanc els valors dels tres components). Per això a la literatura es parla dels primaris additius : Vermell, Verd i Blau (RGB) i dels primaris substractius: Cian, Magenta i Groc (CMY). La relació entre els dos espais ve donada per la següent expressió:

$$(R, G, B) = (1, 1, 1) - (C, M, Y) \quad (2.27)$$

L'espai de colors *CMYK* que s'utilitza en impressió és el mateix que el *CMY* afegint-hi el color negre ja que aquest és molt difícil d'obtenir a partir dels tres primaris.

Espai RGB normalitzat Amb una petita transformació sobre l'espai RGB podem aconseguir un nou espai que sigui independent a variacions uniformes en la intensitat de la font d'il·luminació:

$$\begin{aligned} N_R &= \frac{R}{R + G + B} \\ N_G &= \frac{G}{R + G + B} \\ N_B &= \frac{B}{R + G + B} \end{aligned} \quad (2.28)$$

Aquestes tres expressions anteriors són redundants perquè la suma $N_R + N_G + N_B$ sempre serà 1 i, per tant, podem posar $N_B = 1 - N_R - N_G$. Per tot això, sovint l'espai RGB normalitzat es formula com :

$$\begin{aligned}
 Y &= c_1 R + c_2 G + c_3 B \\
 T_1 &= \frac{R}{R + G + B} \\
 T_2 &= \frac{G}{R + G + B}
 \end{aligned}
 \tag{2.29}$$

on c_1, c_2, c_3 són 3 constants tals que la seva suma és 1. En l'espai representat a 2.29 les components T_1 i T_2 aporten la informació cromàtica, i la component Y aporta la informació d'intensitat, tal i com ho comenta Nevatia en el seu article [Nev77].

2.3.5.2 Models de colors oposats

Són models que expressen la informació cromàtica a partir de diferències entre els colors primaris.

Model YIQ És el model utilitzat pel format de televisió americà NTSC. La component Y aporta la informació de lluminància (informació no cromàtica), mentre que I i Q aporten la informació cromàtica. La transformació lineal que passa d'RGB a YIQ és la següent:

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.273 & -0.322 \\ 0.212 & -0.522 & 0.315 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}
 \tag{2.30}$$

Model I1-I2-I3 Correspon a un model obtingut per Ohta mitjançant mètodes estadístics. La transformació d'RGB a I1, I2 i I3 és:

$$\begin{pmatrix} I1 \\ I2 \\ I3 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & -1/2 \\ -1/4 & 1/2 & -1/4 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}
 \tag{2.31}$$

Model r-g,bl-y,w-bk És un model que vol simular el sistema visual humà i que es defineix, seguint la notació de Ballard i Brown [BB82], com:

$$\begin{pmatrix} R - G \\ Bl - Y \\ W - Bk \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}
 \tag{2.32}$$

2.3.6 Models basats en el To, la Saturació i la Intensitat

Aquest nom designa a tot un conjunt d'espais de color, utilitzats tant en el món dels gràfics per computador (vegeu, per exemple, Foley et al. [FvDFH90], Watt [Wat93] o Hearn i Baker [HB94]) com en el món de la visió per computador (entre altres, Gershon [Ger85], Healey [Hea92], Levkowitz [Lev93], Perez [PK94]). En la literatura referent al color trobem diferents variants d'aquest model, tant pel que fa a la definició de les tres components, com al nom que es dona a la component que fa referència a la informació no cromàtica (Intensitat, Valor, Lluentor)...

Aquests models es basen en intentar ordenar els colors seguint criteris de percepció. A l'hora de parlar d'aquests sistemes cal que fer referència al sistema de colors de Munsell, descrit a la pàgina 18.

Un cas especial el formen les definicions de to, saturació i intensitat definits sobre els espais de color de la CIE, ja explicats en pàgines anteriors.

2.3.6.1 Model HSV (*Hue, Saturation, Value*)

Aquest espai de color el trobem definit a l'article d'Smith [Smi78], o al llibre de Foley, Van Dam et al. [FvDFH90] i es defineix per la següent transformació sobre el cub RGB (suposant que R, G, B pertanyen a l'interval $[0, 1]$):

$$\begin{aligned}
 V &= \max\{R, G, B\} \\
 S &= \frac{\max\{R, G, B\} - \min\{R, G, B\}}{\max\{R, G, B\}} \\
 H &= \begin{cases} \pi/3(B - G)/\Delta & \text{si } R = \max\{R, G, B\} \\ \pi/3(2 + (B - R)/\Delta) & \text{si } G = \max\{R, G, B\} \\ \pi/3(4 + (R - G)/\Delta) & \text{si } B = \max\{R, G, B\} \end{cases} \\
 \Delta &= \max\{R, G, B\} - \min\{R, G, B\}
 \end{aligned} \tag{2.33}$$

De l'expressió anterior es pot veure que el sòlid de colors corresponent al model HSV és una piràmide invertida amb base hexagonal (fig. 2.19.a).

El problema que presenta (i que és comú a tota aquesta família de models) és la inestabilitat del to quan ens apropem a l'eix acromàtic: quan $\max\{R, G, B\} - \min\{R, G, B\}$ tendeix a zero, pot veure's que una petita variació pot provocar una gran variació en la component de to. En el límit (quan $\max\{R, G, B\} - \min\{R, G, B\} = 0$), el to té un valor indefinit, que correspon a un valor gris.

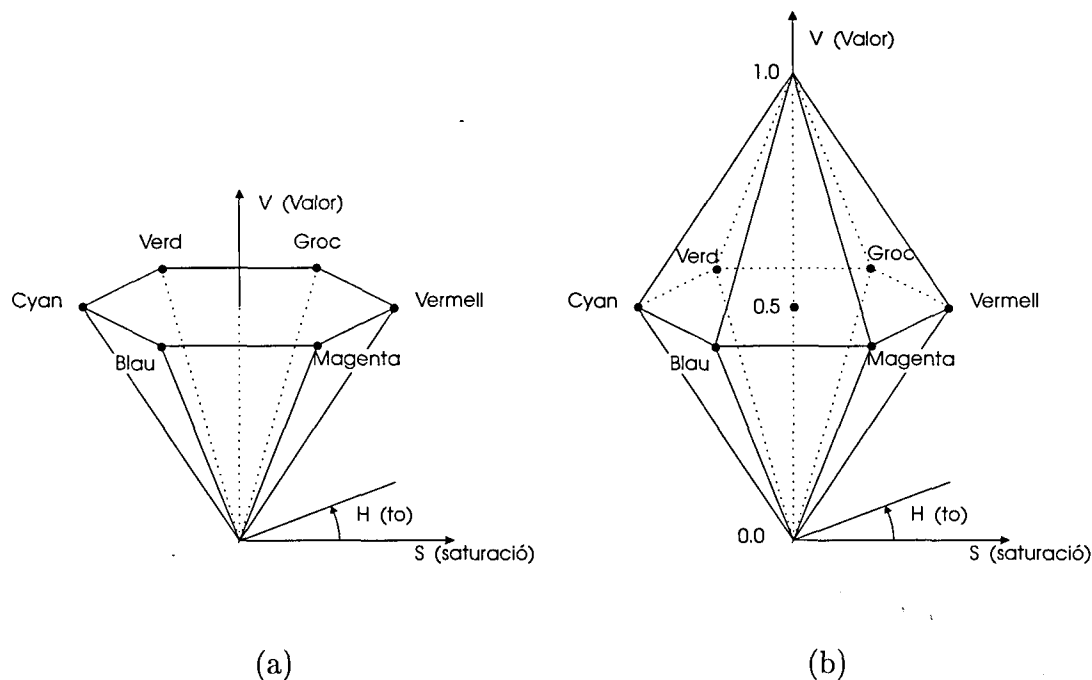


Figura 2.19: Espais de colors HSV (a) i HSL (b).

2.3.6.2 Model HSL (*Hue, Saturation, Lightness*)

És una variant del model HSV. Es manté la definició del to, i canvien les de la saturació i la lluminositat (L):

$$\begin{aligned}
 L &= (\max\{R, G, B\} - \min\{R, G, B\})/2 \\
 S &= \begin{cases} \frac{\max\{R, G, B\} - \min\{R, G, B\}}{\max\{R, G, B\} + \min\{R, G, B\}} & \text{si } L \leq 0.5 \\ \frac{\max\{R, G, B\} - \min\{R, G, B\}}{2 - \max\{R, G, B\} - \min\{R, G, B\}} & \text{si } L > 0.5 \end{cases} \\
 H &= \begin{cases} \pi/3(B - G)/\Delta & \text{si } R = \max\{R, G, B\} \\ \pi/3(2 + (B - R)/\Delta) & \text{si } G = \max\{R, G, B\} \\ \pi/3(4 + (R - G)/\Delta) & \text{si } B = \max\{R, G, B\} \end{cases} \\
 \Delta &= \max\{R, G, B\} - \min\{R, G, B\}
 \end{aligned} \tag{2.34}$$

El sòlid de colors que es genera es mostra a la figura 2.19.b i, com es pot veure, correspon a dues piràmides hexagonals unides per la base.

2.3.6.3 Model HSI *Hue, Saturation, Intensity*)

És una altra variant d'aquests espais (veure, per exemple [PK94]). La definició dels tres components ve donada per:

$$\begin{aligned} I &= \frac{R + G + B}{3} \\ H &= \arctan \frac{\sqrt{3}(G - B)}{(R - G) + (R - B)} \\ S &= 1 - \frac{3 \min\{R, G, B\}}{R + G + B} \end{aligned} \quad (2.35)$$

2.3.6.4 Model HSI pseudo-linealitzat

Kim i Ha [KH93] presenten una modificació al model HSI, passant a una representació pseudo-lineal, treballant amb la relació entre la intensitat i la saturació, que en el model HSI és no lineal:

$$\begin{aligned} I_p &= \frac{1}{3}(R + G + B) \\ H_p &= \arctan \frac{\sqrt{3}(G - B)}{2R - G - B} \\ S_p &= \sqrt{R(R - G) + G(G - B) + B(B - R)} \end{aligned} \quad (2.36)$$

La saturació es normalitza a partir de la següent expressió:

$$S_{p'} = \begin{cases} S_p \frac{1 - I_m}{1 - I_p} & \text{si } I_m \leq I_p < 1 \\ S_p \frac{I_m}{I_p} & \text{si } 0 < I_p < I_m \end{cases} \quad (2.37)$$

I_m és la intensitat corresponent al màxim S_m per un cert valor del to, calculat com es mostra a continuació:

$$S_m = \frac{255 \cos(30)}{\cos(H_d - 30 - 60k)}, k = (\text{int})H_d/60 \quad (2.38)$$

on H_d correspon a H_p expressat en graus.

2.3.6.5 El model LHS generalitzat

El treball unificador de Levkowitz i Herman [Lev93] els va conduir a proposar un nou model de color que anomenen GLHS (*Generalized Lightness, Hue and Saturation*) i que

servirà com a marc de treball de tots els models basats en aquestes tres característiques perceptuals. La seva aportació es concreta amb la definició de 5 propietats comunes a tots els models LHS definits a partir de transformacions sobre les coordenades RGB:

1. Utilitzen coordenades cilíndriques.
2. Als punts amb $r = g = b$ se'ls hi assigna una saturació igual a 0 i un to indefinit (punts acromàtics).
3. Tots els punts del cub RGB que reben una mateixa lluminositat formen una superfície de lluminositat constant (la forma de la qual varia en funció del model concret i determina la forma del sòlid de colors).
4. El to d'un color no acromàtic ve definit per l'angle entre la projecció del vector del color[†] i un vector predefinit que acostuma a correspondre al vermell.
5. La saturació ve donada per la proporció entre la longitud de la projecció del vector de color i la longitud de la projecció més llarga amb la mateixa direcció.

La generalització dels models es fa a partir de la definició d'uns pesos w_{min} , w_{med} , w_{max} no negatius, que compleixen que $w_{max} > 0$ i que $w_{min} + w_{med} + w_{max} = 1$. Es defineix:

$$l(c) = w_{min}min(c) + w_{med}med(c) + w_{max}max(c) \quad (2.39)$$

$$h(c) = (k(c) + f(c))60, \text{ on :} \quad (2.40)$$

$$k(c) = \begin{cases} 0 & \text{si } r > g \geq b \\ 1 & \text{si } g \geq r > b \\ 2 & \text{si } g > b \geq r \\ 3 & \text{si } b \geq g > r \\ 4 & \text{si } b > r \geq g \\ 5 & \text{si } r \geq b > g \end{cases} \quad f(c) = \begin{cases} \frac{med(c) - min(c)}{max(c) - min(c)} & \text{si } k(c) \text{ és parell} \\ \frac{max(c) - med(c)}{max(c) - min(c)} & \text{si } k(c) \text{ és senar} \end{cases} \quad (2.41)$$

[†]Projecció sobre el pla que passa per l'origen i que és perpendicular a l'eix de lluminositat: la projecció del vector (r, g, b) és $((2r - g - b)/3, (2g - b - r)/3, (2b - r - g)/3)$

$$s(c) = \begin{cases} \frac{l(c) - \min(c)}{l(c)} & \text{si } l(c) \leq l(q(c)) \\ \frac{\max(c) - l(c)}{M - l(c)} & \text{si } l(c) > l(q(c)) \end{cases} \quad (2.42)$$

Finalment, $q(c)$ es defineix com (suposant que M és el valor de les coordenades pel color blanc):

$$q(c) = \begin{cases} (M, f(c)M, 0) & \text{si } k(c) = 0 \\ ((1 - f(c))M, M, 0) & \text{si } k(c) = 1 \\ (0, M, f(c)M) & \text{si } k(c) = 2 \\ (0, (1 - f(c))M, M) & \text{si } k(c) = 3 \\ (f(c)M, 0, M) & \text{si } k(c) = 4 \\ (M, 0, (1 - f(c))M) & \text{si } k(c) = 5 \end{cases} \quad (2.43)$$

D'aquestes expressions es pot veure que la transformació que ens dóna el to és equivalent a la de models HSV o HSL, i que w_{\min} , w_{med} i w_{\max} ens permeten variar la saturació i la lluminositat. En concret, $(w_{\min}, w_{\text{med}}, w_{\max}) = (0, 0, 1)$ correspon al model HSV (vegeu equació 2.33) i $(w_{\min}, w_{\text{med}}, w_{\max}) = (0.5, 0, 0.5)$ correspon al model HSL (equació 2.34).

2.3.6.6 Implementacions hardware

La creixent utilització d'aquests models en el camp de la visió per computador ha fet que s'hagi dissenyat hardware especial per conversió d'RGB a HSI, bé sigui implementant la conversió a un model concret, bé sigui oferint una placa més general. Com a exemple de les primeres, trobem el xip de Data Translation, comentat a [Gen93] o [Kay93], que implementa la següent transformació:

$$\begin{aligned} I &= \frac{R + G + B}{3} \\ S &= 1 - \frac{\min(R, G, B)}{I} \\ H &= \arctan \left[\frac{\sqrt{3}(G - B)}{(R - G) + (R + B)} \right] \end{aligned} \quad (2.44)$$

o l'ASIC dissenyat per Andreadis, Iliades i Tsalides [AIT95], que permet la transformació en temps real de l'espai RGB als espais XYZ , YIQ o YUV . Aquests tres espais es

defineixen a partir d'una transformació lineal sobre l'espai RGB i, aquest hardware el que fa és aplicar les corresponents matrius 3×3 sobre els valors RGB obtinguts d'una càmera.

Com a exemple de les segones plaques, trobem el treball de Batlle [Bat93], que dins la seva tesi doctoral sobre la detecció de moviment basant-se en el color, dissenya una placa per a la conversió en temps real d' RGB a qualsevol model tricromàtic que prèviament s'hagi carregat en la memòria. De fet, la placa actua com una LUT que converteix els components R,G,B als tres components del model que s'hagi carregat.

Capítol 3

Treballs relacionats

Presentem tot seguit una sèrie de treballs analitzats que tenen relació amb la tesi que es presenta. Es comenten treballs sobre la il·luminació, la constància del color, el model de la reflexió dicromàtica, l'extracció de contorns i la segmentació d'imatges en color i sobre les possibilitats del color com a mètode d'indexació en bases de dades d'imatges.

3.1 Introducció

En les tècniques de processat d'imatges, el color s'havia considerat inicialment només com una informació addicional que permetria millorar els algoritmes ja existents, afegint-hi noves restriccions, o fent-los més ràpids. Estudis posteriors posaren en evidència que, tot i que l'afirmació anterior és certa en la majoria d'aplicacions, el disposar d'informació cromàtica pot permetre atacar problemes difícils de resoldre en imatges de les que només disposem de la lluminància, tal i com comenten una sèrie d'autors com Gershon, Healey i Binford, Luong o bé Ledley, Buas i Golab en els seus articles genèrics sobre el color aplicat a la visió per computador i que es citen tot seguit.

Gershon [Ger85] escriu:

... recent studies suggest that color contributes to other visual processes, and therefore cannot be considered as a mechanism which merely adds beauty to scenes. It should be an integral part of any computer vision system, ...

Healey i Binford [HB87] comenten que:

... Starting from general physical models, we have shown that color can play an important role in both the classification and in the recognition of objects. An appropriate use of color, therefore, can significantly extend the capabilities of a vision system. ...

Luong [Luo93] diu:

... We emphasize the distinction between tasks that have been extensively studied in monochromatic images and for which the contribution of color is just a quantitative generalization, and tasks where color has a qualitative role. In the first case, additional image features are obtained, and have to be represented and used efficiently. In the latter case, it is hoped that color can help recover the scene's intrinsic physical properties. ...

Ledley, Buas i Golab [LBG90] afirmen que:

... there is a tremendous need by research and clinical pathologists for image analysis systems that can process true-color images as seen through microscope. The interpretation of stained microscope slides by pathologists is generally based on the analysis of true-color patterns. ...

3.2 Il·luminació

És molt difícil aïllar l'estudi d'imatges en color del procés de formació d'aquestes; no es pot oblidar que el valor d'un píxel en una imatge depèn de les característiques de la superfície a la que correspon, de les característiques d'il·luminació de l'escena i de la resposta dels sensors utilitzats per adquirir-la. Per tant, per estudiar les imatges es necessitarà un model de l'escena (tant dels objectes com de les superfícies) i un model del sensor.

En el món de les imatges digitals es troben dues disciplines amb objectius clarament contraposats:

Visió per computador: L'objectiu és extreure informació d'imatges adquirides amb algun sensor. Un dels camps de recerca és, per exemple, el de la interpretació i posterior reconstrucció d'escenes a partir d'imatges d'aquestes.

Gràfics per computador: L'objectiu és construir una imatge digital a partir de certes

dades. Una de les línies de recerca vers aquest objectiu és la de la visualització realista d'escenes, que consisteix en construir una imatge el més real possible a partir del model d'una escena i de paràmetres de l'observador i de la il·luminació.

Tot i que els seus objectius són oposats, ambdues àrees tenen interessos comuns i, de fet, a la literatura apareixen diferents intents de col·laboració entre els dos móns, com per exemple els treballs presentats per Gagalowicz [Gag90] i per Encarnaçao, Groß, Hoffman i Hübner [EaGHH92]. El primer parla d'aquesta col·laboració des d'un punt de vista molt genèric, mentre que els segons plantegen aquesta relació entre gràfics i visió per computador dins el marc del desenvolupament d'aplicacions industrials. De totes maneres aquestes col·laboracions entre els dos camps són esporàdiques i sembla que no acaben de reeixir. Potser una de les línies d'investigació que permetrà una interrelació més forta és la de la *realitat augmentada* que consisteix en introduir elements sintètics en imatges reals, de forma que quedin del tot integrats en aquestes (tal i com comenta R. Azuma al seu treball [Azu95]).

Tant en la visió com en els gràfics es poden distingir dues orientacions en la recerca: una de modelització i una altra d'algorítmica. En la primera, la modelització, és en la que les dues comunitats d'investigadors poden col·laborar, donat que l'objectiu és establir quin és el procés de formació de les imatges. Exemples d'aquestes teories són el *model de reflexió dicromàtica* en el camp de la visió [Sha85, KSK87, KSK88b] o bé el *model de la il·luminació global* [Hec91]. El vessant algorítmic, per altra banda, planteja la resolució de problemes concrets, enfocant-ho cap al problema del cost computacional i, per tant, no permet un intercanvi d'informació massa ric entre les dues comunitats.

3.2.1 Materials i llum

El color de les superfícies observades està condicionat per les característiques cromàtiques i geomètriques dels objectes i dels efectes de la llum incident sobre aquests. Quan la llum incideix sobre una superfície, aquesta pot ésser reflexada, transmesa, o absorbida (generant calor). Quan es parla de reflexió o transmissió, aquestes poden presentar-se de forma més o menys difusa en funció de si la superfície és més o menys rugosa (vegeu, per exemple, el llibre de Hall [Hal89]). La figura 3.1 mostra aquests possibles comportaments.

Klinker proposa a la seva tesi [Kli88] una taxonomia dels materials d'acord amb el comportament de la llum quan incideix en ells. Comença separant els materials entre materials conductors (metalls) i dielèctrics. Pel que fa referència als metalls, la reflexió de la llum que incideix sobre ells és de tipus especular. En els materials dielèctrics, la

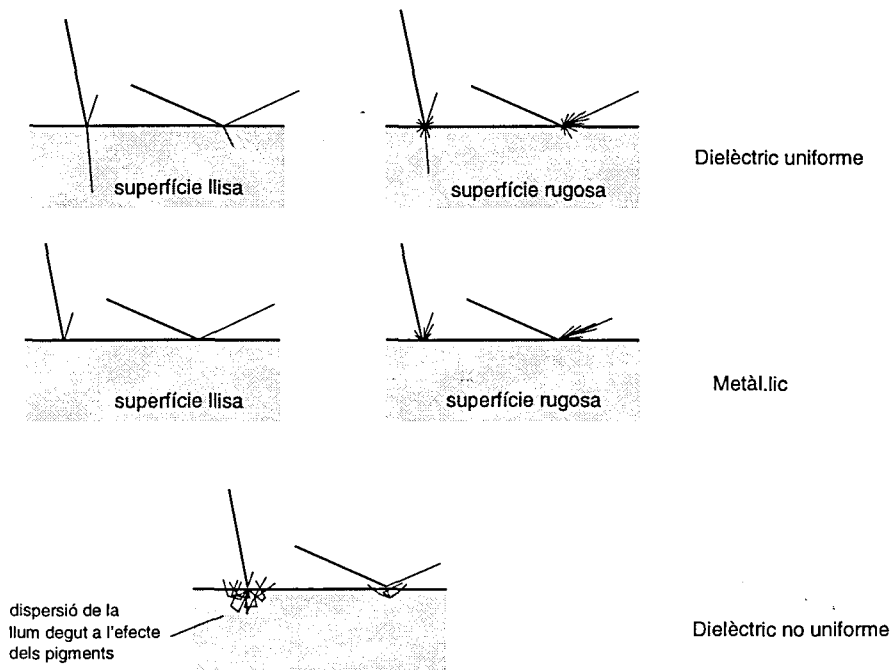


Figura 3.1: Possibles comportaments de la llum incident en funció de la superfície (segons Hall [Hal89]).

reflexió té dues components: especular i difusa. Aquests últims, a la vegada, es divideixen en uniformes i no uniformes segons criteris físics. Els dielèctrics uniformes són aquells materials en els que les partícules que els formen tenen un tamany semblant al de les longituds d'ona de l'espectre visible; són uniformes, per exemple, l'aigua pura o els cristalls de quars. Els no uniformes són aquells en els que les partícules que els formen són més grans que les longituds d'ona en l'espectre visible com ara la majoria de materials que porten pigments com per exemple les pintures, el plàstic, la ceràmica, la roba o el paper. Finalment, els material dielèctrics no uniformes es divideixen en transmissors i opacs en funció de si deixen passar la llum o no. La figura 3.2 presenta aquesta classificació dels materials.

3.2.2 Model de la il·luminació global

L'objectiu del model de la il·luminació global és el de modelar la il·luminació tenint en compte les interrelacions entre diferents superfícies, en contraposició als models d'il·luminació local que consideren que l'aspecte d'una superfície és independent a les altres superfícies (vegeu, per exemple, el llibre de Foley, van Dam, Feiner i Hughes [FvDFH90]).

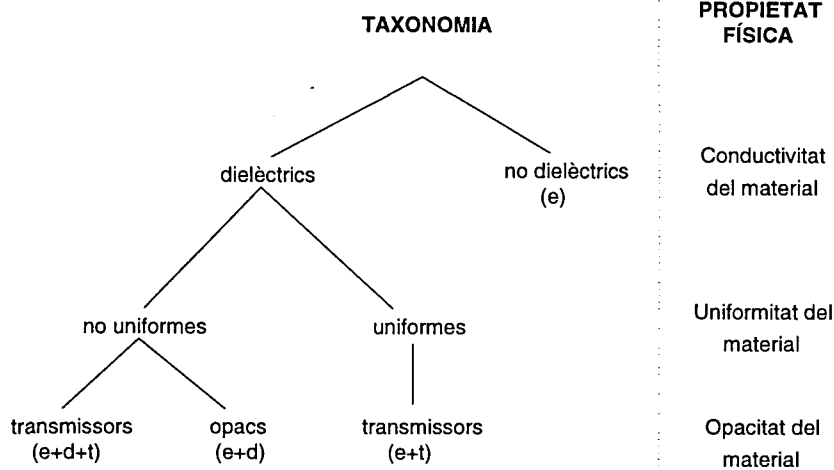


Figura 3.2: Taxonomia dels materials en funció del seu comportament respecte a la llum incident. Les lletres entre parèntesis indiquen si el material presenta reflexió especular (e), difusa (d) o si el material transmet la llum (t).

L'equació que expressa la il·luminació global en medis no participatius és la següent (segons Heckbert [Hec91]):

$$i_o(x, \Theta_o) = i_{emit}(x, \Theta_o) + \int_{\Gamma} dx' \frac{\cos \phi_i \cos \phi'_o}{r^2} v(x, x') \mu(x, \Theta_i, \Theta_o) i_o(x', \Theta'_o) \quad (3.1)$$

on

$$\mu(\Theta_i, \Theta_o) = \rho(\Theta_i, \Theta_o) + \tau(\Theta_i, \Theta_o) \quad (3.2)$$

$$v(x, x') = \begin{cases} 1 & \text{si } x' \text{ és visible des de } x \\ 0 & \text{si } x' \text{ no és visible des de } x \end{cases} \quad (3.3)$$

ρ i τ corresponen a les funcions de distribució bidireccional de reflectivitat i transmissió (BRDF i BTDF), respectivament.

La figura 3.3 mostra la geometria associada a l'equació anterior. Els angles θ_i i ϕ_i descriuen la direcció de la línia que uneix x amb x' i els angles θ_o i ϕ_o descriuen la direcció de la mateixa línia en el sistema de coordenades de x' .

L'expressió mostrada a l'equació 3.1 és el punt de partida pels algorismes de radiositat, una de les tècniques més emprades per la generació d'imatges realistes.

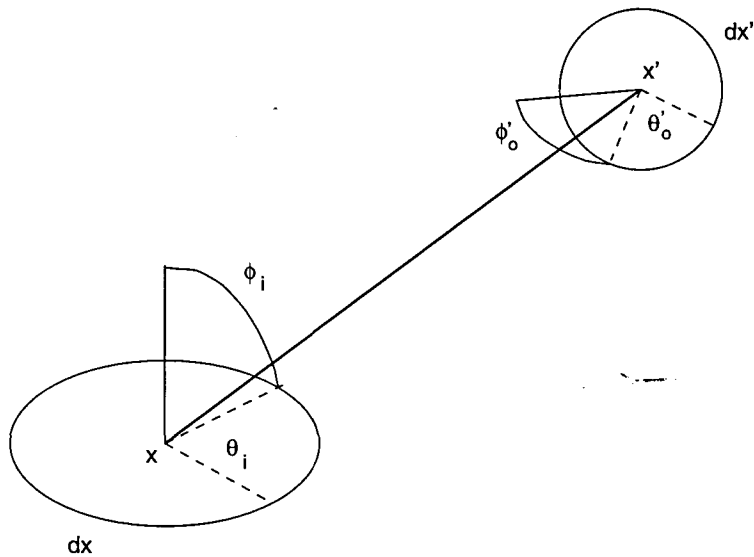


Figura 3.3: Geometria del problema de la il·luminació global per dues superfícies, on la llum surt d' x' i arriba a x .

3.3 Constància del color

Un dels problemes més interessants de la visió en color, estudiat tant des del camp de la visió per computador com des del punt de vista de la percepció humana, és el de la *constància del color*. La constància del color pot definir-se com l'estabilitat del color que es percep d'una superfície davant de canvis en l'illuminació. Com ja hem vist, però, el color adquirit pels sensors variarà en funció de les propietats dels il·luminants i dels objectes, com de la geometria de l'escena (posicions relatives dels sensors, objectes i il·luminants).

El treball de D'Zamura i Lennie [DL86] engloba la visió per computador i la visió humana. El seu objectiu és doble: d'una banda mostrar que és possible recuperar el color d'un objecte independentment de la forma, de la posició i de l'il·luminant. D'altra banda pretenen donar una explicació de quins mecanismes de la percepció humana permeten percebre el color dels objectes de forma independent de l'il·luminant. La recuperació del color dels objectes passa per descomptar l'il·luminant, és a dir, cal obtenir la distribució espectral de la reflectància de les superfícies dels objectes.

Un dels aspectes més desenvolupats en els treballs sobre la constància del color és el de l'estudi del procés físic de formació de les imatges, tant des del punt de vista de la interacció llum/superfícies com des del punt de vista de la influència del sensor en la formació de les imatges. En el primer apartat s'hi inclourien, per exemple, tots els treballs

sobre el model de la reflexió dicromàtica – explicat a partir de la pàgina 53 – començats per Shafer [Sha85], o el treball de Ho, Funt i Drew [HFD90] en els quals presenten un mètode per separar les característiques cromàtiques de l'il·luminant i de les superfícies, a partir de la descripció de la distribució espectral de l'il·luminant i de la reflectància de les superfícies com una combinació lineal d'una base de funcions.

A part de l'estudi de la interacció entre la llum i les superfícies també s'han realitzat treballs modelant l'efecte que el sensor produeix en el color, com queda reflexat a l'article de Healey i Binford [HB88]. Aquests autors caracteritzen els sensors, incloent-hi un model del soroll que es produeix en l'etapa d'adquisició i defineixen una mètrica per mesurar les diferències entre colors. Un altre efecte associat als sensors que ha estat estudiat per altres autors, i que s'explica més endavant, és el de l'aberració cromàtica.

De totes formes, la solució al problema genèric de la constància del color no existeix, i els sistemes implementats ho han estat imposant restriccions a les característiques dels objectes i dels il·luminants. Hi ha estudis, com el de Finlayson [Fin94], en els que es defineixen condicions sota les quals el problema de la constància del color és resoluble; en concret, demostren que si la reflectància s'aproxima a partir d'una base amb tres funcions i els il·luminants s'aproximen a partir d'una base de com a màxim 5 funcions, es pot resoldre el problema de la constància del color a partir de tres superfícies observades sota dos il·luminants diferents.

3.3.1 Models basats en l'aberració cromàtica

Aquest és un camp molt relacionat amb l'òptica. Es presenta tot seguit un treball interessant d'aplicació de l'aberració cromàtica a la resolució del problema de la constància del color.

L'índex de refracció d'un medi (p.e. una lent) és una funció que depèn de la longitud d'ona i per tant, per un certa distribució espectral, només un conjunt molt petit de longituds d'ona estarà correctament enfocat. Això provoca el que es coneix amb el nom d'aberració cromàtica.

Funt i Ho ([FH88, FH89]) demostren que a partir d'una imatge d'intensitat d'un punt en color $\rho(x)$ sobre un fons negre, i tenint en compte l'aberració cromàtica es pot reconstruir la distribució de l'espectre $I(\lambda)$, és a dir, es pot trobar el color del punt. També demostren que a partir de la imatge d'intensitat es pot extreure la diferència entre les distribucions espectrals (SPD) de dues regions de diferent color.

Donat que l'espectre de reflectància de les superfícies $S(\lambda)$ i l'SPD de l'il·luminant

$E(\lambda)$ poden modelar-se com a combinació lineal de dues bases de funcions $S_i(\lambda)$ i $E_i(\lambda)$:

$$S(\lambda) = \sum_{j=1}^n \sigma_j S_j(\lambda) \quad (3.4)$$

$$E(\lambda) = \sum_{i=1}^m \epsilon_i E_i(\lambda) \quad (3.5)$$

per un cert punt de color es té que el seu SPD $I(\lambda)$ és el producte entre $E(\lambda)$ i $S(\lambda)$:

$$I(\lambda) = \sum_{i=1}^m \epsilon_i E_i(\lambda) \sum_{j=1}^n \sigma_j S_j(\lambda) \quad (3.6)$$

A partir de 3.6 i coneixent $I(\lambda)$ es poden generar les equacions necessàries per resoldre el sistema. En el cas ideal de que $S_i(\lambda)$ i $E_i(\lambda)$ siguin totes independents, el nombre d'equacions a generar serà de $n \times m$.

Donat que els coeficients ϵ_i i σ_j que busquem apareixen en forma de producte, podem determinar la informació cromàtica del punt de color i de l'il·luminant, però no la seva intensitat.

3.4. Extracció de contorns i segmentació utilitzant el color

A continuació es presenten tot un seguit de treballs sobre la detecció de contorns i la segmentació en imatges en color que s'han anat fent els darrers anys. La majoria dels treballs presentats no treballen directament sobre els valors *RGB* que proporciona la càmera, sinó que realitzen transformacions a altres espais de color. Tal i com es comenta al final del capítol 5, el mètode que proposem en aquesta tesi és aplicable a una gran majoria d'aquests treballs.

3.4.1 Extracció de contorns

Els treballs realitzats en extracció de contorns en imatges en color parteixen de la recerca feta anteriorment en la detecció de contorns en imatges de nivells de gris. De totes maneres, s'ha vist que el color no és simplement una font d'informació addicional de la que es disposa per la detecció de contorns, sino que permet resoldre millor el problema de la seva detecció en certs tipus d'imatges, especialment aquelles que presenten regions amb diferents intensitats d'il·luminació.

Tot seguit es comenten alguns dels treballs que s'han publicat sobre la detecció de contorns en imatges en color.

Nevatia, [Nev77] És el primer treball sobre detecció de contorns en imatges en color. Defineix la següent transformació sobre l'espai RGB:

$$Y = c_1R + c_2G + c_3B \quad (3.7)$$

$$T_1 = \frac{R}{R + G + B} \quad (3.8)$$

$$T_2 = \frac{G}{R + G + B} \quad (3.9)$$

on $c_1 + c_2 + c_3 = 1$. Aquest espai és equivalent al sistema RGB normalitzat afegint-hi un component d'intensitat. La component Y correspon a la component d'intensitat (és semblant a la Y de l'espai YIQ utilitzat en els televisors en color). Els components T_1 i T_2 aporten la informació cromàtica.

Nevatia estudia tres possibilitats per detectar contorns en una imatge color:

1. Definir una mètrica en l'espai tridimensional.
2. Detectar els contorns en cada component i integrar-los.
3. Suposar que els contorns són independents en els tres components.

Proposa una extensió de l'operador de Hueckel [Hue71] per detectar contorns en color. Conclou que el color pot ajudar a construir sistemes més robustos, però que en general no aporta més informació que la que ja es disposa amb la lluminància.

