

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Departament d'Enginyeria de Sistemes Automàtica i Informàtica Industrial

**APORTACIÓ A LA DESCRIPCIÓ I
SEGUIMENT DE CAMINS
NAVEGABLES EN ENTORNS
NATURALS A PARTIR DE L'ANÀLISI
DE REGIONS EN SEQÜÈNCIES
D'IMATGES**

Autor: Josep Fernández Ruzafa

Director: Alícia Casals Gelpí

Barcelona, febrer del 1998

4. Segmentació de la Imatge

Els sistemes de navegació autònoma a partir de visió per ordinador es basen en l'anàlisi de les imatges subministrades per càmeres. Un dels passos necessaris en aquesta anàlisi és la segmentació de les imatges. La segmentació d'una imatge, és el procés que permet dividir-la en els diferents components o regions que la formen. Cada regió de la imatge segmentada és un conjunt de píxels connexos que tenen en comú certes propietats.

Una de les decisions importants que cal prendre en el moment de definir el procés de la segmentació és els tipus d'informació o característiques de la imatge (monocroma, color, textura, etc.) sobre la que els algorismes de segmentació, avaluaran les propietats. La característica associada a cada un dels píxels o zona de la imatge, és la que es fa servir per determinar a quina regió pertany un píxel. Una característica útil ha d'estar associada a una propietat física dels objectes o components, de forma que sigui fàcil diferenciar-los i assolir una partició de la imatge relacionada amb els objectes o components que la formen. Dit d'una altre manera, és desitjable que cada objecte de l'escena defineixi a la imatge segmentada una o varies regions de forma disjunta, on cada regió de la imatge està associada a un únic objecte.

El problema que ens preocupa, la navegació autònoma en camins feblement estructurats basada en visió per ordinador, requereix de la segmentació d'imatges que han estat captades en entorns naturals. Aquestes escenes es caracteritzen per les seves condicions variables d'il·luminació, la gran quantitat d'objectes presents en l'escena, i pel fet que les matèries que defineixen la superfície del camí també poden formar part dels marges del camí.

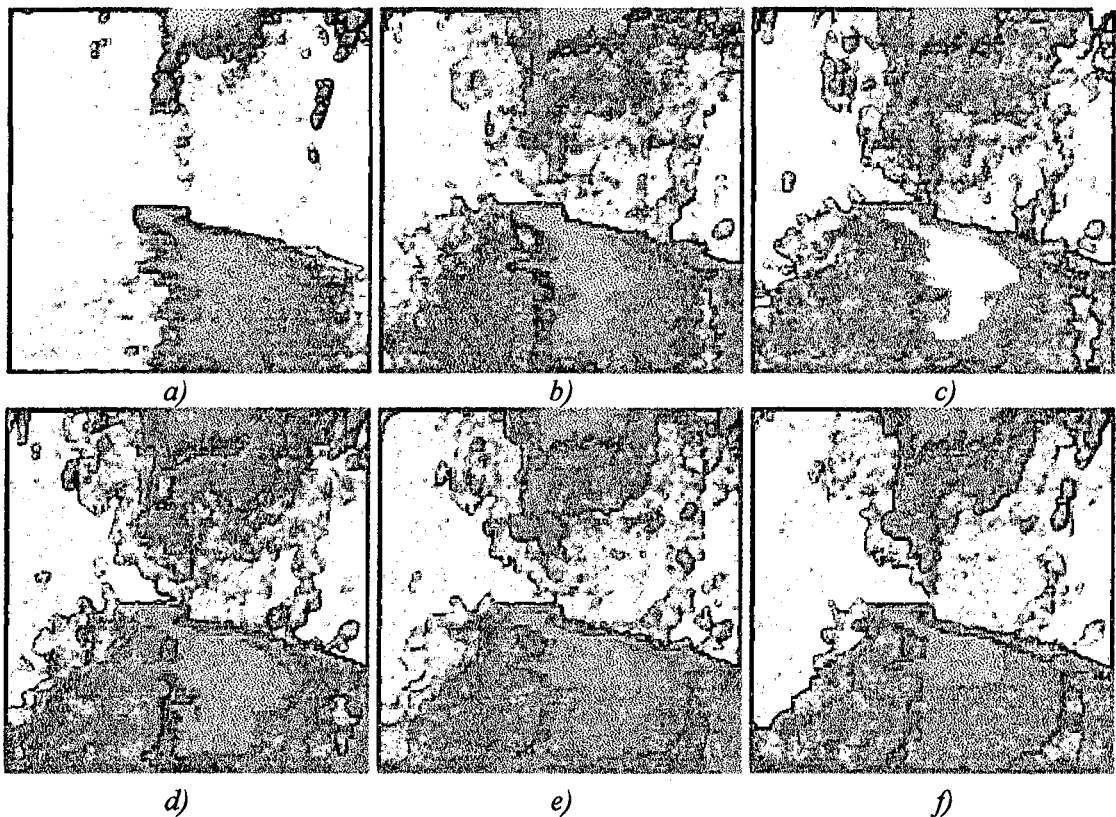
Per tant, en el procés de segmentació d'imatges orientat a la detecció de la posició del camí a la imatge analitzada, resulta apropiat treballar amb una característica independent de les condicions d'il·luminació i lligada a les propietats dels objectes, amb un algorisme de segmentació que tot i la complexitat de la imatge, generi una imatge segmentada simple, i amb una classificació final de les regions, en la que es tindrà en compte tant la característica de la regió, com la posició de la regió a la imatge respecte la posició esperada del camí, per determinar quines regions estan associades als elements d'interès de l'entorn (camí, marge i obstacles).

4.1. Representació de la Informació Visual

La informació més simple d'obtenir i processar, és la intensitat lluminosa o nivell de gris de cada element de la imatge digital. Encara que la percepció visual diürna humana és cromàtica -percebim els colors dels objectes-, la visió monocromàtica -que es dona en els humans durant la percepció d'escenes feblement il·luminades i en la majoria dels animals- resulta suficient, en molts casos, per percebre la informació de l'entorn necessària per interpretar-lo. Per contra hi ha certes situacions on el nivell de gris dels píxels dificulta enormement el segmentar amb encert les imatges analitzades, ja que la il·luminació no uniforme sobre l'escena, les ombres i els reflexos especulars poden fer que un objecte defineixi varies regions a la imatge o que dos objectes diferents defineixin una única regió.

Per a la navegació de vehicles en camins de muntanya, s'han de segmentar imatges que han estat captades en entorns naturals, on s'espera que apareixi un camí, part dels marges del camí i potser, obstacles en el camí. Aquestes imatges (figura 1.2) es caracteritzen per ser complexes, degut al nombre important d'objectes que apareixen i a les condicions, desconegudes i molt variables d'il·luminació, la qual cosa fa que l'aparença dels objectes a la imatge monocromàtica depengui molt de si rep una il·luminació directa, difosa i de la projecció d'ombres sobre ells [Celenk 90]. Aquests fets fan que el nivell de gris no sigui la característica més eficient per segmentar aquest tipus d'imatges.

A la figura 4.1 es pot observar la influència que té la il·luminació no uniforme de l'escena en la segmentació d'un imatge d'intensitat lluminosa. La figura 4.1.a és la imatge RGB original i les figures 4.1.b, c, d, e i f, corresponen a la seva segmentació analitzant el nivell de gris del píxel -quantificat en 128 valors diferents- mitjançant la tècnica de creixement de regions (veure secció 4.2.2), amb diferents criteris de creixement (la diferència màxima entre el valor del píxel i la mitja de la regió per a que el píxel sigui agregat a la regió és de 6, 8, 12, 16 i 20 respectivament). En tots els casos, la zona de la carretera de la imatge es veu associada a diferents regions en la imatge segmentada. El nombre de regions de les imatges segmentades és de 212, 106, 74, 72 i 58 respectivament. En el pitjor cas, és possible que el nivell de gris del marge i de la carretera siguin semblants, i com a conseqüència obtindríem una segmentació que no es correspon amb les àrees significatives de l'escena.



*Figura 4.1. Segmentació d'imatges en entorns naturals a partir de l'anàlisi de la intensitat.
a) Imatge RGB original. b, c, d, e i f) Segmentació mitjançant la tècnica de creixement de regions en diferents condicions.*

Com es pot veure a la figura 4.1, el problema de que un objecte/estructura estigui associat a varies regions en la imatge segmentada, és independent de les condicions en les que es realitza la segmentació, ja que es deu principalment al fet que la intensitat lluminosa reflectida pels objectes sobre la imatge ve influenciada per la presència de reflexes especulars i d'ombres. D'aquesta forma, el nivell de gris dels pixels pot dependre més de les condicions d'il·luminació que de les propietats dels objectes presents a l'escena. En aquestes situacions, és preferible l'ús d'una informació de l'escena menys sensible a les condicions d'il·luminació, com ara el color o la textura, per obtenir-ne una bona segmentació.

En imatges multicromàtiques, es pot obtenir informació del color, que és una propietat intrínseca dels objectes i, si s'utilitza de forma adequada, bastant independent de les variacions d'il·luminació. Per tant, l'anàlisi d'escenes d'entorns naturals a partir d'imatges color ens ha de permetre obtenir millors resultats que l'anàlisi d'imatges monocromàtiques, tal com queda recollit en [Otha 80] [Celenk 90] [Tseng 92] [Batlle 93] o [Pujas 95], ja que el color permet una millor discriminació dels objectes de l'escena, això si, amb un augment de la complexitat del sistema de percepció.

Un altre tipus d'informació que es pot obtenir de les imatges és la textura. Aquesta característica, que està associada a propietats dels objectes, i la seva utilització en la segmentació d'imatges han estat estudiades intensament [Laws 80] [Tomita 90] [Grau 97]. Alguns autors utilitzen les textures presents a les imatges monocromàtiques com l'única característica o com una dintre d'un conjunt de característiques, per segmentar imatges captades en entorns naturals, incloent-hi algun sistema de navegació autònom per a carreteres i/o autopistes [Thorpe 88,91].

Els avantatges que presenta el color en front de la textura, en el context de la segmentació i anàlisi d'una seqüència d'imatges captades en un entorns naturals, són:

- L'obtenció de la informació color requereix, normalment, d'un processament menys costós que les característiques de textura.
- Alguns descriptors de textura divideixen la imatge en zones, a les que s'anomena texels *-texture element-*, a l'hora de calcular i representar la textura. El resultat d'aquest procés és una imatge de resolució inferior, que, en la tasca que ens interessa, redueix la precisió de posicionament al món 3D dels elements que apareixen en la imatge, i dificulta la detecció d'objectes (obstacles) petits a la imatge.
- El color es pot utilitzar com a propietat que ens permet identificar o etiquetar els objectes que s'analitzen. Això pot ser d'interès, entre altres casos, quan s'analitza una seqüència d'imatges i es necessita establir relacions interimatges amb els objectes, ja que el color d'un objecte a la imatge no es veu afectat o modificat en el procés de formació de la imatge (és independent al factor d'escala, la rotació i la perspectiva). Molts dels descriptors de textura, tenen una dependència de la posició relativa càmera-objecte.

Malgrat tots aquests avantatges, també és cert que la textura pot aportar informació per distingir objectes del mateix color, o prou semblant, quan les seves característiques superficials són prou diferenciades (per exemple, es podria distingir una pedra entre grava encara que els seus colors fossin similars).

Un altre aspecte tant important com l'elecció de les característiques de la imatge a utilitzar a l'hora de segmentar-la, és la tècnica utilitzada per definir els segments o regions de la imatge. Les tècniques de segmentació d'imatges es poden agrupar en dues categories [Gonzalez 87]:

1. *Localització dels límits de les regions.* En aquest cas s'utilitzen operadors de tipus gradient per detectar les discontinuïtats del valor de la característica

associada a cada pixel o conjunt de pixels. Els pixels on aquest gradient sobrepassa un llindar, són marcats com punts frontera entre dues regions diferents de la imatge. Aquesta tècnica és coneguda amb el nom d'extracció de contorns.

2. *Determinació dels pixels que formen la regió.* A diferència de la tècnica anterior, aquest mètode es basa en agrupar aquells pixels que a més de ser veïns, tenen un valor de la característica similar. D'aquesta forma es defineixen zones connexes i homogènies a la imatge. La binarització, l'agrupació de pixels, el creixement de regions i la divisió i fusió¹ són mètodes que s'inclouen en aquesta categoria.

Com que la zona interior dels punts contorn defineix una regió, i la frontera entre regions no és més que una imatge de contorns, aquests mètodes no són més que dos camins per arribar al mateix objectiu. Les imatges contorn són molt atractives ja que contenen en una imatge de complexitat reduïda -normalment binària- tota la informació necessària. En canvi no són útils per escenes complexes o en aquelles on la variació de la característica considerada no sempre està relacionada amb un canvi d'objecte, com és el cas de les imatges captades en entorns naturals. En aquests casos s'obtenen un gran nombre de segments de contorn, molts d'ells oberts -la frontera dels objectes a la imatge no està tancada- éssent difícil determinar les accions a prendre per a tancar-los. La figura 4.2 mostra el resultat de l'extracció de contorns de les imatges de la figura 1.2, utilitzant l'operador de Sobel i binaritzant la imatge gradient obtinguda (el llindar de binarització ha estat fixat a un valor que permetia la detecció de contorns en el conjunt de les 9 imatges considerades. El valor utilitzat és de 74).

Com es pot observar en la figura 4.2, en aquestes imatges apareixen una gran quantitat de contorns, sense que necessàriament apareixin els contorns associats als límits del camí. En molts casos, els elements que formen el camí (sorra, petites pedres, gespa, ...) o les ombres projectades sobre el camí generen contorns a la imatge que dificulten enormement la determinació de la posició a la imatge dels límits del camí. En aquests casos, la selecció de quin contorn defineixen els límits del camí pot passar per la utilització d'una tècnica de detecció de contorns globals o envolvents [Cufi 96]. Aquests contorns són aquells que presenten certes propietats com ara que la seva llargada sigui gran, comparativament amb els altres contorns de la imatge, i que separin regions de la imatge que presentin una variació important en un conjunt de característiques de color i/o textura. Si bé aquesta tècnica permet detectar, en certes

¹ En anglès, *thresholding*, *clustering*, *region growing* i *split-and-merge*, respectivament.

situacions, els contorns més " importants " de la imatge, el seu cost computacional resulta molt elevat.