Cumani, Grattoni i Guiducci, [CGG91] Els punts de contorn en imatges monocromàtiques es poden definir a partir del creuament per zero de la segona derivada. Aquests autors presenten una extensió d'aquesta definició per imatges multiespectrals i, en concret, per imatges en color.

Si considerem que una imatge amb m components està constituïda per m funcions $f_i(x, y)$, la funció de contrast multiespectral en la direcció $u = (u_1, u_2)$ ve definida com:

$$S(x, y, u) = Eu_1^2 + 2Fu_1u_2 + Gu_2^2 \quad (3.10)$$

on

$$E = \sum_{i=1}^m \left(\frac{\delta f_i}{\delta x} \right)^2, \quad F = \sum_{i=1}^m \frac{\delta f_i}{\delta x} \frac{\delta f_i}{\delta y}, \quad G = \sum_{i=1}^m \left(\frac{\delta f_i}{\delta y} \right)^2 \quad (3.11)$$

Si $n(x, y)$ és el vector propi corresponent al valor propi més gran, el màxim de S respecte a u es donarà quan $u = n$, i aquest màxim de S s'expressa com:

$$\lambda(x, y) = \left(E + G + \sqrt{(E - G)^2 + 4F^2} \right) / 2 \quad (3.12)$$

i finalment els contorns es defineixen com els creuaments per zero de la segona derivada $D_S(x, y)$, que correspon a la següent expressió:

$$D_S(x, y) = E_x n_1^3 + (2F_x + E_y) n_1^2 n_2 + (G_x + 2F_y) n_1 n_2^2 + G_y n_2^3 \quad (3.13)$$

en la que els subíndexs x i y indiquen les derivades parcials respecte a x i respecte a y .

Carron i Lambert, [CL94, CLM93] Proposen dos detectors de contorns en imatges color a partir dels tres components cromàtics d'una variant del model HSI , definint la saturació com:

$$S = \sqrt{(R - G/2 - B/2)^2 + (\sqrt{3}B/2 - \sqrt{3}G/2)^2} \quad (3.14)$$

Proposen una mesura de la diferència entre dos valors de to tenint en compte que es tracta d'una representació circular (és a dir, que la diferència entre els valors de to 245 i 5 és 10 en lloc de 235) i que l'assignació de valors numèrics als colors és arbitrària (no té massa sentit, per exemple, considerar que la diferència entre el groc i el blau és més gran que la diferència entre el groc i el verd). A partir d'aquestes dues premisses defineixen la diferència entre dos valors de to H_1 i H_2 com:

si $|[H_1 - H_2]_{255}| \leq \Delta H_{max}$ llavors

$$\delta(H_1, H_2) = [H_1 - H_2]_{255}$$

sino

si $[H_1 - H_2]_{255} \geq 0$ llavors

$$\delta(H_1, H_2) = \Delta H_{max}$$

sino

$$\delta(H_1, H_2) = -\Delta H_{max}$$

fsi

fsi

on ΔH_{max} correspon a la màxima diferència que els autors permeten entre dos colors qualssevol, i que agafen com la diferència de to entre dos vèrtexs consecutius del cub *RGB*, que és aproximadament 42.5 (que correspon a un angle de $\pi/3$ radians expressat en el rang $[0, 255]$).

Com a detector de contorn proposen una modificació a l'operador de Sobel i consideren que la informació de gradient G^C en un píxel $C = (c_x, c_y)$ es pot calcular a partir de les imatges obtingudes aplicant l'operador de Sobel sobre els components de saturació i intensitat i l'operador de Sobel modificat sobre el component de to. Com a resultat de l'aplicació d'aquests operadors sobre els components *HSI* de la imatge s'obtenen les imatges G_x^H i G_y^H pel to, G_x^S i G_y^S per la saturació i G_x^I i G_y^I per la intensitat. La integració dels resultats obtinguts per cada component es pot fer de dues maneres:

- Considerant que els tres components aporten la mateixa quantitat d'informació, la magnitud del gradient G^C es pot obtenir com:

$$\begin{aligned} G_x^C(c_x, c_y) &= |G_x^H| + |G_x^S| + |G_x^I| \\ G_y^C(c_x, c_y) &= |G_y^H| + |G_y^S| + |G_y^I| \\ G^C(c_x, c_y) &= \sqrt{G_x^C(c_x, c_y)^2 + G_y^C(c_x, c_y)^2} \end{aligned} \quad (3.15)$$

La orientació Φ^C del gradient es pot aproximar a partir de la orientació del més gran dels tres gradients G^H , G^S i G^I . Així, l'expressió per obtenir Φ^C és:

$$\Phi^C = \begin{cases} \arctan(G_y^H(c_x, c_y)/G_x^H(c_x, c_y)) & \text{si } G^H = \max\{G^H, G^S, G^I\} \\ \arctan(G_y^S(c_x, c_y)/G_x^S(c_x, c_y)) & \text{si } G^S = \max\{G^H, G^S, G^I\} \\ \arctan(G_y^I(c_x, c_y)/G_x^I(c_x, c_y)) & \text{si } G^I = \max\{G^H, G^S, G^I\} \end{cases} \quad (3.16)$$

- Considerant que la informació del to és la més important, el gradient dels components corresponents a la saturació i a la intensitat es multiplica pel coeficient $(1-p)$, essent p un indicador de la importància que es vol donar al component corresponent al to: p està entre 0 i 1, dependent de si es desitja donar poc pes ($p = 0$) o molt pes ($p = 1$) al to. Es pot veure que el cas $p = 0$ equival al cas anterior (els tres components aporten la mateixa informació).

Pietikäinen i Harwood, [PH86] El mètode proposat per aquests autors es basa en la utilització d'histogrames de diferències. En imatges de grisos, l'histograma de diferències s'obté calculant el valor absolut de la diferència entre cada píxel i els seus veïns. La

figura 3.4 mostra l'aspecte general d'un d'aquests histogrames. El pic de l'histograma que comença al 0 és fruit dels píxels que pertanyen a regions uniformes i el pic que s'observa més cap a la dreta és fruit dels píxels corresponents a regions de contorn.

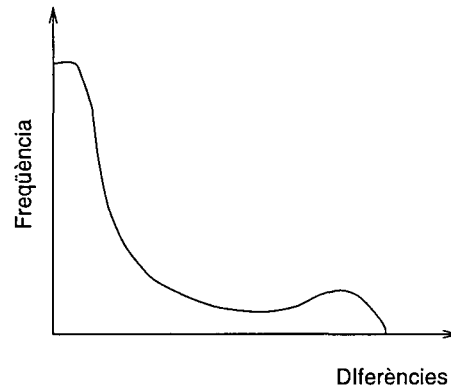


Figura 3.4: Histograma de diferències en imatges monocromàtiques.

Els autors extenen aquest concepte d'histograma de diferència per imatges amb més d'un component i, en concret, a imatges en color, utilitzant-lo per definir un mètode de suavitzat de la imatge en color preservant els contorns. El valor del contorn a cada píxel s'agafa com el màxim de les diferències entre el píxel en qüestió i els seus vuit veïns. Cal comentar que en aquest treball només obtenen la magnitud del contorn, prescindint de la seva orientació.

Cufí i Casals, [CC96] Seguint en la línia de treballs que hem presentat, aquests autors proposen un mètode d'integració de la informació de contorns obtinguda sobre diferents components cromàtics amb l'objectiu de trobar els *contorns envolupants* de la imatge. Defineixen els contorns envolupants com aquells que corresponen als contorns perceptualment més significatius d'una imatge. A més de la informació de color, utilitzen informació sobre les textures de les regions que hi ha a banda i banda dels possibles contorns.

Per l'obtenció de contorns treballen sobre l'espai de color *HLS* i calculen per cada component els contorns locals, definint-los a partir de la seva magnitud i la seva orientació: $C^H(M^H, O^H)$ pel to, $C^L(M^L, O^L)$ per la lluminància i $C^S(M^S, O^S)$ per la saturació.

Per la integració de les tres imatges de contorn obtingudes es defineix la funció $f(C^H, C^L, C^S)$ de la següent forma:

$$f(C^H, C^L, C^S) = \max(\gamma_H C^H, \gamma_L C^L, \gamma_S C^S) \quad (3.17)$$

on els γ_H , γ_L i γ_S són uns pesos calculats localment per cada píxel en funció dels píxels veïns, a diferència de la proposta de Carron i Lambert [CL94] en la que el pes era el mateix per la saturació i la intensitat i, a més, era el mateix per tota la imatge.

Un cop obtinguda la imatge amb els contorns ja integrats $f(C^H, C^L, C^S)$, apliquen un algoritme d'aprimament seguit d'un algoritme d'unió de segments de contorn per obtenir els contorns envolupants de la imatge. Dos segments s'uneixen si la seva distància en un espai de paràmetres que els autors defineixen és prou petita. L'espai de paràmetres té en compte la distància entre els extrems dels segments, la orientació i magnitud del gradient en els tres components cromàtics i la colinealitat.

3.4.2 Segmentació

Igual que passava amb la detecció de contorns, els treballs existents sobre la segmentació d'imatges en color es basen en tècniques ben conegudes de segmentació sobre imatges d'intensitat. La majoria de treballs es basen en tècniques de *clustering* o tècniques d'*split and merge*.

Respecte als espais de color utilitzats, es veu que hi ha una gran varietat, tot i que la majoria es decanten cap a espais de color perceptualment uniformes. En aquest sentit, l'espai *CIELAB* és utilitzat per Celenk [Cel94], Zheng [ZVG93] o Tominaga [Tom90], l'espai *CIEUVW* ho és per Schettini [Sch94b, Sch93] o Tseng i Chan [TC92]. L'espai *HSI* és utilitzat, entre altres, per Perez i Koch [PK94] i l'espai *CIEUVW* és utilitzat per Lim [LL90]. Lee, Kim i Kim utilitzen l'espai *HLS* en el seu article sobre el reconeixement de matrícules [LKK94]. Per la segmentació també s'han utilitzat altres espais de color no perceptualment uniformes com el cas de Lim [LL90] que compara els espais *RGB*, *YIQ*, *CIEUVW* i *I1, I2, I3* [OKS80]. Aquest últim espai també és utilitzat per Sun, Wu, Lin i Chou [SWLC93] en el seu treball sobre la classificació d'imatges microscòpiques de teixit de fetge. Thorpe, Hebert, Kanade i Shafer [THKS88] utilitzen l'espai *RGB* per la segmentació de carreteres dins del projecte NavLab per la conducció automatitzada de vehicles. Aquest espai també l'han utilitzat Zhong, Karu i Jain [ZKJ95] per la segmentació de caràcters en imatges en color.

A continuació es comenten alguns dels treballs que s'han anat fent els darrers anys. El conjunt de treballs sobre el model de la reflexió dicromàtica es comenten al final de forma conjunta degut a la seva proximitat al tema de la present tesi.

Lim i Lee, [LL90] Presenten un algoritme per segmentació d'imatges en color basat en la segmentació a partir dels histogrames dels components cromàtics seguit d'un procés

de refinament utilitzant tècniques de *clustering* difós.

La segmentació a partir de l'histograma consisteix en dividir aquest en regions significatives, a partir de la recerca de regions amb un màxim local (pics) i de regions amb un mínim local (valls). La búsqueda de màxims i mínims es realitza trobant els creuaments per zero de la segona derivada. Abans, però, suavitzen l'histograma convolucionant-lo amb una gaussiana amb una amplada calculada a partir de tècniques d'espai escala (Witkin [Wit84]).

El resultat d'aquesta primera segmentació és una partició sobre l'espai de color en un conjunt de classes que seran els valors d'entrada del mòdul de refinament, basat en un procés de *fuzzy clustering*.

El mètode que proposen l'apliquen utilitzant quatre espais de color diferents *RGB*, *YIQ*, *CIEUVW* i *I1, I2, I3* i conclouen que el que dona millors resultats és l'espai *I1, I2, I3* proposat per Ohta. La comparació entre els quatre espais es fa a partir de la similitud entre la segmentació obtinguda aplicant l'algoritme i la segmentació feta a mà.

Tomiyasu i Nishio, [TN91] En zones de la imatge amb colors poc saturats, la segmentació basada en els components cromàtics de to i saturació no és viable. Per això, els autors plantegen la solució al problema de la segmentació d'imatges en color a partir de la solució de dos subproblemes: la segmentació en zones de la imatge on els colors tenen una saturació molt pobre i la segmentació en zones de la imatge amb colors més saturats. En les primeres zones la segmentació es basarà en la informació d'intensitat mentre que en les segones la segmentació es basarà en el to.

L'espai de colors que utilitzen és, curiosament, el definit pel sòlid de colors de Munsell i, el motiu pel qual l'utilitzen és que en la seva definició es van tenir en compte les sensacions que els colors provoquen a les persones. Aquest argument no ha estat trobat a cap altre article consultat.

Bartneck i Ritter, [BR92] Proposen un sistema de segmentació basat en la classificació polinòmica que apliquen a la lectura d'adreces en objectes postals i a l'anàlisi de senyals de trànsit en escenes del món natural. Tant en un cas com en l'altre suposen que el sistema té una fase d'aprenentatge guiat.

El model de color amb el que treballen és una variant del model HSI:

$$H = \arctan\left(\frac{g - 1/3}{r - 1/3}\right) \quad (3.18)$$

$$S = 1 - 3 \min\{r, g, b\} \quad (3.19)$$

on r, g, b són els valors d' RGB normalitzats.

Pel reconeixement dels senyals de trànsit dissenyen un classificador polinòmic capaç d'etiquetar els píxels com a pertanyents a una de les quatre classes *vermell, blau, groc, fons*. En la fase d'aprenentatge arriben a la conclusió que el conjunt de valors d'entrada que discrimina millor entre aquestes quatre classes és el format pels valors H, R, G i B i que cal un classificador polinòmic amb polinomis de grau 2.

Pel que fa a l'anàlisi d'objectes postals, els valors d'entrada que permeten una millor discriminació segueixen essent H, R, G i B , però ara és suficient utilitzar un classificador lineal (polinomi de grau 1).

Taylor i Lewis, [TL92] Estudien dues formes d'atacar el problema de la segmentació sobre un espai de color perceptualment uniforme. La primera, basada en l'*split and merge*, els hi produeix resultats poc contínuos, cosa que solucionen lleugerament amb un processat a posteriori per suavitzar el resultat.

L'altra alternativa es basa en un procés de creixement de regions, basat en un algoritme de la llavor. S'agafa com a llavor el conjunt de quatre píxels adjacents amb la diferència de color entre ells més petita de tota la imatge. A continuació es mira quin dels píxels adjacents a la regió té el color més proper a la mitjana aritmètica dels colors de la regió i, si la diferència de color entre la mitjana aritmètica de la regió i el nou píxel és inferior a un cert umbral, aquest s'afegeix a la regió. Aquest procés es va repetint fins que no es pot afegir a la regió cap dels píxels que hi són adjacents.

L'espai de color perceptualment uniforme utilitzat és el TekHVC, desenvolupat per Tektronix [TMM89] i il·lustrat a la figura 3.5. Sobre aquest espai es defineix la diferència entre dos colors (h_1, v_1, c_1) i (h_2, v_2, c_2) com:

$$\Delta C = \sqrt{(v_1 - v_2)^2 + c_1^2 + c_2^2 - 2c_1c_2 \cos(h_1 - h_2)} \quad (3.20)$$

El fet que l'espai de colors sigui perceptualment uniforme permet definir també les mitjanes aritmètiques $(\bar{h}, \bar{v}, \bar{c})$ d'un conjunt de N colors (h_i, v_i, c_i) de la següent manera:

$$\bar{h} = \arctan\left(\frac{\bar{p}}{\bar{q}}\right) \quad (3.21)$$

$$\bar{v} = \frac{1}{N} \sum_{i=1}^N v_i \quad (3.22)$$

$$\bar{c} = \sqrt{\bar{p}^2 + \bar{q}^2} \quad (3.23)$$

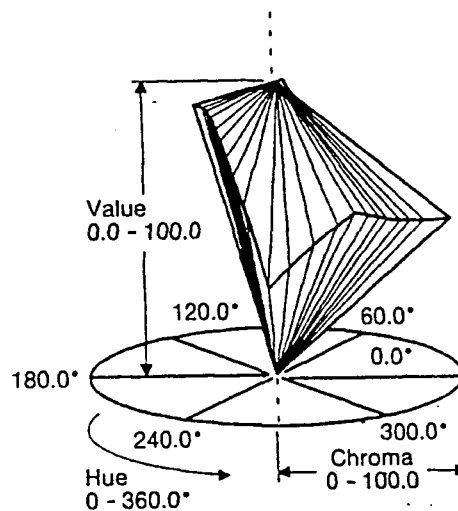


Figura 3.5: Esquema de l'espai de colors TekHVC de Tektronix.

essent \bar{p} i \bar{q} la mitjana aritmètica del conjunt de vectors (h_i, c_i) :

$$\bar{p} = \frac{1}{N} \sum_{i=1}^n c_i \cos h_i \tag{3.24}$$

$$\bar{q} = \frac{1}{N} \sum_{i=1}^n c_i \sin h_i \tag{3.25}$$

Tseng i Chang, [TC92] Presenten un mètode de segmentació d'imatges en color basat en un model de color uniforme. El model escollit per treballar és el $CIEL^*u^*v^*$, utilitzant les seves coordenades cilíndriques L^*, H^*, S^* . Aquests autors defineixen la diferència D entre dues coordenades L_1, H_1, S_1 i L_2, H_2, S_2 com:

$$D = \sqrt{d_L^2 + d_C^2} \tag{3.26}$$

$$d_L = |L_1 - L_2| \tag{3.27}$$

$$d_C = \sqrt{S_1^2 + S_2^2 - 2S_1S_2 \cos \theta} \tag{3.28}$$

$$\theta = \begin{cases} |H_1 - H_2| & \text{si } |H_1 - H_2| \leq \pi \\ 2\pi - |H_1 - H_2| & \text{si } |H_1 - H_2| > \pi \end{cases} \tag{3.29}$$

Paral·lelament, proposen una separació de l'espai LHS en dues regions: la cromàtica i l'acromàtica al voltant de l'eix L , tal i com es mostra a la figura 3.6.

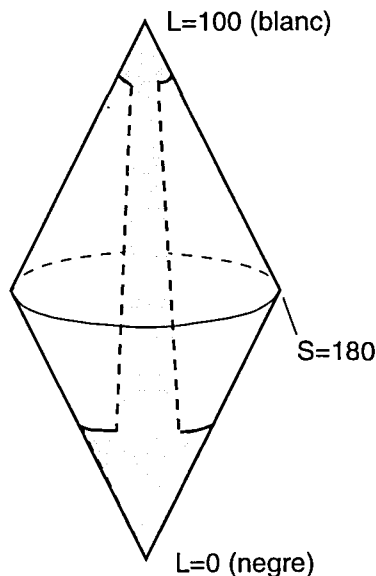


Figura 3.6: Separació de l'espai *IHS* en regions cromàtiques i acromàtiques (segons Tseng i Chang [TC92]). La zona ombrejada correspon a les dues regions acromàtiques del sòlid de colors: els dos extrems dels cons i un cilindre al voltant de l'eix *L*.

A diferència del treball de Tomiyasu i Nishio [TN91], la separació entre zones cromàtiques i acromàtiques es basa en criteris geomètrics. La zona acromàtica dins del sòlid de colors *LHS* es divideix en dues regions:

- Regió 1: (intensitat > 95) o (intensitat < 25) *
- Regió 2: (81 < intensitat < 95) i (saturació < 18) o bé
 (61 < intensitat < 80) i (saturació < 20) o bé
 (51 < intensitat < 80) i (saturació < 30) o bé
 (41 < intensitat < 80) i (saturació < 40) o bé
 (25 < intensitat < 80) i (saturació < 60)

El procediment que segueixen per la segmentació és el següent: en un primer pas classifiquen els píxels en acromàtics o cromàtics. A continuació, sobre cada una de les dues classes de píxels apliquen, per separat, una segmentació basada en la detecció de pics en els histogrames. Pels píxels de les zones cromàtiques la segmentació es basa en el to i pels píxels de les zones acromàtiques la segmentació es basa en la intensitat. Acaben amb un

*S'han mantingut els rangs de valors utilitzat pels autors que és de 0 a 100 per la intensitat i de 0 a 180 per la saturació, tot i que els rangs habituals en la resta de treballs van entre 0 i 1 o entre 0 i 255.

algoritme de fusió de regions per eliminar la possible sobresegmentació. Aquest algoritme de fusió és un algoritme del tipus *region growing* basat en el concepte de *clustering* sobre l'espai *IHS*. La idea és detectar les subregions formades en el primer pas de l'algoritme que estiguin formades per un conjunt de píxels amb un valor en el seu to inestable, i es mira si poden ser fusionades amb subregions veïnes.

Okii, Kaneki, Hara i Ono, [OKHO94] Presenten un mètode de segmentació basat en xarxes neuronals. El mètode és específic per la segmentació d'imatges de teixits cel·lulars tenyits amb hematoxilina-eosina, substàncies que tenyeixen el nucli de les cèl·lules d'un color porpra i l'interstici d'un color rosat.

La xarxa neuronal, que és del tipus *back-propagation*, classifica cada píxel com a nucli, interstici o fons. La xarxa neuronal té tres capes, amb sis nodes a la capa d'entrada: R, G, B, R^2, G^2, B^2 , 40 nodes en una capa oculta i tres nodes a la capa de sortida: un node pel color negre (nuclis), un pel color rosa (interstici) i un pel color blanc (fons).

Han fet proves amb altres espais de color, com per exemple amb l'espai *CIELAB* o l'espai *HVC*, però han desestimat la seva utilització perquè han vist que el color porpra del nucli i el color rosat de l'interstici tenien unes coordenades cromàtiques molt properes i, per tant, era molt difícil decidir a quina de les dues classes pertanyien els píxels.

Schettini, [Sch94b, Sch93] En el primer article [Sch93] es presenta un mètode de segmentació aplicant la tècnica de *split and merge* basat en una segmentació inicial a partir dels histogrames, seguit d'un procés d'unió de les regions obtingudes. Treballa sobre l'espai perceptualment uniforme *CIELUV*, aplicant la transformació de l'espai *RGB* del sensor a l'espai *CIELUV* descrita per Strachan en el seu article sobre la calibració de sistemes de visió en color [SNA90].

La segmentació inicial es basa en l'anàlisi d'histogrames utilitzant tècniques d'espai escala, anant convolucionant l'histograma amb gaussianes d'amplada gradualment incrementada. El procés de *merging* es basa en criteris de semblança i connectivitat.

Aquest primer treball és continuat pel mateix autor a [Sch94b] amb alguna modificació: s'aplica un filtratge de mediana abans de processar la imatge, es proposa una lleugera modificació a l'espai *CIELUV* dividint els components cromàtics per 2, i el procés de partició es fa ara a partir de l'obtenció del diagrama de Voronoi. Donats n píxels d'una imatge, el diagrama de Voronoi d'aquesta és l'esquelet de la partició de la imatge en n regions poligonals, cada una associada a un dels n píxels (anomenats llavors). Per cada regió poligonal es compleix que tots els seus píxels tenen unes característiques més properes

al píxel llavor de la regió en qüestió que no pas als píxels llavor de les regions adjacents.

Huang, Dom, Steele, Ashley i Niblack, [HDS⁺95] Aquests autors proposen un mètode per segmentar les imatges en dues classes: objectes i fons. Es basen en la fusió d'informació de les regions obtingudes a partir d'un algoritme simple de *clustering* i de les regions obtingudes utilitzant contorns.

Els contorns s'obtenen a partir de la imatge gradient $g(x, y)$, l'expressió de la qual és:

$$g(x, y) = \max(g_R(x, y), g_G(x, y), g_B(x, y)) \quad (3.30)$$

essent $(g_R(x, y), g_G(x, y), g_B(x, y))$ els gradients per cada un dels tres components. Els píxels considerats del contorn s'obtenen a partir d'una llindarització adaptativa de la imatge $g(x, y)$. El llindar a aplicar es determina per cada píxel (x, y) a partir de la següent expressió:

$$T_{x,y} = w_1(m_B + 3s_B) + w_2(m_{x,y} + 3s_{x,y}) \quad (3.31)$$

on m_B i s_B són una estimació del valor del gradient del seu fons (m_B és el valor estimat i s_B és la variança). Aquests valors s'obtenen a partir dels píxels que es troben dins de quatre finestres situades a les quatre cantonades de la imatge. Els valors w_1 i w_2 són uns pesos i els valors $m_{x,y}$ i $s_{x,y}$ corresponen a la mitjana i la desviació del valor del gradient de tots els píxels que pertanyen a la mateixa fila o columna del píxel estudiat. En el cas d'una fila, pel píxel l dins d'aquesta fila, $m_{x,y}$ i $s_{x,y}$ serien:

$$m_{x,y} = \frac{\sum_{i=0}^l g(i, y)}{l} \quad (3.32)$$

$$s_{x,y} = \sqrt{\frac{\sum_{i=0}^l g(i, y)^2 - m_{x,y}^2}{l - 1}} \quad (3.33)$$

Celenk, [Cel95, Cel94, Cel93, Cel91, Cel90] Mehmet Celenk en els seus treballs proposa un sistema molt interessant per la segmentació d'imatges en color basat en l'espai de color perceptual $CIEL^*a^*b^*$. El sistema es basa en la detecció de *clusters* en les coordenades cilíndriques de l'espai $CIEL^*a^*b^*$ anomenades L^*, H°, C^* (on H° correspon al to i C^* a la cromaticitat).

A partir dels histogrames de cada un dels tres components construeix uns volums de decisió acotats per dos cilindres, dos plans horitzontals i dos plans verticals que passen

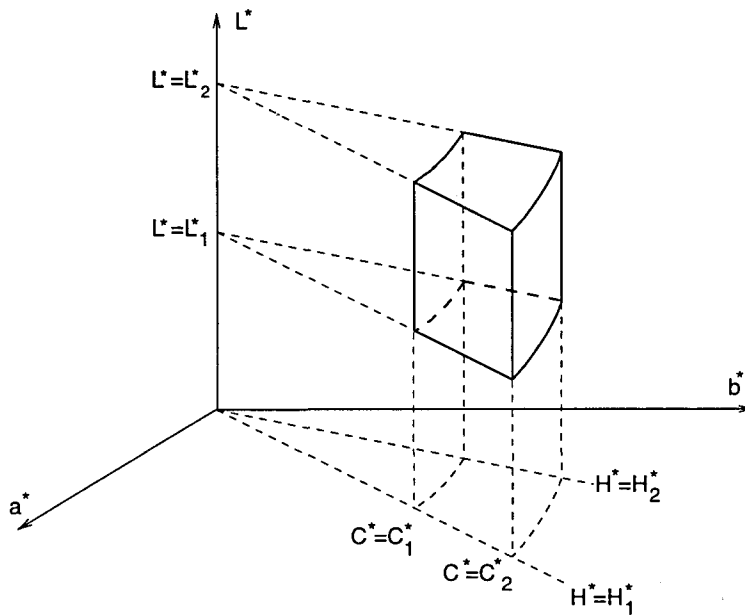


Figura 3.7: Volum de decisió proposat per Celenk.

per l'eix L^* . La figura 3.7 ho mostra gràficament: els dos cilindres tenen radis C_1^* i C_2^* (que corresponen a dues superfícies amb cromaticitat constant), els dos plans horitzontals estan situats a $L^* = L_1^*$ i $L^* = L_2^*$ (que corresponen a dues superfícies amb lluminositat constant) i els dos plans verticals són dos plans paral·lels a l'eix L^* , les interseccions dels quals amb el pla a^*b^* corresponen a les rectes H_1^* i H_2^* .

El pas de l'espai RGB al CIE XYZ el fa suposant que l'espai RGB de la càmera és equivalent al definit per la CIE, sense fer cap tipus de calibració.

3.5 Model de la reflexió dicromàtica

Aquest model fou proposat per Shafer [Sha85] i posteriorment desenvolupat conjuntament amb Klinker i Kanade. Una col·lecció molt interessant de treballs sobre aquest model és la formada pels articles de Klinker, Shafer i Kanade entre els anys 1987 i 1990 [KSK87, KSK88b, KSK88a, KSK90] i la tesi doctoral de Klinker [Kli88]). Més recentment Maxwell i Shafer ([MS93, MS94]) continuen els treballs, incloent-los en un model més ampli pel que fa al domini d'escenes considerat.

El model proposat inicialment per Shafer és vàlid per escenes en les que les superfícies siguin dielèctriques no homogènies i opaques – tals com plàstics o pintures –, amb els

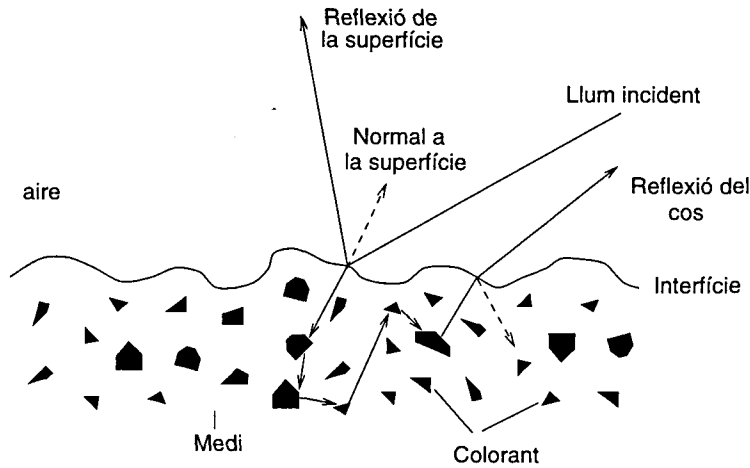


Figura 3.8: Model Dicromàtic: Reflexió de la llum.

pigments distribuïts aleatòriament, i que continguin una única font d'il·luminació (sense llum ambient).

El mètode es basa en separar la reflexió en dos components: la reflexió de superfície L_s (parts llents), amb una distribució espectral molt semblant a la llum incident, i la reflexió del cos L_b (parts no llents), que depèn dels colorants que porti l'objecte i que és la que dona el color a la superfície. La figura 3.8 mostra aquest comportament de la llum i la figura 3.9 il·lustra esquemàticament la geometria sobre la que treballa el model dicromàtic: i és l'angle d'incidència entre la direcció de l'il·luminant I i la normal a la superfície N , e correspon a l'angle de sortida entre N i la direcció d'observació V i, finalment, g és l'angle de fase entre I i V .

A partir d'aquests dos components es generen les que anomenen imatges intrínseques: una formada per la reflexió del cos (en la qual s'han eliminat les llentors) i l'altra que només inclou les llentors.

En un punt de la superfície, la llum emesa vindrà donada per la següent expressió:

$$L(\lambda, i, e, g) = L_s(\lambda, i, e, g) + L_b(\lambda, i, e, g) \tag{3.34}$$

El model separa les propietats de reflexió espectrals de les propietats de reflexió geomètriques, quedant l'expressió 3.34 com:

$$L(\lambda, i, e, g) = m_s(i, e, g)c_s(\lambda) + m_b(i, e, g)c_b(\lambda) \tag{3.35}$$

on $c_s(\lambda)$ i $c_b(\lambda)$ corresponen, respectivament, a la distribució espectral de la llum emesa

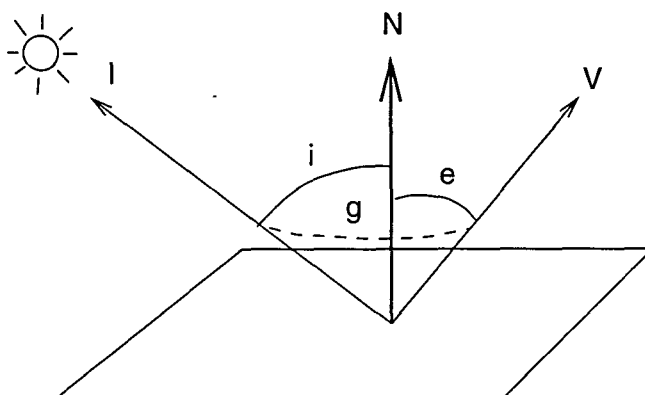


Figura 3.9: Model Dicromàtic: notació.

per la reflexió de la superfície i per la reflexió del cos, i són constants per tota la superfície, degut a que les superfícies s'han considerat dielèctriques. Els factors m_s i m_b escalen c_s i c_b en funció de la geometria descrita a la figura 3.9 ($c_s, c_b \in [0, 1]$). Això ens defineix un espai d'infinites dimensions. Donada una longitud d'ona λ la figura 3.10 mostra el comportament dels píxels d'un objecte. En aquesta figura es pot veure que els píxels no lluent segueixen la direcció de $c_b(\lambda)$ i els brillants segueixen la de $c_s(\lambda)$.

Aquesta representació en forma de T , la forma de la qual dependrà dels angles (i, e, g), permet separar els components no lluent i els components lluent, o el que és el mateix, permet eliminar les lluentors de l'escena i trobar una aproximació de la distribució espectral de l'illuminant.

Finalment, i degut a la impossibilitat de tractar aquest espai d'infinites dimensions, el que proposa Shafer és limitar l'estudi a només tres longituds d'ona diferents (vermell, verd i blau), quedant l'expressió 3.35 que dona el color de cada píxel de la imatge de la següent manera:

$$C(x, y) = m_s(i, e, g)C_s + m_b(i, e, g)C_b \quad (3.36)$$

on $C_s \in \{R_s, G_s, B_s\}$ i $C_b \in \{R_b, G_b, B_b\}$.

Tot aquest estudi és continuat en anys posteriors, juntament amb Klinker i Kanade, en treballs sobre la segmentació d'imatges en color, automatitzant el procés introduint-hi la utilització d'hipòtesis sobre l'escena i la seva posterior validació [KSK88b, KSK88a, Kli88].

Una part d'aquests treballs està molt lligada a aquesta tesi. Es tracta de l'apartat dedicat a la recuperació del color en les zones amb lluentors (que solen coincidir amb

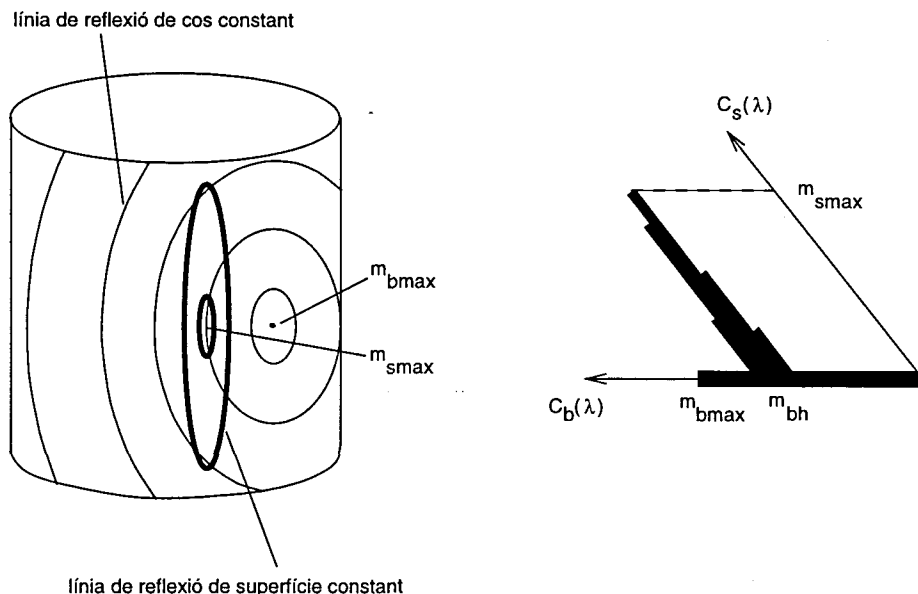


Figura 3.10: Representació gràfica d'una part de l'espai de reflexió dicromàtica.

zones on els sensors han quedat saturats). Klinker proposa corregir aquesta informació cromàtica de la següent manera: si p és un píxel amb dos components saturats, el tercer component és el que aporta la informació cromàtica. Es busca en el *cluster* al que pertany p un píxel q amb el mateix valor en el component no saturat i s'assigna a p el valor de q . En el cas de píxels amb un canal saturat es fa el mateix. Fixem-nos, però, que això serveix per casos en els que la saturació dels components sigui deguda a l'efecte de la propagació de la saturació d'algun píxel veí (*blooming*). Si la saturació fos deguda a un increment de la intensitat de l'il·luminant (que és el problema amb el que treballem nosaltres) no es trobaria cap píxel q amb un mateix valor als components no saturats, doncs tots serien inferiors.

El model dicromàtic tal i com va ser formulat originalment no tenia en compte les interreflexions. Tjahjadi, Litwin i Yang [TLY95] proposen una extensió al model per tal d'incloure-hi l'efecte d'aquestes, afegint a l'equació dos termes per cada superfície que interactuï amb el punt considerat: un per la reflexió de superfície i l'altre per la reflexió del cos.

Healey, al seu article [Hea88] presenta un model de reflexió que anomena *model aproximat de la reflectància del color* vàlid tant per materials dielèctrics com per materials metàl·lics. En el cas de materials dielèctrics, el model presentat per Healey és equivalent al de la reflexió dicromàtica.

3.6 Indexació en Bases de Dades d'Imatges

L'abaratiment continuat dels dispositius relacionats amb l'adquisició i emmagatzemament, ha portat a la proliferació de grans bases de dades d'imatges. Aquest creixement ha comportat l'aparició d'un problema nou: *com recuperar la imatge desitjada?*. En bases de dades petites és suficient mantenir una llista amb la descripció textual del contingut de les imatges, però quan el nombre d'imatges creix això és inviable.

En aquest context han anat apareixent els darrers anys una sèrie de treballs orientats cap a la recuperació d'imatges a partir del seu contingut (recuperació per contingut), i no a partir de descripcions textuales.

A l'hora de descriure el contingut d'una imatge el color hi juga un paper molt important, i molts dels treballs que s'han fet sobre la indexació en bases de dades d'imatges el té en compte. Tot seguit es presenten alguns dels treballs en el camp de la indexació en bases de dades d'imatges que utilitzen el color com a característica important d'aquestes.

Swain i Ballard, [SB91] Utilitzen la descripció dels objectes a partir dels seus histogrames de color com a mecanisme d'indexació en una base de dades de models d'objectes.

L'histograma de color consisteix en comptar quantes vegades apareix cada combinació de valors (r, g, b) en la imatge. Afirment que els histogrames són invariants a l'escalat i a la translació i que varien d'una forma suau quan es varia l'angle d'observació. Això permet que amb un nombre reduït d'histogrames es pugui caracteritzar correctament un objecte per tal de ser reconegut des de qualsevol angle.

Estudien dos problemes:

Identificació. A partir de la imatge d'un objecte veure a quin objecte correspon.

Localització. Donat un objecte, veure en quina posició d'una imatge apareix.

Pel procés d'identificació, aquests autors proposen la comparació entre els histogrames dels models i els de les imatges utilitzant el mètode de la *intersecció d'histogrames*. Donat l'histograma I de la imatge i l'histograma M del model, la intersecció entre els dos histogrames es defineix com:

$$\sum_{j=1}^n \min(I_j, M_j) \quad (3.37)$$

on n és el número d'elements dels histogrames. Per obtenir una mesura de la semblança entre els dos histogrames $H(I, M)$ es normalitza el valor de la intersecció obtingut a l'equació 3.37:

$$H(I, M) = \frac{\sum_{j=1}^n \min(I_j, M_j)}{\sum_{j=1}^n M_j} \quad (3.38)$$

Els experiments els varen portar a terme utilitzant un model de color oposat semblant al que té el sistema visual humà descrit a l'equació 2.32. Per la discretització de l'histograma agafaren $n = 2048$ repartits de la següent manera: 8 particions a l'eix acromàtic i 16 a cada un dels dos eixos cromàtics. Les proves van ser fetes utilitzant 32 imatges amb un objecte a cada una i una base de dades de 66 models. Cada model es va comparar amb cada imatge obtenint una matriu de diferències $H(I_i, M_j)$, $1 \leq i \leq 32$, $1 \leq j \leq 66$. En 29 dels 32 casos l'objecte correcte va obtenir la màxima puntuació i en els altres tres casos l'objecte correcte va obtenir la segona millor puntuació.

Segons els autors, aquest mètode és robust davant de: variacions en el fons de les imatges, canvis en el punt de vista de l'observador, oclusions i variacions en la resolució de les imatges. Com a aspecte negatiu, el mètode no és robust a les variacions en l'il·luminant, deixant els autors la solució a aquest problema en mans del desenvolupament d'un mòdul que s'incorporaria a l'etapa de preprocessat del sistema. Pel que fa referència als models de color, Swain i Ballard no plantegen l'estudi d'altres models que no sigui el model *RGB* normalitzat.

Pel procés de localització utilitzen el següent algoritme de *backprojection* de l'histograma color:

1. Per cada element j de l'histograma calcular $R_j := \min(M_j/I_j, 1)$
2. Per cada x, y fer $b_{x,y} := R_{h(c_{x,y})}$
3. $b := D^r * b$
4. $(x_t, y_t) := \text{loc}(\max_{x,y} b_{x,y})$

on $c_{x,y}$ representa el color del píxel (x, y) , $h(c)$ és la funció que assigna al color de c un element de l'histograma, D^r és un disc de radi r , l'operador $*$ és l'operador de convolució i la funció *loc* retorna el píxel que té per valor el seu argument.

Aquest treball fou posteriorment complementat amb un estudi teòric fet per Stricker i Swain [SS94] sobre la capacitat d'indexació que tenen els histogrames del color, ja que fins aleshores les proves s'havien realitzat sempre amb bases de dades relativament petites (amb menys de 1000 imatges). En aquest treball es dona una cota inferior del número que indica la quantitat de models que es poden representar sense ambigüetat mitjançant histogrames.

Schettini, [Sch94a] El treball d'Schettini es centra en el tema de la localització d'un objecte dins d'una imatge utilitzant informació sobre la forma de l'objecte i el seu color. Primer estableix hipòtesis de localització a partir de la forma de l'objecte i llavors aquestes hipòtesis són validades utilitzant les descripcions de forma i color de l'objecte en qüestió.

Pel que respecte al color, aquest autor utilitza el mateix mètode de la intersecció d'histogrames utilitzat per Swain i Ballard [SB91]. La diferència principal entre ells és la utilització, per part d'Schettini, d'un espai de color perceptualment uniforme com és el cas de l'espai *CIELUV*, més robust davant de canvis en l'illuminant.

Mehre, Kankanhalli, Narasimhalu i Man, [MKNM95] Presenten dues alternatives per al mètode de la intersecció d'histogrames proposat per Swai i Ballard: el mètode de la distància i el mètode del color de referència.

El *mètode de la distància* consisteix en caracteritzar les imatges a partir dels valors corresponents a la mitjana aritmètica dels histogrames 1D de cada un dels components cromàtics *RGB*. Una imatge es caracteritza amb un vector de tres components \bar{f} :

$$\bar{f} = (\mu_R, \mu_G, \mu_B) \quad (3.39)$$

essent μ_R , μ_G i μ_B les mitges de cada un dels tres histogrames. La distància entre una imatge *I* i el model *M* serà la distància euclídea entre els seus vectors de característiques \bar{f}_I i \bar{f}_M :

$$D = \sqrt{(\bar{f}_I - \bar{f}_M)^2} = \sqrt{\sum_{R,G,B} (\mu_I - \mu_M)^2} \quad (3.40)$$

En el *mètode del colors de referència* el que proposen és convertir els colors de la imatge a un conjunt reduït de colors que ha estat prèviament definit. Concretament, els autors proposen un conjunt de 27 colors, que corresponen als colors que poden obtenir-se utilitzant tres valors per cada component *RGB*. En aquest mètode, cada imatge es caracteritza per un vector amb *n* components, una per cada un dels colors del conjunt reduït:

$$\bar{f} = (l_1, l_2 \dots l_n) \quad (3.41)$$

on l_i és la freqüència relativa del color *i* i la distància entre la imatge i el model és:

$$D_{q,i} = \sum_{k=1} n\omega_k \sqrt{(l_k^I - l_k^M)^2} \quad (3.42)$$

$$\omega_i = \begin{cases} l_k^I & \text{si } l_k^I, l_k^M > 0 \\ 1 & \text{si } l_k^I \text{ o } l_k^M = 0 \end{cases} \quad (3.43)$$

Com es pot veure aquest dos mètodes són molt més simples que el proposat per Swain i Ballard, però, tal com diuen els autors, estan pensats per treballar amb imatges sintètiques, en les que no cal la precisió que ofereix el mètode de la intersecció dels histogrames. De fet, dels experiments dels autors es desprèn que per imatges sintètiques els dos mètodes proposats funcionen millor que el de Swain i Ballard. Cal dir, però, que les proves han estat realitzades utilitzant logotips de companyies aèries i banderes i no sembla un conjunt de proves massa general.

Jain i Vailaya, [JV96] Presenten un mètode de recuperació per contingut d'imatges corresponents a anagrames d'empreses. El mètode de recuperació es basa, igual que proposava Schettini [Sch94a], en el color i la forma de l'objecte a recuperar. La diferència amb el treball d'Schettini és que aquest primer identifica l'objecte per la forma i utilitza el color en l'etapa de validació, mentre que Jain i Vailaya identifiquen els objectes utilitzant conjuntament la forma i el color. L'índex de similitud entre una imatge i el model es defineix com:

$$S_t = \frac{w_c S_c + w_s S_s}{w_c + w_s} \quad (3.44)$$

essent S_c i S_s els índexs de similitud de color i de forma respectivament. w_c i w_s són pesos associats a les dues característiques. Els autors treballen amb $w_c = w_s = 1$.

L'espai de color utilitzat per les comparacions és l'espai RGB . Donada una imatge I i un model M la similitud entre els dos es calcula com:

$$S_c(M, I) = \frac{\sum_r \min(M_R(r), I_R(r)) + \sum_g \min(M_G(g), I_G(g)) + \sum_b \min(M_B(b), I_B(b))}{3 \min(|M|, |I|)} \quad (3.45)$$

on M_R, M_G, M_B i I_R, I_G, I_B són els histogrames de cada component del model i de la imatge respectivament.

3.7 Conclusions

De tots aquests treballs presentats, els més propers a la tesi que aquí es presenta són els treballs sobre el model de la reflexió dicromàtica iniciats per Shafer [Sha85] i continuats per ell mateix conjuntament amb altres autors com Klinker o Kanade. Aquests autors treballen

amb hipòtesis sobre la distribució dins de l'espai *RGB* dels valors dels píxels associats a un mateix objecte, i es centren en l'estudi dels objectes que presenten distribucions especulars. L'objectiu dels seus estudis és el de la determinació de les característiques de l'il·luminant a partir de la detecció de les lluentors. Pel que fa al problema de la saturació dels CCDs, ells resolen l'efecte de la saturació deguda al *blooming*, però no tracten el tema d'increments en la intensitat de l'il·luminant.

Pel que respecta als treballs presentats sobre la detecció de contorns i la segmentació, s'han extret idees per la part de la validació del mètode que proposem com a etapa de preprocessat en algorismes ja existents. S'han extret les idees del volum de decisió sobre un espai cromàtic basat en el to, la saturació i la intensitat com el proposat per Celenk [Cel90], o les idees de la separació de la imatge entre regions cromàtiques (píxels amb un valor alt corresponent a la saturació) i regions acromàtiques (píxels amb una saturació baixa) present en el treballs de Tseng i Chang [TC92].

En l'estudi del comportament dels components cromàtics corresponents al to davant de canvis en la intensitat de l'il·luminant, s'han tingut en compte les definicions de diferència en variables circulars que apareixen als treballs de Perez i Koch [PK94] o Carron i Lambert [CL94].

Els treballs sobre la indexació de les bases de dades s'han inclòs com un exemple d'un camp d'investigació que s'ha obert els darrers anys en el que les idees d'aquesta tesi es podrien aplicar d'una forma natural, com un component més dels sistemes que s'estan desenvolupant.

Capítol 4

Estudi del problema de la saturació dels CCDs

Molts treballs de processat d'imatges en color es basen en la utilització de components cromàtics que es mantenen estables a les variacions en la intensitat de l'illuminant. Aquesta estabilitat es fonamenta sobre la linealitat de la variació de la resposta del sensor respecte a canvis en la intensitat de l'illuminant. Els sensors, però, només tenen un comportament lineal fins que queden saturats. En aquest capítol mostrem com afecta el fenomen físic de la saturació del sensor a diferents components cromàtics que han estat utilitzats en treballs relacionats amb la visió per computador.

4.1 Introducció

Molts dels treballs existents sobre la segmentació o detecció de contorns en imatges en color utilitzen components cromàtics que són estables a les variacions en la intensitat de l'illuminant. Aquesta invariança, però, a la pràctica no sempre es compleix i depèn, en gran mesura, del comportament dels sensors utilitzats davant els canvis en l'illuminant. Tal i com es mostra tot seguit, la modelització de la resposta de les càmeres reflexa que la variació d'aquesta respecte a la intensitat de l'illuminant és lineal.

Seguint la nomenclatura de Shafer [Sha85], donat un cert píxel x la resposta p_x d'una càmera monocroma ve expressada per la següent integral:

$$p_x = \int_{\lambda_{\min}}^{\lambda_{\max}} X(\lambda)s(\lambda)d\lambda \quad (4.1)$$

on $X(\lambda)$ representa la quantitat de llum incident per longitud d'ona, i $s(\lambda)$ correspon a la sensibilitat de la càmera per cada longitud d'ona. L'interval $[\lambda_{min}, \lambda_{max}]$ correspon a l'interval de longituds d'ona on la càmera és sensible.

En les càmeres en color la resposta per cada píxel ve expressada per un vector amb tres components: vermell, verd i blau. Així, la resposta C_x per un cert píxel x ve expressada per:

$$C_x = \begin{bmatrix} r_x \\ g_x \\ b_x \end{bmatrix} = \begin{bmatrix} \int_{\lambda_{min}}^{\lambda_{max}} X(\lambda)\tau_r(\lambda)s(\lambda)d\lambda \\ \int_{\lambda_{min}}^{\lambda_{max}} X(\lambda)\tau_g(\lambda)s(\lambda)d\lambda \\ \int_{\lambda_{min}}^{\lambda_{max}} X(\lambda)\tau_b(\lambda)s(\lambda)d\lambda \end{bmatrix} \quad (4.2)$$

on $\tau_r(\lambda), \tau_g(\lambda), \tau_b(\lambda)$ corresponen a la transmitància dels filtres associats als components vermell ($\tau_r(\lambda)$), verd ($\tau_g(\lambda)$) i blau ($\tau_b(\lambda)$).

Observant les equacions anteriors es pot veure que, efectivament, la resposta d'una càmera varia linealment en relació als canvis en la intensitat de l'il·luminant, donat que una variació en la intensitat equival a multiplicar la funció $X(\lambda)$, que expressa la quantitat de llum que arriba al sensor per longitud d'ona, per un cert valor a . Òbviament, els valors de r_x, g_x i b_x queden també multiplicats per a .

Les càmeres, però, són dispositius físics, i com a tals tenen uns límits a l'hora d'adquirir les imatges. En concret, la resposta d'una càmera en color està circumscrita a un cub a l'espai RGB que nosaltres suposarem definit pels punts $(0, 0, 0), (1, 1, 1)$. Qualsevol valor r_x, g_x o b_x de l'equació 4.2 que sigui superior a 1, queda retallat a 1 i es diu que el component en qüestió ha quedat saturat i que el color del píxel ha quedat retallat.

Molts autors de treballs en el camp de la visió per computador utilitzant el color han buscat transformacions cap a espais de color en els que alguns components fossin estables a variacions en la intensitat de l'il·luminant. Aquesta estabilitat, però, tal com demostrem més endavant, només existeix mentre la resposta del sensor té una variació lineal respecte a la intensitat, linealitat que es perd quan algun dels components del sensor queda saturat.

Es pot afirmar, doncs, que la resposta dels CCDs té, de forma inherent, un comportament lineal respecte a la intensitat de la llum rebuda. Hi ha, però, un conjunt de paràmetres que poden trencar aquesta linealitat (segons Novak, [NSW90]):

Correcció γ . Moltes de les càmeres estan pensades per ser utilitzades conjuntament amb un monitor de video. Els monitors no tenen una resposta lineal al voltatge d'entrada, sinó que la intensitat B percebuda en el monitor per un cert voltatge d'entrada V segueix la relació:

$$B = \alpha V^\gamma \quad (4.3)$$

on α i γ estan relacionats amb els controls de contrast i d'intensitat del monitor. Per això, els fabricants de càmeres corregeixen la resposta d'aquestes de forma que quan la imatge adquirida sigui visualitzada en im monitors el seu comportament sigui lineal. És a dir, modifiquen la resposta m dels CCDs obtenint la resposta corregida m' que serà enviada als monitors:

$$m' = m^{1/\lambda} \quad (4.4)$$

Retall del color. Es produeix quan per un píxel arriba a la càmera una intensitat lluminosa més gran que la que pot acceptar el sensor, com ja ha estat exposat anteriorment.

Blooming. Si la intensitat que arriba per un píxel és molt més gran que la màxima admesa, els píxels veïns poden veure's afectats amb un increment en la resposta de la càmera.

Aberració cromàtica. L'índex de refracció de les lents de les càmeres és una funció que depèn de la longitud d'ona. Això provoca que les diferents longituds d'ona d'un senyal lluminós quedin enfocats en profunditats diferents, provocant que una part d'aquests colors quedin desenfocats.

D'aquests quatre problemes l'estudi que presentem es centra en el problema del retall del color. Pel que respecta a la correcció γ moltes càmeres permeten desactivar-la, desapareixent, per tant, aquest problema. Si la càmera utilitzada aplica la correcció γ i no és possible desactivar-la, el que cal és estimar el paràmetre γ de l'equació 4.4 adquirint una imatge d'una escala de grisos amb una reflectància coneguda. En relació al problema del *blooming* no té una solució clara, ja que implica fer hipòtesi sobre què ha passat en els píxels veïns. El problema de l'aberració cromàtica és més proper a treballs sobre el desenfocament o a treballs d'òptica, que no formen part dels objectius d'aquesta tesi.

4.2 Variació dels components cromàtics

En aquesta secció es presenten exemples de com varien, en funció de la intensitat de l'il·luminant, alguns dels components cromàtics més utilitzats. Les gràfiques corresponen



Figura 4.1: Carta de colors *Macbeth COLORCHECKER*.

a la representació de les funcions matemàtiques que defineixen les transformacions de l'espai *RGB* a aquests components. Els colors representats són els 12 colors que es troben a les dues files centrals de la carta de colors *The Macbeth COLORCHECKER* i que descrivim tot seguit.

4.2.1 La carta de colors *The Macbeth COLORCHECKER*

La carta de colors *The Macbeth COLORCHECKER Color rendition Chart* [Mac91, MMD76], fou creada a mitjans dels anys 70 per fer comprovacions de qualitat del color en pel·lícules fotogràfiques, però el seu ús s'ha anat extenent i últimament ha estat utilitzada també en processos de calibració ([SNA90]) o en estudis sobre la constància del color ([NS90]).

La *Macbeth COLORCHECKER* té un tamany de 13.5 per 9.25 polzades (34.3 per 23.5 cm) i està formada per 24 mostres de colors distribuïdes en 4 files de 6 colors cada una. Les proporcions són les mateixes que les d'un negatiu fotogràfic de 35mm i estan pensades de forma que un cop fotografiada la carta, cada mostra ocupi sobre la pel·lícula un quadrat de 5mm, tamany suficient per ser analitzats amb un colorímetre. Les tres primeres files corresponen a mostres cromàtiques i l'última fila correspon a 6 mostres acromàtiques (vegeu figura 4.1).

La descripció dels 24 colors es mostra a la taula 4.1. El nom dels colors és una tra-

ducció dels noms que apareixen a la documentació que acompanya a la carta [Mac91]. Les coordenades cromàtiques es donen agafant com a referència l'il·luminant estàndard C de la CIE, il·luminant que fou definit de forma que la seva distribució espectral fos similar a la de la llum diürna. Es donen les coordenades dels colors en dos espais de color: el CIE_{xyY} (descriu a la pàgina 20) i el de Munsell (descriu a la pàgina 18). Els 24 colors foren escollits, segons comenten McCamy, Marcus i Davidson [MMD76], de forma que incloguessin:

- Una escala de colors neutres des del blanc fins el negre. Són els 6 colors de la fila inferior.
- Els tres primaris additius (vermell, verd i blau) i els tres substractius (cian, magenta i groc). Aquest sis colors estan situats a la tercera fila.
- Un conjunt de colors representatius dels que es poden trobar en escenes naturals. Corresponen als 12 colors que es troben a la primera i a la segona fila.

4.2.2 Gràfiques del comportament dels components cromàtics

Per cada un dels dotze colors estudiats es coneixen les seves coordenades cromàtiques sobre el diagrama cromàtic CIE-1931. Pels espais que no es defineixen sobre el sistema de coordenades XYZ de la CIE, aquestes coordenades han estat convertides a l'espai RGB i, a partir d'aquest, als espais cromàtics desitjats. A efectes il·lustratius s'ha escollit treballar a l'espai RGB de la CIE, deixant els espais RGB propis de les càmeres pels estudis experimentals. En qualsevol llibre sobre el color es pot trobar la transformació de l'espai RGB a l'espai XYZ (per exemple a [WS82]):

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (4.5)$$

i, per tant, per passar d' XYZ a RGB només cal invertir la matriu anterior:

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1.1456 & -0.4344 & -0.2268 \\ -0.2496 & 1.3257 & 0.0362 \\ 0.0025 & -0.0134 & 1.0097 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (4.6)$$

No.	Nom	CIE (1931)			Notació Munsell	
		x	y	Y	Hue	Value/Chroma
1	Pell fosca	0.400	0.350	10.1	3 YR	3.7/3.2
2	Pell clara	0.377	0.345	35.8	2.2 YR	6.47/4.1
3	Blau cel	0.247	0.251	19.3	4.3 PB	4.95/5.5
4	Fulles	0.337	0.422	13.3	6.7 GY	4.2/4.1
5	Blau flor	0.265	0.240	24.3	9.7 PB	5.47/6.7
6	Verd-blau	0.261	0.343	43.1	2.5 BG	7/6
7	Taronja	0.506	0.407	30.1	5 YR	6/11
8	Blau-violeta	0.211	0.175	12.0	7.5 PB	4/10.7
9	Vermell suau	0.453	0.306	19.8	2.5 R	5/10
10	Violeta	0.285	0.202	6.6	5 P	3/7
11	Groc-verd	0.380	0.489	44.3	5 GY	7.1/9.1
12	Taronja-groc	0.473	0.438	43.1	10 YR	7/10.5
13	Blau	0.187	0.129	6.1	7.5 PB	2.9/12.7
14	Verd	0.305	0.478	23.4	0.25 G	5.4/8.65
15	Vermell	0.539	0.313	12.0	5 R	4/12
16	Groc	0.448	0.470	59.1	5 Y	8/11.1
17	Magenta	0.364	0.233	19.8	2.5 RP	5/12
18	Cian	0.196	0.252	19.8	5 B	5/8
19	Blanc	0.310	0.316	90.0	N	9.5/
20	Neutre 8	0.310	0.316	59.1	N	8/
21	Neutre 6.5	0.310	0.316	36.2	N	6.5/
22	Neutre 5	0.310	0.316	19.8	N	5/
23	Neutre 3.5	0.310	0.316	9.0	N	3.5/
24	Negre	0.310	0.316	3.1	N	2/

Taula 4.1: Valors dels colors de la Macbeth COLORCHECKER [Mac91].

No.	Nom	CIE xyz (1931)			CIE rgb (1931)		
		x	y	z	r	g	b
7	Taronja	0.506	0.407	0.087	0.383	0.416	0.084
8	Blau-violeta	0.211	0.175	0.614	0.027	0.202	0.618
9	Vermell suau	0.453	0.306	0.241	0.332	0.301	0.240
10	Violeta	0.285	0.202	0.514	0.122	0.215	0.517
11	Groc-verd	0.380	0.489	0.131	0.193	0.558	0.128
12	Taronja-groc	0.473	0.438	0.090	0.331	0.465	0.086
13	Blau	0.187	0.129	0.685	0.000	0.149	0.690
14	Verd	0.305	0.478	0.217	0.092	0.566	0.214
15	Vermell	0.539	0.313	0.149	0.447	0.286	0.147
16	Groc	0.448	0.470	0.082	0.290	0.515	0.077
17	Magenta	0.364	0.233	0.404	0.224	0.232	0.406
18	Cian	0.196	0.252	0.552	0.000	0.305	0.555

Taula 4.2: Coordenades xyz i rgb dels dotze colors centrals de la *Macbeth* COLORCHECKER [Mac91].

A la taula 4.2 es mostren les coordenades obtingudes per cada un d'aquests dotze colors.

Per cada un dels dotze colors s'ha estudiat el comportament previst d'un sensor respecte a variacions en la intensitat de l'illuminant i com aquest comportament afecta a un conjunt de components cromàtics. S'ha suposat que el sensor té un rang dinàmic entre $[0, 255]$ per cada un dels tres components RGB . Les figures 4.2 a 4.25 mostren les gràfiques obtingudes. En aquestes, el nom dels colors són els introduïts a la taula 4.1, i a més, per facilitar la seva localització, s'hi ha afegit la fila i la columna que ocupen dins de la carta de colors *Macbeth* COLORCHECKER.

Per cada un dels 12 colors, s'han analitzat 15 components cromàtics: 8 pertanyents a espais de color definits per la CIE, i els altres 7 pertanyents a espais definits per diversos autors.

Pel que fa als components d'espais que no són de la CIE, cada una de les dotze figures conté vuit gràfiques, el contingut de les quals es mostra a la taula 4.3.

Cal comentar que de les set transformacions, només dues les dues últimes no són estables a les variacions en la intensitat i s'hi han posat com a exemple il·lustratiu. Per

la resta de transformacions, es pot veure que mantenen el comportament estable per l'interval d'intensitats en el que cap dels tres components RGB està saturat.

$R(i)G(i)B(i)$	
$H = \tan^{-1} \frac{\sqrt{3}(G - B)}{(R - G) + (R - B)}$	
$H =$	$\begin{cases} \pi/3(G - B)/\Delta & \text{si } R = \max \\ \pi/3(2 + (B - R)/\Delta) & \text{si } G = \max \\ \pi/3(4 + (R - G)/\Delta) & \text{si } B = \max \end{cases}$
$\Delta = \max\{R, G, B\} - \min\{R, G, B\}$	
$r = R/(R + G + B)$	
$g = G/(R + G + B)$	
$b = B/(R + G + B)$	
$S = 1 - \frac{\min\{R, G, B\}}{(R + G + B)/3}$	
$S = \sqrt{C_1^2 + C_2^2}$	
$S = \frac{\max\{R, G, B\} - \min\{R, G, B\}}{\max\{R, G, B\}}$	
$I1 = (R + G + B)/3$	
$I2 = (R - B)/2$	
$I3 = G/2 - (R + B)/4$	

Taula 4.3: Components cromàtics no pertanyents a models de la CIE estudiats a les figures de les pàgines següents (indicades d'esquerra a dreta i de dalt a baix).

El component de θ és una variable circular i s'expressa entre $[0, 2\pi]$, essent indefinit quan els tres components RGB tenen el mateix valor (que en els exemples següents només es dona quan els tres components queden saturats, situació en la que s'obté el color blanc). La saturació i els components RGB normalitzats es presenten dins l'interval $[0, 1]$. Els components del model $I1I2I3$ es presenten en l'interval $[0, 255]$.

Pels càlculs dels models de color de la CIE, hem considerat que el valor d' Y del blanc de referència és de $Y_n = 255$ i que els seus components cromàtiques són $(x, y) = (0.33, 0.33)$. Amb aquests valors podem calcular les coordenades XYZ del blanc de referència aplicant les relacions següents:

$$X = \frac{Y}{y}x \quad Y = \frac{Y}{y}y \quad Z = \frac{Y}{y}z \quad (4.7)$$

Les coordenades que s'obtenen són $X_n = 255.0, Y_n = 255.0, Z_n = 255.0$. Aquests valors són necessaris per la transformació al model $CIEL^*a^*b^*$. Per la transformació al model $CIEL^*u^*v^*$ cal conèixer les coordenades u, v del blanc de referència i aquestes corresponen, en el nostre cas, a $u_n = 0.2105, v_n = 0.4737$.

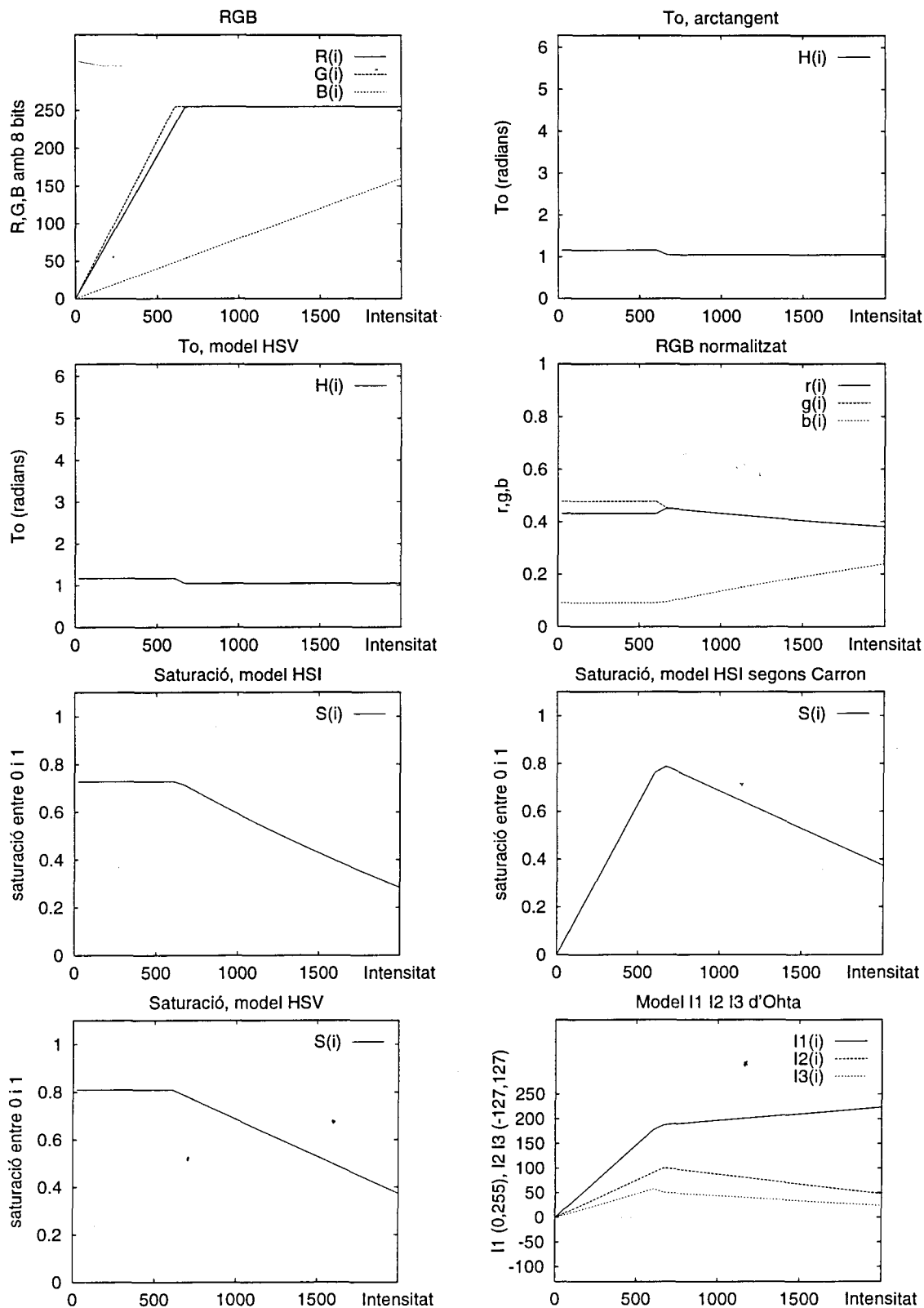


Figura 4.2: Comportament de diferents components cromàtics corresponents al color taronja: $(r, g) = (0.3832, 0.4164)$, (fila 2, columna 1 de la *Macbeth COLORCHECKER*).

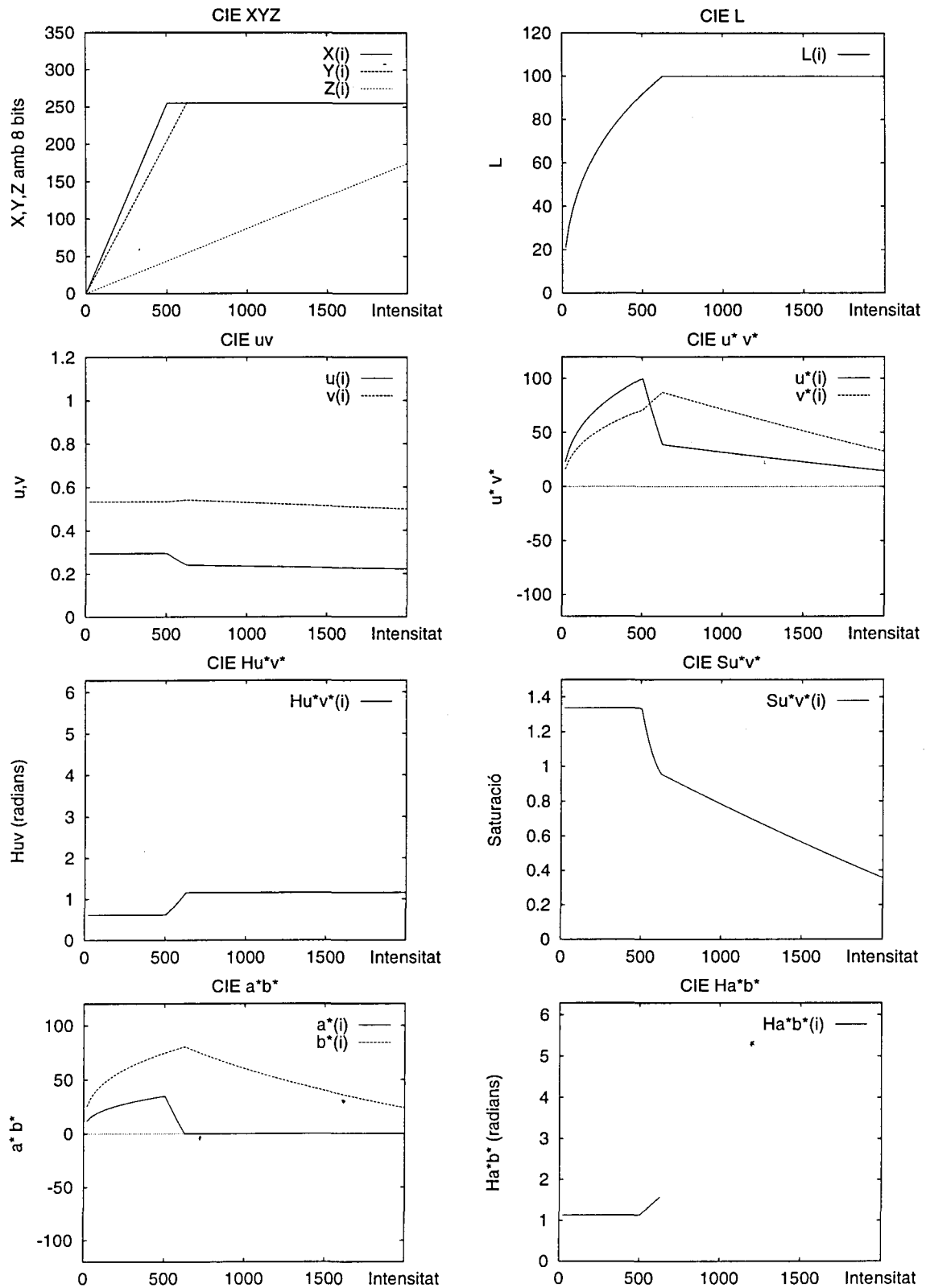


Figura 4.3: Comportament de diferents components cromàtics de la CIE corresponents al color taronja: $(x, y) = (0.5060, 0.4070)$, (fila 2, columna 1 de la *Macbeth COLORCHECKER*).