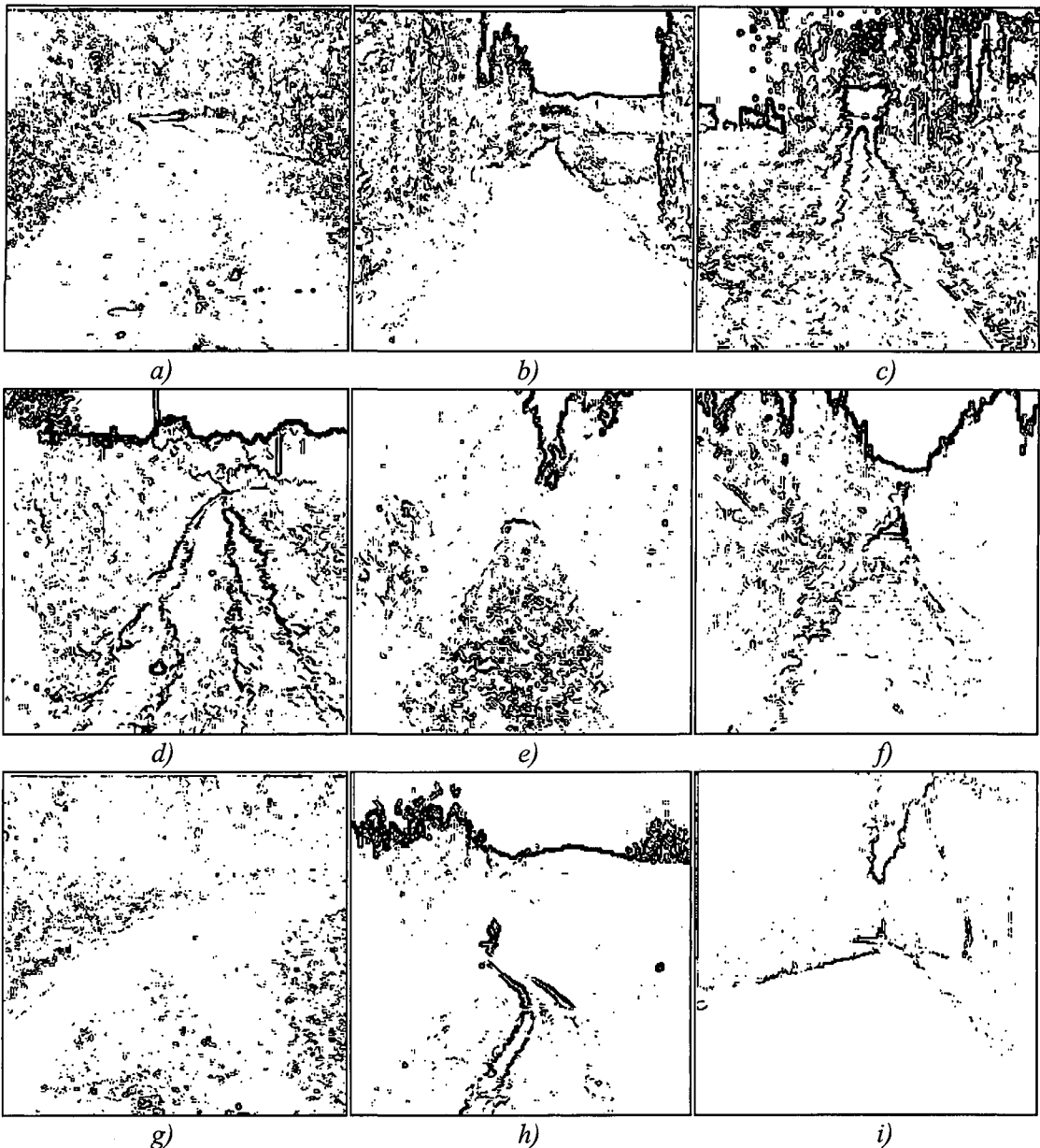


Figura 4.2. Extracció de contorns en les imatges de la figura 1.2.

Així doncs, el nivell de gris no és la característica més adequada per a la segmentació d'imatges captades en entorns exteriors (la imatge intensitat es veu fortament condicionada per les condicions d'il·luminació i els descriptors de textura, que normalment s'obtenen a partir de la imatge intensitat, resulten computacionalment més costosos d'obtenir que la característica color i/o estan associats a imatges de baixa resolució), i la segmentació per extracció de contorns no és escaient donada la complexitat de les imatges captades en entorns naturals. Aquestes dues circumstàncies

es donen en el tipus d'escenes que hem d'analitzar, la qual cosa ens porta a utilitzar un altre tipus de característica, i una tècnica de segmentació basades en la semblança.

4.1.1. El Color a la Visió per Ordinador

El color d'un pixel a la imatge està associat a la longitud d'ona de l'energia lluminosa captada. Aquest valor depèn bàsicament de la distribució espectral de la llum que incideix sobre els objectes de l'escena, de la distribució espectral de la reflectància d'aquest objecte, i de la sensibilitat espectral de la càmera. La reflectància d'un objecte és una de les seves característiques intrínseques. Aquest és l'atribut que ens permet dir si un objecte és blau, groc o vermell. Igual que pels humans, el color pot ser una característica molt interessant per un sistema de visió per ordinador a l'hora de segmentar, identificar i reconèixer els objectes. Per contra, el fet de que el fenomen de la percepció del color sigui tant sofisticat, fa que una tasca trivial pels humans com és la identificació del color d'un objecte, amb certa independència de les condicions d'observació -de la il·luminació de l'escena i de la càmera utilitzada-, sigui una tasca sense una solució eficient, en molts casos, en l'àmbit de la visió per ordinador.

Al llarg de la història dels sistemes de navegació basats en visió, la utilització d'imatges color s'ha revelat molt útil. Com queda reflectit al segon capítol, un nombre important de sistemes de navegació autònoms es basen en l'anàlisi d'imatges color per realitzar les tasques de detecció i seguiment de carreteres, aprofitant que el color d'una calçada pavimentada és ben diferent al color dels marges. En canvi, en les imatges monocromàtiques l'aspecte de la calçada depèn de la presència de reflexes especulars i d'ombres. D'aquesta forma, la segmentació de la carretera en la imatge color, que conté més informació, pot resultar més robusta que en una imatge monocroma, això sí, al preu d'un processat més complexa.

Sistemes de Representació del Color

Si bé la forma bàsica de representar el color d'un pixel és mitjançant els valors RGB que subministren les càmeres color, aquest no és el sistema més adient pel processament de la informació color. El processament d'imatges RGB presenta alguns problemes, com ara:

- El gran volum d'informació (típicament 24 bits/pixel) que s'ha de tractar.
- La mètrica euclidiana no és adequada per mesurar la diferència de color a l'espai RGB.

- La forta correlació entre els components RGB i que els seus valors no tenen una relació directa amb les característiques perceptual humanes com ara la cromàtica i la lluminositat.

Aquests inconvenients fan que s'utilitzin altres sistemes de representació. Els més utilitzats en l'àmbit de la colorimetria i la visió per ordinador són:

- **Y,I,Q.** El model YIQ és un sistema definit per la NTSC (National Television Systems Committee). La transformació RGB/YIQ ve donada per:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

$$I = 0.500 * R - 0.230 * G - 0.270 * B$$

$$Q = 0.202 * R - 0.500 * G + 0.298 * B$$

Aquest sistema es va definir per a la transmissió de la TV color amb l'objectiu de minimitzar l'ample de banda del senyal. El component Y està pensat per obtenir la compatibilitat amb els televisors blanc i negre, ja que és la lluminància. Els components I i Q, mitjançant una transformació lineal simple, contenen la informació cromàtica i són independents de la lluminància.

- **I1,I2,I3.** Aquest sistema proposat per Otha [Otha 80] parteix de l'anàlisi de vuit imatges, a les quals ha aplicat la transformació de Karhunen-Loewe. Per tant, el sistema proposat és el millor, des del punt de vista estadístic, per representar la diferència de colors per aquest conjunt de vuit imatges utilitzats per definir el sistema. Els paràmetres I1,I2,I3 s'obtenen mitjançant una transformació lineal dels valors R,G,B, però per una imatge donada els altres sistemes de coordenades poden ser més eficaços a l'hora de realitzar la segmentació. Els valors d'aquests paràmetres són:

$$I1 = (R+G+B) / 3$$

$$I2 = (R-G) / 2$$

$$I3 = (2 * G - R - B) / 4$$

- **U*,V*,W* i L,a,b.** Aquests models de representació del color, establert per la CIE (Commission Internationale de l'Eclairage) a l'any 1964 a partir del model XYZ, tenen com a objectiu definir un sistema més proper a la percepció humana, són perceptualment uniformes i en ells la distància euclidiana en el diagrama cromàtic associat, és significativa de la diferència entre colors observats. Malgrat que són bastant utilitzats en el camp de la colorimetria, són poc utilitzats en la visió per ordinador.

$$\begin{aligned}
a &= 500 * ((100 * X / X_0)^{1/3} - (Y / Y_0)^{1/3}) & X &= 0.49 * R + 0.31 * G + 0.2 * B \\
b &= 200 * ((Y / Y_0)^{1/3} - (Z / Z_0)^{1/3}) & Y &= 0.18 * R + 0.81 * G + 0.011 * B \\
L = W^* &= 25 * (100 * Y / Y_0)^{1/3} - 16 & Z &= 0.00001 * R + 0.01 * G + 0.99 * B
\end{aligned}$$

X_0, Y_0, Z_0 són les coordenades a l'espai XYZ del blanc de referència

$$\begin{aligned}
U^* &= 13 * W^* (u - u_0) & u &= 4 * X / (X + 15 * Y + 3 * Z) \\
V^* &= 13 * W^* (v - v_0) & v &= 6 * Y / (X + 15 * Y + 3 * Z) \\
W^* &= 25 * (100 * Y / Y_0)^{1/3} - 16 & u_0 &= 0.199 & v_0 &= 0.308
\end{aligned}$$

- **HSI.** El sistema HSI (croma -o Hue-, Saturació i Intensitat) posseeix propietats tant interessants com la independència entre la croma, la saturació i la intensitat, i té una interpretació directa en termes de percepció humana. La croma (hue) és el valor que diferencia els colors (verd, blau, vermell, groc, etc. ..), la saturació fa referència a la puresa del color o la quantitat de color blanc que en conté (és el matis entre el blau marí i el blau cel o entre el vermell i el rosa) i la intensitat fa referència a la quantitat de llum captada. Els valors HSI s'obtenen a partir dels valors RGB, obtinguts per un captador d'informació color. Una de les transformacions utilitzades per obtenir els valors HSI és:

$$I = (R + G + B) / 3 \quad (4.1)$$

$$S = 1 - \min(R, G, B) / I \quad (4.2)$$

$$H = \begin{cases} \arccos \left(\frac{\frac{1}{2}((R - G) + (R - B))}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right) & \text{si } G \geq B \\ 2\pi - \arccos \left(\frac{\frac{1}{2}((R - G) + (R - B))}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right) & \text{si } B > G \end{cases} \quad (4.3)$$

Malauradament la transformació RGB a HSI no és lineal i presenta alguns inconvenients, examinats en detall per Kender [Kender 76]:

1. La funció H no està definida pels punts acromàtics ($R=G=B$).
2. La funció S no està definida pel color negre ($R=G=B=0$). Petites variacions de R, G o B quan es representa el color negre provoquen variacions importants de H i/o de S.
3. El valors H i S són poc estables si la intensitat lluminosa, I, és molt gran o molt petita.
4. La distribució dels valors transformats no és uniforme.

Malgrat els inconvenients que presenta el sistema HSI, aquest és molt utilitzat per a la segmentació d'imatges color, ja que els seus avantatges (alt poder discriminador del hue, independència entre els seus valors H, S i I, interpretació directa en termes de percepció humana, obtenció dels valors HSI a partir dels RGB amb menys càlculs que el sistema Lab i $U^*V^*W^*$) no els ofereixen els altres sistemes de representació del color.

Tots aquests sistemes són equivalents ja que tots representen en espais diferents la mateixa informació, i es pot passar sense gran dificultat d'una a l'altre. No hi ha cap d'ells que sigui clarament millor que els altres, i fins i tot, hi ha estudis [Otha 80] que arriben a la conclusió que el resultat de la segmentació d'una imatge color es veu poc influenciada pel sistema de representació del color utilitzat. A la secció 4.2.2, es presenten alguns treballs de segmentació d'imatges color d'interès.

4.1.2. L'Espai de Color H/I Proposat

Per a la resolució del problema que ens plantegem, la segmentació d'imatges captades en entorns naturals, s'ha mostrat la conveniència d'utilitzar imatges cromàtiques, encara que queda per triar la forma en la que es representarà la informació color. Ens interessa representar el color amb uns atributs altament discriminadors i independents de la lluminància (que depèn bàsicament de la il·luminació ambiental). Així mateix, ens interessa que la característica sigui el més estable possible davant de variacions de la il·luminació, continua i amb una resolució cromàtica adequada pel tipus d'escenes que s'ha d'analitzar. Com s'ha indicat anteriorment, el sistema RGB -en el qual s'obtenen les imatges subministrades per una càmera color- no presenta aquestes característiques. D'altra banda, la resolució colorimètrica més habitual de les imatges RGB (2^{24} colors diferents) resulta molt gran per a la majoria d'aplicacions.

Diferents autors [Tominaga 87] [Luong 90] [Tseng 92] [Pujas 95] han arribat a la conclusió que el paràmetre hue (o cromà) és el que presenta per si sol un major poder discriminador entre tots aquells que estan associats amb el color de la llum captada (R, G, B, I, X, Y, Z, I1, I2, I3, Q, U, V, W, L, a, b, H, S, etc.). Altres propietats d'interès del hue en front dels altres paràmetres són que té una interpretació directa en termes de la visió humana i que resulta bastant independent de les condicions d'il·luminació. Aquestes raons fan que l'espai HSI sigui un dels més utilitzats en els sistemes de visió per ordinador que han d'analitzar el color. Entre les diferents transformacions existents

per obtenir els valors HSI a partir dels valors RGB subministrats per la càmera, la més utilitzada és la indicada a les equacions 4.1, 4.2 i 4.3.

El fet que s'utilitzi l'espai HSI (3 valors) per representar i processar les imatges color, en comptes d'únicament el Hue és degut, principalment, a dues raons:

- La necessitat d'una gran resolució coloromètrica. Si es vol que el sistema pugui detectar canvi en la il·luminació de l'escena o petits canvis de matís del color (entre un verd oliva i un verd ampolla o entre un groc lluminós i un groc llimona), és necessari l'anàlisi de la saturació i la intensitat lluminosa.
- El hue no està definit o és inestable per certs valors de RGB (veure secció 4.1.1).

Com es pot veure, l'espai de color HSI presenta propietats interessants per la segmentació d'imatges color. Per contra, seria desitjable eliminar o minimitzar els efectes dels inconvenients que presenta aquest espai, així com reduir el nombre de color a considerar per adequar la complexitat coloromètrica de la imatge amb la tasca que es vol realitzar, la segmentació d'imatges captades en entorns naturals.

Per solucionar aquests problemes, proposem un nou espai de representació del color, el **H/I (Hue/Intensitat)**, derivat del l'espai HSI. Amb aquesta representació, es pot treballar de forma eficient amb un volum de dades inferior, triant per a cada pixel el paràmetre més adequat, és a dir, aquell que resultarà més útil durant el procés de segmentació. La conversió RGB a H/I utilitzada es realitza en dos passos:

1. Determinar per a cada pixel si la conversió RGB a hue és inestable, és a dir, si el hue està definit i resulta un bon discriminador per poder determinar a quina regió pertany el pixel.
2. Si el hue és inestable (no és un bon discriminador), s'assigna al paràmetre H/I el valor de I (equació 4.1), sinó s'assigna al paràmetre H/I el valor de H (equació 4.3).

Donat el valor (R,G,B) d'un pixel p , el *hue* serà el valor més discriminant si està definit pel valor (R,G,B), i si tots els valors $(R\pm i, G\pm j, B\pm k)$ $-0\leq i,j,k\leq M$ que representen un color similar tenen un valor de hue similar $(H\pm 1 -0\leq 1\leq N)$. Quan es dona aquesta situació es diu que la transformació RGB a H és estable pel valor (R,G,B)). En cas contrari es selecciona la *Intensitat* com paràmetre més discriminant pel pixel p . El valor M està relacionat amb el soroll present a la imatge RGB i el valor N de la distància màxima entre dos hues que perceptivament es consideren diferents.

Una forma d'observar la inestabilitat de la transformació RGB a Hue és analitzant la imatge hue obtinguda d'una carta o gamma de colors que representa tot l'espai RGB. La figura 4.3.a mostra una imatge sintètica d'una carta de colors representatius de tot l'espai RGB, on apareixen a prop de 3000 colors o divisions. Cada divisió està formada per 125 pixels de color ($A \cdot m + E_R$, $A \cdot n + E_G$, $A \cdot o + E_B$), on A és una constant que indica quin és el mínim increment del valor d'una component a l'espai RGB per apreciar perceptivament un color diferent i $0 \leq m, n, o \leq D$ són constants per tots els pixels de la divisió -els valors ($A \cdot m, A \cdot n, A \cdot o$) fixen el color bàsic de la divisió-, i $0 \leq E_R, E_G, E_B \leq S$ és un "soroll" aleatori de distribució uniforme diferent per a cada pixel de la divisió -S representa el soroll present en una imatge RGB digitalitzada². La figura 4.3.b mostra la imatge hue obtinguda aplicant l'equació 4.1. Com es pot observar per la majoria de colors (divisions), la presència del "soroll" en la imatge RGB no provoca variacions significatives del valor H obtingut (la variació és inferior o igual a N). Per aquestes divisions diem que la transformació RGB a H és estable. Per contra, hi ha un nombre important de divisions en les quals la petita variació del valor (R,G,B) provoca una variació molt important del valor H en la imatge transformada, éssent en aquests casos la transformació RGB a H inestable.

Les divisions que mostren ser inestables (per $N=1$) en la transformació RGB a H, tenen com característica comú que $I=(R+G+B)/3$ és inferior a 15 o que la mínima diferència entre els valors de R,G i B és inferior a 10^3 .

Per tant la transformació RGB a H/I queda definida per:

$$H / I = \begin{cases} H & \text{Si la transformació RGB a Hue es estable} \\ I & \text{Si la transformació RGB a Hue es inestable} \end{cases}$$

El hue és un paràmetre cíclic que normalment s'expressa en graus -entre 0° i 360° -. Una decisió important a prendre és la resolució coloromètrica amb la que es defineix l'espai H/I, és a dir, el nombre de valors de H i valors de I utilitzats a l'espai H/I. Aquesta decisió està molt lligada a l'aplicació a la que està destinada la imatge H/I.

² Donat que la resolució de l'espai RGB utilitzat per representar la imatge 4.3.a és de 256 valors per component, i que per aquest cas el mínim increment que produeix una variació perceptiva del color és de $A=20$. Per obtenir tota la gamma de color cal que $D=256/20=13$, i es fixa $S=5$. Les divisions veïnes en el recorregut esquerra-dreta/inferior-superior és diferencien en A unitats d'una de les tres components RGB (són colors semblants). Les discontinuïtats de color que apareixen són degudes a la representació bidimensional d'un espai tridimensional.

³ L'espai RGB considerat té una resolució de 256 valors per component.

En el nostre cas, la descripció de camins en entorns naturals mitjançant la segmentació de la imatge color, com no es requereix d'una gran resolució coloromètrica (com es podrà veure a la secció 4.2.3), resulta interessant treballar amb un conjunt reduït de colors per minimitzar el volum de memòria necessària per representar la informació i reduir la complexitat i el temps d'execució dels algorismes utilitzats per analitzar-la. Per aquesta raó, s'ha optat per representar l'espai H/I utilitzant 1 byte per pixel. Per reduir el nombre d'operacions necessàries per distingir els valors H dels I, i calcular la diferència de color entre dos valors a l'espai H/I de forma ràpida, s'ha utilitzat 128 hues diferents (representats amb els valors de 0 a 127) i 64 nivells de gris diferents (representats amb els valors de 128 a 191). A la figura 4.4 es mostren els 192 colors utilitzats en l'espai H/I. Les 8 files superiors corresponen als hues (fila superior esquerra H/I=0 i fila 8 dreta H/I=127) i les 4 inferiors els nivells de gris (fila 9 esquerra H/I=128 i fila 12 dreta H/I=191).

En la figura 4.5 es pot observar els efectes de la reducció coloromètrica de l'espai H/I. La figura 4.5.a mostra la imatge RGB d'una carta amb 60 mostres de color diferents adquirida mitjançant un "scanner" i la figura 4.5.b mostra la seva transformació a l'espai H/I. En la imatge 4.5.b es pot observar com malgrat la pèrdua de resolució coloromètrica de la imatge H/I (en la imatge RGB de la figura 4.5 es fan servir més de 16 milions de valors⁴, i en la H/I només s'utilitzen 192 valors), aquesta conserva prou informació per indicar quin és el color dels objectes, i permet, en determinades aplicacions, obtenir una segmentació de la imatge coherent amb l'escena (veure secció 4.2.3). Aquest nivell de discretització ha demostrat ser suficient per segmentar correctament imatges captades en entorns naturals [Fernández 97a].

La figura 4.6 mostra el resultat de la conversió de l'espai RGB al H/I, de les imatges RGB corresponents a camins feblement estructurats, de la figura 1.2. Les imatges H/I resultants, es representen amb una resolució de 192 valors.

⁴ El nombre de colors més habitual d'una imatge RGB obtinguda amb un sistema d'adquisició d'imatges és de 8 bits per component, el que fa un total de 2^{24} colors. Ja que aquesta resolució és superior a la de la visió humana, hi ha sistemes que amb l'objectiu de reduir el cost del sistema d'adquisició, treballen amb un nombre inferior de colors (com ara 4 o 6 bits per component, el que fa, respectivament, 2^{12} i 2^{18} colors).

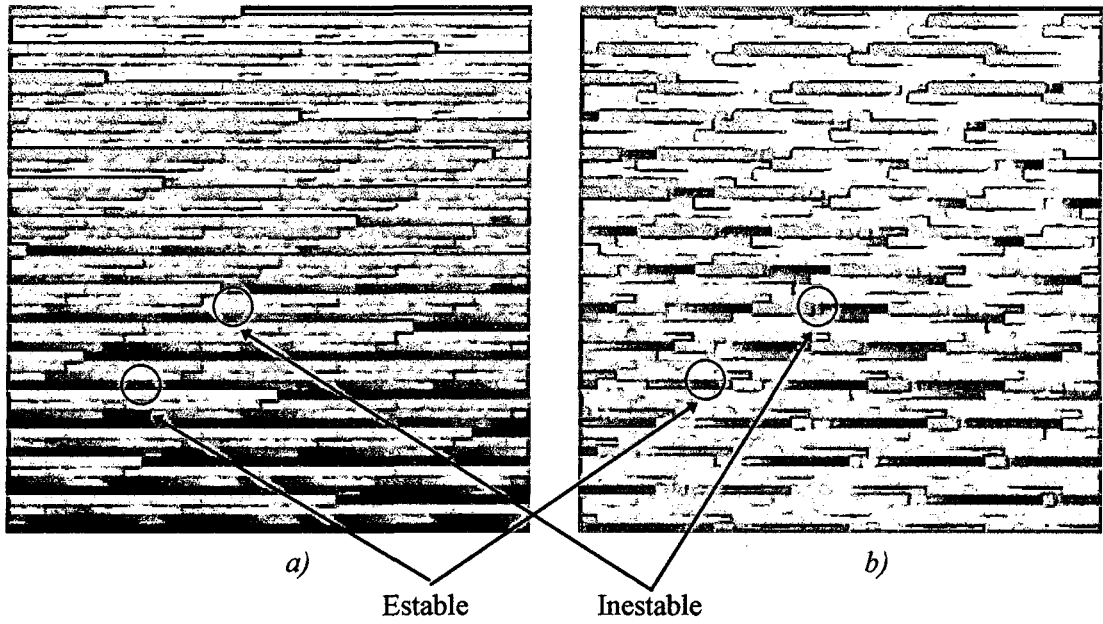


Figura 4.3. Imatge sintètica d'una carta amb 2916 colors.
 a) Imatge RGB amb soroll aleatori b) Imatge Hue

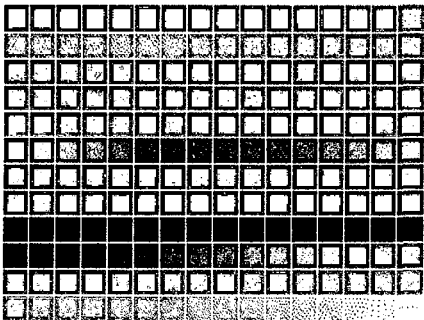


Figura 4.4. Colors que formen l'espai H/I amb una resolució de 192 valors. (128 Hues i 64 grisos).

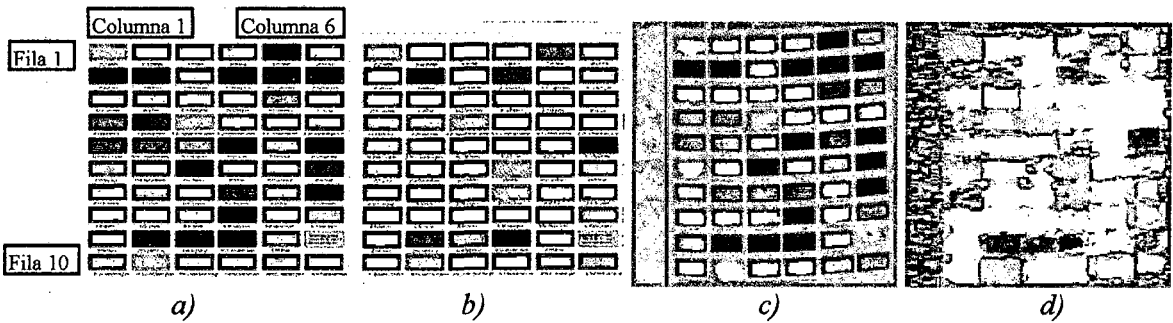


Figura 4.5. Imatges d'un mapa de 60 colors. a) Imatge RGB captada amb un scanner.
 b) Imatge H/I de (a). c) Imatge RGB captada amb una vídeo càmera. d) Imatge H/I de (c).

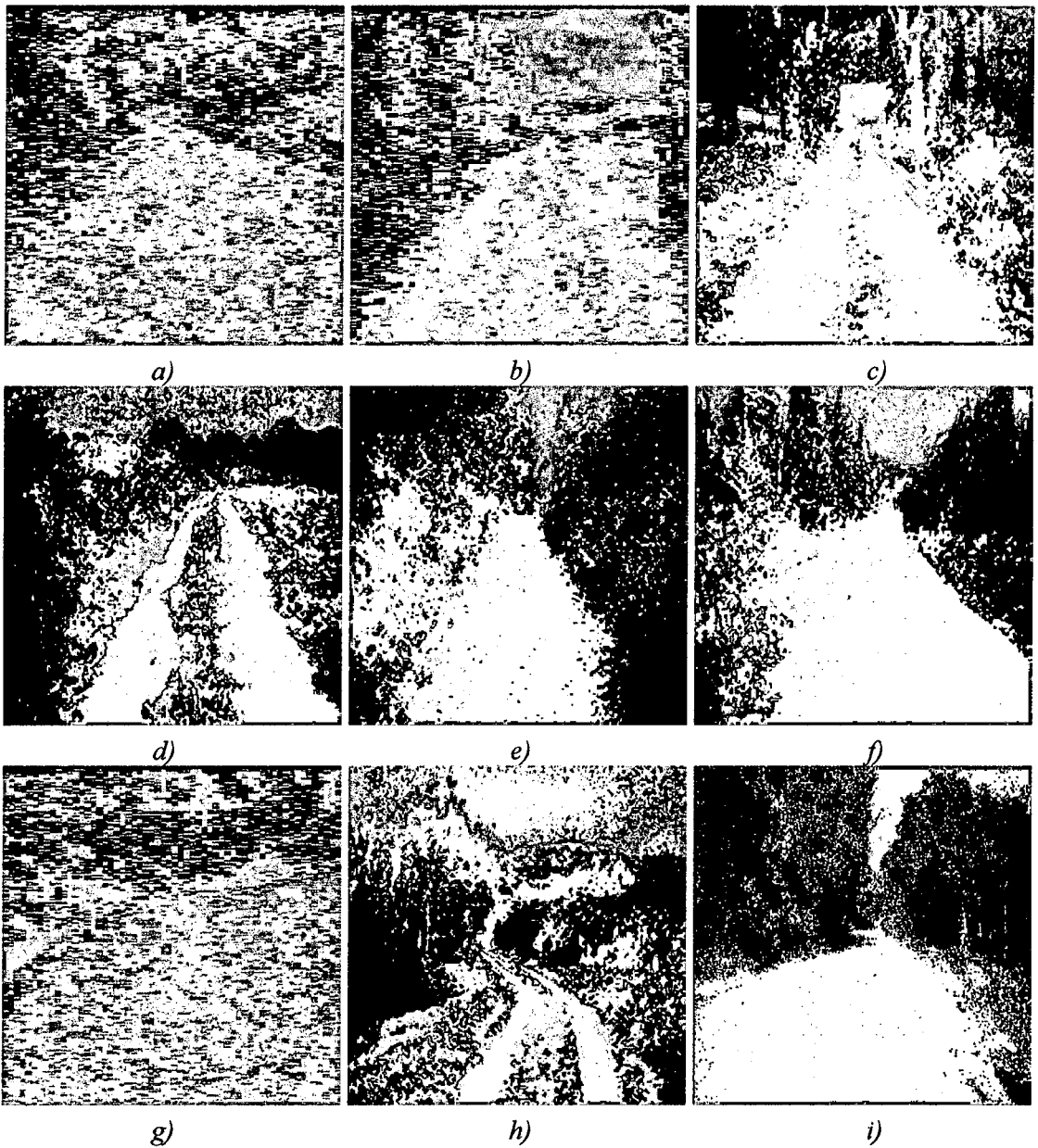


Figura 4.6. Resultat de la representació a l'espai H/I de les imatges RGB de la figura 1.2.