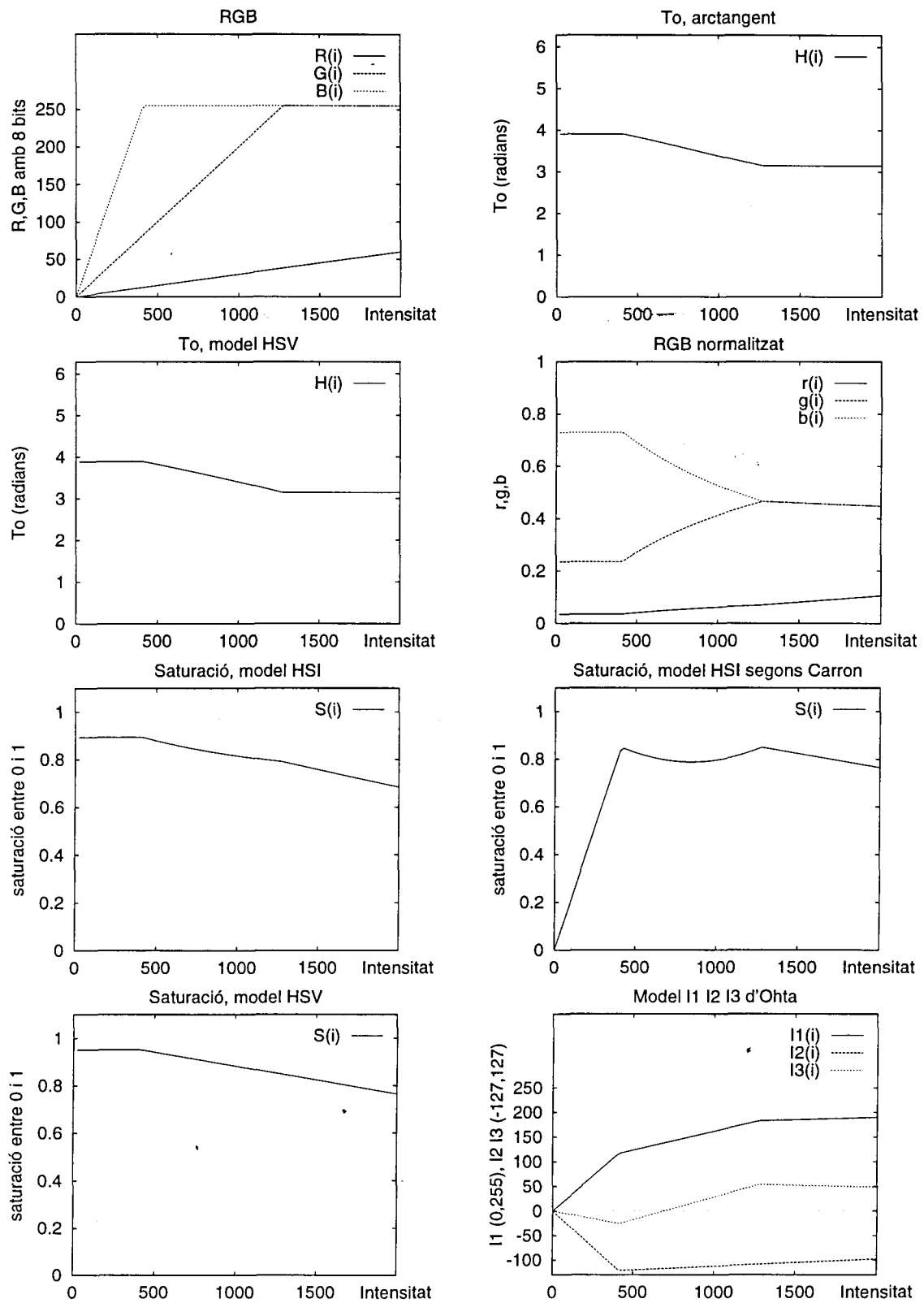


Figura 4.4: Comportament de diferents components cromàtics corresponents al color blau purpuri: $(r, g) = (0.0265, 0.2016)$, (fila 2, columna 2 de la *Macbeth COLORCHECKER*).

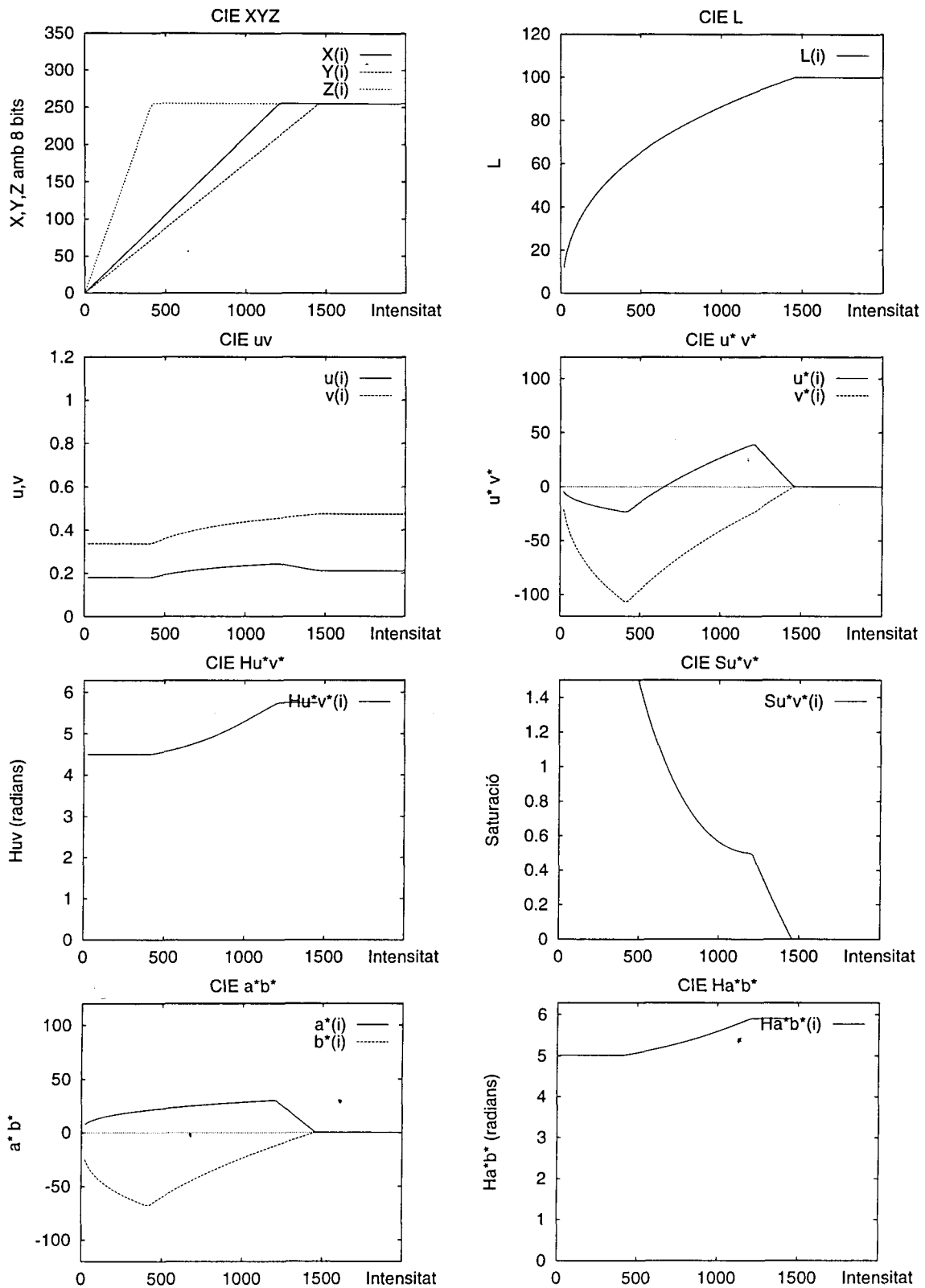


Figura 4.5: Comportament de diferents components cromàtics de la CIE corresponents al color blau purpuri: $(x, y) = (0.2110, 0.1750)$, (fila 2, columna 2 de la *Macbeth COLORCHECKER*).

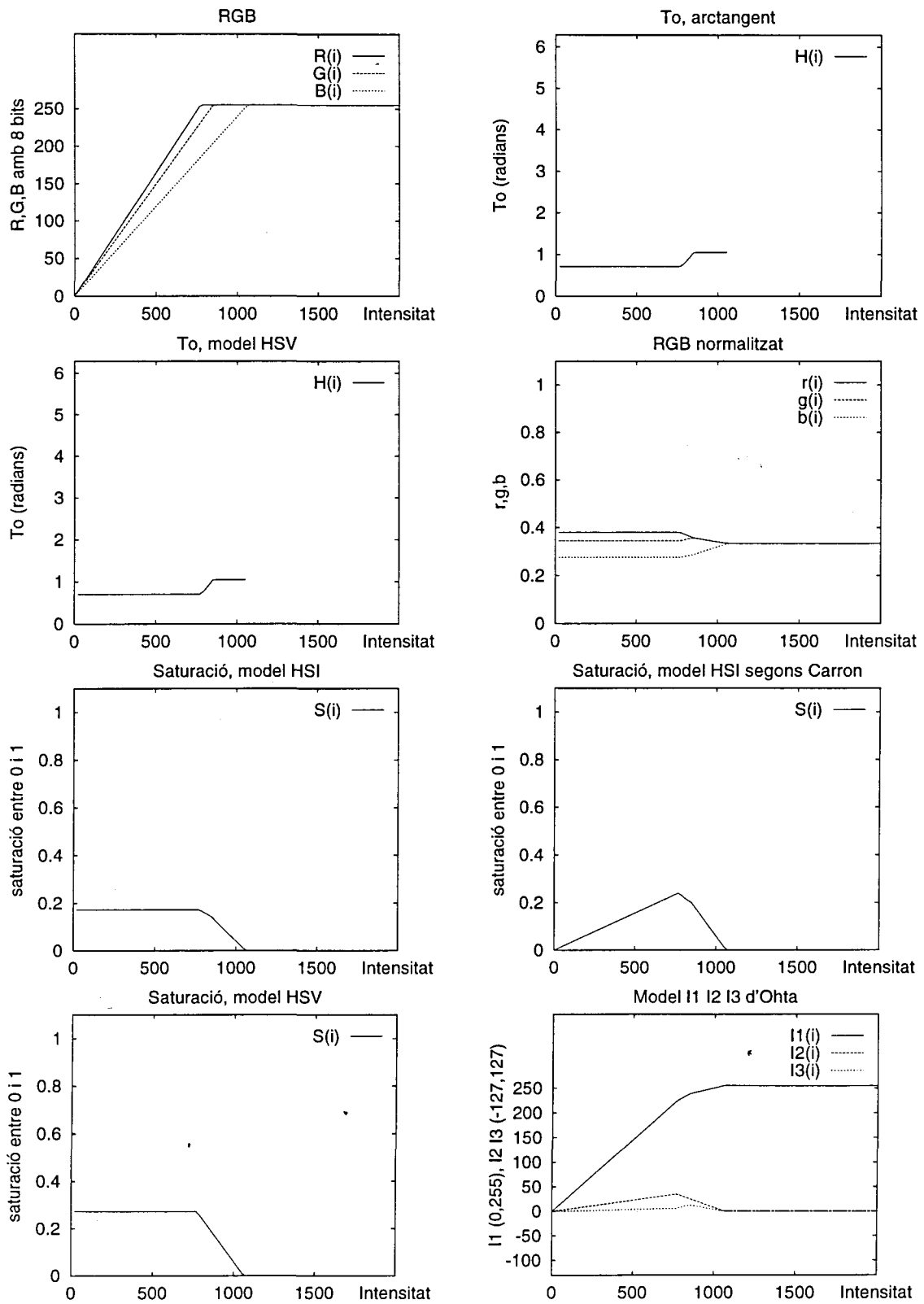


Figura 4.6: Comportament de diferents components cromàtics corresponents al color vermell rosat: $(r, g) = (0.3318, 0.3010)$, (fila 2, columna 3 de la *Macbeth COLORCHECKER*).

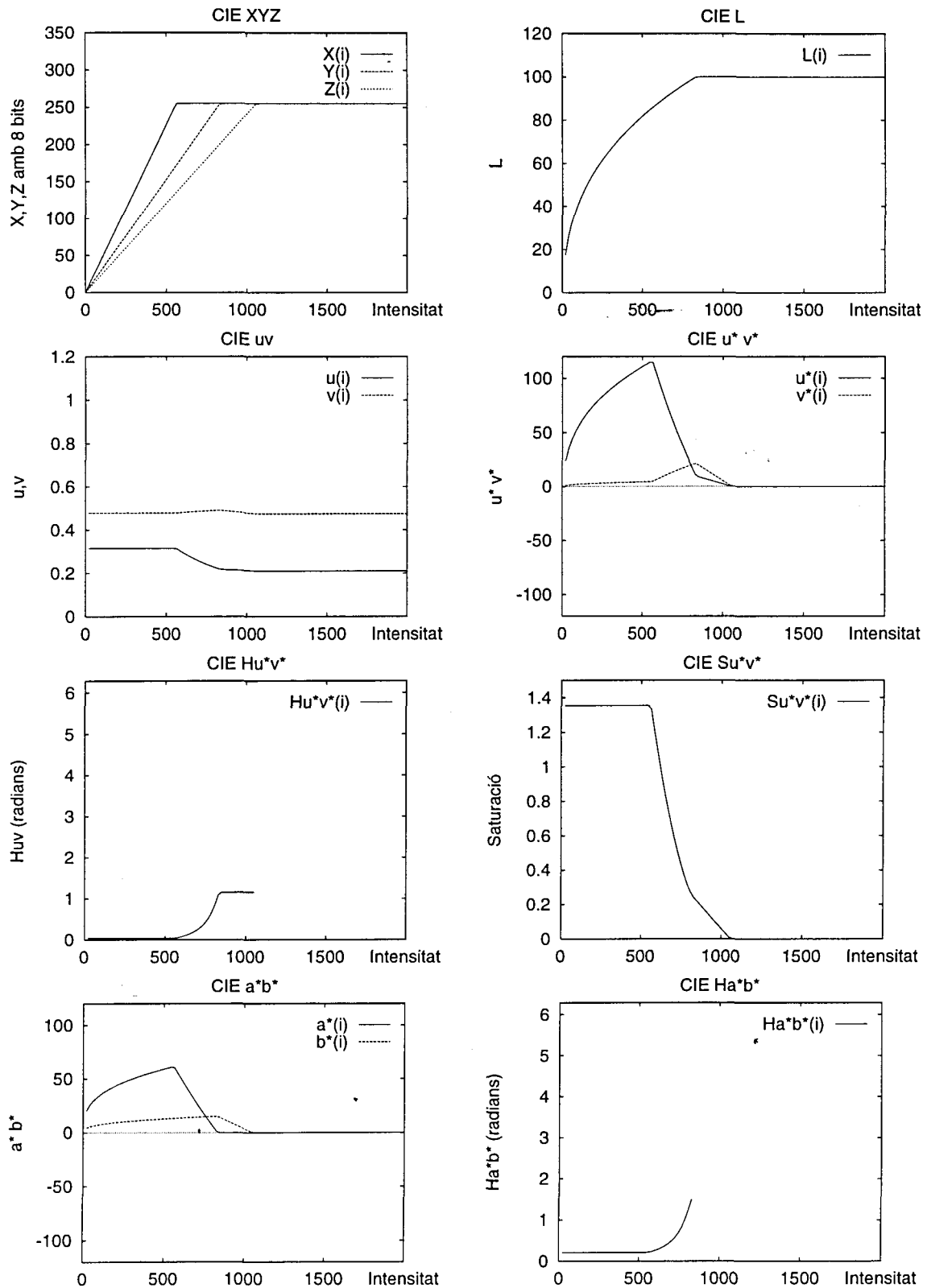


Figura 4.7: Comportament de diferents components cromàtics de la CIE corresponents al color vermell rosat: $(x, y) = (0.4533, 0.3058)$, (fila 2, columna 3 de la Macbeth COLORCHECKER).

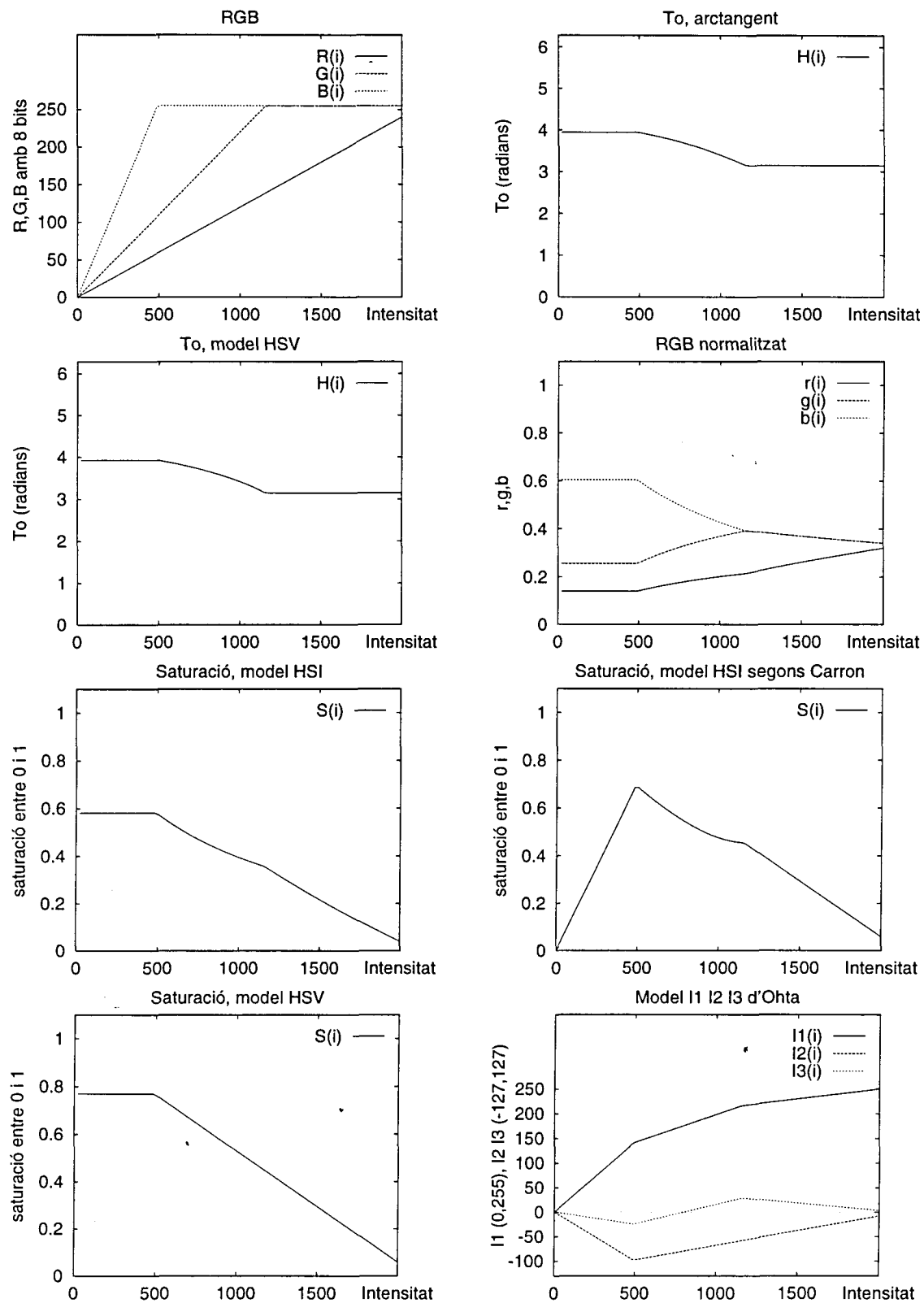


Figura 4.8: Comportament de diferents components cromàtics corresponents al color porpra: $(r, g) = (0.1217, 0.2154)$, (fila 2, columna 4 de la *Macbeth COLORCHECKER*).

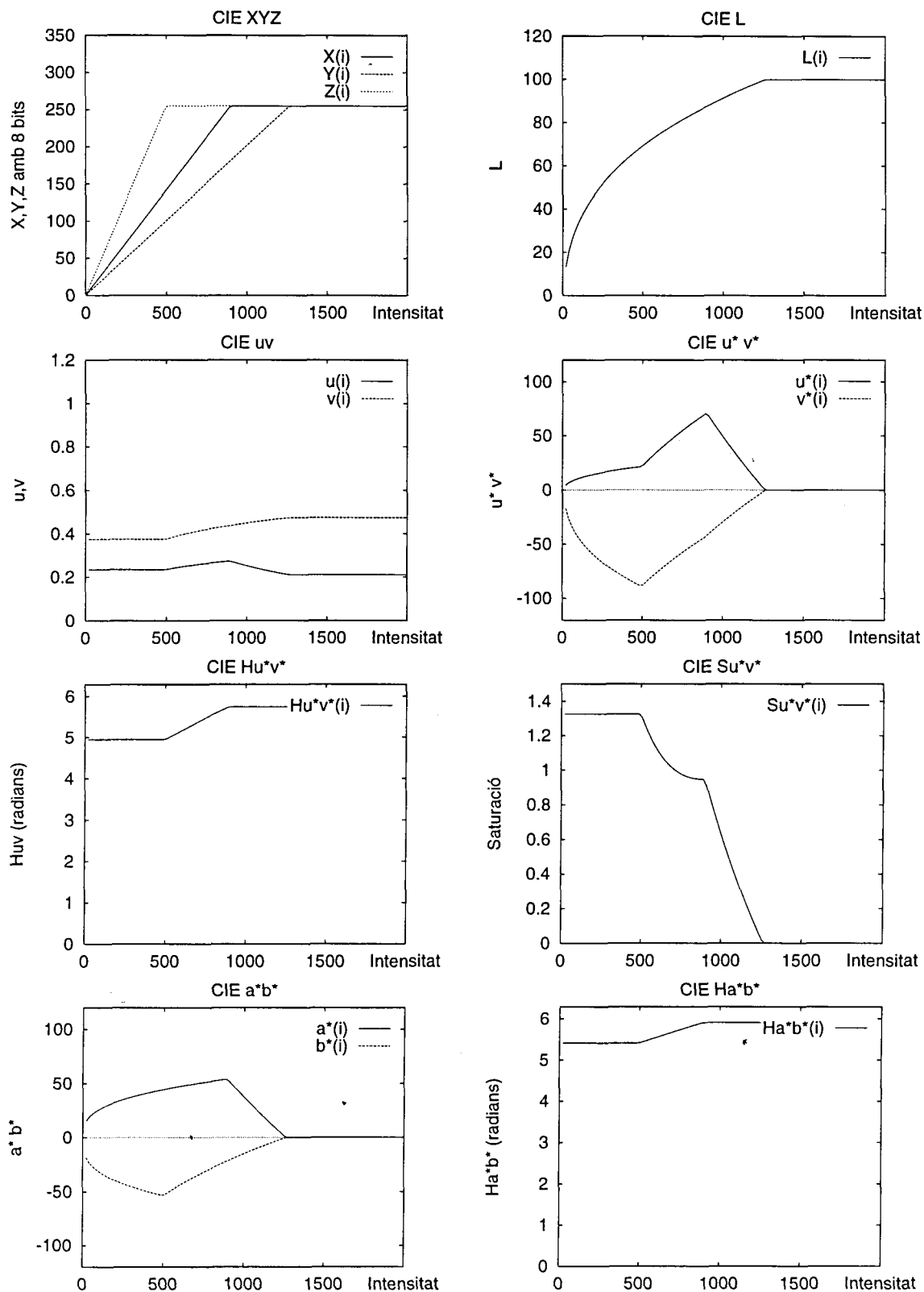


Figura 4.9: Comportament de diferents components cromàtics de la CIE corresponents al color porpra: $(x, y) = (0.2845, 0.2020)$, (fila 2, columna 4 de la *Macbeth COLORCHECKER*).

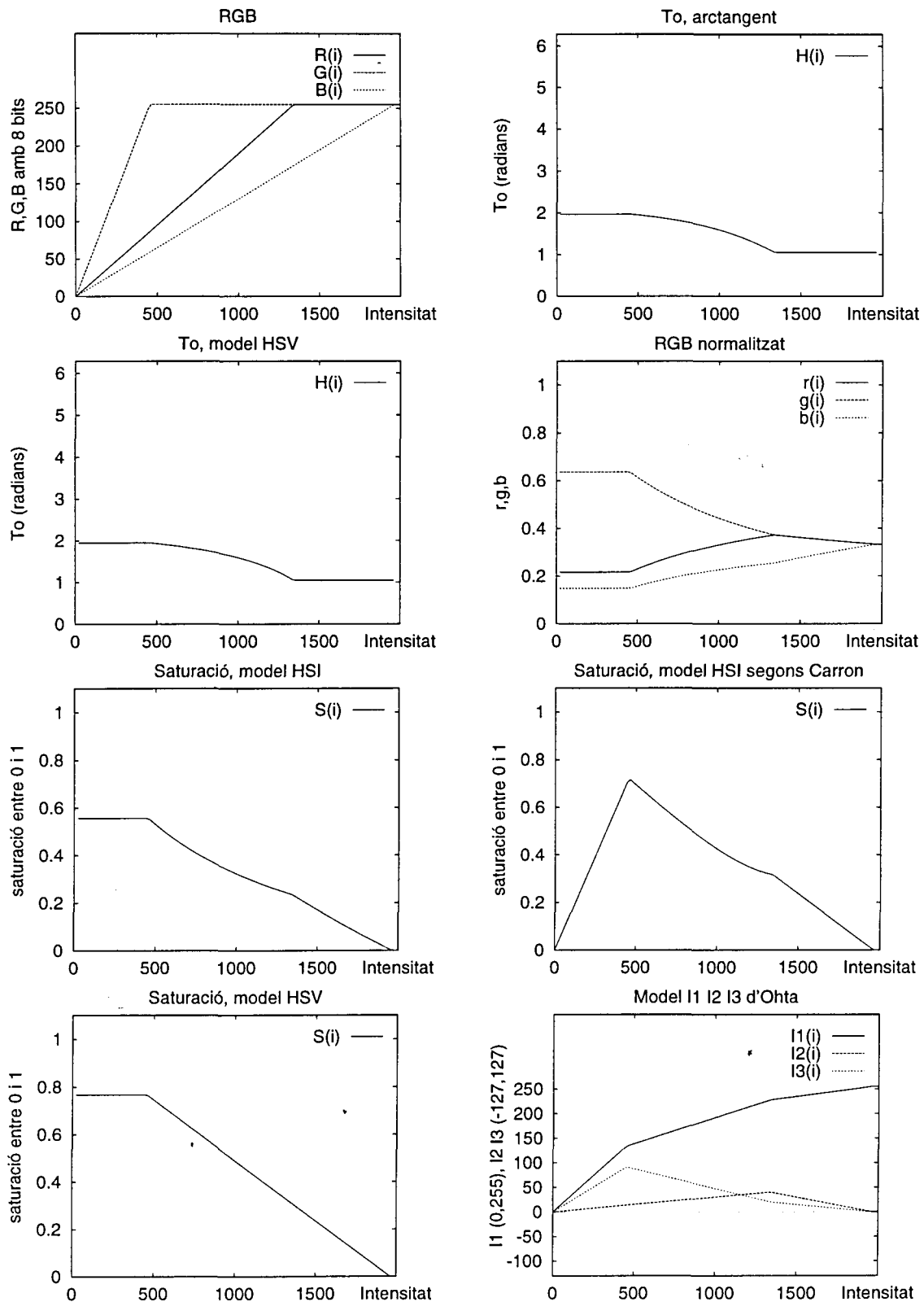


Figura 4.10: Comportament de diferents components cromàtics corresponents al color groc verdós: $(r, g) = (0.1932, 0.5578)$, (fila 2, columna 5 de la *Macbeth COLORCHECKER*).

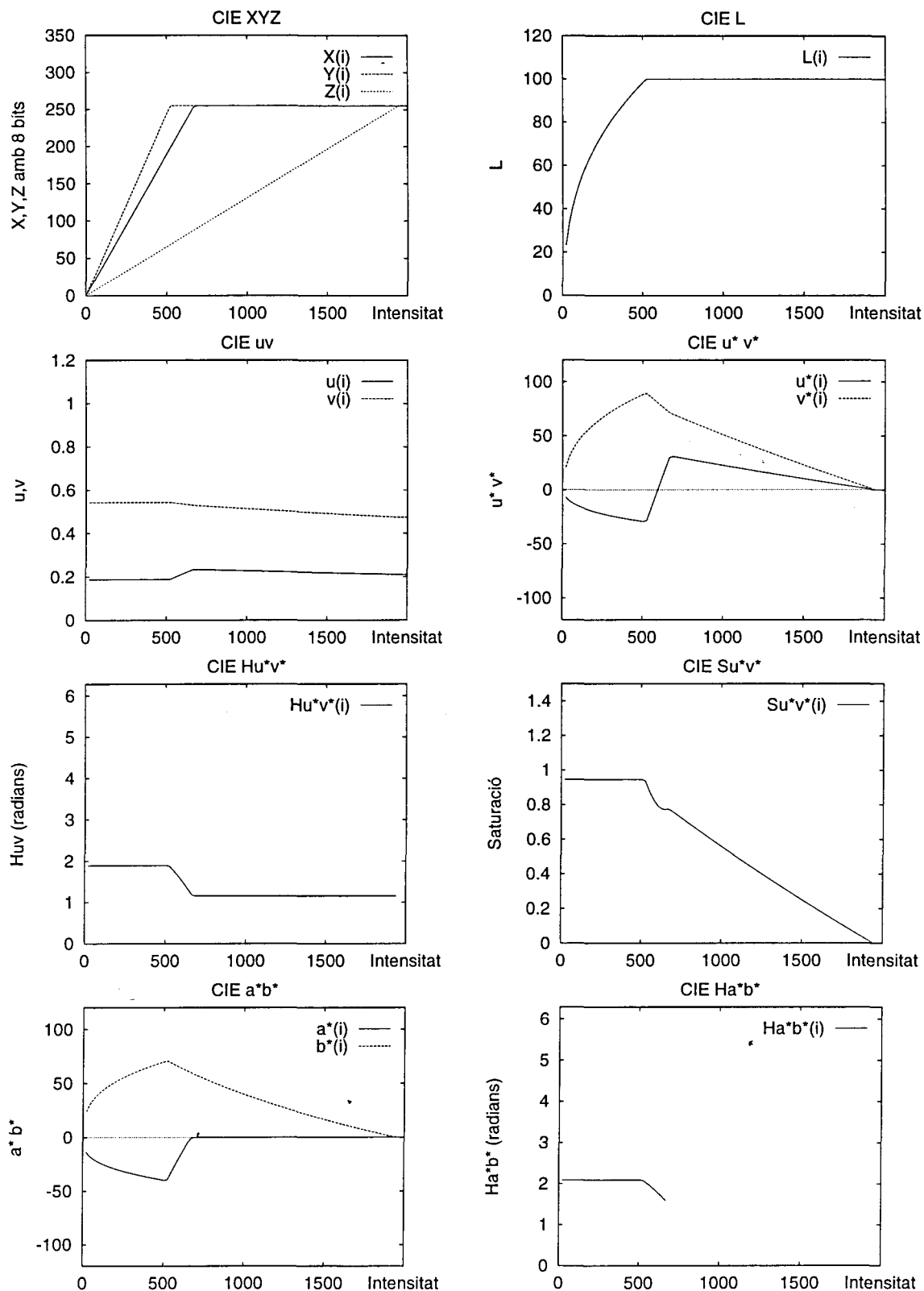


Figura 4.11: Comportament de diferents components cromàtics de la CIE corresponents al color groc verdós: $(x, y) = (0.3800, 0.4887)$, (fila 2, columna 5 de la *Macbeth COLORCHECKER*).

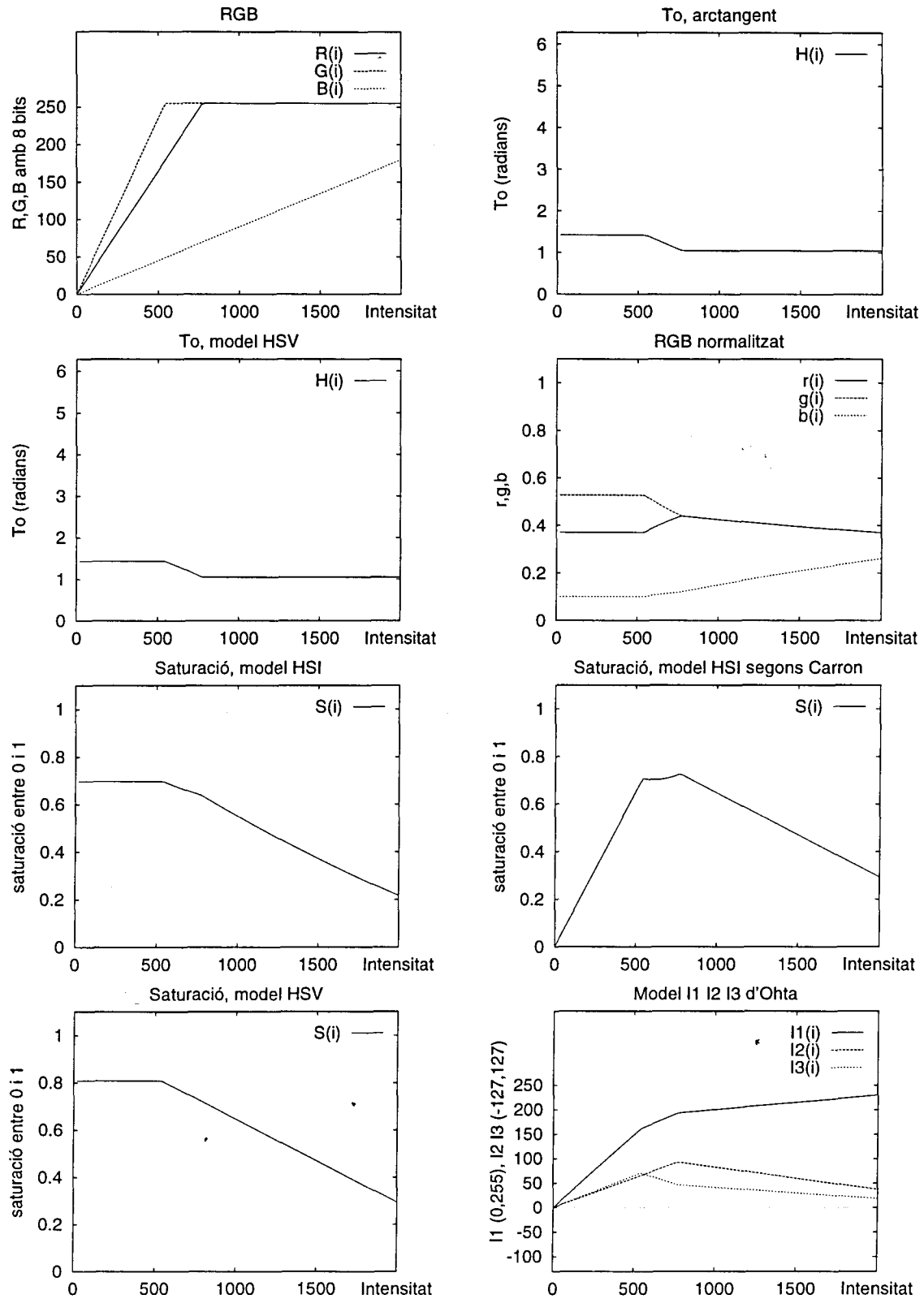


Figura 4.12: Comportament de diferents components cromàtics corresponents al color taronja groguenc: $(r, g) = (0.3314, 0.4652)$, (fila 2, columna 6 de la Macbeth COLORCHECKER).

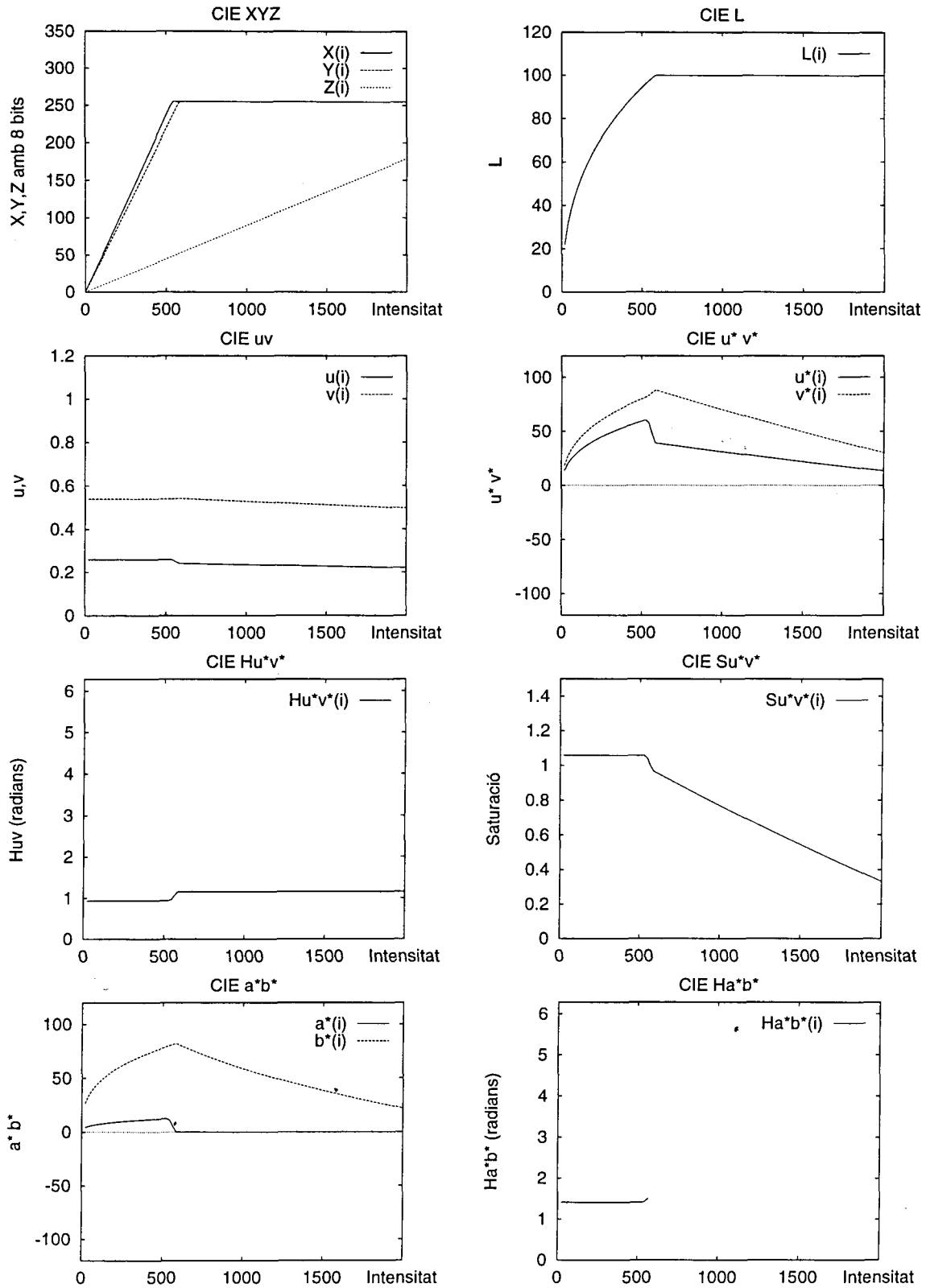


Figura 4.13: Comportament de diferents components cromàtics de la CIE corresponents al color taronja groguenc: $(x, y) = (0.4729, 0.4375)$, (fila 2, columna 6 de la *Macbeth COLORCHECKER*).

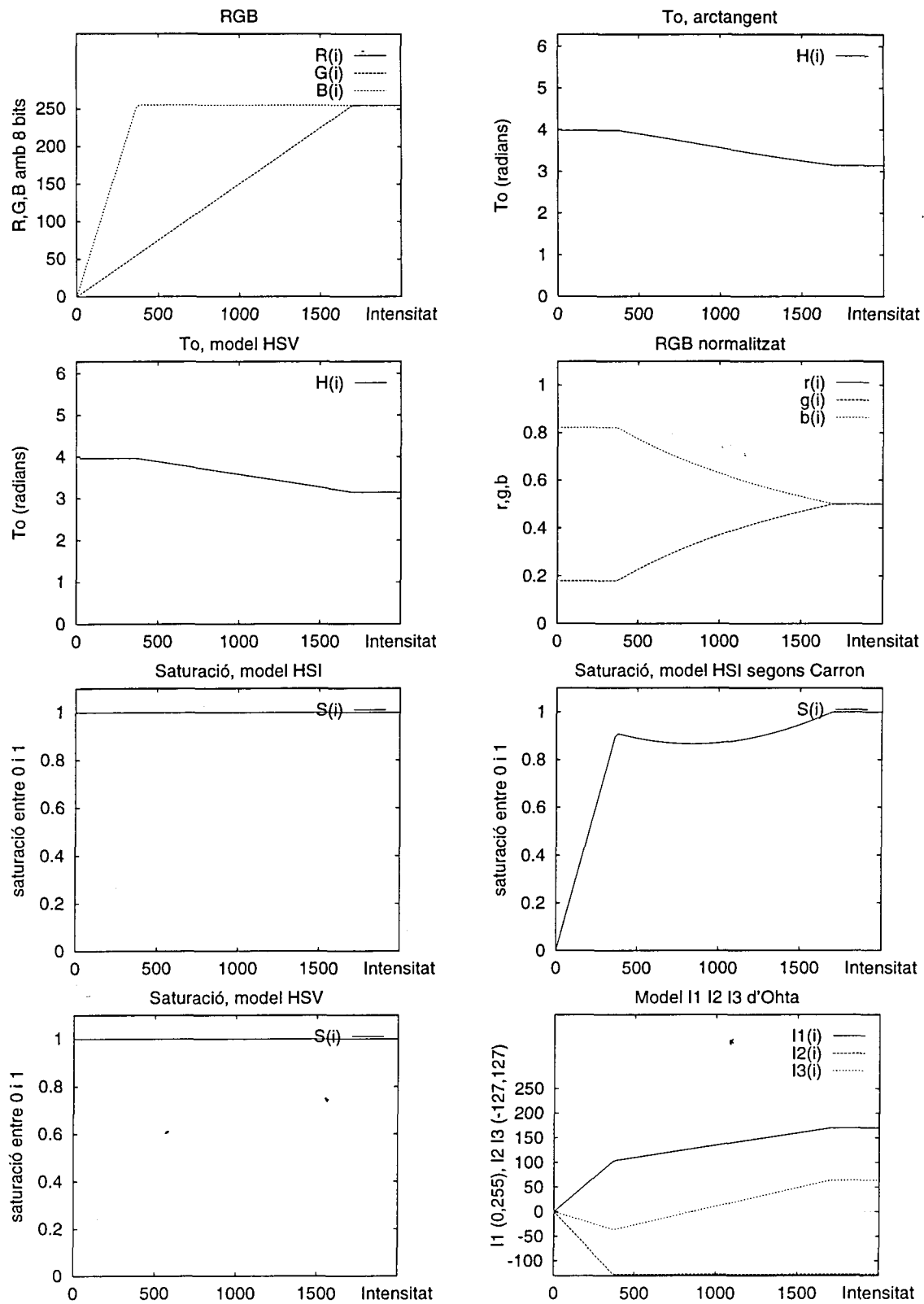


Figura 4.14: Comportament de diferents components cromàtics corresponents al color blau: $(r, g) = (0.0026, 0.1486)$, (fila 3, columna 1 de la *Macbeth COLORCHECKER*).

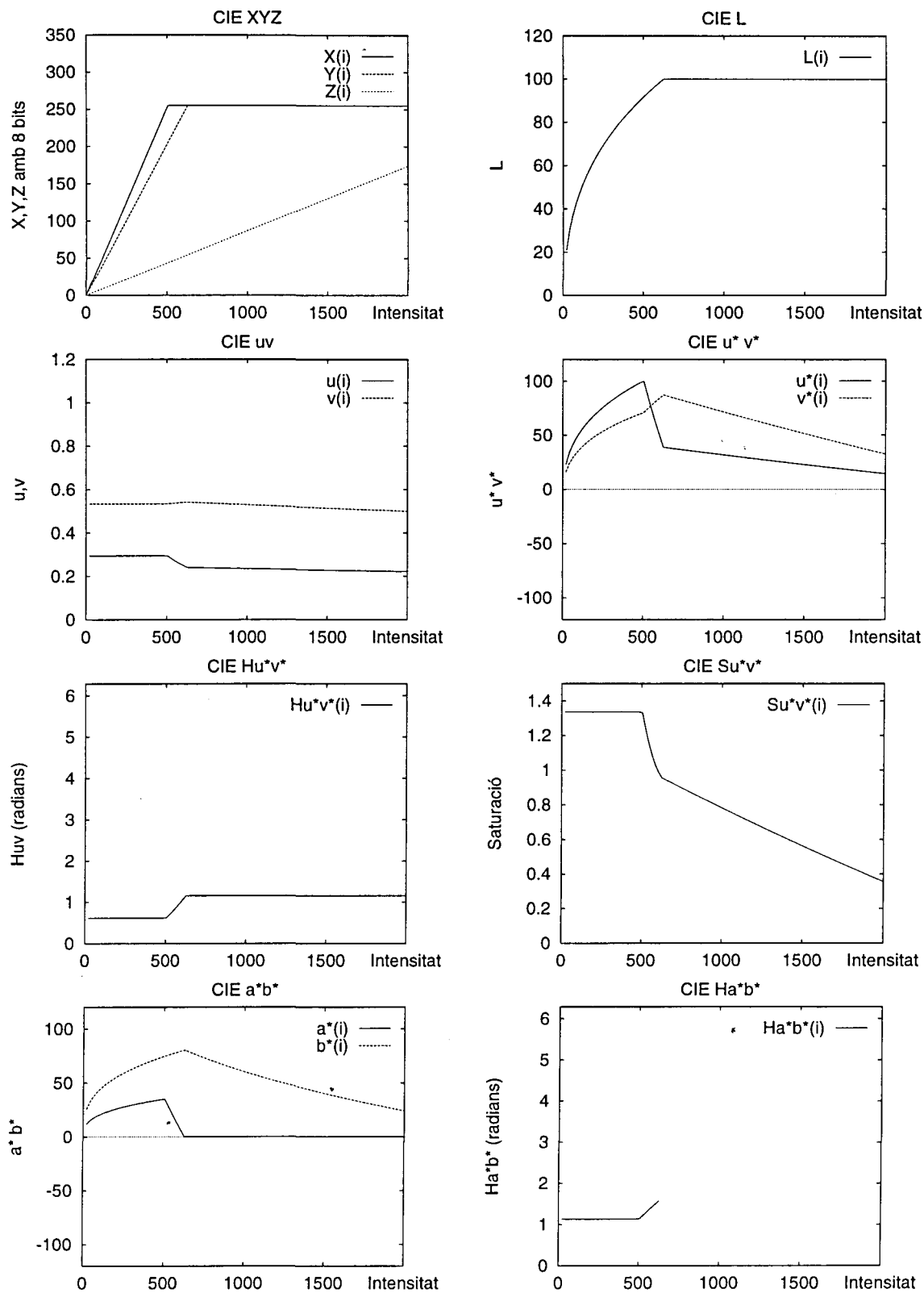


Figura 4.15: Comportament de diferents components cromàtics de la CIE corresponents al color blau: $(x, y) = (0.1866, 0.1285)$, (fila 3, columna 1 de la *Macbeth COLORCHECKER*).

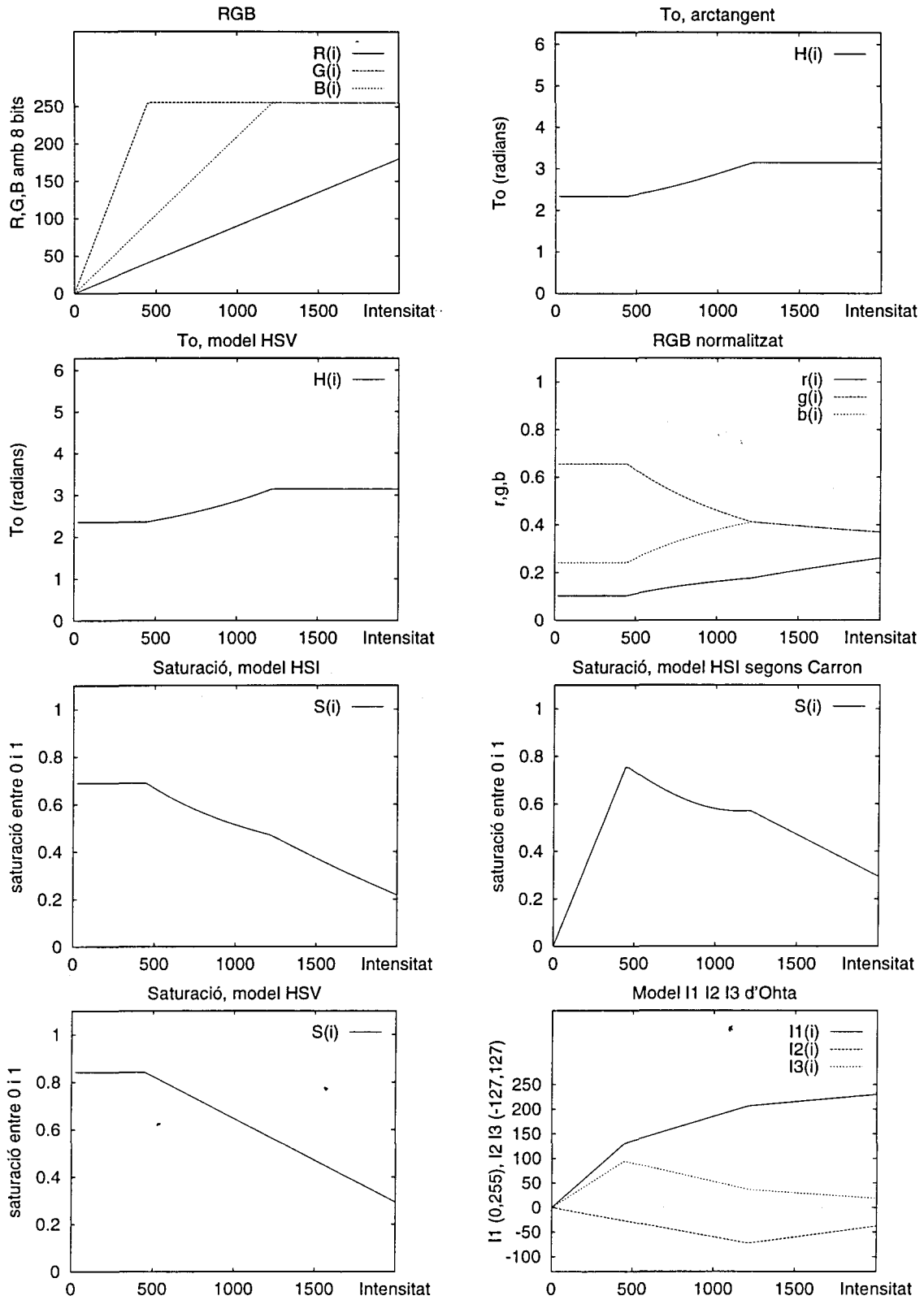


Figura 4.16: Comportament de diferents components cromàtics corresponents al color verd: $(r, g) = (0.0920, 0.5660)$, (fila 3, columna 2 de la *Macbeth COLORCHECKER*).

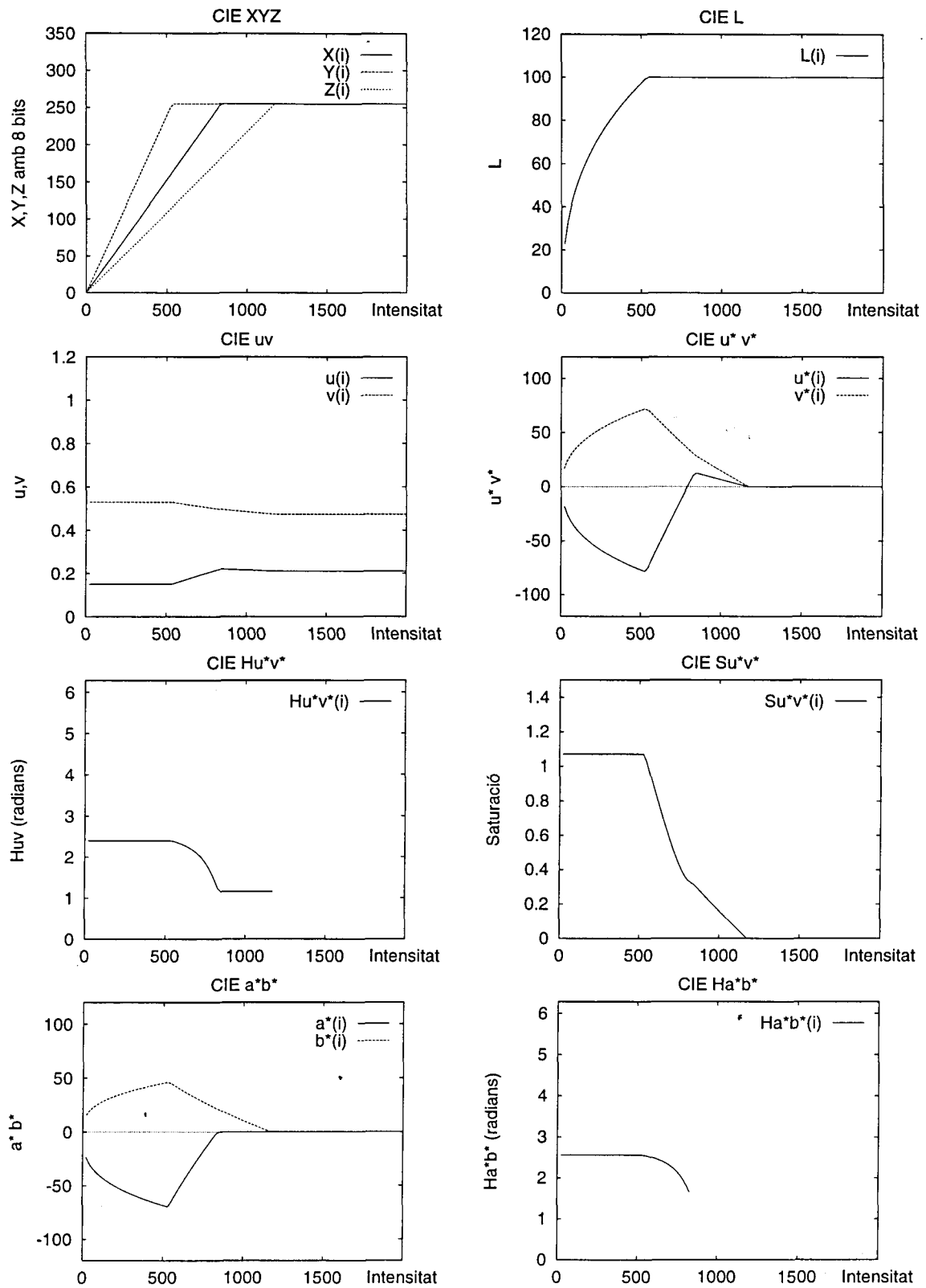


Figura 4.17: Comportament de diferents components cromàtics de la CIE corresponents al color verd: $(x, y) = (0.3046, 0.4782)$, (fila 3, columna 2 de la Macbeth COLORCHECKER).

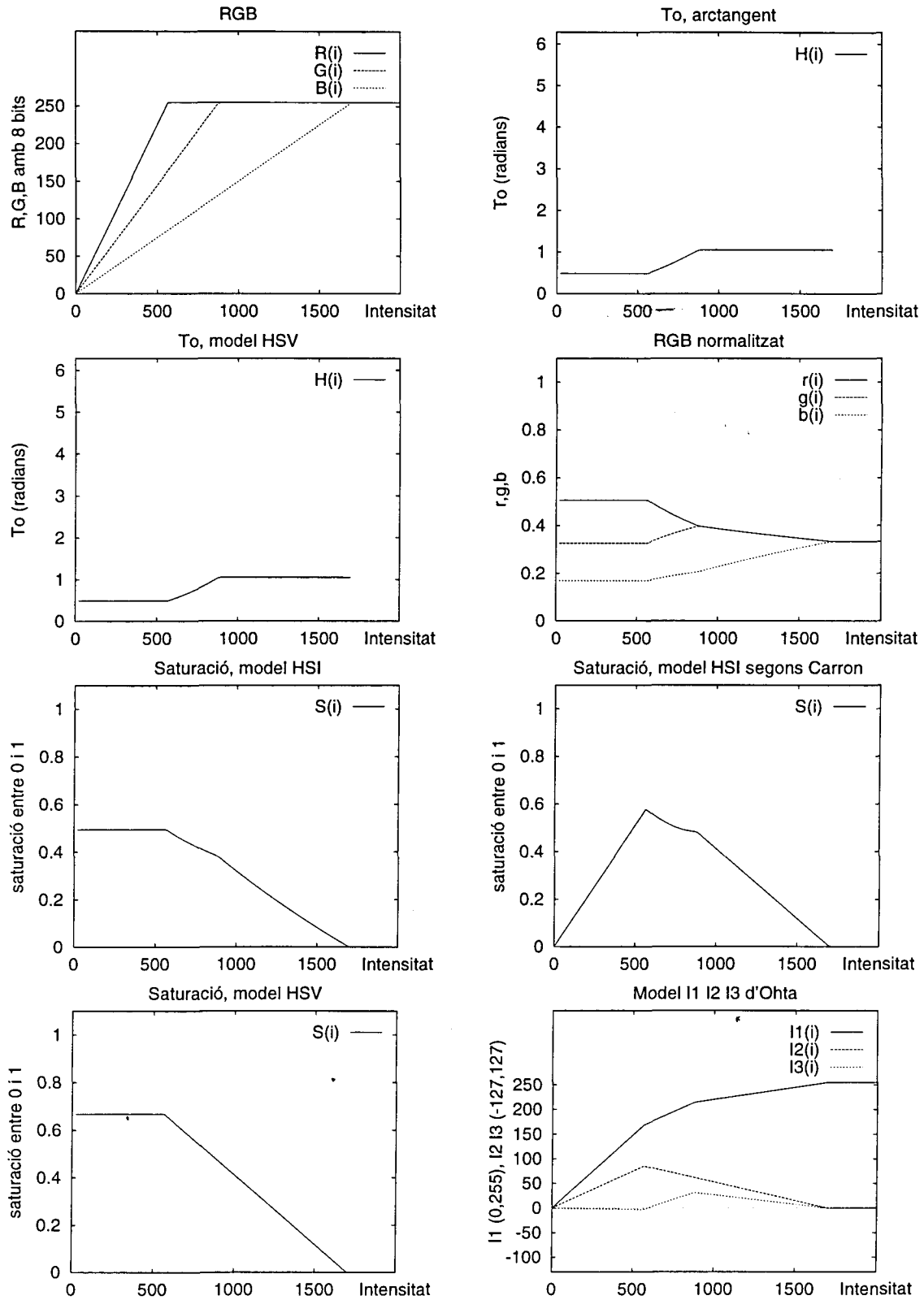


Figura 4.18: Comportament de diferents components cromàtics corresponents al color vermell: $(r, g) = (0.4473, 0.2860)$, (fila 3, columna 3 de la *Macbeth COLORCHECKER*).

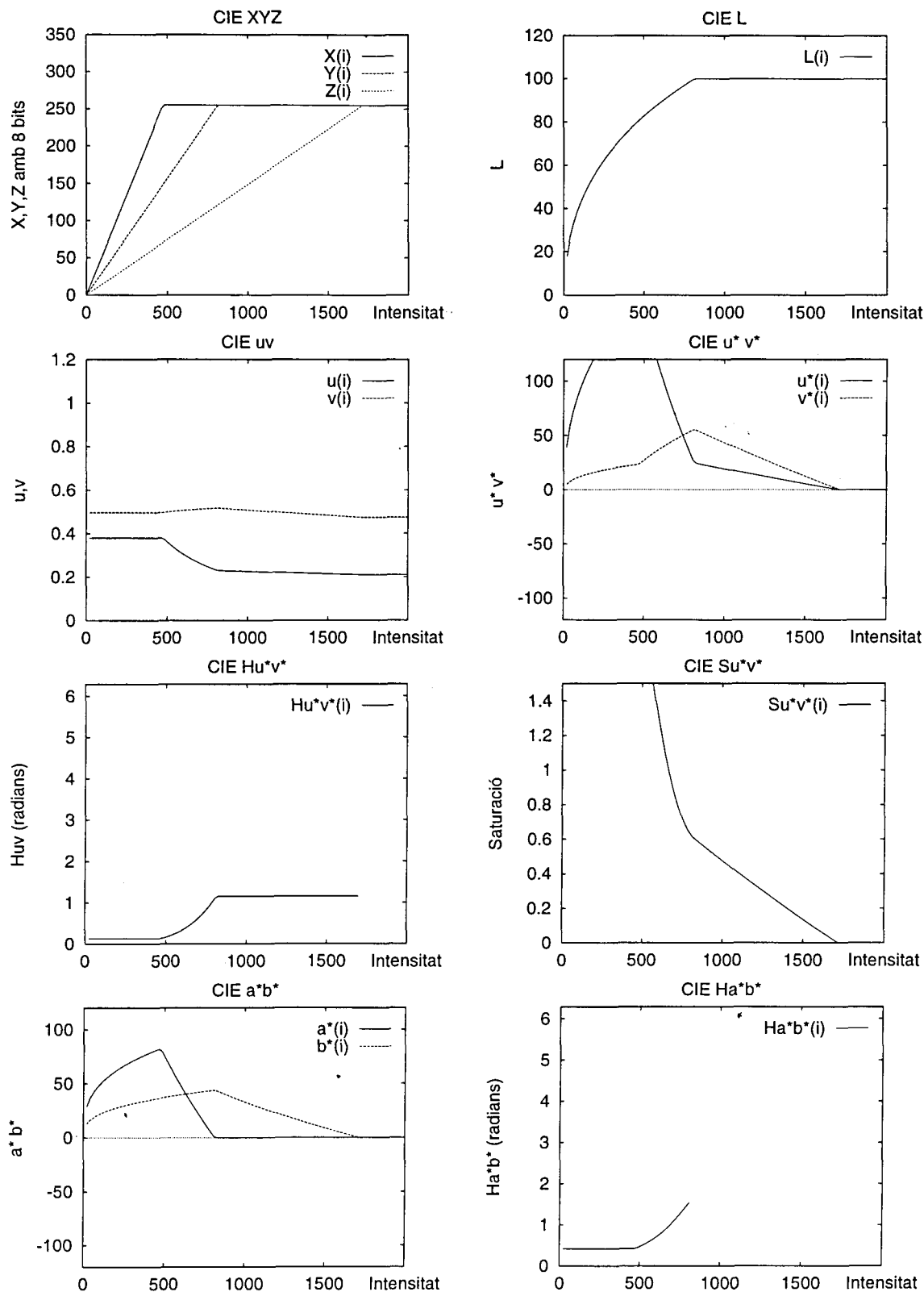


Figura 4.19: Comportament de diferents components cromàtics de la CIE corresponents al color vermell: $(x, y) = (0.5385, 0.3129)$, (fila 3, columna 3 de la *Macbeth COLORCHECKER*).

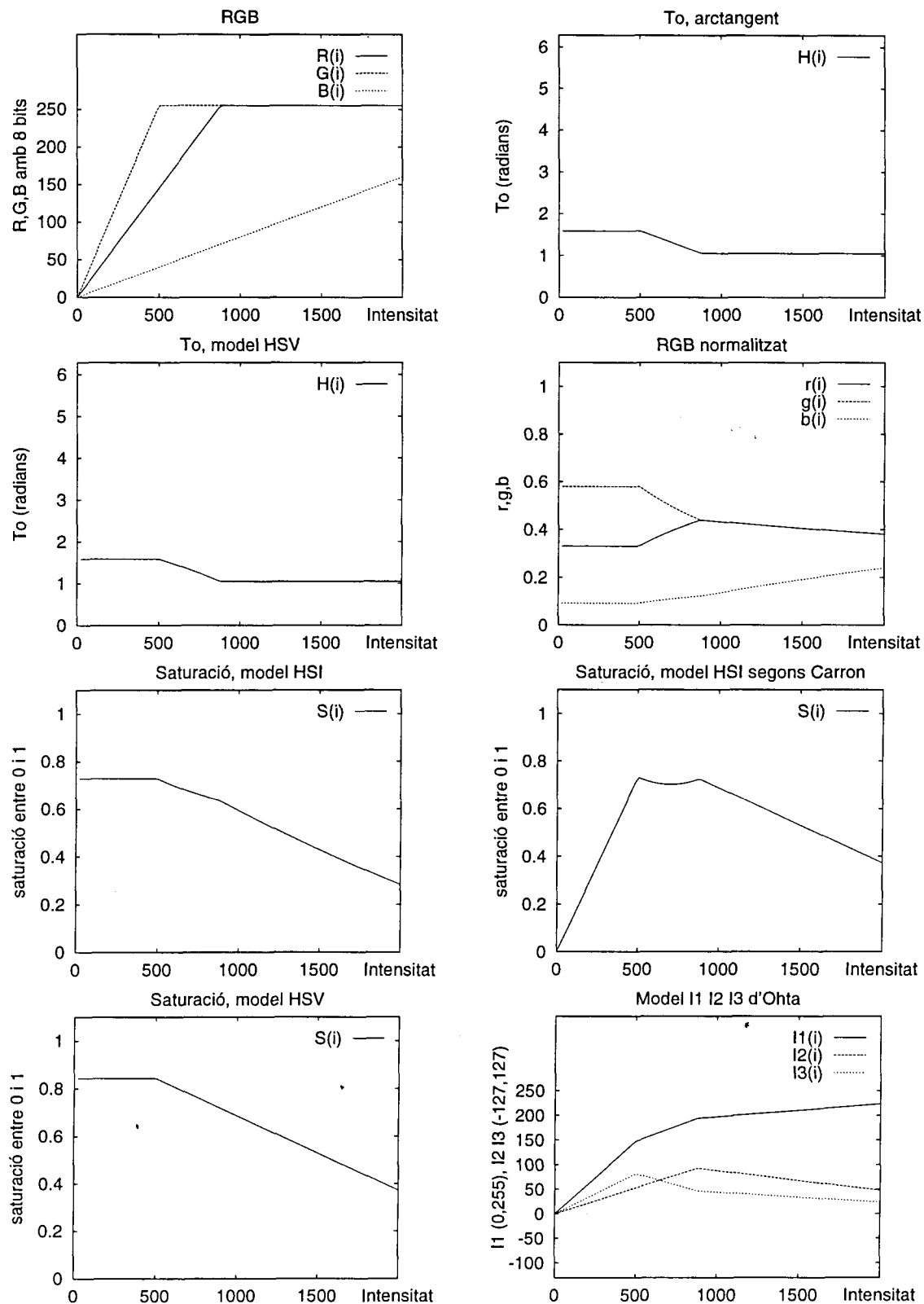


Figura 4.20: Comportament de diferents components cromàtics corresponents al color groc: $(r, g) = (0.2904, 0.5146)$, (fila 3, columna 4 de la *Macbeth COLORCHECKER*).