En la figura 4.7 es poden veure alguns dels avantatges de l'espai H/I per la segmentació d'imatges captades en entorns naturals. En la figura 4.7.b (imatge hue) es pot observar com en certes zones de la imatge (part dreta i esquerra superior que corresponen a les zones més fosques de la imatge RGB), el valor del hue varia de forma considerable encara que correspon a valors RGB semblants en la imatge original (figura 4.7.a). Aquest fet, degut a la inestabilitat de la transformació RGB a hue, no es dona a la imatge H/I -figura 4.7.c-, ja que en aquestes zones s'utilitza la intensitat en lloc del hue. Les figures 4.7.d, 4.7.e i 4.7.f corresponen a la segmentació utilitzant una tècnica de creixement de regions (veure secció 4.2.2), si s'analitza la intensitat o nivell de gris, el hue o el valor H/I respectivament. Els inconvenients de la transformació RGB a hue fan que el nombre de regions que apareixen en la

segmentació obtinguda mitjançant l'anàlisi del hue (figura 4.7.e) sigui elevat, ja que un mateix objecte/estructura genera diverses regions a la imatge segmentada quan el hue de l'objecte a la imatge és inestable. Com la transformació RGB a H/I evita aquest problema (figura 4.7.c) -valors (R,G,B) similars donen valors H/I similars-, es pot obtenir una segmentació de la imatge més coherent amb l'escena. Així mateix, es pot observar la millora obtinguda en la figura 4.7.f -segmentació a partir de l'anàlisi del H/I- respecte la figura 4.7.d -segmentació a partir de l'anàlisi de la intensitat o I-. La independència del hue respecte les variacions d'il·luminació, fa possible que la zona de carretera quedi inclosa pràcticament en una única regió en la imatge H/I segmentada. El nombre de regions generades per la segmentació de la imatge de la figura 4.7.a analitzant la intensitat, el hue o el H/I són respectivament, 106, 97 i 30. Com es pot veure comparant aquestes tres segmentacions, el paràmetre H/I resulta més eficient pel tipus d'escena que considerem. En la secció 4.2 s'analitzarà en detall l'ús de la característica H/I per a la segmentació d'imatges captades en entorns naturals.

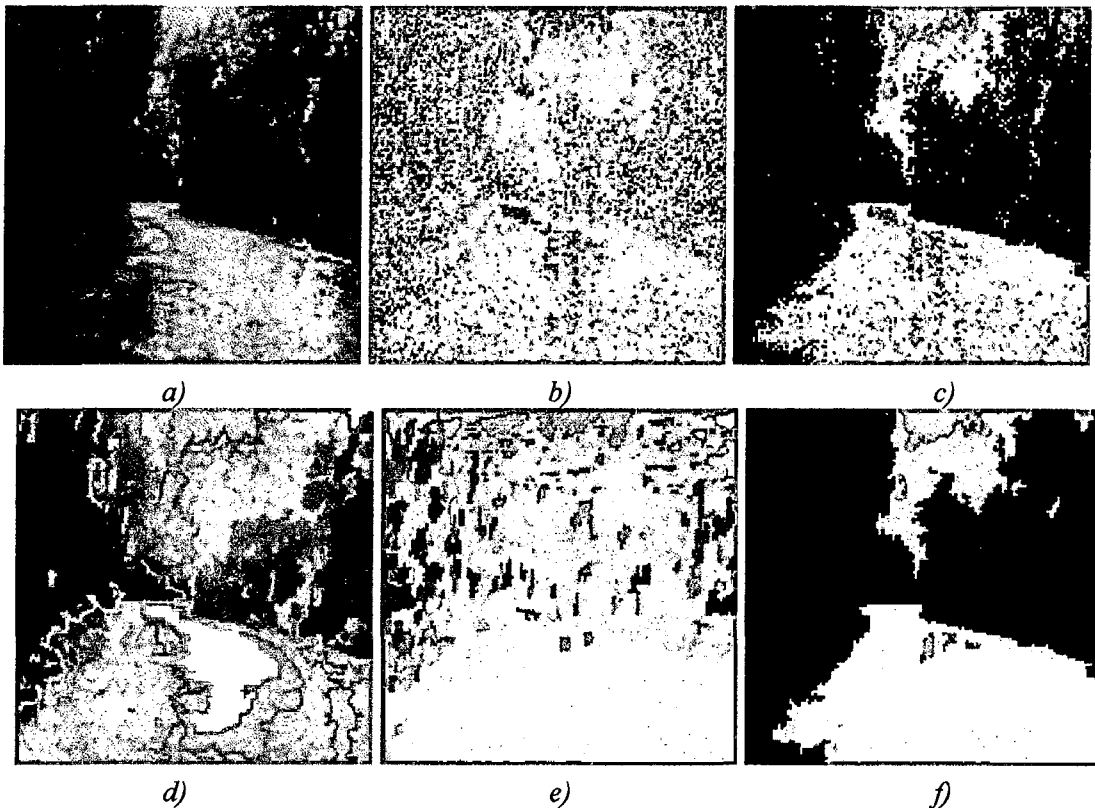


Figura 4.7. Segmentació d'una imatge color amb diferents característiques.

a) Imatge RGB. b) Imatge Hue. c) Imatge H/I.

d) Segmentació a partir de l'anàlisi de la intensitat. e) Segmentació a partir de l'anàlisi del Hue. f) Segmentació a partir de l'anàlisi del H/I.

A la secció 4.1.3, es realitza un estudi de la validesa de l'espai H/I per a la segmentació d'imatges de camins feblement estructurats, i un anàlisi de la variació del

valor H/I en imatges captades en aquests entorns en condicions d'il·luminació variables i desconegudes.

Existeixen dues alternatives per a la implementació de la transformació RGB a H/I, la primera mitjançant programari, on per cada pixel de la imatge de valor (R,G,B) es calcula quin es el valor H/I associat, i la segona, incorporant dintre de la placa d'adquisició d'imatges un mòdul dedicat a la transformacions dels valors dels pixels, de forma que sigui el *hardware* d'adquisició el que subministri directament la imatge H/I. En l'annex B es presenten en detall les diferents alternatives per a la implementació de la transformació RGB a H/I.

4.1.3. Estudi de la Validesa de l'espai H/I

En aquesta secció es fa un estudi empíric de la capacitat discriminadora del hue, de l'estabilitat d'aquest paràmetre en front de les variacions de la il·luminació a l'escena i de la poca informació que aporta la saturació per a la segmentació d'imatges color. Finalment s'analitzarà quina variació presenta el paràmetre H/I quan representa imatges captades amb una càmera en un entorn exterior.

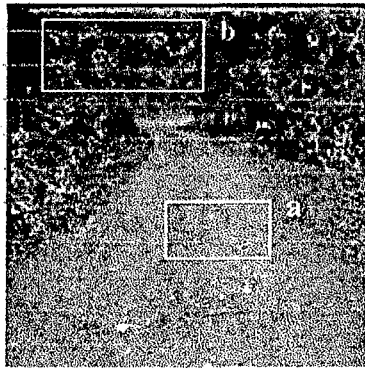
A més de la constatació perceptiva que s'observa a les figures 4.4, 4.5 i 4.6, de que les imatges H/I conserven prou informació per identificar i diferenciar els colors bàsics presents en una imatge, resulta interessant fer un estudi del comportament del valor H/I en el tipus d'imatge considerades. En les figures 4.8, 4.9, 4.10, 4.11 i 4.12 es poden veure les variacions dels valors RGB, HSI i H/I en les zones emmarcades. Cada una de les zones emmarcades correspon a una àrea de l'entorn formada per una única matèria (terra, gespa, arbres).

En aquestes zones, el paràmetre hue és el més estable entre els sis representats, en front dels canvis del nivell de saturació (o matís) del color propis dels elements naturals, i dels canvis de la il·luminació propis dels entorns exteriors. En la taula 4.1 es recull uns quants valors estadístics de la distribució del H/I, en les zones emmarcades de les figures de 4.8 a 4.12, com són la moda (valor més freqüent) del valor H/I, la variació entre la moda i el valor mínim, i la variació entre la moda i el valor màxim.

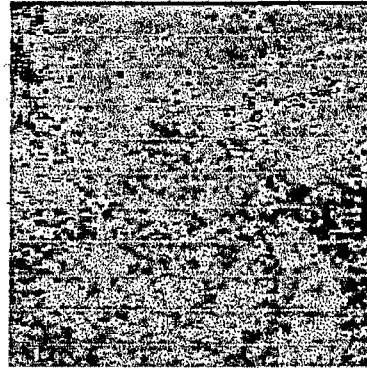
Encara que la matèria en les zones emmarcades sigui única, el color de les zones no és homogeni, sinó que en elles pot aparèixer una variació del color de la matèria. D'aquí la variació significativa dels valors H/I respecte de la moda (en mitja en les zones indicades a la taula 4.1, entre -5.0 i +6.3). Així mateix, en la taula 4.1 es pot veure la diferència significativa entre el valor de la moda del paràmetre H/I en les diferents zones emmarcades (zones a i b de la imatge que delimiten zones de diferents matèries) d'una mateixa imatge. Com conclusió de l'estudi realitzat, es pot dir que el valor H/I resulta un paràmetre força estable en front a les variacions de la il·luminació i permet diferenciar entre els diferents elements que normalment apareixen en els entorns naturals. Per contra la saturació és un paràmetre molt menys estable que depèn de la il·luminació i de les petites variacions de matis del color dels materials.

Zona	Moda	Var.Mín.	Var.Màx.
4.8.a	15	-8	+7
4.8.b	27	-7	+10
4.9.a	12	-7	+12
4.9.b	27	-3	+7
4.10.a	15	-8	+7
4.10.b	9	-2	+2
4.11.a	18	-5	+7
4.11.b	9	-2	+2
4.12.a	8	-3	+3

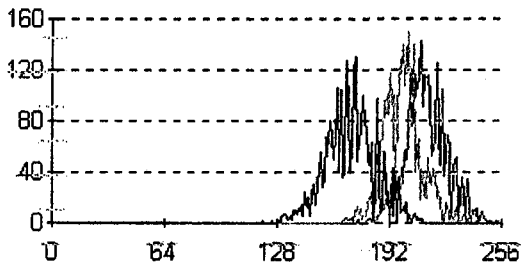
Taula 4.1. Variació del paràmetre H/I en les zones emmarcades de les figures 4.8 a 4.12.



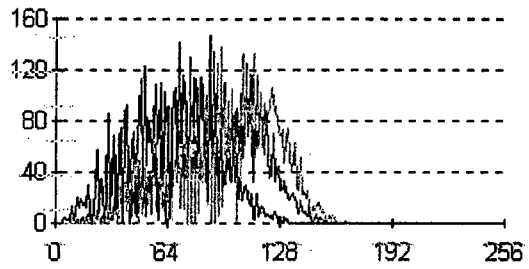
a)



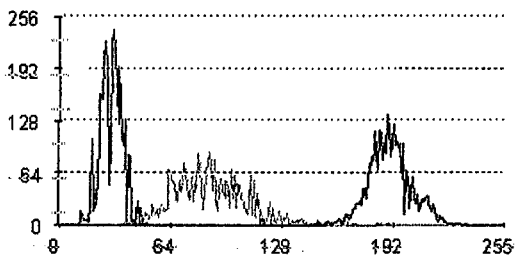
b)



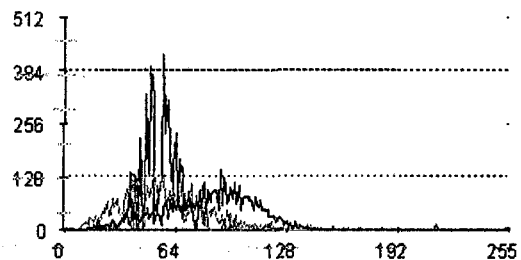
c)



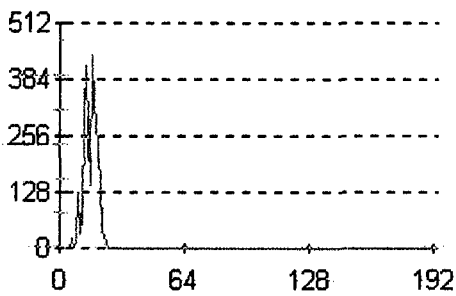
d)



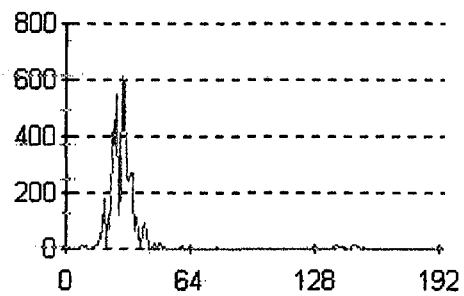
e)



f)

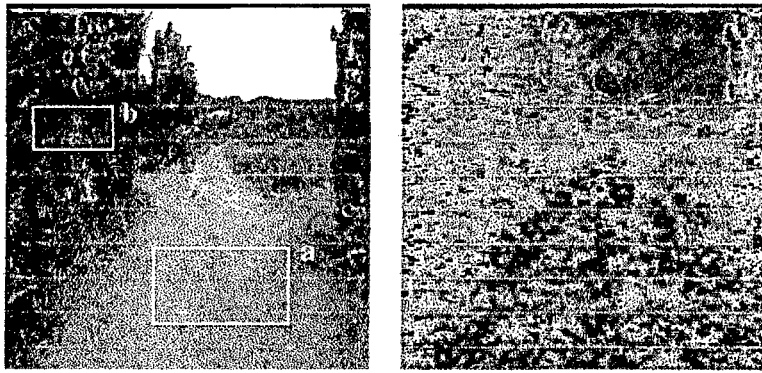


g)



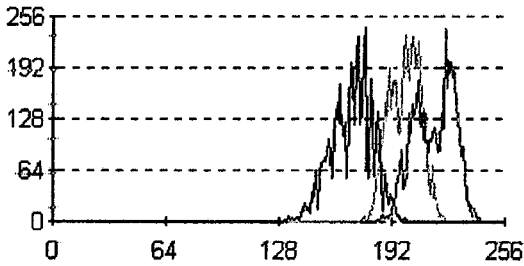
h)

Figura 4.8. Anàlisi de la variació dels valors RGB, HSI i H/I -I.
a) Imatge RGB amb les zones a i b marcades. b) Imatge H/I de a.
c) Histograma RGB de la zona a. d) Histograma RGB en la zona b.
e) Histograma HSI de la zona a. f) Histograma HSI en la zona b.
(en vermell R,H. en verd G,S i en blau B,I)
g) Histograma H/I de la zona a. h) Histograma H/I de la zona b.