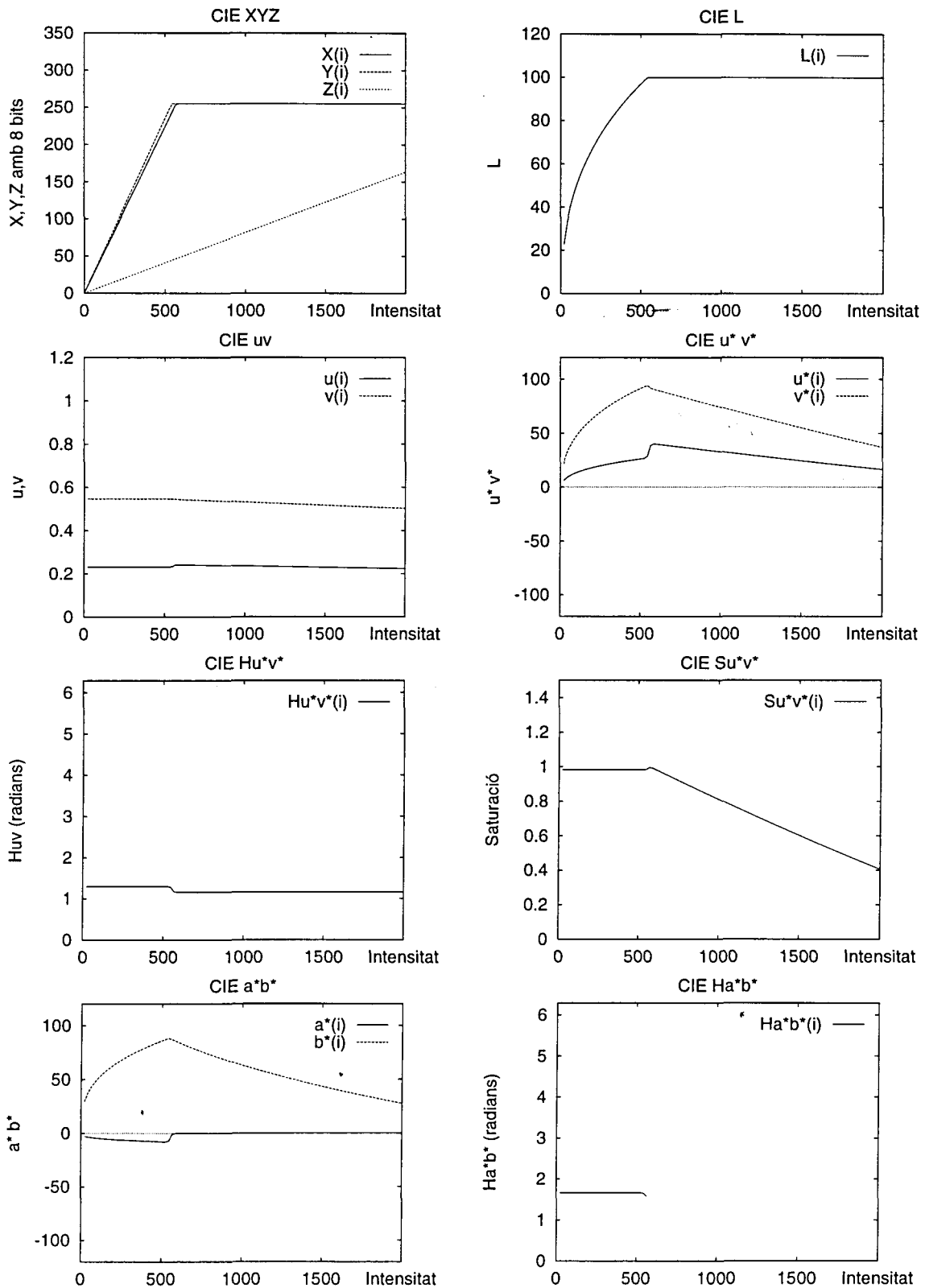


Figura 4.21: Comportament de diferents components cromàtics de la CIE corresponents al color groc: $(x, y) = (0.4480, 0.4703)$, (fila 3, columna 4 de la Macbeth COLORCHECKER).

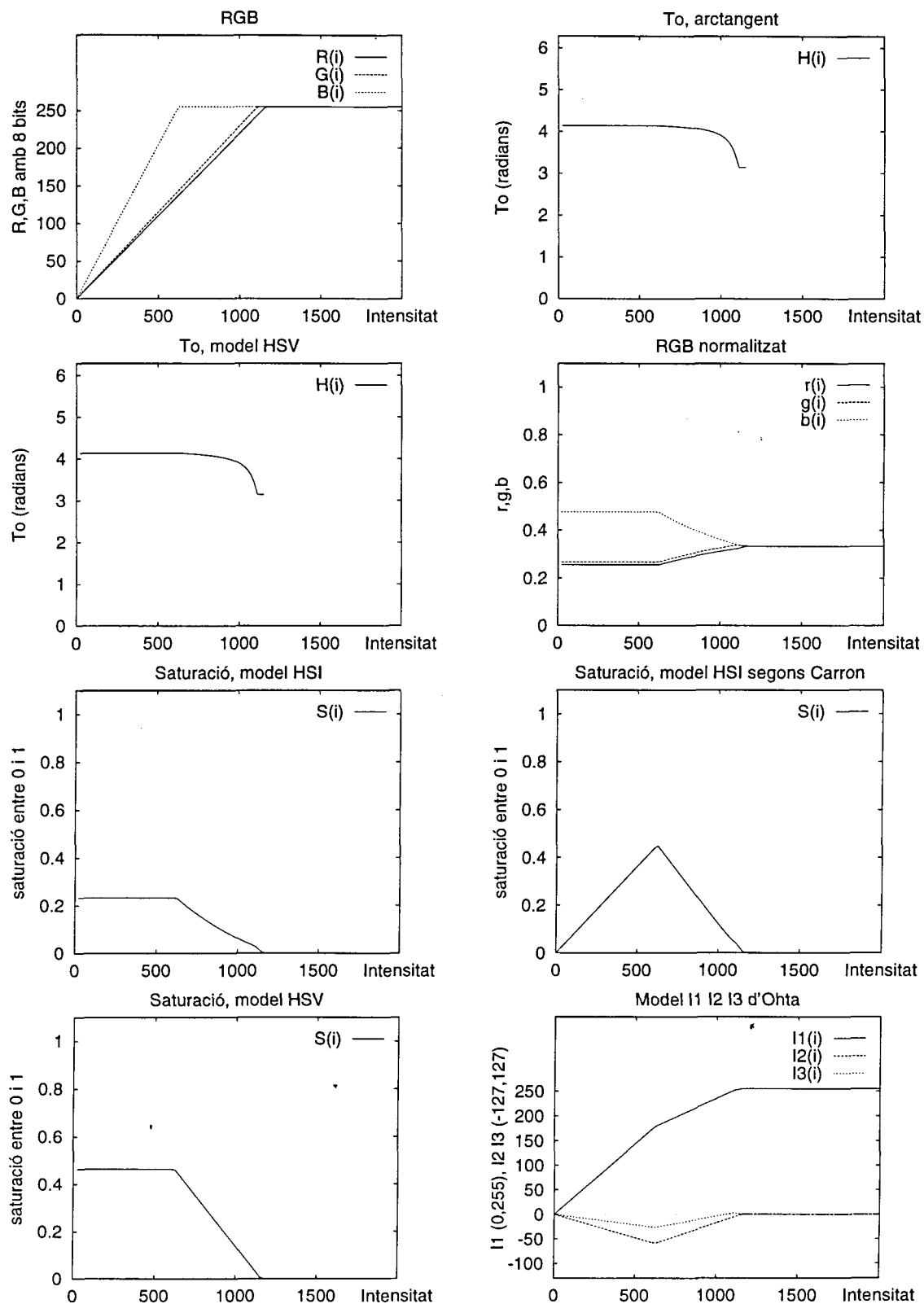


Figura 4.22: Comportament de diferents components cromàtics corresponents al color magenta: $(r, g) = (0.2238, 0.2321)$, (fila 3, columna 5 de la *Macbeth COLORCHECKER*).

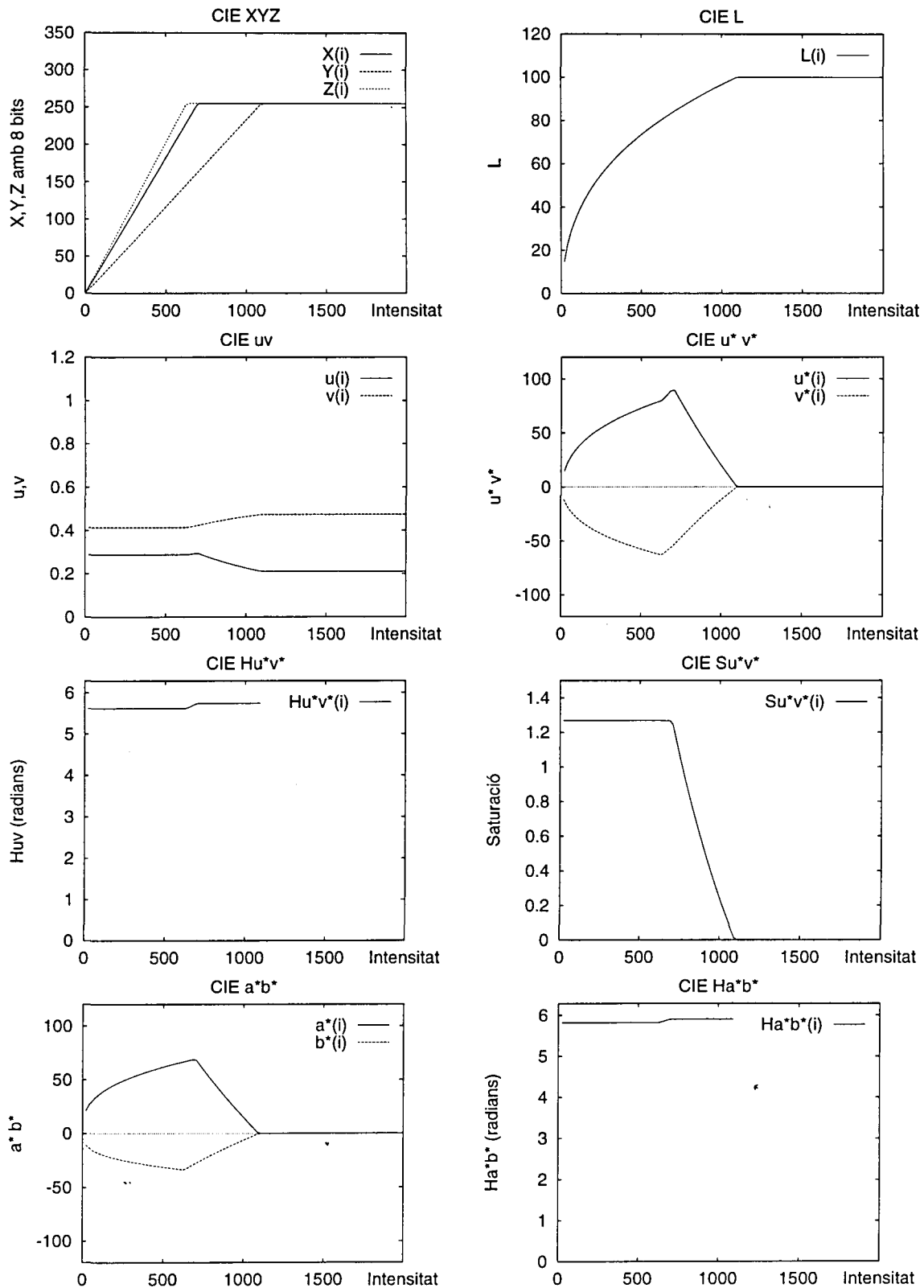


Figura 4.23: Comportament de diferents components cromàtics de la CIE corresponents al color magenta: $(x, y) = (0.3635, 0.2325)$, (fila 3, columna 5 de la Macbeth COLORCHECKER).

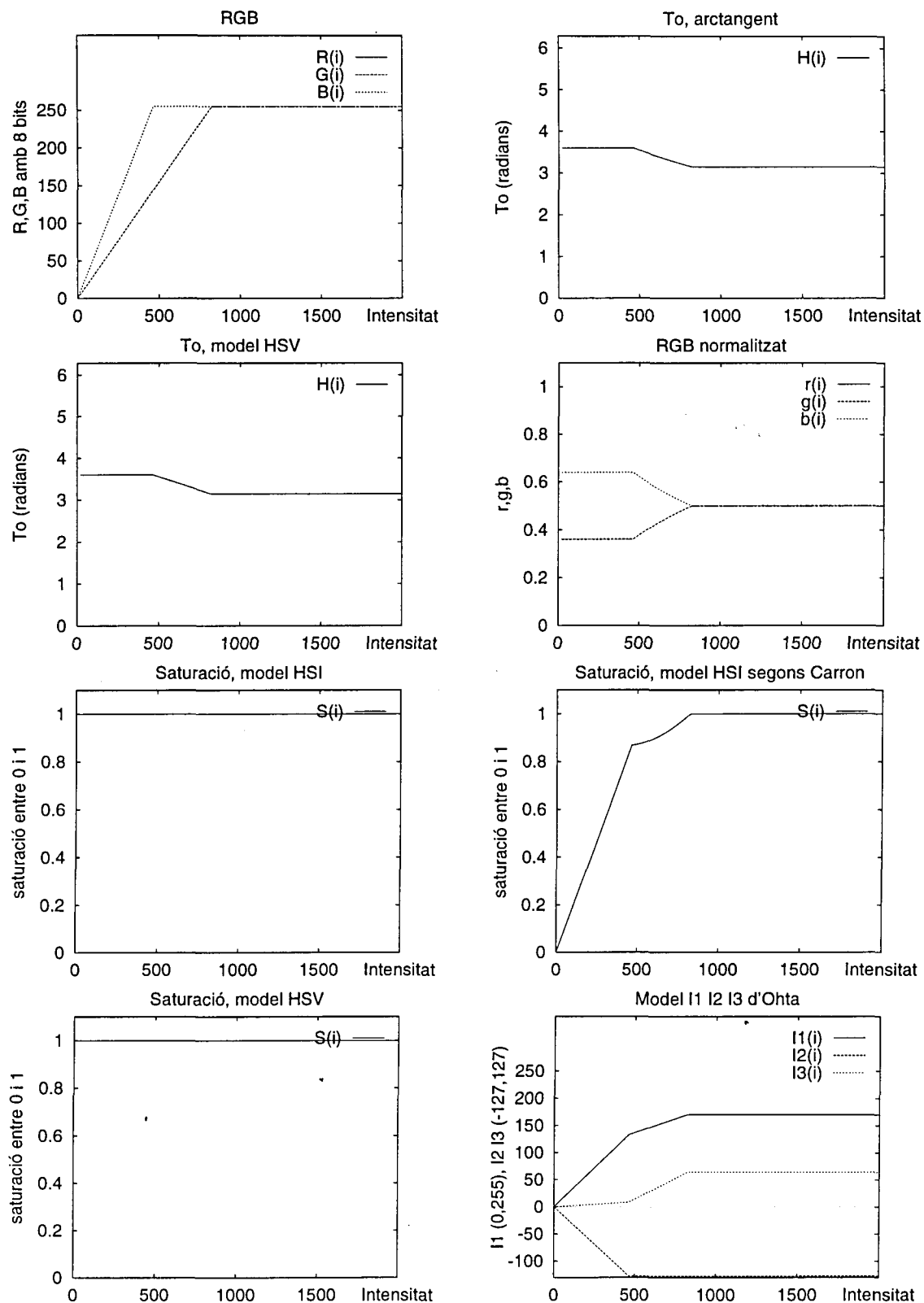


Figura 4.24: Comportament de diferents components cromàtics corresponents al color cian: $(r, g) = (0.0000, 0.3051)$, (fila 3, columna 6 de la *Macbeth COLORCHECKER*).

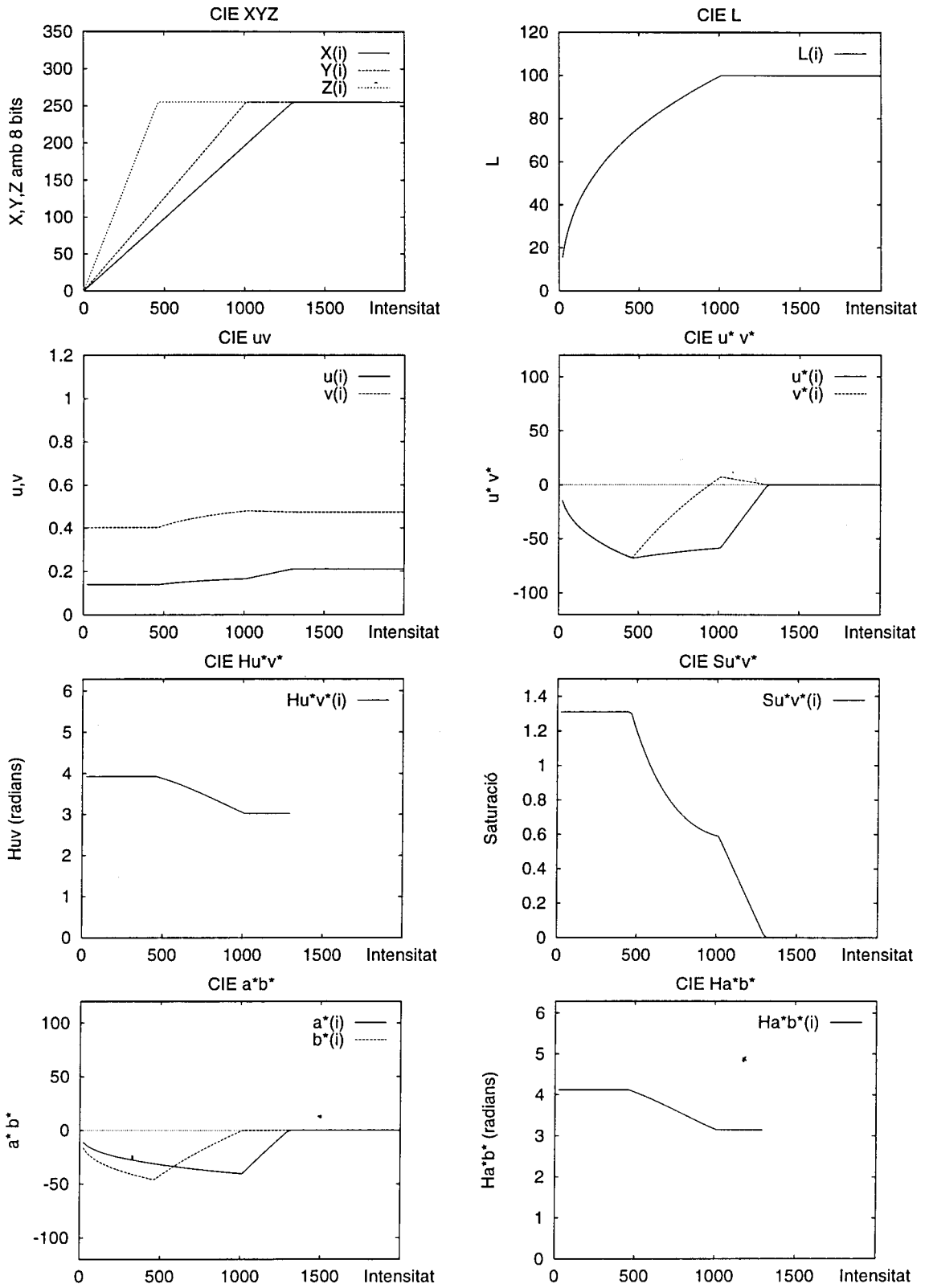


Figura 4.25: Comportament de diferents components cromàtics de la CIE corresponents al color cian: $(x, y) = (0.1958, 0.2519)$, (fila 3, columna 6 de la *Macbeth COLORCHECKER*).

A partir de les gràfiques anteriors es pot veure que alguns dels components cromàtics representats són invariants a l'increment de la intensitat. Pel que fa als espais no definits per la CIE, es tracta dels components corresponents al to, dels components de l'espai *RGB* normalitzat i de dos dels tres components que representen la saturació. Pel que respecta als espais de la CIE, són invariants els components u i v , la saturació definida sobre l'espai *CIELUV* i el to definit sobre els espais *CIELAB* i *CIELUV*.

Si mirem, per exemple, el comportament dels components cromàtics corresponents al color *groc verdós* (figures 4.10 i 4.11) es pot veure que:

- Dels tres components *RGB*, el predominant és el verd, seguit del vermell; essent el menys predominant el blau. La relació R:G:B és de 0.33:1.00:0.23.
- Per una intensitat de 455 el component verd queda saturat. El vermell queda saturat a 1340, mentre que el blau i queda a 1960. Aquestes tres intensitats marquen quatre intervals $\{[0, 455], (455, 1340], (1340, 1960], (1960, \infty)\}$, en els quals la resta de components cromàtics tenen comportaments diversos.
- Pels components de to es pot veure que aquest comença amb un valor d'aproximadament 2 radians, corresponent al valor de to d'un color situat entre els grocs i els verds. Aquest valor és constant en tot el primer interval $[0, 455]$, però a partir d'aquí comença a disminuir, i decreix durant tot el segon interval $(455, 1340]$, fins arribar a un valor d'aproximadament 1.05 radians. Cal fixar-se en que si es discretizés el to utilitzant 8 bits, això implicaria que es passaria d'un valor de to de 79 a un valor de 42, que és una variació important. Pels altres dos intervals restants, el valor del to es manté constant, però és molt diferent del valor que hauria de tenir.
- El cas de l'espai *RGB* normalitzat es pot veure que els tres components tenen un comportament diferent per cada un dels quatre intervals, essent constants i amb el valor correcte només al primer interval. A partir del segon interval, els tres components tendeixen cap al valor $(1/3, 1/3, 1/3)$, que és el que correspon als píxels acromàtics. Aquesta afirmació és vàlida per tots els colors representats, excepte aquells en els que manca algun dels tres components R, G, B com per exemple el color blau (figura 4.14) o cian (figura 4.24).
- Pel que fa a les saturacions es pot veure que n'hi ha dues d'estables al primer interval i que llavors el valor de la saturació va disminuint fins arribar a 0 quan els tres components estan saturats (color blanc). N'hi ha una que no és estable a cap interval, i es pot veure que al principi, va creixent a mesura que van creixent

els valors d'*RGB* i que, a partir del segon interval també comença a decreixer fins arribar a 0.

- En l'espai *I1I2I3* es pot veure que els components *I1*, *I2*, *I3* no són invariants als canvis en la intensitat: *I1* augmenta fins que els tres components queden saturats, *I2* és creixent fins que queda el component verd saturat, i *I3* és creixent mentre no hi cap component saturat, començant a decreixer quan el vermell queda saturat. Els components *I2* i *I3* arriben al valor 0 quan els tres components *R*, *G*, *B* prenen el mateix valor (que en aquest cas correspon al color blanc assolit quan queden tots tres saturats).
- Analitzant les gràfiques corresponents als components d'espais definits per la CIE es pot veure que els que són invariants són els components *u*, *v*, el to definit sobre el *CIELUV* i *CIELAB* i la saturació definida sobre l'espai *CIELUV*. La invariança es manté en l'interval en el que cap dels tres components *X*, *Y*, *Z* ha assolit el seu valor màxim. El component de lluminància *L* no és invariant i, tal i com es veu, la forma de la gràfica reflexa el fet que *L* està definida a partir d'una arrel cúbica.

Si ens fixem ara amb el color *blau* (figures 4.14 i 4.15), veiem que:

- No té component vermell. Per això només queden definits tres intervals: $[0, 370]$ on no hi ha cap component saturat, $(270, 1700]$ on està saturat només el blau i $(1700, \infty)$ on estan saturats el blau i el verd. Una cosa semblant passa amb el cian (figures 4.24 i 4.25): a l'interval $[0, 460]$ no hi ha cap component saturat, a l'interval $(460, 834]$ està saturat només el component blau i a l'interval $(834, \infty)$ ho estan el blau i el verd.
- Les dues saturacions que són constants tenen el mateix valor per tots tres intervals, i aquest valor és el màxim que pot tenir la saturació: 1. L'explicació d'això és senzilla si s'observen les expressions pel càlcul de la saturació que es donen a la taula 4.3. Les dues saturacions en qüestió assoleixen el valor 1 quan almenys un dels tres components val 0, i per aquest color concret hi ha un component que sempre val 0. L'altra saturació, per contra, només assoleix el màxim quan al menys dos components queden saturats. Això també es pot observar amb el color *cian* (figura 4.24).
- Pel que respecte al model *I1I2I3* es pot veure que la manca d'un dels components *RGB* provoca que els components *I2* i *I3* no acabin a zero, a diferència del que passa

amb els colors que tenen els tres components. Igual que amb les saturacions, aquest fet també es pot observar amb el color *cian* (figura 4.24).

- S'ha vist fins ara que la manca d'un dels tres components fa que el comportament dels components cromàtics sigui diferent del de la resta de colors. Pel que fa als espais de la CIE, es pot veure que tot i la manca del component vermell, els tres components *XYZ* apareixen en la definició del color. Això permet que, per aquests colors, el comportament dels components cromàtics que s'han definit a partir de l'espai *XYZ* segueixi el mateix patró que en la resta de colors.

4.3 Exemples pràctics

Presentem tot seguit uns quants exemples de l'efecte que té el retall del color no ja des d'un punt de vista teòric, sinó a partir de mesures fetes sobre imatges adquirides amb una càmera.

4.3.1 Colors en superfícies planes

Primer il·lustrem el comportament d'un sensor concret quan s'adquireixen diferents imatges d'una superfície plana d'un mateix color. Les proves s'han fet amb els 12 colors de les dues files centrals de la *Macbeth COLORCHECKER*. El sensor utilitzat ha estat una càmera Sony 3CCD model DXC-930P, amb la correcció gamma desactivada. Com a font d'il·luminació s'ha usat una làmpara halògena de tungstè de 100W i un conjunt de 5 filtres neutres (amb densitats 0.1, 0.2, 0.3, 0.4 i 1.0) per tal d'obtenir diferents intensitats en l'il·luminant sense variar la seva distribució espectral i la intensitat s'ha mesurat amb un luxòmetre. S'han adquirit 15 imatges de cada un dels 12 colors. De cada imatge s'ha extret una mostra del color (un cercle de radi 20 píxels). En la figura 4.26, que mostra els 180 cercles obtinguts, cada fila correspon a una intensitat i cada columna a un color.

Per cada mostra de color i d'intensitat s'ha calculat el valor corresponent a la mitjana aritmètica de cada un dels tres components (vermell, verd i blau) i les corresponents desviacions. La gràfica de dalt a l'esquerra de les figures 4.28 a 4.39 mostren els resultats obtinguts (les desviacions no es mostren a les gràfiques perquè no són significatives). Tal i com es pot veure la resposta de la càmera per cada component és lineal fins que aquest queda saturat.

A la secció 4.2.2 hem descrit com afectava la saturació dels CCDs a la informació cromàtica a partir d'un model teòric. Ara, però, podem veure aquesta distorsió en un



Figura 4.26: Mostres de color extrems d'imatges de la *Macbeth COLOR-CHEKER* variant la intensitat de l'illuminant.

cas real. Per fer-ho, hem transformat les 120 mostres de color als espais estudiats HSI , I_1, I_2, I_3 i RGB normalitzat, representat gràficament les coordenades obtingudes. Les figures 4.28 a 4.39 mostren els valors obtinguts. A cada una de les figures, la gràfica (a) correspon a la representació de les mitjanes dels valors R, G, B obtinguts, la gràfica (b) representa les mitjanes dels valors obtinguts de la transformació a l'espai HSI , la gràfica (c) correspon a les mitjanes dels valors del model I_1, I_2, I_3 i, finalment, la gràfica (d) correspon a les mitjanes dels components de l'espai RGB normalitzat.

Pel que fa al model HSI , cal tenir en compte que el θ és una variable circular i que, com a tal, la seva mitjana aritmètica no la podem calcular simplement com la suma de les mesures dividida pel seu nombre. A continuació exposem com s'ha calculat la mitjana aritmètica pel component del θ .

4.3.1.1 Mitjana i desviació en variables circulars

Una bona referència sobre aquest tema és el llibre de K.V. Mardia [Mar72]. Sense perdre generalitat, es pot suposar que les mesures angulars es prenen sobre un cercle de radi 1, i que estan en l'interval $[0, 2\pi]$. Partim d'una col·lecció de valors angulars $\{\theta_1, \dots, \theta_n\}$,

mesurats respecte al vector OX i en sentit contrari a les agulles del rellotge, tal i com es mostra a la figura 4.27, i anomenem $\{P_1, \dots, P_n\}$ als punts sobre el cercle unitari que corresponen als angles mesurats.

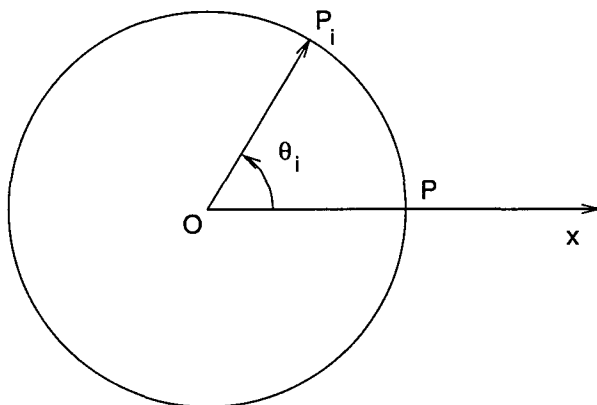


Figura 4.27: Representació de l'observació θ_i en variables angulars

La direcció mitjana. Sigui P_i el punt en el cercle unitari corresponent a l'angle θ_i per $i = 1, \dots, n$. La mitjana aritmètica de la direcció \bar{x}_0 corresponent als $\theta_1, \dots, \theta_n$ es defineix com la direcció del vector suma dels vectors $\overline{OP}_1, \dots, \overline{OP}_n$. Les coordenades cartesianes del vector \overline{OP}_i vénen donades per $(\cos \theta_i, \sin \theta_i)$. El centre de gravetat de tots aquests punts (\bar{C}, \bar{S}) ve definit per les expressions:

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n \cos \theta_i \quad (4.8)$$

$$\bar{S} = \frac{1}{n} \sum_{i=1}^n \sin \theta_i \quad (4.9)$$

Si definim \bar{R} com el mòdul del vector (\bar{C}, \bar{S}) :

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2} \quad (4.10)$$

podem obtenir \bar{x}_0 deslliurant-lo de les següents expressions:

$$\bar{C} = \bar{R} \cos \bar{x}_0 \quad (4.11)$$

$$\bar{S} = \bar{R} \sin \bar{x}_0 \quad (4.12)$$

finalment obtenim \bar{x}_0 a partir de l'arc tangent i dels signes de \bar{S} i \bar{C} :

$$\bar{x}_0 = \begin{cases} \bar{x}_0' & \text{si } \bar{C} > 0 \text{ i } \bar{S} > 0 \\ \bar{x}_0' + \pi & \text{si } \bar{C} < 0 \\ \bar{x}_0' + 2\pi & \text{si } \bar{C} > 0 \text{ i } \bar{S} < 0 \end{cases} \quad (4.13)$$

$$\bar{x}_0' = \arctan(\bar{S}/\bar{C}) \quad (4.14)$$

La variança circular. Un cop calculada la mitjana aritmètica, un altre paràmetre que ens aporta informació sobre un conjunt de valors és la variança. En el cas de variables angulars, la variança S_0 es defineix com:

$$S_0 = 1 - \frac{1}{n} \sum_{i=1}^n \cos(\theta_i - \bar{x}_0) \quad (4.15)$$

4.3.1.2 Variació dels component cromàtics

A les pàgines següents mostrem com varien els components cromàtics corresponents a quatre espais de color pels dotze colors estudiats, amb mesures fetes sobre les imatges adquirides amb la càmera. Els tres components de l'espai *HSI* han estat escalats a l'interval $[0, 255]$. Pel to, el valor 0 equival a 0 radians i el valor 255 equival a 2π radians. Per la saturació el 0 equival al cas acromàtic, mentre que el 255 correspon a la saturació màxima i, finalment, per la intensitat el 0 equival al negre i el 255 al blanc. Els tres components de l'espai *RGB* normalitzat, amb valors que van entre 0 i 1, també han estat escalats a l'interval $[0, 255]$. L'espai $I1, I2, I3$ s'ha deixat en els intervals de valors sobre els que està definit per tal de mostrar gràficament les transicions entre valors negatius i positius; aquests intervals són $[0, 255]$ per $I1$ i $[-127, 127]$ pels components $I2$ i $I3$.

Per facilitar l'apreciació dels resultats, els punts mesurats consecutivament s'han unit per línies rectes. No s'ha pretès aquí fer cap ajustament de les mesures obtingudes.

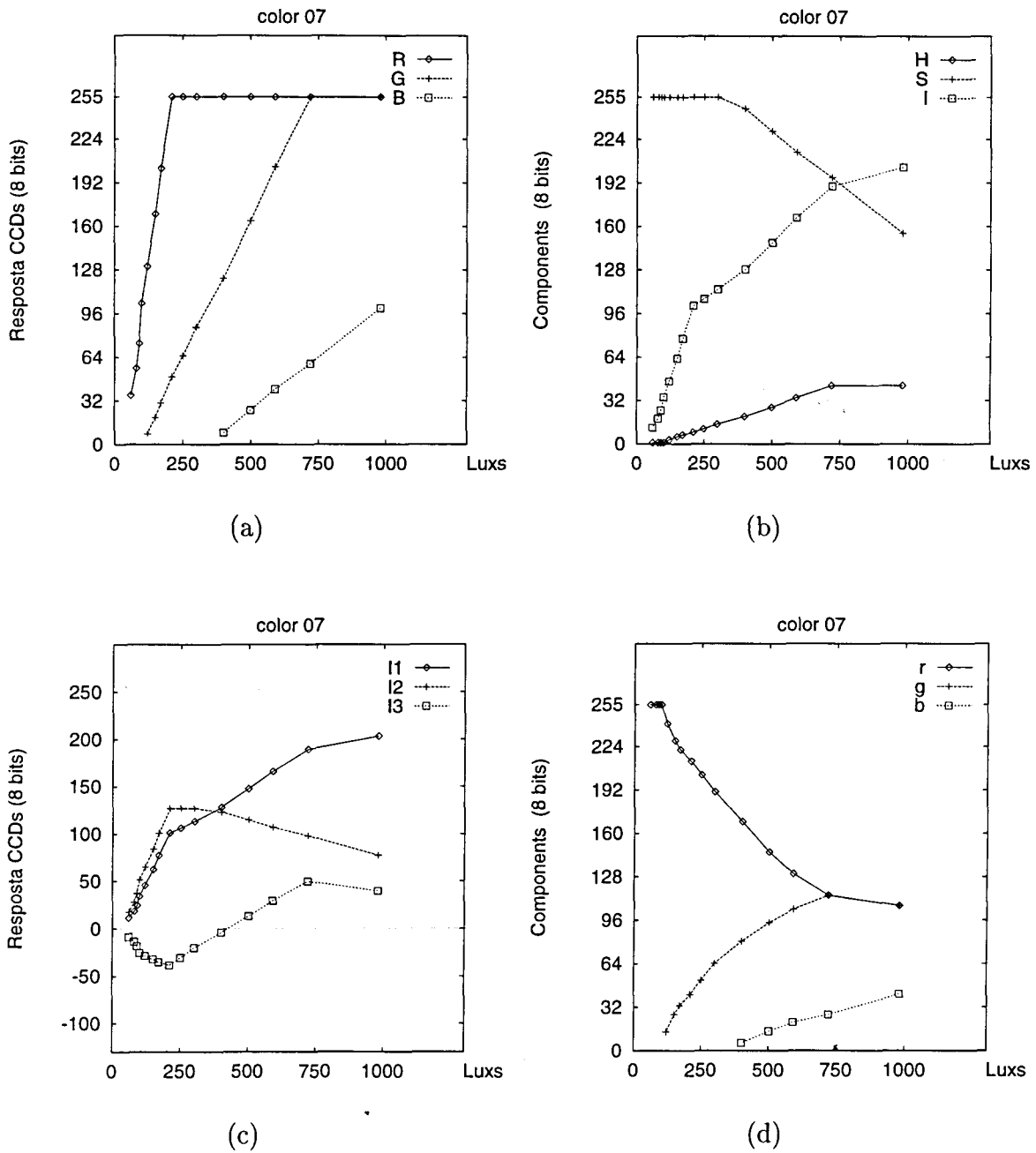


Figura 4.28: Comportament dels components *RGB* (a), *HSI* (b), *I1I2I3* (c) i *RGB* normalitzat (d) pel color 7 (Taronja) de la *Macbeth COLORCHECKER* variant la intensitat.