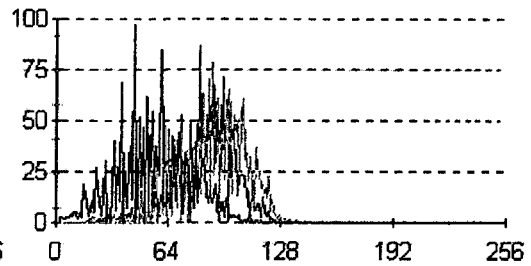


a)

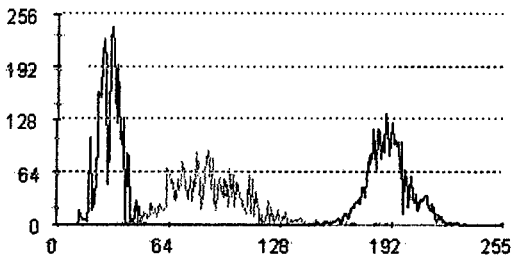
b)



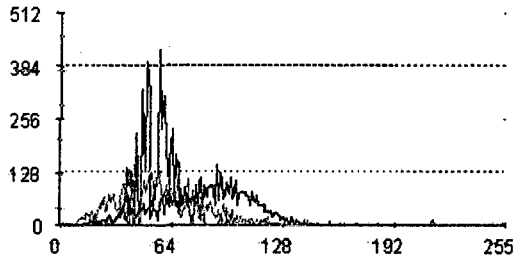
c)



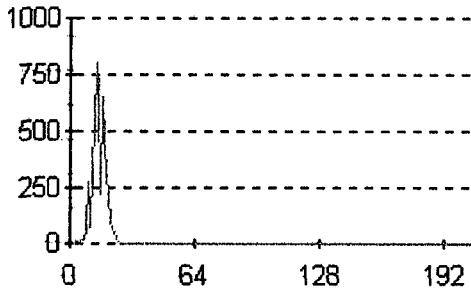
d)



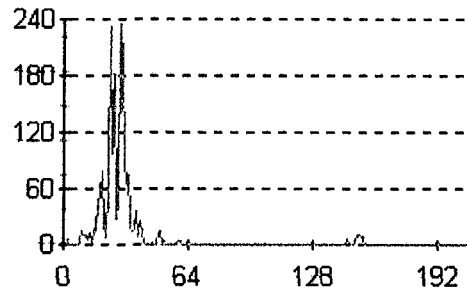
e)



f)

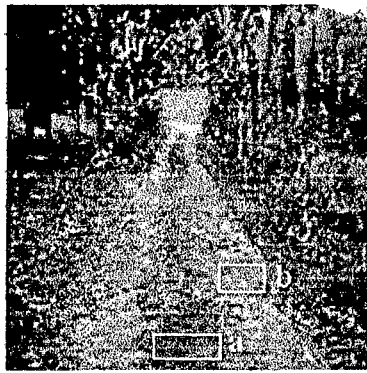


g)



h)

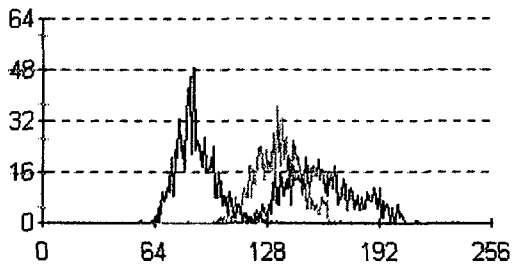
Figura 4.9. Anàlisi de la variació dels valors RGB, HSI i H/I -II.
a) Imatge RGB amb les zones a i b marcades. b) Imatge H/I de a.
c) Histograma RGB de la zona a. d) Histograma RGB en la zona b.
e) Histograma HSI de la zona a. f) Histograma HSI en la zona b.
(en vermell R,H. en verd G,S i en blau B,I)
g) Histograma H/I de la zona a. h) Histograma H/I de la zona b.



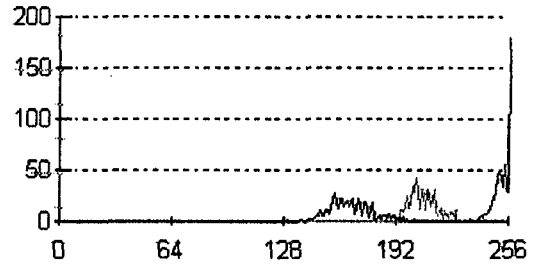
a)



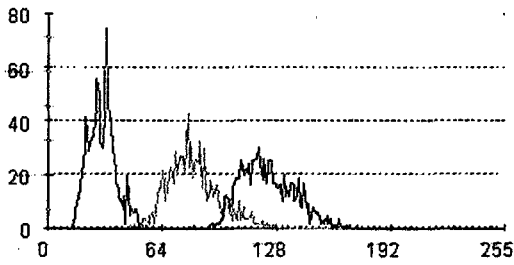
b)



c)



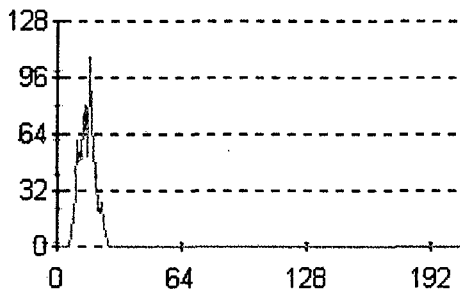
d)



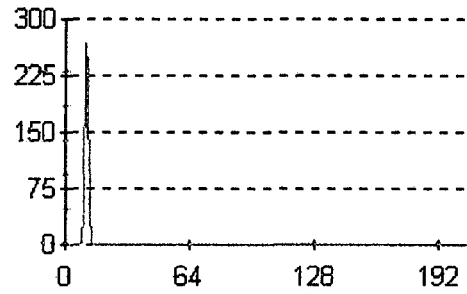
e)



f)

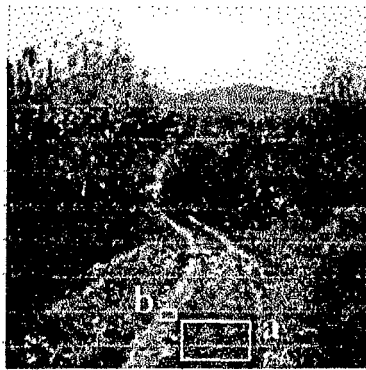


g)

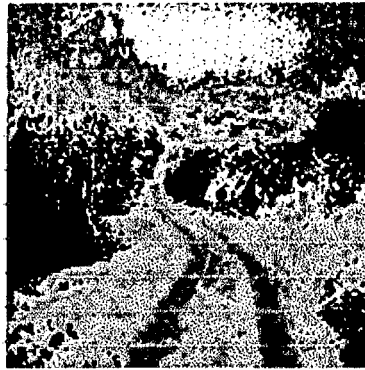


h)

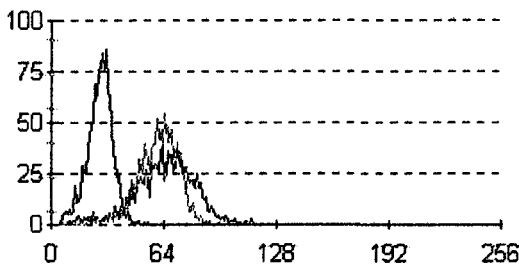
Figura 4.10. Anàlisi de la variació dels valors RGB, HSI i H/I -III.
a) Imatge RGB amb les zones a i b marcades. b) Imatge H/I de a.
c) Histograma RGB de la zona a. d) Histograma RGB en la zona b.
e) Histograma HSI de la zona a. f) Histograma HSI en la zona b.
(en vermell R,H. en verd G,S i en blau B,I)
g) Histograma H/I de la zona a. h) Histograma H/I de la zona b.



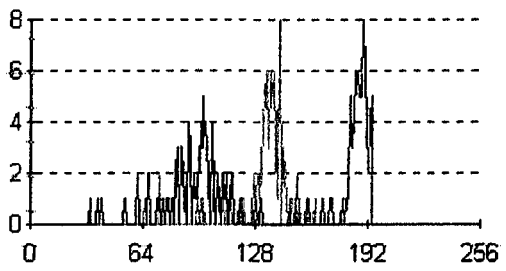
a)



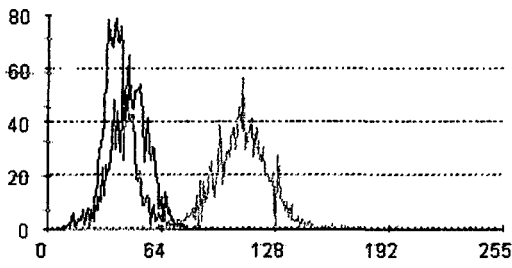
b)



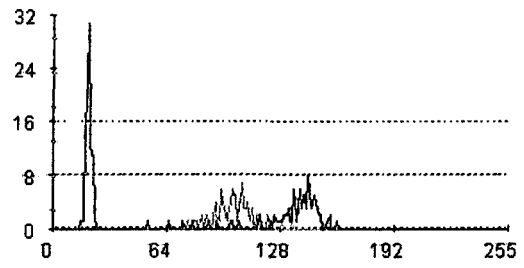
c)



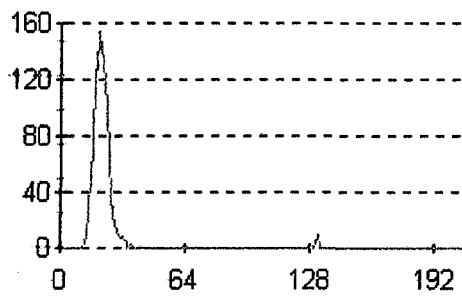
d)



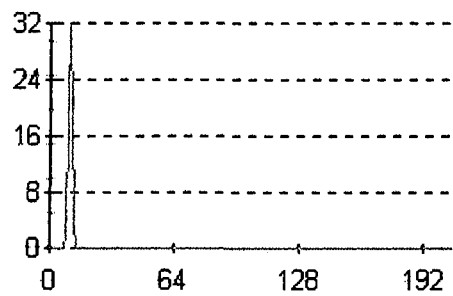
e)



f)

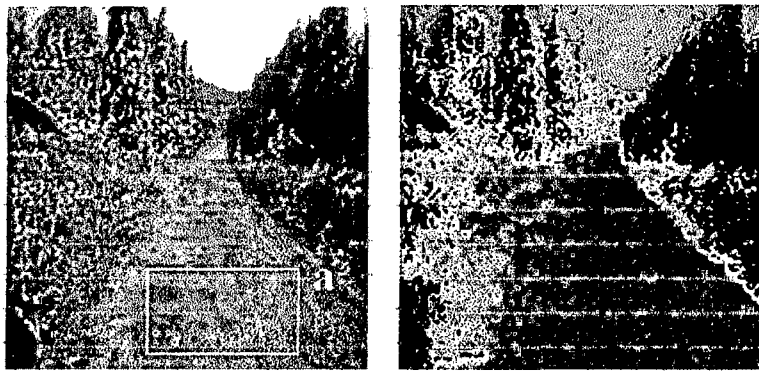


g)



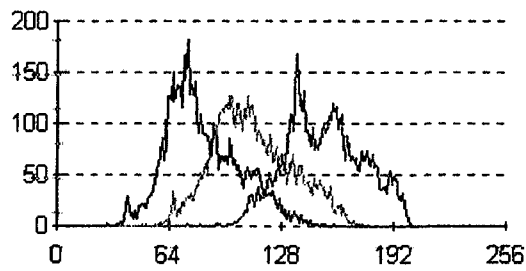
h)

Figura 4.11. Anàlisi de la variació dels valors RGB, HSI i H/I -IV-.
a) Imatge RGB amb les zones a i b marcades. b) Imatge H/I de a.
c) Histograma RGB de la zona a. d) Histograma RGB en la zona b.
e) Histograma HSI de la zona a. f) Histograma HSI en la zona b.
(en vermell R, H. en verd G, S i en blau B, I)
g) Histograma H/I de la zona a. h) Histograma H/I de la zona b.

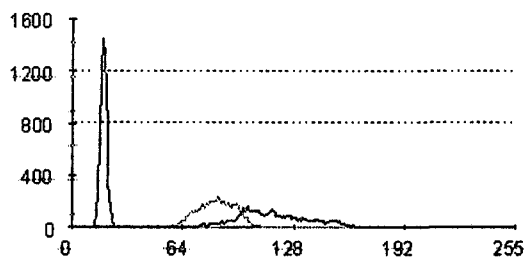


a)

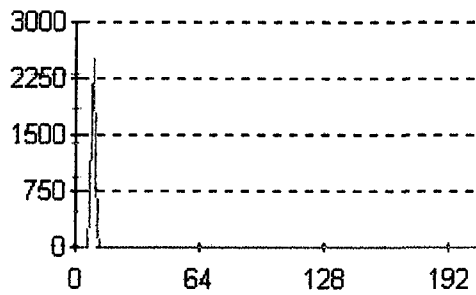
b)



c)



d)



e)

Figura 4.12. Anàlisi de la variació dels valors RGB i H/I -V.
a) Imatge RGB amb les zones a i b marcades. b) Imatge H/I de a.
c) Histograma RGB de la zona a. d) Histograma HSI de la zona a.
(en vermell R,H. en verd G,S i en blau B,I)
e) Histograma H/I de la zona a.

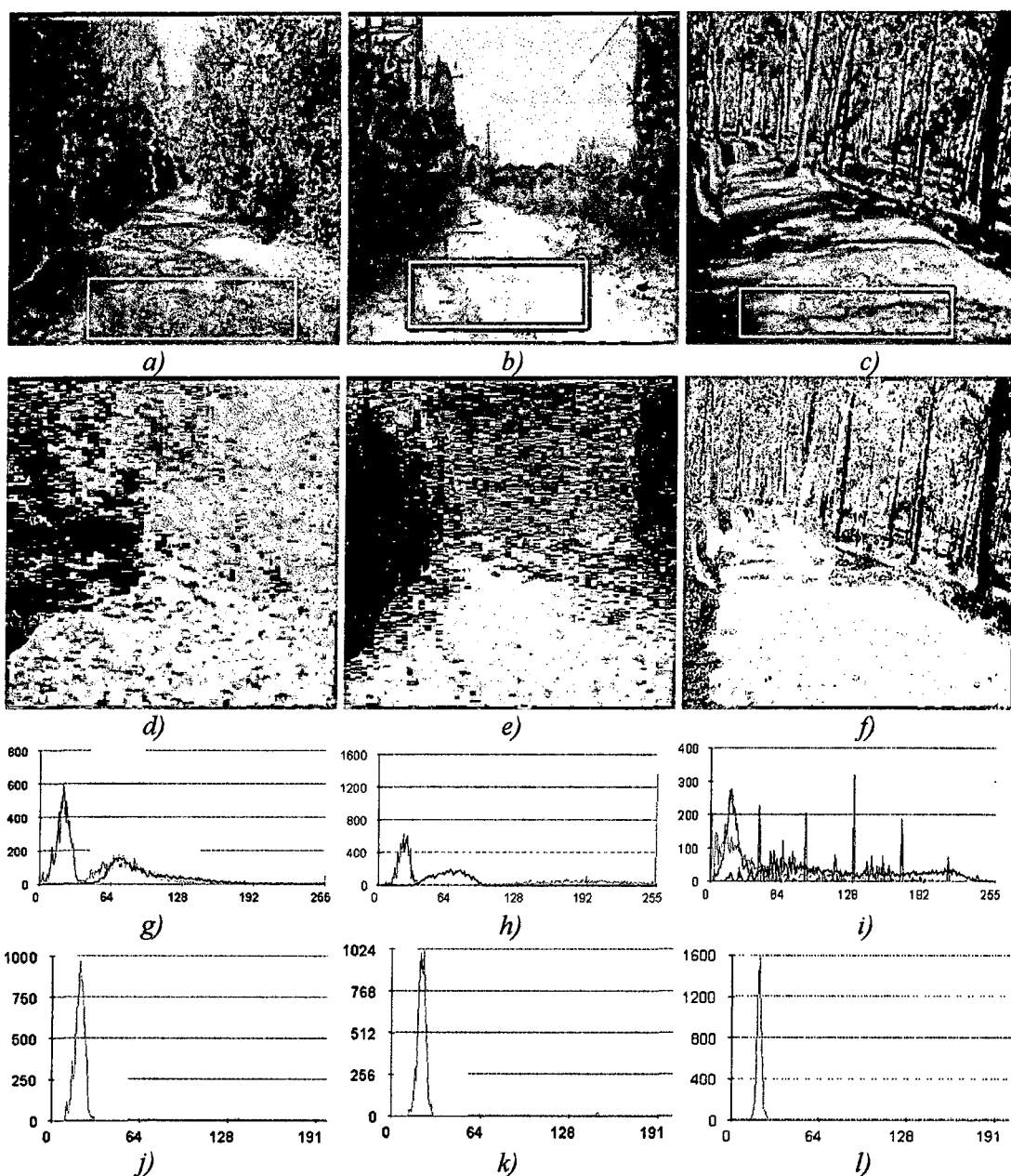


Figura 4.13. Comportament dels valors HSI i H/I en situacions on la il·luminació de l'escena és no uniforme. a,b,c) Imatges RGB. d,e,f) Imatges H/I. g, h, i) Histogrames HSI de les zones emmarcades (en vermell H, en verd S i en blau I). j, k, l) Histogrames H/I de les zones emmarcades.

A la figura 4.13 s'observa clarament la gran dependència de les característiques R, G, B, S i I en les situacions on la il·luminació de l'escena és no uniforme -apareixen ombres i reflexes especulars (zones en les que incideixen un nivell de llum molt superior a la resta de l'entorn) a l'interior del camí-. Malgrat la gran variació del valor RGB en aquestes zones, es pot observar que el valor del hue es manté força constant (en contra del que passa amb la saturació). Únicament, a la figura 4.13.i es pot observa com el color (hue) captat per la càmera és inestable en condicions extremes d'il·luminació. Aquest inconvenient queda en part resolt a l'espai H/I (figura 4.13.l). Aquest fet fa que un mateix objecte o estructura de color constant (en aquest cas el

camí) presenti un valor quasi constant en la imatge H/I, amb una certa independència de les condicions d'il·luminació.

En la figura 4.14, es pot observar com el hue permet diferenciar entre el color de les dues matèries presents en les zones emmarcades (histograma bimodal), i com la distribució dels valors de la saturació no presenta tant clarament aquest aspecte discriminador.

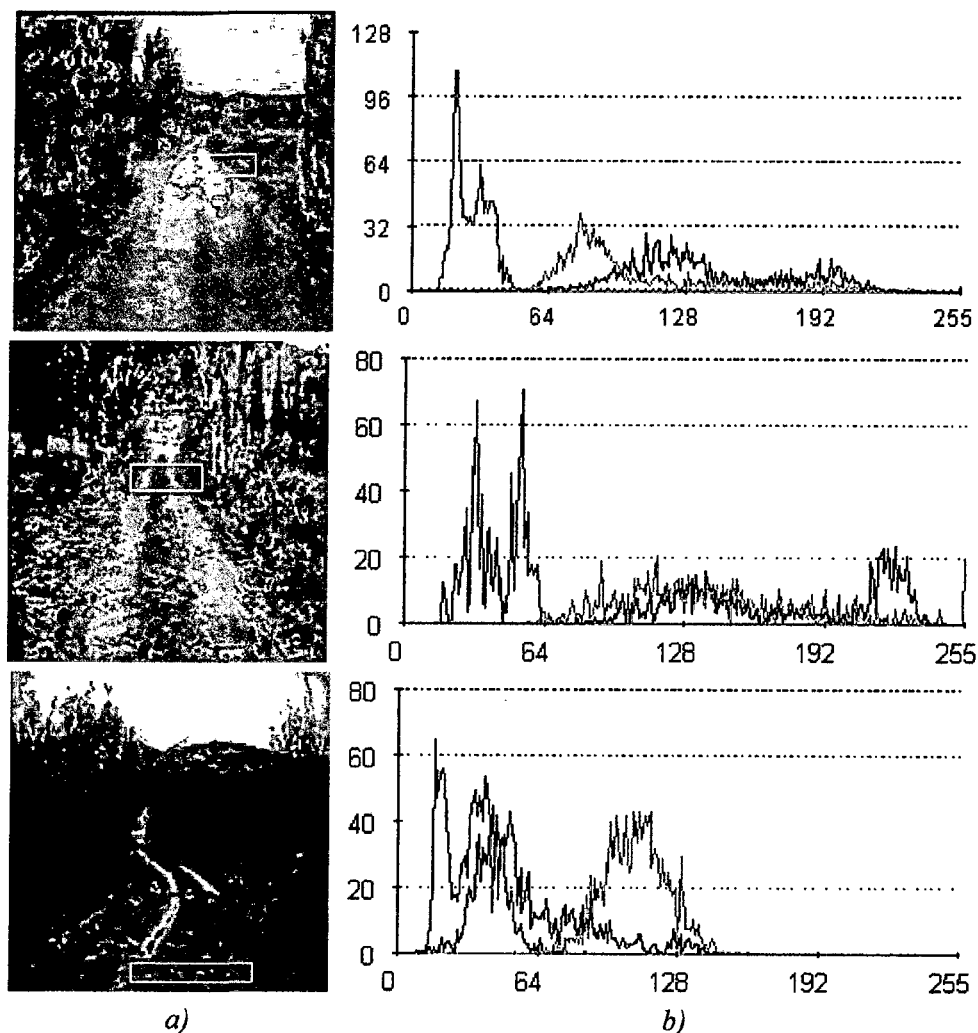


Figura 4.14. Anàlisi de la variació dels valors HSI entre el color de diferents matèries.
 a) Imatges RGB. b) Histograma HSI de les corresponents zones emmarcades
 -en vermell el Hue, en verd la Saturació i en blau la Intensitat-.

Finalment s'analitzarà quin és l'efecte del soroll present a la imatge RGB sobre els valors H/I obtingut en la transformació de l'espai de color. La figura 4.5.c mostra una imatge RGB de la carta de 60 colors que ha estat adquirida utilitzant una vídeo càmera enregistradora sobre una cinta en format VHS, i digitalitzada (la imatge enregistrada) posteriorment amb una placa d'adquisició. La degradació de la informació color (que es pot associar a la presència de soroll a la imatge RGB) és evident si es comparen les

imatges obtingudes amb el "*scanner*" (figura 4.5.a) i amb la vídeo càmera (figura 4.5.c).

La figura 4.5.d mostra la transformació a l'espai H/I de la figura 4.5.c. Com es pot veure, la major variació dels valors RGB en aquest cas es reflecteix en una major variació dels valors H/I, com queda en evidència si es comparen les imatges 4.5.b i d. A la imatge 4.5.d, les variacions dels valors H/I no són degudes a la inestabilitat de la transformació RGB a H/I sinó a la gran variació (soroll) dels valors RGB, i per tant no són imputables a l'espai H/I sinó a la qualitat del sensor utilitzat per captar la imatge RGB.

En les taules 4.2 i 4.3 es pot observar la influència del soroll present en la imatge RGB, en la imatge H/I. En aquest cas es considera una imatge amb poc soroll (-figura 4.5.a- obtinguda amb el *scanner*) i una amb un gran soroll (-figura 4.5.c- imatge enregistrada amb llum natural en vídeo i posteriorment digitalitzada). A la figura 4.15 es presenta la variació que presenta el valor H/I per les 6 mostres de color de la fila superior de la carta a la imatge 4.5.d, i en la taula 4.2 es presenten els valors estadístics de la mitjana, la moda, la diferència entre la moda i el valor mínim, i la diferència entre la moda i el valor màxim, pels 6 colors de la primera fila de la carta. En la taula 4.3 es mostren la moda, la diferència entre la moda i el valor mínim i màxim d'altres 15 colors de la carta.

Com es pot apreciar a les figures 4.5.d, 4.15 i les taules 4.2 i 4.3, malgrat la degradació de la imatge RGB, els valors a la imatge H/I resulten força estables i permeten diferenciar entre els colors bàsics, i fins i tot, colors propers com són el vermell i el taronja presenten modes diferents (els valors de les modes són 0 i 3 respectivament, treballant amb una resolució de l'espai H/I de 128 hues). En mitjana, el valor H/I que descriu el color d'un objecte -en aquest cas una mostra de la carta captada amb la càmera- presenta una variació d'aproximadament ± 4.5 respecte la moda. Aquesta variació del valor H/I (que és deguda bàsicament al soroll inherent al procés de captació, enregistrament, reproducció i digitalització de la imatge) dificulta la segmentació en regions diferents, d'objectes veïns en la imatge que presenten un valor H/I molt semblant. Això és degut a que no és possible diferenciar quan el valor H/I varia per un canvi de color o pel soroll present a la imatge. La variació del paràmetre H/I deguts al soroll present a la imatge ha de ser tinguda en compte a l'hora de segmentar la imatge.

A la vista del comportament dels paràmetres hue i saturació en les imatges color captades en camins feblement estructurats, podem concloure que el hue és el que resulta més discriminador i presenta un comportament més estable davant de la il·luminació no homogènia de l'escena. Així mateix, els fets que la saturació depengui de la il·luminació, i que en molts casos no aporti informació per la discriminació d'elements de color diferent, fan que la saturació sigui una informació que més que millorar, dificulta i fa augmentar la complexitat del procés de segmentació del tipus d'imatges considerades.

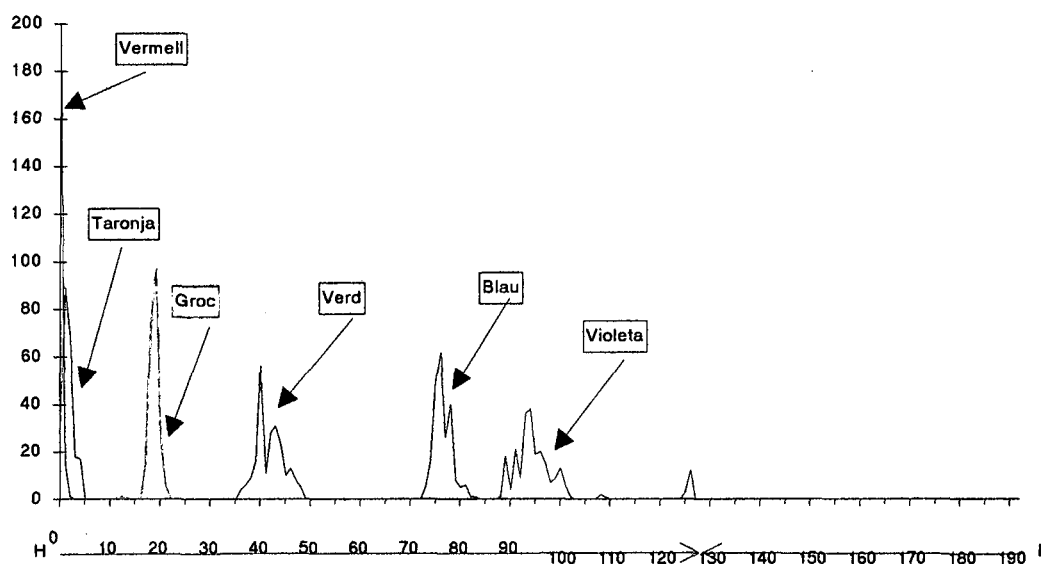


Figura 4.15. Histogrames del valor H/I pels 6 colors de la primera fila de la figura 4.5.d.

Color	Scanner				Càmera			
	Mitiana	Moda	Min.	Max.	Mitiana	Moda	Min.	Max.
Groc (1.1)	18	18	0	0	18.7	19	-2	+3
Taronja (1.2)	3	3	0	0	1.57	1	-1	+3
Vermell (1.3)	0	0	0	0	0.1	0	-3	+2
Blau (1.4)	75.03	75	0	+1	76.38	76	-3	+5
Violeta (1.5)	92.19	92	0	+1	94.34	94	-4	+8
Verd (1.6)	41.0	41	0	0	41.85	40	-4	+6
Mitiana			0	+0.33			-2.83	+4.5

Taula 4.2. Variació del valor H/I pels 6 colors de les figures 4.5.b i d.

Color (fila,col.)	Scanner			Càmera		
	Moda	Var.Mín.	Var.Màx.	Moda	Var.Mín.	Var.Màx.
Marró Pard (2,1)	1	0	+1	0	-8	+9
Marró Clar (2,5)	2	0	0	0	-4	+6
Verd Fosc (2,6)	50	0	+1	49	-3	+6
Bordeus (3,5)	122	0	0	119	-5	+5
Lila (3,6)	113	0	0	113	-5	+6
Verd Oliva (4,2)	35	0	0	40	-5	+6
GrocLlimona (4,3)	18	0	0	19	-3	+3
Blau Reial (5,6)	87	0	0	89	-7	+5
Blau Cel (6,1)	72	0	0	69	-4	+4
Castany (6,3)	0	0	0	5	-7	+6
Terracota (6,5)	2	0	0	2	-3	+3
Violeta (6,6)	100	0	0	101	-6	+4
Oxid (7,2)	1	0	0	0	-3	+5
Esmeralda (7,6)	50	0	+1	47	-4	+4
Albaricoc (8,2)	5	0	0	5	-1	+2
Verd Flor (8,6)	29	0	+1	31	-7	+3
Mitjana		0	+0.25		-4.69	+4.53

Taula 4.3. Variació del valor H/I per 16 colors de les figures 4.5.b i d.