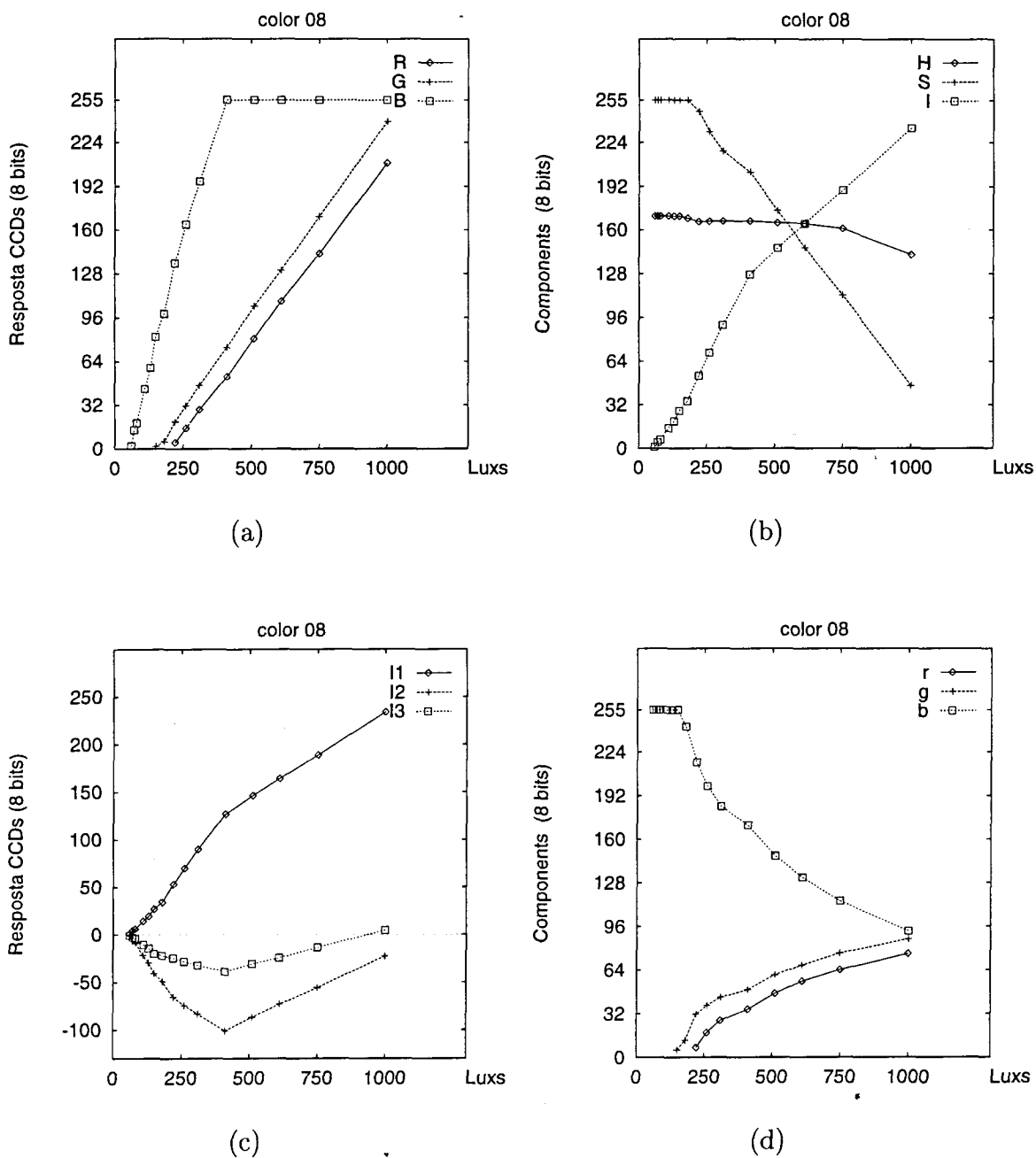


Figura 4.29: Comportament dels components *RGB* (a), *HSI* (b), *I1I2I3* (c) i *RGB* normalitzat (d) pel color 8 (Blau violeta) de la *Macheth COLORCHECKER* variant la intensitat.

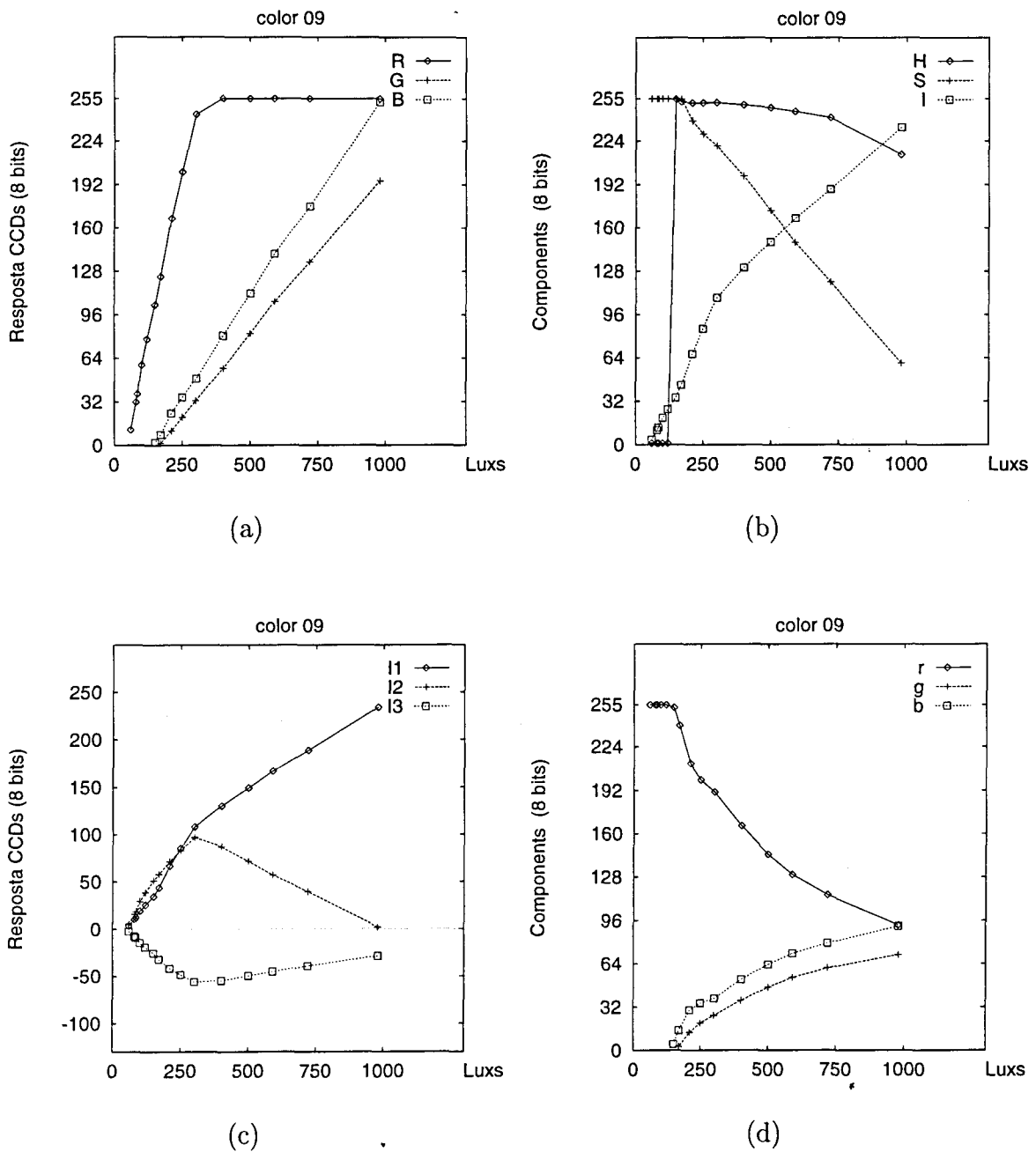


Figura 4.30: Comportament dels components *RGB* (a), *HSI* (b), *I1I2I3* (c) i *RGB* normalitzat (d) pel color 9 (Vermell suau) de la *Macbeth COLORCHECKER* variant la intensitat.

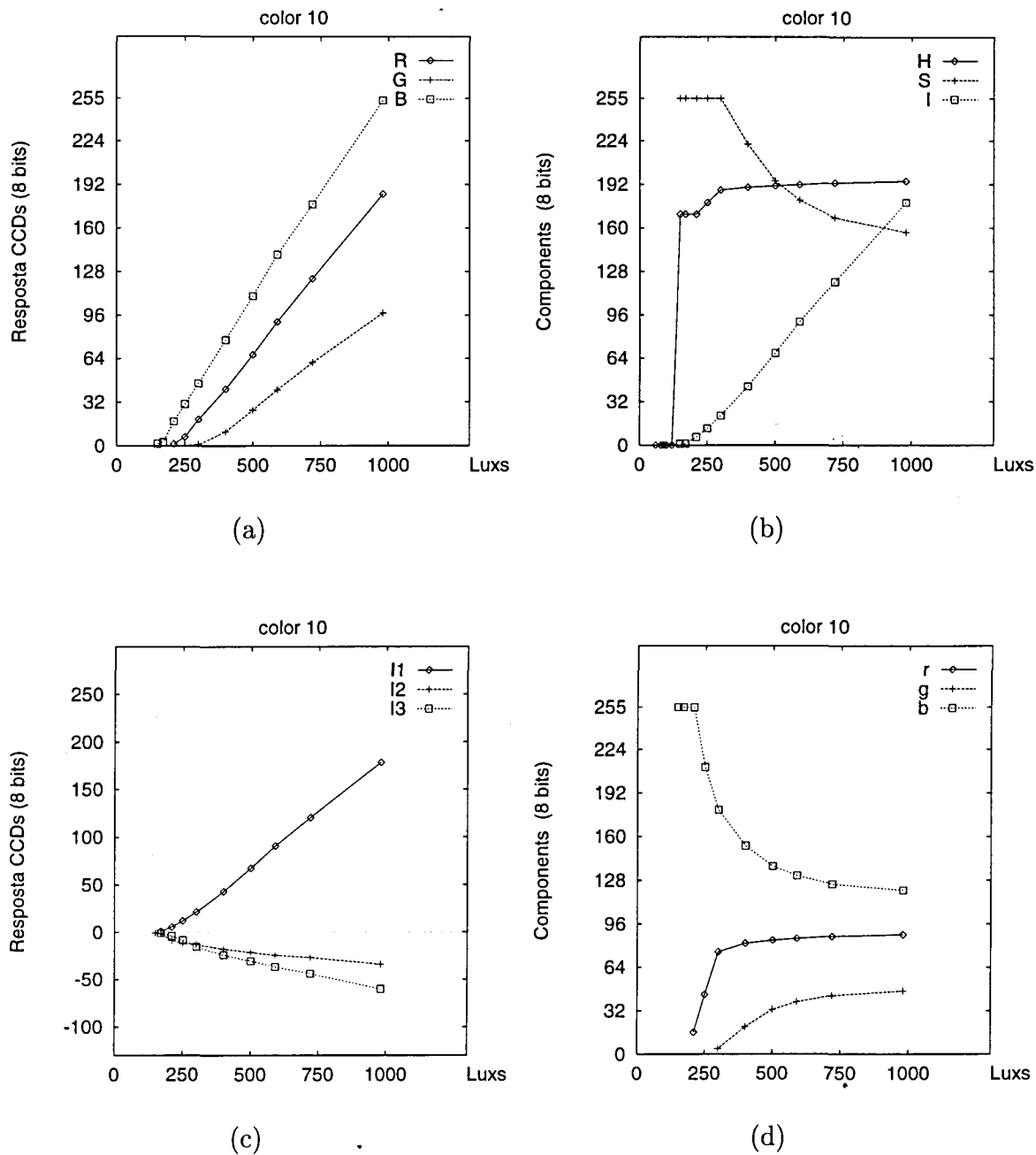


Figura 4.31: Comportament dels components *RGB* (a), *HSI* (b), *I1I2I3* (c) i *RGB* normalitzat (d) pel color 10 (Violeta) de la *Macbeth COLORCHECKER* variant la intensitat.

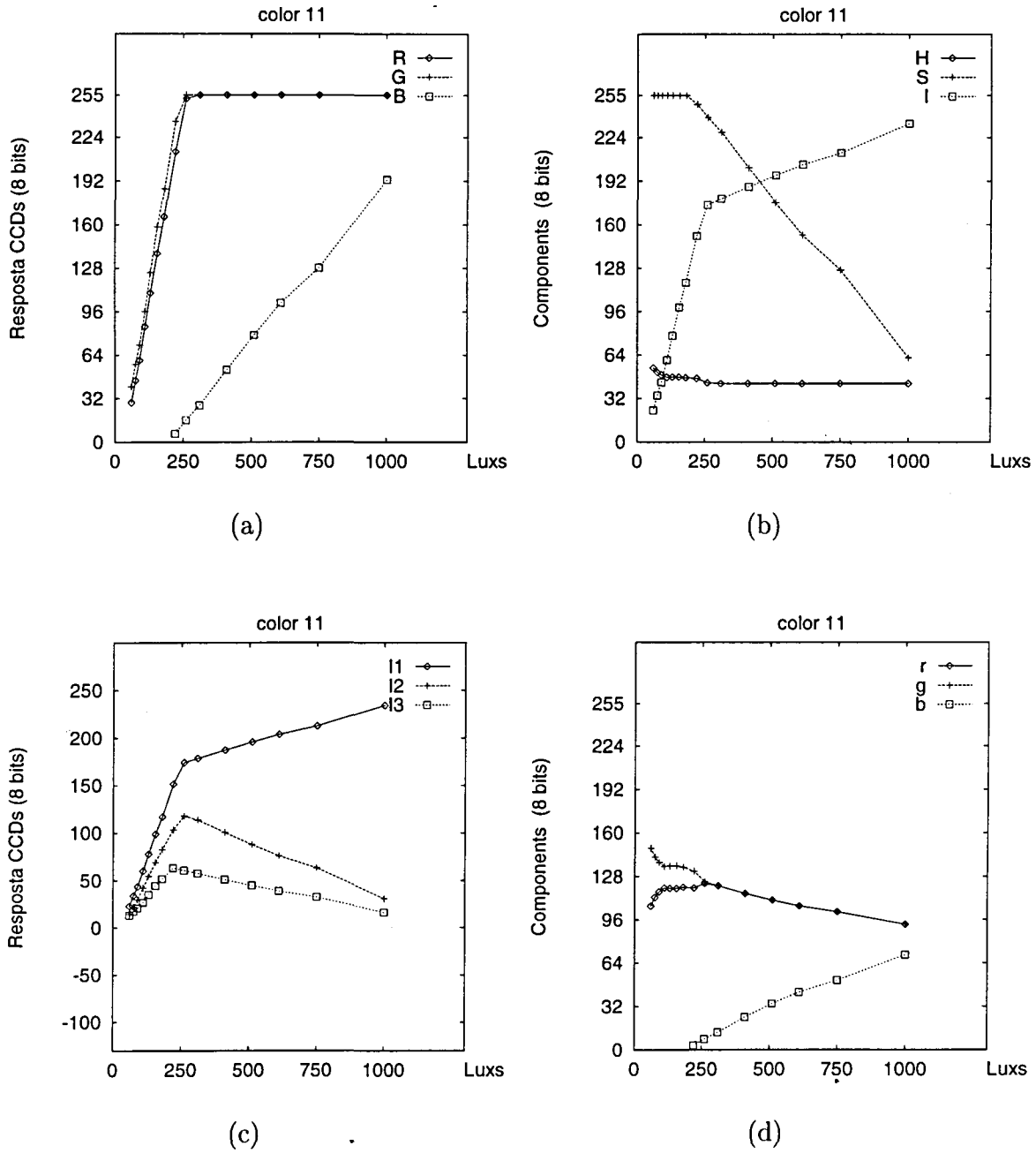


Figura 4.32: Comportament dels components RGB (a), HSI (b), I1I2I3 (c) i RGB normalitzat (d) pel color 11 (Groc verd) de la Macbeth COLORCHECKER variant la intensitat.

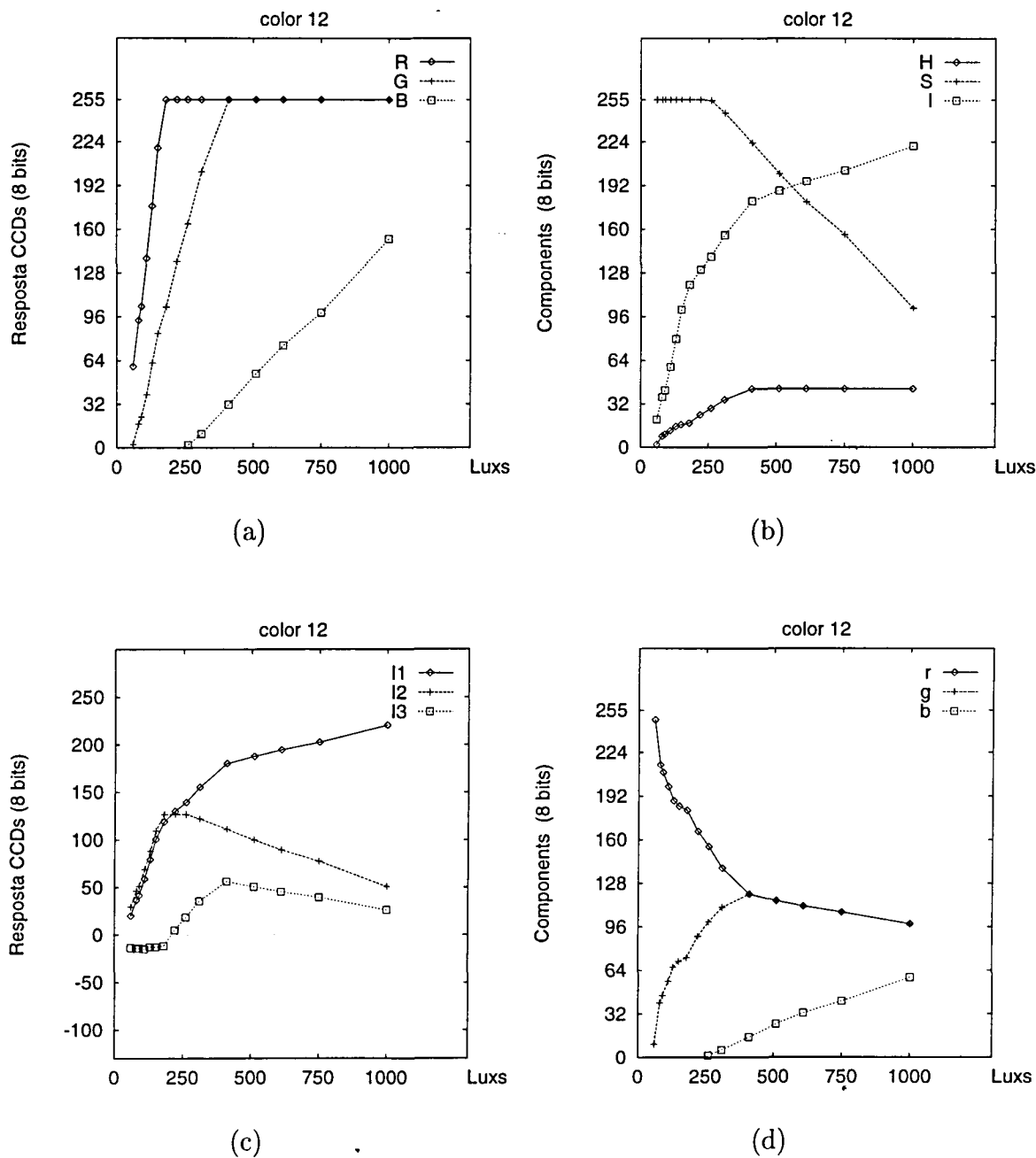


Figura 4.33: Comportament dels components *RGB* (a), *HSI* (b), *I1I2I3* (c) i *RGB* normalitzat (d) pel color 12 (Taronja groc) de la *Macbeth* COLORCHECKER variant la intensitat.

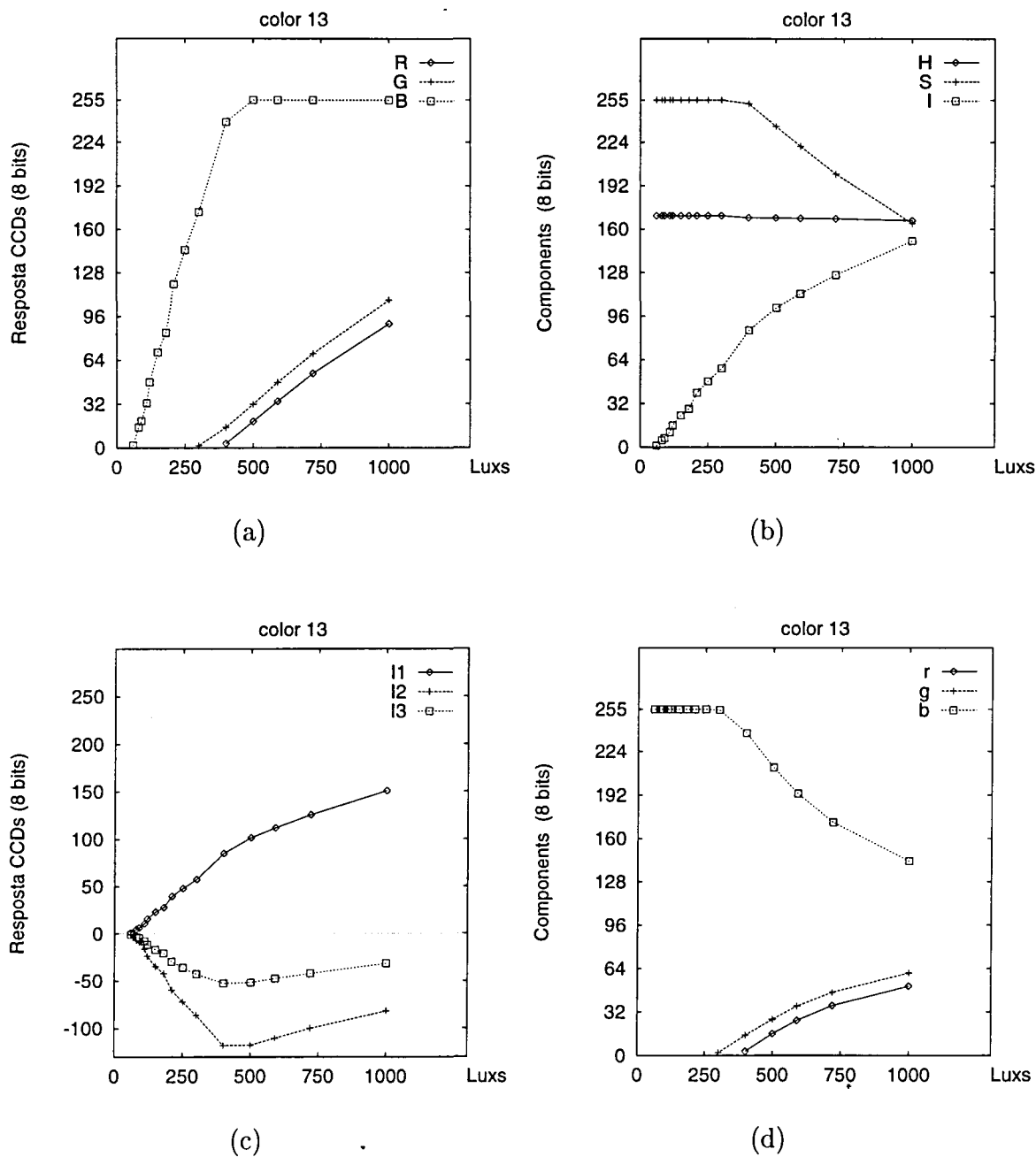


Figura 4.34: Comportament dels components RGB (a), HSI (b), I1I2I3 (c) i RGB normalitzat (d) pel color 13 (Blau) de la Macbeth COLORCHECKER variant la intensitat.

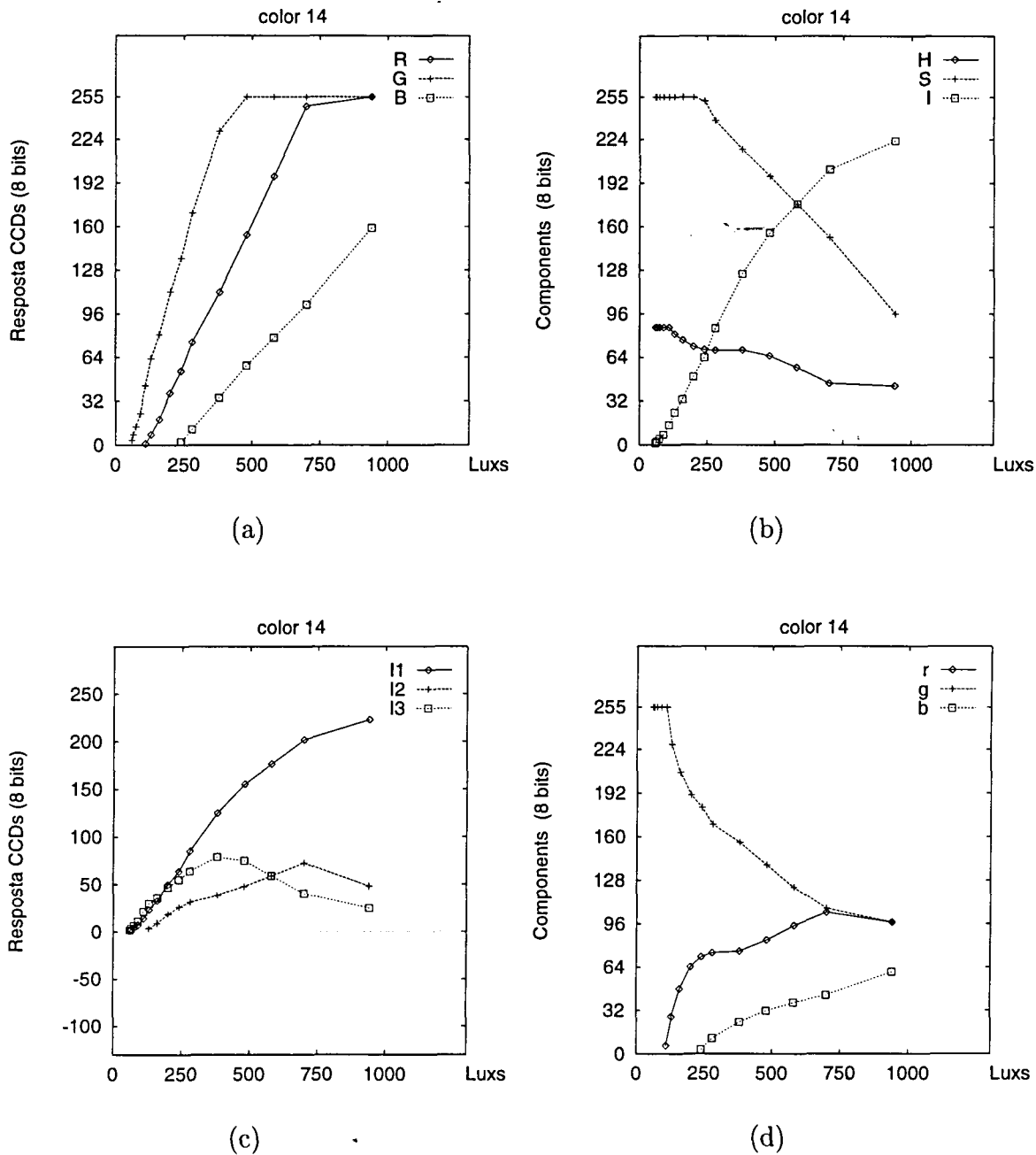


Figura 4.35: Comportament dels components RGB (a), HSI (b), I1I2I3 (c) i RGB normalitzat (d) pel color 14 (Verd) de la Macbeth COLORCHECKER variant la intensitat.

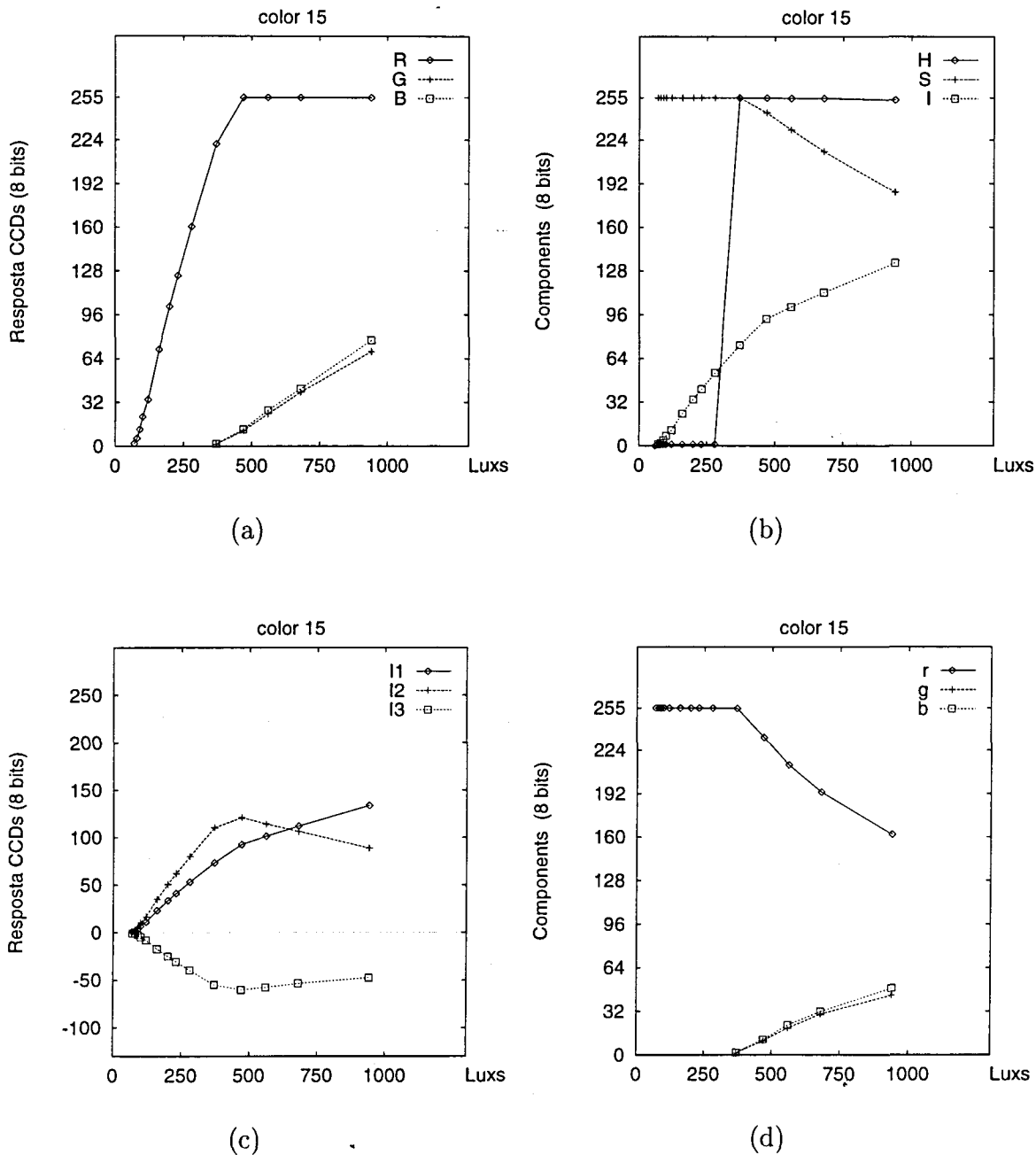
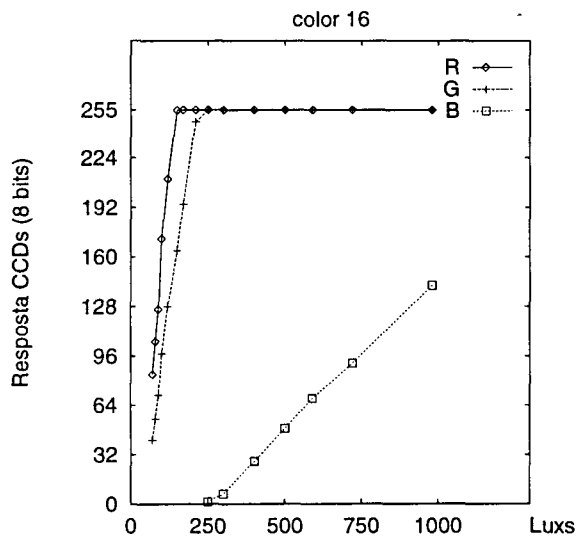
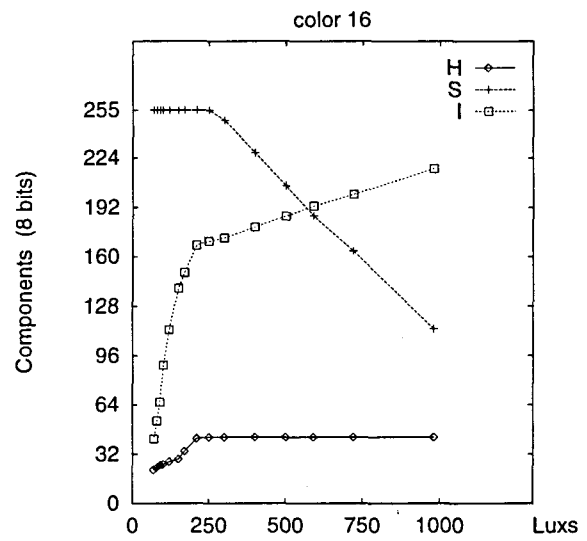


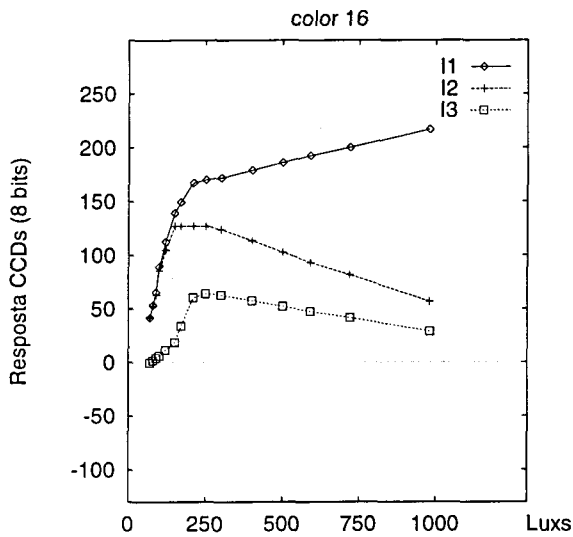
Figura 4.36: Comportament dels components RGB (a), HSI (b), I1I2I3 (c) i RGB normalitzat (d) pel color 15 (Vermell) de la Macbeth COLORCHECKER variant la intensitat.