4.2. Segmentació d'imatges Captades en Entorns Naturals

"La segmentació és el procés que divideix una imatge en les seves parts constituents o objectes. La segmentació és un dels passos més importants en l'anàlisi automàtic d'imatges ja que aquest és el pas on els objectes o altres entitats d'interès són extretes de la imatge per al seu processament posterior, com ara la descripció o el reconeixement" [Gonzalez 87].

Les regions en la imatge corresponen a les superfícies dels objectes en l'escena. Junt amb els contorns, les regions són els elements bàsics a partir dels quals es descriuen les imatges.

Les regions resultants d'una bona segmentació han de ser uniformes i homogènies respecte alguna propietat com ara el nivell de gris, el color o la textura. L'interior dels segments han de ser simples, sense molts petits forats. Les regions adjacents han de tenir una diferència significativa respecte a la propietat que determina les regions. El límit dels segments han de ser simples, continus i espacialment precisos.

El procés de la segmentació es pot definir de la següent forma:

Sigui R la regió formada per tota la imatge. Es pot veure la segmentació com el procés que parteix R en n subregions R_1, R_2, \dots, R_n tal que:

$$(a) \bigcup_{i=1}^n R_i = R.$$

(b) R_i és una regió connexa, $i=1, 2, 3, \dots, n$.

(c) $R_i \cap R_j = \emptyset$ per tot i i j , $i \neq j$.

(d) $P(R_i) = \text{CERT}$ per tot $i=1, 2, \dots, n$.

(e) $P(R_i \cup R_j) = \text{FALS}$ per totes les regions R_i, R_j veïnes.

on $P(R_i)$ és un predicat lògic definit sobre els píxels que formen R_i que indica si es compleix una certa propietat i \emptyset és el conjunt buit.

La condició (a) indica que la segmentació ha de ser completa, és a dir, que tots els píxels de la imatge formen part d'una regió. La segona condició requereix que tots els punts d'un segment han d'estar connectats. La condició (c) indica que les regions han de ser disjunctes. La condició (d) fa referència a les propietats que han de complir els píxels que formen una regió. L'última condició indica que les regions veïnes R_i i R_j , són diferents en el sentit que indica el predicat P , és a dir, tenen diferents propietats.

4.2.1. Tècniques de Segmentació d'Imatges en Regions

Com ja s'ha presentat a la secció 4.1, les tècniques de segmentació d'imatges es classifiquen en dos grans grups: *detecció de discontinuïtats* i *detecció de semblança*. Donades les característiques de les imatges captades en entorns naturals i el problema que s'ha de resoldre, les tècniques basades en la detecció de similitud són més adequades per obtenir una segmentació de la imatge profitosa. A continuació presentarem i compararem alguns dels mètodes de segmentació basats en la similitud, com són: la lliandarització o agrupament, el creixement de regions i la divisió/fusió.

1. Lliandarització o Agrupament. Aquesta tècnica utilitza el procés d'agrupament de valors pròxims en l'espai de representació de la característica per segmentar la imatge. Aquest procés es divideix en els següents passos: transformació de la informació de la imatge a l'espai de representació, obtenció d'una partició en classes de l'espai de representació, assignació a cada píxel de l'etiqueta de la classe a la que pertany en l'espai de representació i finalment, definició de les regions a la imatge com els conjunts de píxels connexos amb la mateixa etiqueta. El cas més

simple de llindarització és la binarització, on l'espai de representació es divideix en només dues classes.

La bondat d'aquest tipus de segmentació depèn directament de com els objectes d'interès a la imatge es separen o no en diferents classes a l'espai de mesura. Normalment aquest procediment dona bons resultats en situacions on hi ha pocs objectes, que presenten a la imatge nivells de gris, color o textura ben diferents i aquests apareixen sobre un fons gairebé uniforme o simple. Aquestes situacions es donen poc sovint, ja que les il·luminacions no uniformes fan que el fons, encara que sigui homogeni, doni valors diferents, i per tant diferents zones del fons seran assignats a diferents classes, i que objectes diferents amb diferent reflectància presentin valors semblants de nivell de gris a la imatge, éssent per tant assignats a la mateixa classe.

La forma més habitual de mostrar l'espai de representació és l'histograma. La segmentació es realitza a partir de la determinació dels valors (els llindars) que permeten separar les diferents agrupacions de valor que és donen a l'espai de representació per definir les classes. Un pixel, amb valor de la seva característica assignat a la classe i , és etiquetat amb l'índex i i pertany a la regió definida pel conjunt connex de pixels amb índex i . El principal problema d'aquest mètode és la determinació dels llindars que permeten que els diferents objectes de la imatge apareixin en diferents classes a l'espai de representació.

Diferents autors han proposat variacions sobre aquest esquema bàsic, on cada una de elles obté millors resultats per a un tipus d'imatges de característiques determinades. Una característica comú a tots aquests mètodes, és el gran cost computacional necessari per a la seva realització, ja que el pas de la imatge a l'espai de representació i , especialment, la definició de les classes a l'espai de representació són processos molt lents. El cost computacional d'aquest algorismes és $O(n^2)$, on n és el nombre de pixel de la imatge -cada pixel s'ha de processar una primera vegada per obtenir l'espai de representació, cal calcular la distància entre totes les parelles de pixels per a definir les classes, indicar per a cada pixel de la imatge a quina classe pertany, i finalment, cal processar tots els pixels, al pla imatge, per definir les regions-.

2. *Creixement de Regions*. El creixement de regions és un procediment que uneix pixels i subregions en regions més grans. El mètode més simple és l'agregació de pixels, on es comença amb un conjunt de pixels-llavor cada un com element inicial de les regions i a partir d'ells es van afegint a cada llavor els pixels veïns que tenen propietats similars. La segmentació obtinguda depèn de dos paràmetres: la selecció de les llavors inicials i la selecció de la propietat que permet l'agregació de pixels

durant el procés de creixement. La determinació dels valors idonis per aquests paràmetres es basa en les característiques de les imatges que s'analitzen. La propietat més simple és la diferència de nivell de gris entre pixels. Aquest mètode, al que anomenarem bàsic, resulta molt atractiu, quan és aplicable, per la seva simplicitat.

En el cas del creixement de regions híbrid, per a cada pixel es calcula un vector de característiques a partir del veïnatge $k \times k$ del pixel. En aquest cas, els pixels que són similars ho són en el sentit que els pixels del seu veïnatge són similars. Aquesta tècnica presenta un major cost computacional, però resulta més eficaç en imatges sorolloses.

En el creixement de regions per mitjana, a diferència del mètode bàsic, no es comparen la similitud de les parelles de pixels veïns. En aquest cas, la imatge s'explora en un ordre determinat, el més habitual és l'esquerra-dreta superior-inferior. El valor d'un pixel es compara amb la mitjana dels valors dels pixels que formen una regió veïna ja existent però no necessàriament completa. Si el seu valor i la mitjana de la regió són prou propers, llavors s'afegeix el pixel a la regió i es modifica el valor de la mitjana de la regió tenint en compte aquest nou component. Si hi ha més d'una regió veïna que té una mitjana similar al valor del pixel, aquest s'afegirà a la regió de mitjana més propera. Així mateix, si la mitjana de dues regions que competeixen per absorbir un pixel són properes, llavors es fusionen les dues regions i s'afegeix el pixel a la nova regió generada per la fusió. Si cap regió veïna té una mitjana suficientment propera, llavors es crea una nova regió amb el pixel considerat com primer membre. Un dels problemes que presenta aquesta tècnica és la dependència de l'ordre de recorregut de la imatge, ja que el recorregut esquerra-dreta, superior-inferior pot no donar el mateix resultat que el recorregut dreta-esquerra, inferior-superior. Normalment, però, les diferències que es presenten són petites.

El mètode anomenat *agrupament espacial*, combina l'obtenció d'informació a l'espai de representació i el creixement de regions. En aquest cas, la segmentació es realitza localitzant els valors de pic a l'histograma de la imatge, prenent aquests valors com llavors per realitzar el creixement de regions. El criteri de fusió és que l'entorn d'un pixel ja assignat a una regió, s'afegeix a la regió si un dels pixels d'aquest entorn té un valor proper al pixel amb el que estem creixent, i la seva freqüència d'aparició a la imatge és inferior a la freqüència d'aparició del valor del pixel que està creixent.

El cost computacional d'aquesta tècnica de segmentació és més reduït que en el cas de l'agrupament. El cost de les més simples, creixement de regions bàsic i per

mitjana, és de $O(n)$ -cal processar una primera vegada els n píxels per generar les regions i una segona per generar la imatge segmentada. En el cas de la tècnica d'agrupació espacial, el cost és de $O(n)$ -ja que a més dels processos descrits anteriorment, cal obtenir l'espai de representació de la imatge- i el cost en el creixement de regions híbrid és de $O(k^2*n)$ -ja que cada píxel de la imatge forma part de k^2 finestres diferents, de mida $k \times k$ píxels, de la imatge.

3. *Divisió i Fusió*. Aquesta tècnica de segmentació s'inicia amb una partició de la imatge en un conjunt de regions quadrades definides per l'usuari segons algun criteri. Posteriorment, aquestes regions són subdividides o fusionades amb l'objectiu que satisfacin les cinc condicions que hem definit a l'inici de la secció 4.2. Aquest procediment es repeteix fins satisfer unes condicions d'aturada.

Una formalització d'aquest mètode és:

Donada una regió quadrada R i una propietat P aplicable sobre les regions de la imatge, la regió R és divideix en quadrants R_i , $i=1, 2, 3$ i 4 , si $P(R)$ és fals (la regió no és homogènia). Si per alguna regió R_i , $P(R_i)$ és fals, es torna a subdividir en quadrants. Si només fem servir el procediment de divisió, al final ens trobarem amb regions adjacent amb propietats idèntiques o similars. Per tant hem de fusionar les regions adjacents que al combinar els seus píxels donin una regió que encara sigui homogènia, és a dir, donades dues regions adjacents R_i i R_j es fusionaran si i només si $P(R_i \cup R_j)$ és cert.

El cost computacional del mètode de segmentació de divisió i fusió, és en el pitjor cas de $O(n^k)$, on n és el nombre de píxels de la imatge i k és el nombre de vegades que és subdivideixen, com a màxim, les regions inicials.

Aquesta tècnica presenta diversos problemes, com ara la determinació de la mida inicial de les regions, el criteri d'aturada per la divisió de regions, i que la segmentació obtinguda, donat que sempre es manipulen regions quadrades, té un aspecte quadriculat artificial.

Com es pot veure, hi ha diverses alternatives en la segmentació d'imatges. Per a cada aplicació, una d'aquestes tècniques es mostrarà com més interessant fent que el procediment de segmentació acabi éssent dissenyat a mida. A continuació, anem a centrar-nos en l'estudi i implementació de quina és la tècnica més adequada pel tipus de tasca que hem de realitzar, la segmentació d'imatges captades en entorns naturals orientada a la navegació autònoma en camins feblement estructurats.

4.2.2. Segmentació d'Imatges Color

Des de l'inici de la visió per ordinador, s'ha utilitzat el color amb l'esperança de que les regions o fronteres obtingudes en una imatge segmentada siguin més significatives. En efecte el color és un atribut més estable que la lluminància i d'altre banda constitueix una informació més rica. Dos aspectes a considerar en la segmentació d'imatges color, són l'escala de cromacitat uniforme⁵ i la percepció humana. Un espai de color d'escala de cromacitat uniforme presenta la característica que la diferència de color en l'espai de representació (tal com ho percebim els humans) pot ser expressada directament amb la distància euclidiana. Els espais Lab, Luv i HSI són uniformes. En canvi, els espais RGB, YIQ, XYZ i I1I2I3 no són uniformes. Els espais uniformes faciliten la segmentació de la imatge color en regions similars a les percebudes per l'ull humà. L'espai HSI s'ha definit per representar el comportament psicològic del sistema visual humà, com ja s'ha indicat anteriorment.

La majoria dels algorismes de segmentació d'imatges color són una generalització dels mètodes monocromàtics amb la realització dels càlculs a l'espai tridimensional. El problema que apareix és la búsqueda d'una mètrica adequada que indiqui de forma constant, a tot l'espai de color, la diferència de colors. Aquesta mètrica està molt lligada a l'elecció de l'espai de representació del color. Alguns dels treballs de segmentació d'imatges color més interessants són presentats a continuació.

En [Otha 80] fan servir l'espai de representació del color (I1,I2,I3) que proposen (veure secció 4.1.1) per a la segmentació d'imatges utilitzant un algorisme de llindarització recursiu proposat a [Ohlander 75]. Els autors realitzen un estudi comparatiu de la segmentació obtinguda de vuit imatges -que són les utilitzades per definir l'espai I1I2I3 mitjançant la transformació de Karhunen-Loeve-, utilitzant diferents espai de representació del color (I1I2I3, RGB, XYZ, YIQ, Lab, U*V*W*, Irg i HSI). Les conclusions d'aquest estudi és que les diferències entre les segmentacions obtingudes amb els diferents espais de color no són significatives, i que per a la segmentació d'imatges captades en entorns naturals hi ha prou amb la utilització de dos components de l'espai color (un relacionat amb la intensitat lluminosa i l'altre relacionat amb la cromà). Cal remarcar que les vuit imatges amb les que es realitza l'estudi presenten una il·luminació homogènia.

⁵ en anglès Uniforme Chromaticity Scale -UCS-.

En el treball presentat en [Tominoga 87], es fa una anàlisi coloromètrica d'objectes naturals (plantes i flors) a partir d'una imatge multicromàtica (analitza quatre components: R, G, B i I). L'autor conclueix que el hue és la característica més discriminadora, i fa ús de la llindarització de l'histograma del hue per segmentar la imatge.

En [Celenk 90] es presenta un algorisme de segmentació per escenes naturals basat en l'espai de color Lab i la tècnica d'agrupament (*clustering*). La detecció de les classes en compte de realitzar-se sobre l'histograma tridimensional, es realitza sobre els tres histogrames (L, a i b) unidimensionals, amb la conseqüent reducció de temps de procés. Les classes definides en els tres histogrames són finalment combinades per segmentar la imatge.

En [Tomiyasu 91], els autors proposen un mètode orientat a la segmentació d'imatges color basat en la determinació de quines són les regions de la imatge que presenten una saturació significativa i quines no. El primer pas consisteix en la transformació de la imatge RGB a HSI, posteriorment es decideix si un pixel (o regió) presenta una saturació significativa, a partir del càlcul d'un llindar que depèn dels valors H i I - $Ll = F(H, I)$. Si la saturació del pixel (o regió) és superior al llindar ($S > Ll$), es considera que la saturació del pixel (o regió) és significativa, en cas contrari es considera que la saturació no és significativa. Les regions de saturació significativa són posteriorment segmentades a partir de l'anàlisi del hue, i les regions de saturació no significativa, són segmentades mitjançant l'anàlisi de la intensitat. Els autors no avaluen el funcionament del sistema amb imatges captades en escenes on la il·luminació varia de forma important.