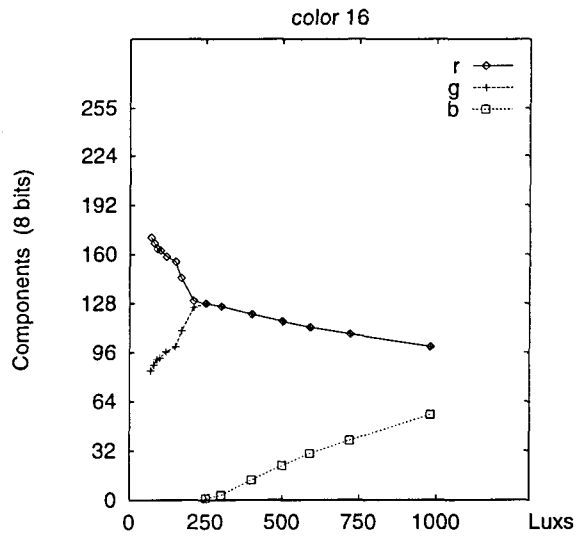
(a)



(b)



(c)



(d)

Figura 4.37: Comportament dels components *RGB* (a), *HSI* (b), *I1I2I3* (c) i *RGB* normalitzat (d) pel color 16 (Groc) de la *Macbeth COLORCHECKER* variant la intensitat.

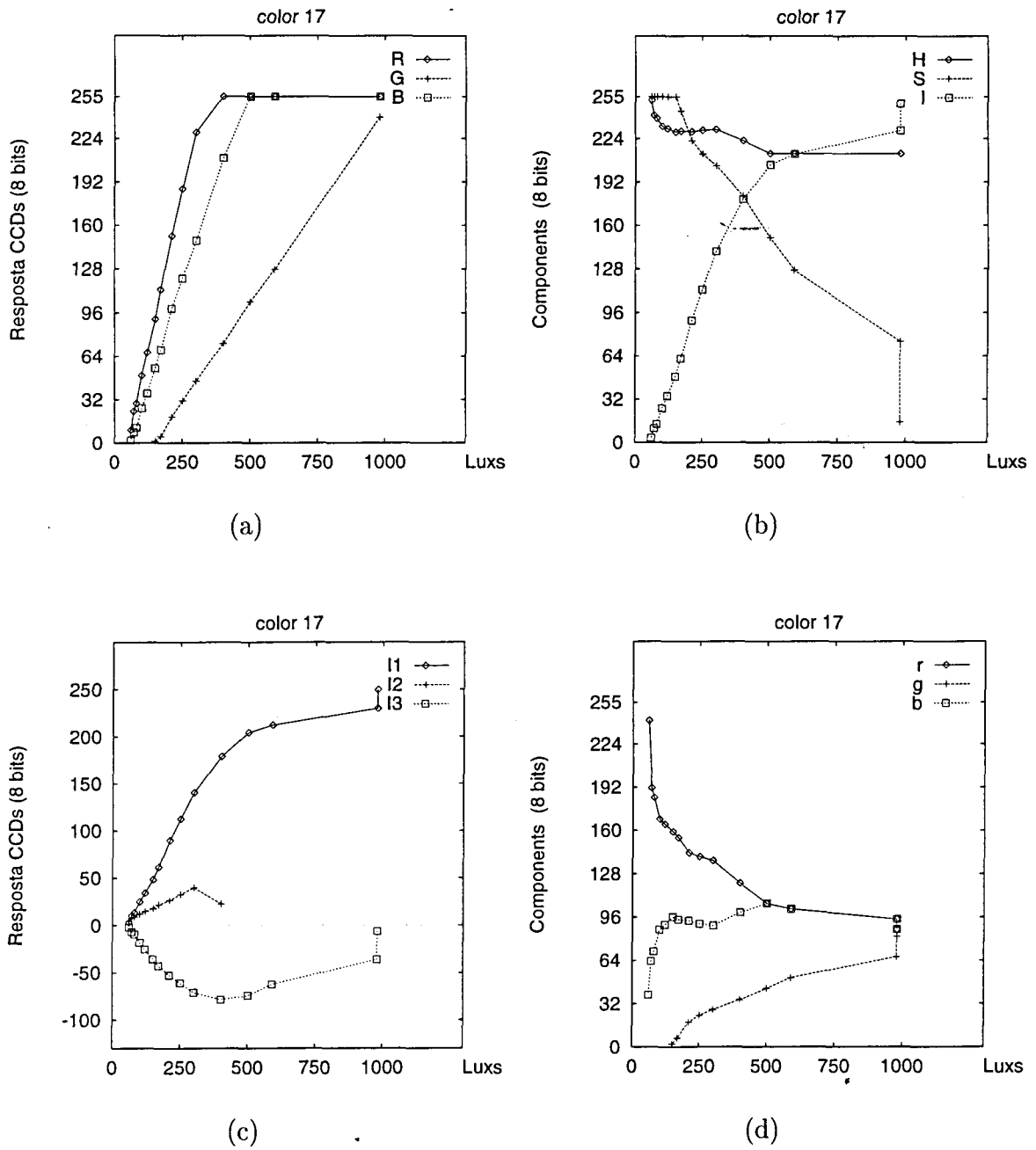


Figura 4.38: Comportament dels components *RGB* (a), *HSI* (b), *I1I2I3* (c) i *RGB* normalitzat (d) pel color 17 (Magenta) de la *Macbeth COLORCHECKER* variant la intensitat.

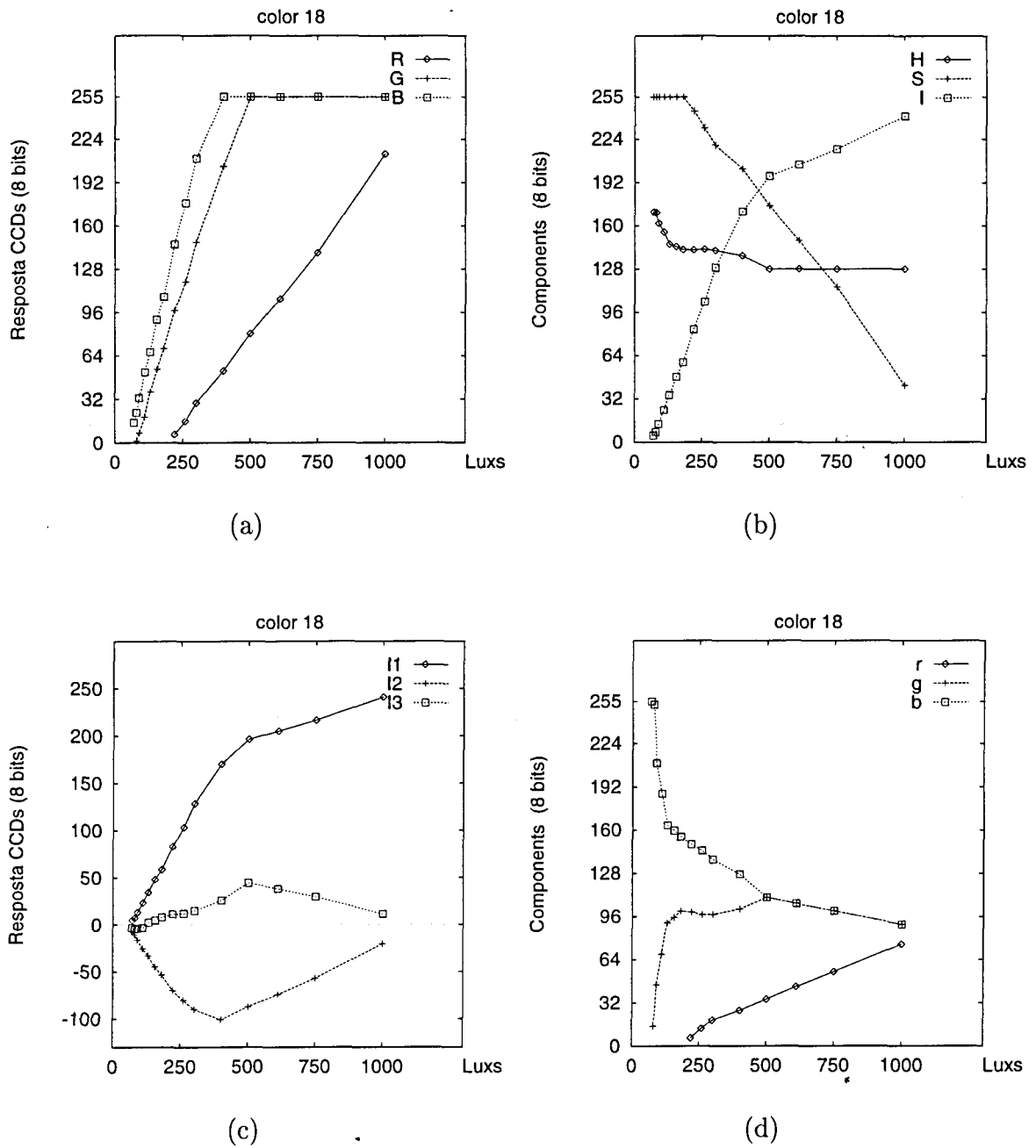


Figura 4.39: Comportament dels components RGB (a), HSI (b), I1I2I3 (c) i RGB normalitzat (d) pel color 18 (Cian) de la Macbeth COLORCHECKER variant la intensitat.

Observant aquestes gràfiques es pot concloure que:

- El comportament dels sensors respecte a variacions de la intensitat és realment lineal pels tres components R , G i B fins que queden saturats.
- Els components $I1$, $I2$ i $I3$ no són estables a variacions en l'intensitat, tal i com es despenia de les gràfiques teòriques. Tampoc és invariant el component I de l'espai HSI (que és equivalent al component $I1$ de l'espai $I1I2I3$).
- En la resta de components representats H , S , r , g i b s'observa que tendeixen a ser invariants a variacions en la intensitat dins de l'interval en que no hi ha cap component R, G, B saturat. El fet que no siguin estrictament invariants és degut a que les rectes que segueixen els components R , G i B no passen per l'origen de coordenades, tal i com s'exposa al final del proper capítol.
- Una observació important és que el component més estable és el component de to, mentre que la saturació i els components RGB normalitzats són molt sensibles al fet que les rectes que descriuen els components RGB no passin per l'origen.

4.3.2 Colors en objectes

Quan estudiem les superfícies planes, per una certa intensitat en l'illuminant tota la superfície té el mateix color. Quan tractem amb objectes tridimensionals això no és cert, doncs la geometria de l'objecte fa que la quantitat de llum que rebí en unes zones o unes altres sigui diferent. Aquesta diferència pot portar a que els píxels corresponents a l'objecte es puguin classificar entre píxels amb la informació cromàtica correcta i píxels en els que el color ha quedat retallat.

4.3.2.1 Exemple 1

Per exemple, les figures 4.40.a i 4.40.b mostren dues imatges de la mateixa escena d'un tren de juguina adquirides amb diferents intensitat de l'illuminant.

La imatge de la figura 4.40.a ha estat adquirida amb una intensitat de l'illuminant tal que molts pocs píxels corresponents a la màquina pateixen l'efecte del retall del color, mentre que la imatge de la figura 4.40.b ha estat adquirida forçant la intensitat de la il·luminació, per mostrar l'efecte de la saturació dels CCDs.

La figura 4.41 mostra els perfils dels components de to corresponents als píxels que es troben al llarg de la línia entre els punts A i B de la figura 4.40. El perfil il·lustrat a la



(a)

(b)

Figura 4.40: Imatge adquirida amb la intensitat de l'il·luminant adient (a) i imatge adquirida amb una intensitat massa elevada (b).

figura 4.41.a correspon a la imatge adquirida amb una intensitat de llum tal que molts pocs píxels quedin saturats i, com es pot veure, el to és quasi constant per tots els píxels que es troben al llarg de la línia (el forat que apareix prop del punt *B* no és tal, doncs no cal oblidar que el to és una variable circular). Per altra banda, la figura 4.41.b mostra el perfil al llarg de la mateixa línia quan la imatge és adquirida amb una il·luminació més alta. Es pot veure que en aquest segon cas el to és molt variable, trobant les màximes variacions en les zones més il·luminades de la imatge. Dels dos forats que apareixen, el primer començant per l'esquerra sí que és real i correspon a una zona de la màquina de tren (sobre la roda) on el to és indefinit. Les diferències entre els valors del to dels píxels amb la informació cromàtica correcte i els que han sofert el retall del color és d'aproximadament un 14% (que si representem el to utilitzant un byte per píxel, representa una diferència de 35 nivells de to).

La figura 4.42 mostra els perfils del component corresponent a la saturació pels píxels que es troben al llarg de la línia *A, B* per les dues imatges de la figura 4.40. Com es pot veure la saturació, tot i ser més variable que el to, també és més inestable quan hi ha píxels que han sofert l'efecte del retall del seu color.

Les mateixes observacions poden fer-se sobre el to i la saturació en el cas del vagó de

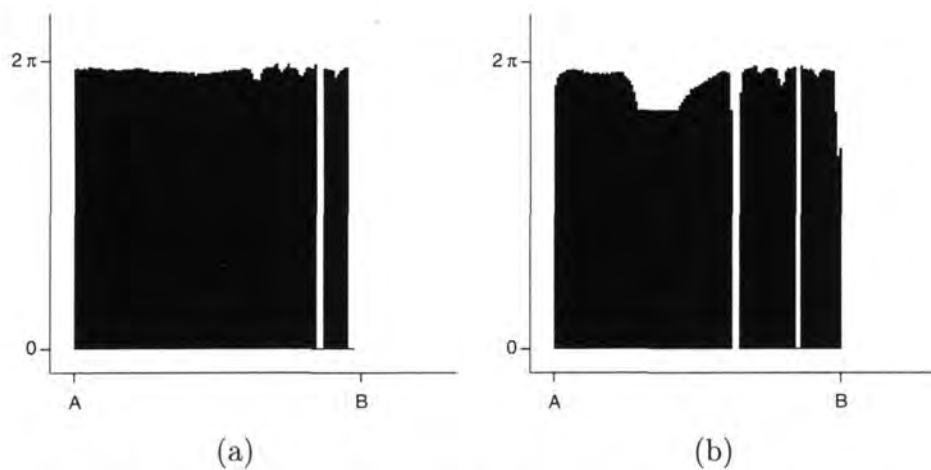


Figura 4.41: Perfils del to al llarg de la línia A, B sobre la màquina de les imatges de la figura 4.40.

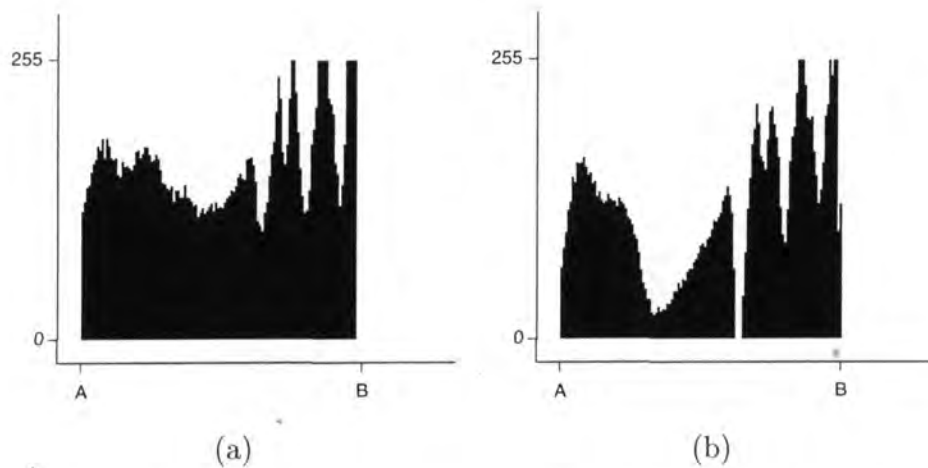


Figura 4.42: Perfils entre els punts A i B del component corresponent a la saturació per les dues imatges de la figura 4.40.

color verd clar i del vagó de color blau. La figura 4.43 conté els perfils del component de to sobre la línia C, D i la figura 4.44 conté els perfils sobre la mateixa línia corresponents a la saturació. Es pot veure que tot i que la saturació contínua essent molt diferent entre les dues imatges, els dos perfils corresponents al to no són gaire diferents. Això és degut a que el CCD corresponent al component blau no ha quedat gaire saturat en comparació amb els CCDs corresponents als components vermell i verd.

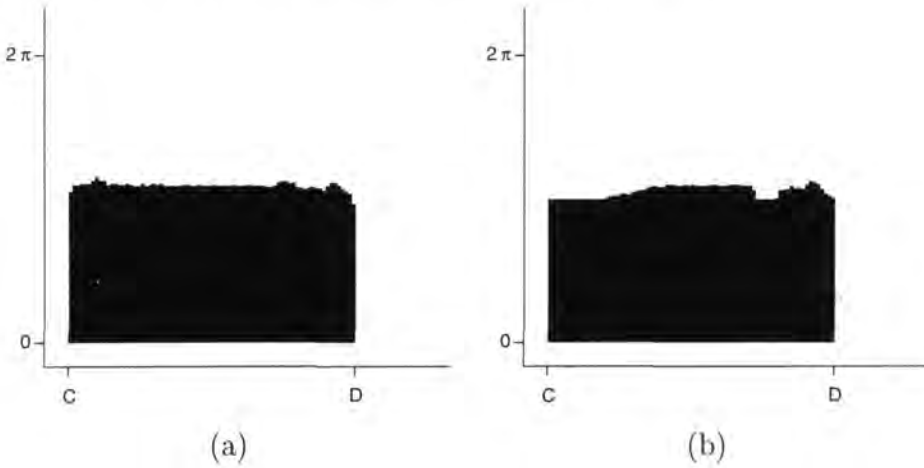


Figura 4.43: Perfils del to al llarg de la línia C, D sobre el vagó blau de les imatges de la figura 4.40.

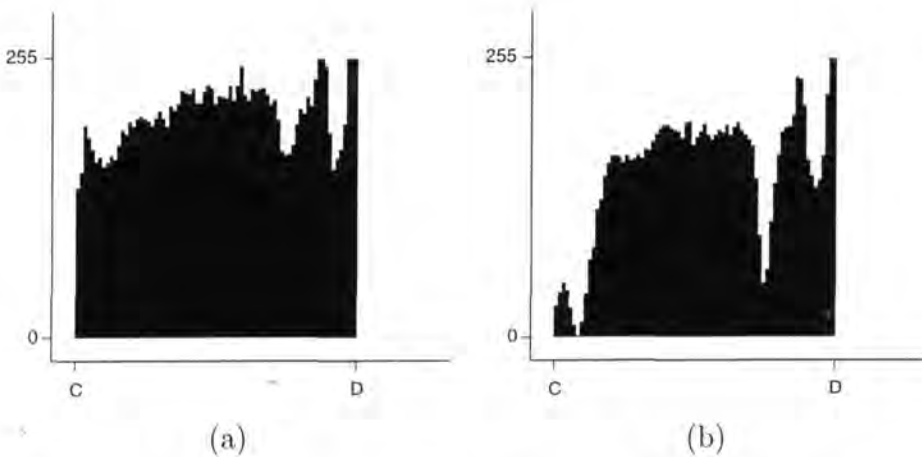


Figura 4.44: Perfils de la saturació al llarg de la línia C, D sobre el vagó blau de les imatges de la figura 4.40.

Pel que respecte al vagó de color verd clar, les figures 4.45 i figura 4.46 contenen els

perfils corresponents al to i a la saturació pels píxels que es troben sobre la línia E, F de les imatges de la figura 4.40. Per aquest color, igual que passa amb el color de la màquina, la variació en el component del to és important. Es pot veure que hi ha tot un conjunt de píxels al principi del perfil (píxels del sostre del vagó) que han passat a tenir un valor indeterminat de to.

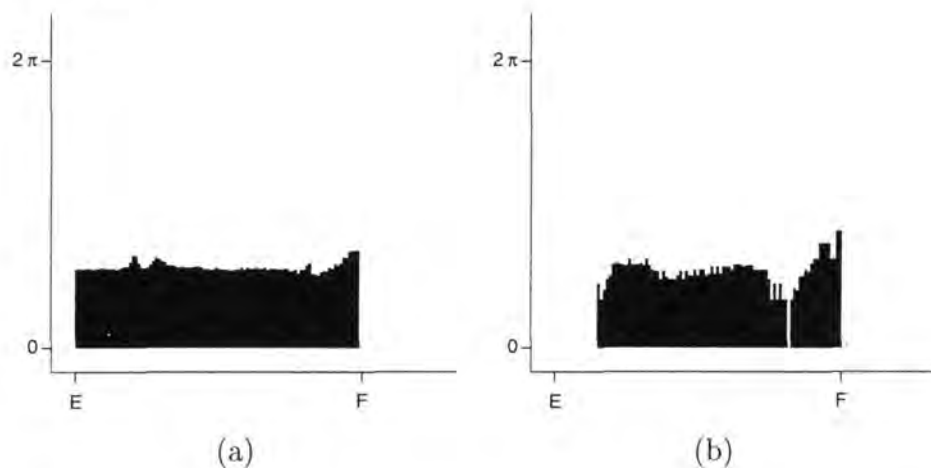


Figura 4.45: Perfils del to al llarg de la línia E, F sobre el vagó verd de les imatges de la figura 4.40.

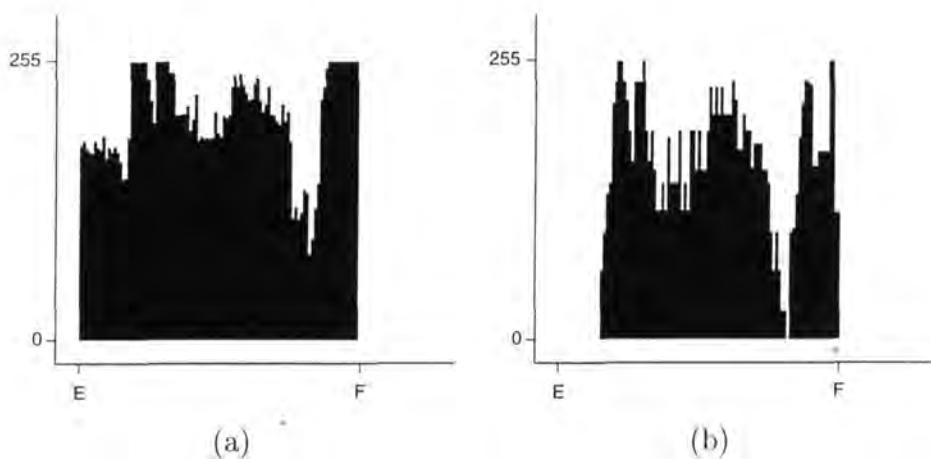


Figura 4.46: Perfils de la saturació al llarg de la línia E, F sobre el vagó verd de les imatges de la figura 4.40.

Si sobre la imatge correctament il·luminada trobem els intervals de to i saturació que corresponen a la màquina i als dos vagons, i els utilitzem per segmentar les dues imatges, obtenim el resultat que es mostra a la figura 4.47. Es pot observar que el resultat de l'esquerra (corresponent a la figura 4.40.a) s'ajusta molt més al resultat desitjat que el de la dreta (corresponent a la figura 4.40.b)

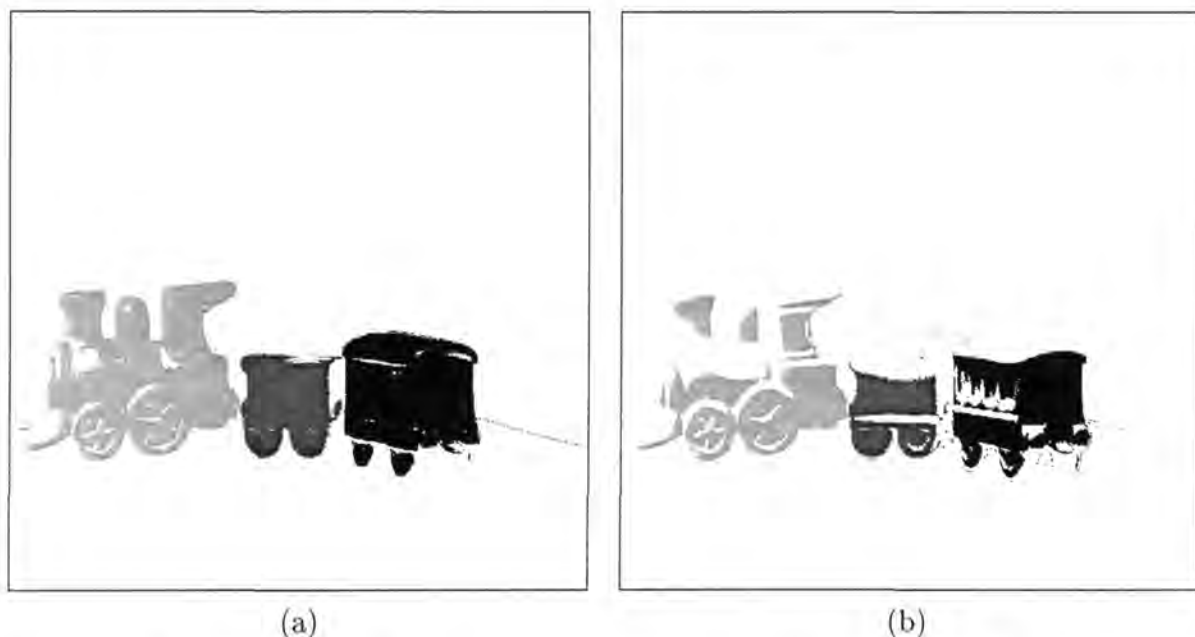


Figura 4.47: Resultat de la segmentació basada en el to i la saturació per les imatges de les figures 4.40.a i 4.40.b . En Gris clar la màquina, en gris més fosc el vagó blau i en negre el vagó verd.

4.3.2.2 Exemple 2

El segon exemple que presentem correspon a unes imatges d'unes pilotes de goma, i l'objectiu és segmentar correctament els píxels de la pilota taronja. La figura 4.48 conté les dues imatges que es comenten; a l'esquerra hi ha la imatge amb la intensitat de llum adient i a la dreta la imatge amb massa llum. En aquest cas, la il·luminació és la mateixa per les dues imatges, i el que canvia és que en la primera imatge hi ha una altra pilota que tapa una mica la llum que arriba a la pilota taronja.

Els perfils del to entre els punts *A*, *B* de les imatges de la figura 4.48 es mostren a les figures 4.49. Es pot veure que per la imatge amb la il·luminació correcta, els píxels pertanyents a la pilota taronja tenen un to molt semblant, mentre que per la imatge amb

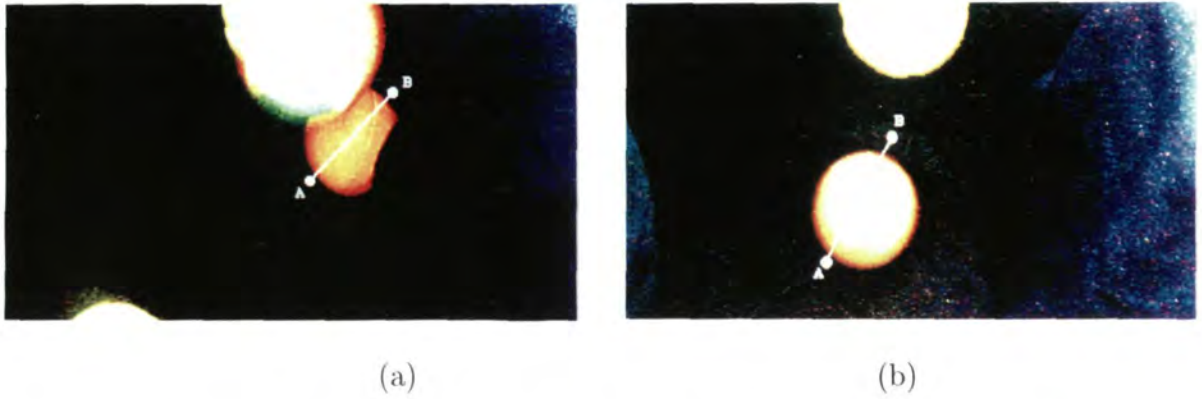


Figura 4.48: Imatge adquirida amb la intensitat de l'illuminant adient (a) i imatge amb una intensitat massa elevada a la zona on es troba la pilota taronja (b).

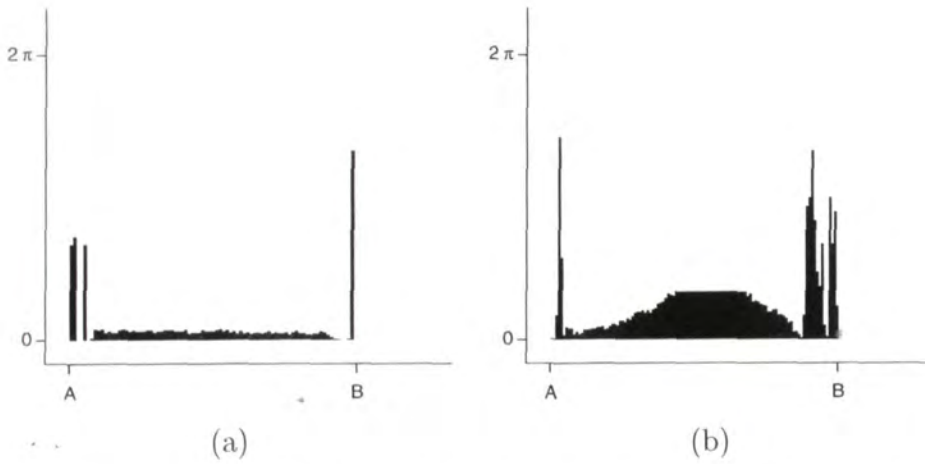


Figura 4.49: Perfils del to al llarg de la línia A, B sobre la pilota taronja de les imatges de la figura 4.48.

més il·luminació el valor del t_o va creixent a mesura que els components queden saturats (cap al centre de la pilota). Cal fer un comentari sobre els píxels de la línia A, B que cauen fora de la pilota: en aquests, el component de t_o té un comportament aleatori, i això es degut a que el fons utilitzat en l'adquisició de les imatges és una roba negra i, tal com comenta Perez a la seva tesi [Per94], el t_o té un comportament aleatori pels píxels amb una intensitat molt baixa.

Pel que fa als perfils del component corresponent a la saturació es mostren a la figura 4.50. Igual que en l'exemple del tren, la saturació està més afectada pel soroll que el t_o . De totes maneres, les gràfiques indiquen una major variació de la saturació en la imatge massa il·luminada. Pels píxels de fora la pilota s'observa el mateix efecte del soroll que en el t_o .

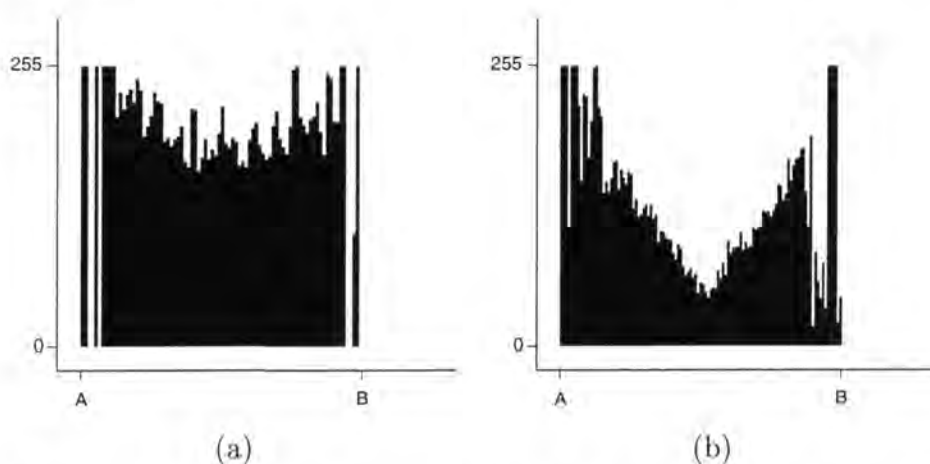


Figura 4.50: Perfils de la saturació al llarg de la línia A, B sobre la pilota taronja de les imatges de la figura 4.48.

El resultat de la segmentació de la pilota taronja està il·lustrada a la figura 4.52. La segmentació s'ha fet utilitzant els tres components HSI , donat que en aquests tipus d'imatges que tenen un fons molt afectat pel soroll és necessari utilitzar el component d'intensitat per eliminar-lo (la figura 4.51 mostra el perfil de la intensitat de les línies A, B de la figura 4.48). Igual que en l'exemple del tren, l'efecte del retall del color sobre el resultat és evident.

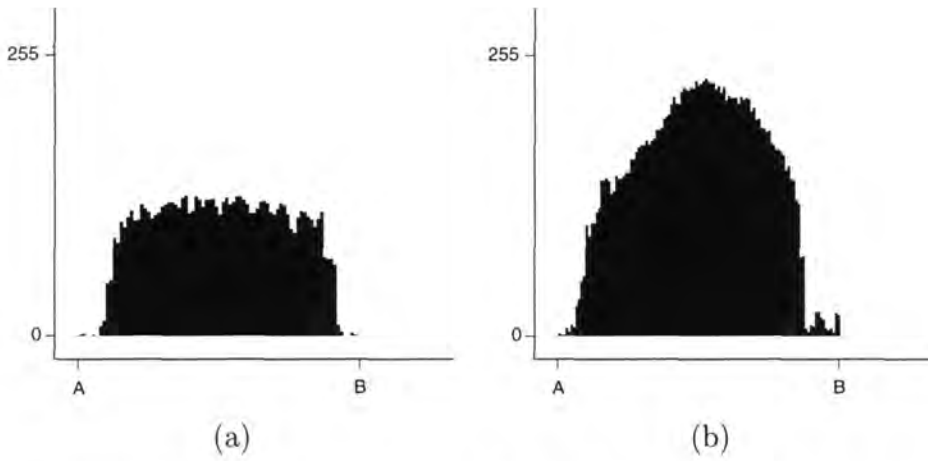


Figura 4.51: Perfils de la intensitat al llarg de la línia A, B sobre la pilota taronja de les imatges de la figura 4.48.

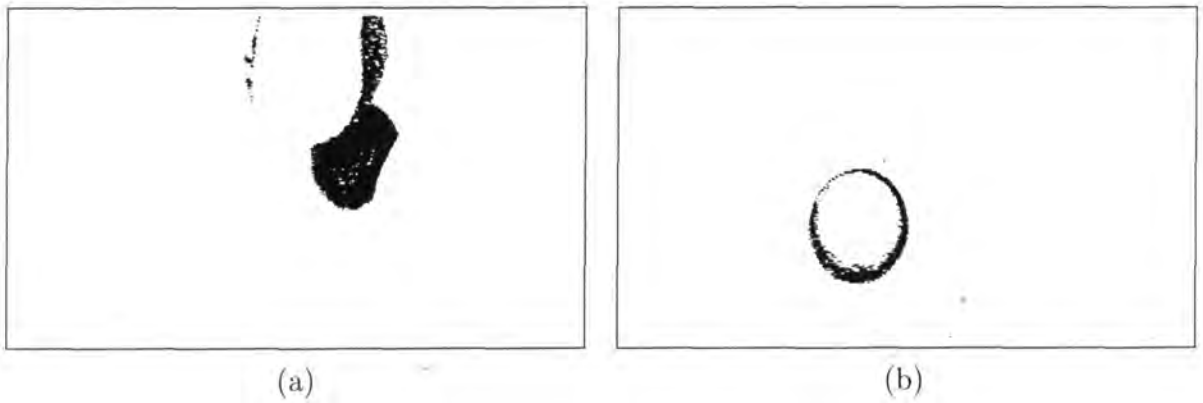


Figura 4.52: Resultat de la segmentació basada en el to, la saturació i la intensitat per les imatges de les figures 4.48.a i 4.48.b.