Tseng i Chang proposen una tècnica de segmentació basada en els paràmetres HSI, però per aquelles zones de la imatge en que el valor H i/o el valor S no són estables, no seran utilitzats per segmentar aquella zona [Tseng 92]. Una primera segmentació es realitza a partir de l'agrupament a l'histograma del hue o del histograma de la intensitat -si el hue i/o la saturació són inestables-. Per a les regions de hue similar s'analitza si l'histograma de la saturació presenta una variació significativa, en aquest cas, la regió es divideix en funció de les classes definides a l'histograma de S. Finalment, realitzen una eliminació de les regions no significatives de la imatge, aplicant un algorisme de creixement de regions en aquelles zones de mida 8x8 pixels en que apareixen pixels simultàniament amb hue i saturació significatius (estables) i no significatius.

La tècnica de segmentació d'imatges color proposada en [Pujas 95] es basa en l'espai HSI i el creixement de regions. A l'igual que el treball anterior, una vegada transformada la imatge RGB a HSI, s'avalua el grau de significació del hue a partir dels valors de S i I (la significació és un valor continu de 0 a 1) que els autors proposen calcular de forma analítica o mitjançant una xarxa neuronal. Per a que un pixel p sigui agregat a una regió R s'han de satisfer dues condicions: que el pixel p i la regió R , i el pixel p i el pixel q (q és un pixel veí de p que ja està inclòs en R), tinguin un grau de significació del hue (o d'intensitat) i un valor de hue (o d'intensitat) semblants. Com es pot veure, la saturació únicament s'utilitza per avaluar la significació del hue, ja que les components de color utilitzades per definir les regions són el hue i la intensitat. El temps de càlcul per la segmentació d'una imatge de 256x256 pixels amb un processador Sparc-20 és de 35 segons.

4.2.3. Segmentació a Baix Nivell Proposada Basada en l'anàlisi del Color

Les característiques principals dels camins feblement estructurats, com ja s'ha introduït a l'inici de la secció 3.1, són que els seus marges estan poc definits, la seva superfície esta formada per diferents matèries i la il·luminació és molt variable i no homogènia. La diferència més important entre les característiques d'aquest tipus de vies i les carreteres o autopistes, és que la seva superfície pot estar formada per diferents matèries. Aquest fet ha de ser considerat a l'hora de segmentar la imatge en regions associades al camí, als marges i, si n'hi ha, als obstacles.

La segmentació mitjançant l'extracció de contorns d'imatges captades en entorns naturals, resulta poc atractiva donada la gran quantitat de contorns que poden aparèixer en aquest tipus d'imatges, degut a la il·luminació no uniforme i la complexitat de l'escena. La segmentació mitjançant l'extracció de contorns ha estat utilitzada amb èxit en sistemes de navegació per carreteres i autopistes, ja que en aquests entorns la gran diferència existent entre la calçada i els marges, i la senyalització horitzontal fan aparèixer uns contorns rellevants i molts útils per a la determinació de l'espai travessable. En canvi, s'ha de realitzar un esforç considerable per a no confondre els contorns que apareixen a la imatge generats per les ombres que es projecten sobre la calçada, amb els contorns generats pels límits de la calçada i la senyalització horitzontal.

Així doncs, les tècniques de segmentació d'imatges basades en la similituds (agrupament, creixement de regions i, divisió i fusió) faciliten l'obtenció d'una segmentació coherent amb l'escena, especialment quan aqueta és complexa. Això si, a condició d'analitzar una característica de la imatge independent de les condicions d'il·luminació, com ara el color, i d'orientar el procés de segmentació per obtenir únicament les regions de major dimensions de la imatge.

En la metodologia proposada, en un primer pas, es vol segmentar la imatge color en regions que presenten un color homogeni (segmentació a baix nivell). En un processat posterior, amb la classificació de regions, es completarà la segmentació de la imatge, indicant per a cada regió a quina categoria -entre camí, marge o obstacle potencial- s'associa en funció del seu color i la seva posició a l'entorn (es defineix el camí com un conjunt de regions que compleixen una certa propietat).

La segmentació utilitzant la tècnica d'agrupament requereix de tres processos força lents com són la transformació de la imatge a l'espai de representació, la partició de l'espai de representació mitjançant un procediment recursiu, i la definició de les regions a la imatge. La realització d'aquests processos en un interval curt de temps només és possible amb imatges de baixa resolució. A més, si es vol obtenir una segmentació coherent amb l'escena cal definir correctament el nombre de classes a considerar a l'espai de representació. Això només és possible amb un coneixement a priori de les característiques -nivells de gris, color o textura dels objectes de l'escena- de les imatges a segmentar, restricció que no es vol imposar en el nostre cas.

Així mateix, la segmentació mitjançant la tècnica de divisió i fusió resulta poc adequada per utilitzar-la en imatges amb un gran nombre d'objectes, ja que això implica un procés de divisió llarg. Així mateix el fet que les fronteres de les regions generades no corresponguin a les dels objectes, ni en el seu aspecte quadrícula ni en la seva posició, fan que es generin errors importants en el procés de posicionament al món dels elements presents a la imatge segmentada.

La tècnica de creixement de regions, permet generar la partició de la imatge en un únic pas, en el que es recorren, seguint un ordre, tots els pixels de la imatge. Aquest procés requereix un temps més curt que els descrits anteriorment. A més, la utilització d'aquesta tècnica ens permet integrar en el procés de segmentació el coneixement a priori que es té de les imatges amb l'objectiu de simplificar i millorar el procés de segmentació. Amb aquest objectiu es pot indicar al procés de creixement de regions en quina posició de la imatge apareix la zona del camí més propera al vehicle, quin són

els colors de les matèries que formen la superfície del camí en aquesta zona, i quina és la mida mínima d'una regió per que hagi de ser considerada.

Tot i que les tres tècniques de segmentació per similitud poden generar una segmentació correcta del tipus d'imatges que es consideren, la de creixement de regions permet indicar-li quines han de ser les llavors a utilitzar en el procés de creixement per obtenir les regions associades al camí, obtenir una segmentació acurada, en el grau que ho exigeixi la resolució de la descripció de l'entorn, i generar la imatge segmentada en un interval més curt de temps.

El procés de segmentació a baix nivell proposat assumeix que:

1. El nombre de matèries que formen el camí en una zona és petit
2. El vehicle sempre es troba a l'interior del camí, i per tant la zona inferior central de la imatge correspon a la projecció de l'àrea del camí més propera al vehicle.

La segmentació de les imatges, utilitzant la característica H/I presentada a l'apartat 4.1.2 per representar la informació color, es realitza en dos passos:

1. Determinació dels colors de les matèries que formen el camí en la zona analitzada, als que anomenarem **colors-camí**.
2. Creixement de regions per mitjana (veure secció 4.2.1). Definició de les regions de la imatge, càlcul de certes propietats (color, mida, posició a la imatge) i relacions de veïnatge entre les regions, que seran utilitzades en la posterior classificació de les regions.

Com s'utilitza el color per distingir els diferents components de la imatge, resulta interessant identificar els colors de les regions que formen el camí. Per determinar els colors predominants del camí en l'àrea que s'analitza, s'utilitza el coneixement que es té de la posició relativa entre el mòbil i el camí.

Assumint que el mòbil es troba a l'interior del camí, i conegudes les característiques de la càmera i la disposició (desplaçament i orientació relatiu) de la càmera respecte el camí; la regió inferior central de la imatge (figura 4.16.a) correspon majoritàriament a la projecció de la part del camí més propera al vehicle. Per tant els colors més freqüents d'aquesta àrea, que anomenarem els **colors-camí**, es considera que corresponen als colors de les matèries que formen la superfície del camí en la zona observada en la imatge. Com el nostre model de camí estableix que la seva superfície

és heterogènia, aquells pixels que tenen un color semblant (i per tant un valor H/I semblant) a un dels colors-camí seran considerats com possible part del camí.

El nombre de colors-camí a considerar per assolir una segmentació de la forma més eficient, ha de coincidir amb el nombre de matèries que formen el camí en la imatge analitzada. Un nombre inferior suposarà que regions associades al camí siguin inicialment considerades, al realitzar la classificació, com obstacles potencials (com es podrà veure a la secció 4.3), i un nombre superior fa més complexa i lent el procés de segmentació.

El procediment que es segueix per determinar els colors-camí és :

1. Calcular l'histograma dels valors H/I en la zona inferior central (figura 4.16.a).
2. Buscar a l'histograma els N valors (colors) més freqüents tenint en compte que la diferència de color entre aquests valors ha de ser significativa (els valors representen colors diferents).
3. Verificar que cap dels N colors seleccionats (amb valor color-camí 1, ... , color-camí N) no correspon al color d'un obstacle (la determinació de quines de les regions de la imatge estan associades a un obstacle es realitza mitjançant l'anàlisi de les imatges anteriors de la seqüència, veure secció 5.1.2.)

En l'exemple de la figura 4.16 apareixen dos colors-camí que corresponen al marró clar de la sorra del camí i al verd de la gespa situada en la part central del camí. Com es pot observar comparant les figures 4.16.b i d, la gran variació dels valors RGB dificulta el càlcul de quin són els colors més freqüents a l'àrea analitzada.

Per segmentar la imatge H/I utilitzant la tècnica de creixement de regions per mitjana, es recorra tota la imatge en sentit esquerra-dreta, inferior-superior, partint d'un pixel de la zona inferior central que té com valor el color-camí 1. Aquest ordre de recorregut es prefereix a altres, ja que la part d'informació més rellevant, el camí, es troba a la part inferior de la imatge. Per a cada pixel (i, j) de la imatge es comprova si el seu color i els colors de les regions veïnes (en l'ordre de recorregut indicat, els pixels veïns esquerra (i-1, j) i inferior (i, j+1) -figura 4.17- ja han estat analitzats i assignats a les regions R_a i R_b respectivament, on R_a i R_b poden ser la mateixa regió) tenen un color similar. El pixel s'afegeix a aquella regió que tingui una mitjana del valor H/I més similar, i a continuació s'actualitza tota la informació associada a la regió que s'ha engrandit. Si les dues regions tenen característiques (colors) similars, es fusionen. En el cas que la diferència de color entre el pixel i les regions sigui més gran

que un cert llindar, es crea una nova regió que té el color d'aquest pixel com valor inicial.

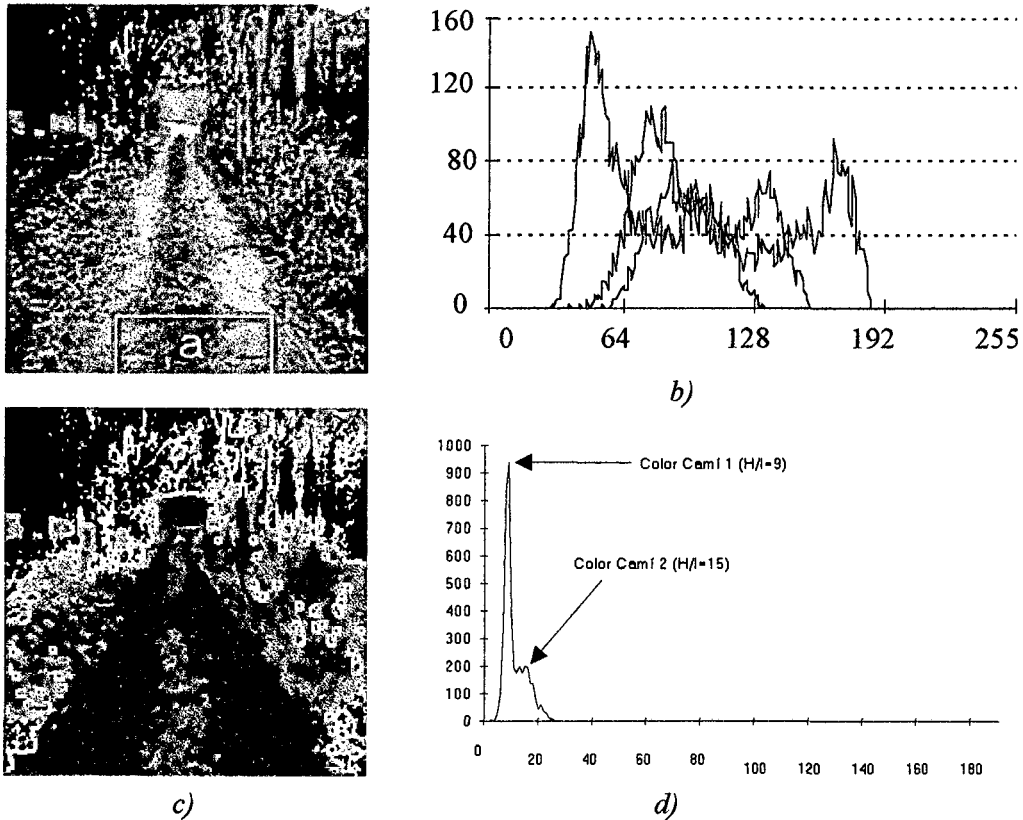


Figura 4.16. Exemple de camí de muntanya amb la determinació dels colors-camí.
 a) Àrea inferior central. b) Histograma de la zona a de la imatge RGB
 c) Imatge H/I. d) Histograma de la zona a de la imatge H/I

Per tal que la segmentació no es vegi afectada pel soroll que pugui haver a la imatge, només es crea una nova regió si es veu que aquesta pot créixer posteriorment. Per verificar-ho, es compara el color del pixel inicial de la regió -el pixel (i, j)- amb el color dels pixels veïns (pixels dret (i+1, j) i superior (i, j-1) -figura 4.17-) que encara no han estat analitzats (i per tant no pertanyen a cap regió). Si almenys un d'aquests dos pixels té un color semblant al valor del pixel (i, j), la regió que té com element inicial el pixel (i,j) podrà créixer, sinó la regió generada amb el pixel (i,j) tindrà una grandària d'un pixel. Com una bona segmentació no ha d'estar formada amb regions d'aquestes característiques, es força l'agregació del pixel que s'analitza a la regió veïna, R_a o R_b , amb color més proper.

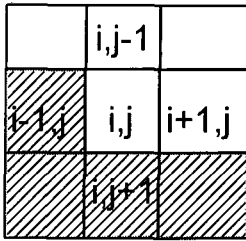


Figura 4.17. Situació del procés de segmentació a la posició (i, j) . L'àrea marcada ja ha estat analitzada i assignada a una regió. Els pixels $(i-1, j)$ i $(i, j+1)$ han estat assignat a les regions R_a i R_b respectivament.

Donades les característiques de l'escena i la tasca que s'ha de realitzar, es desitjable que el resultat de la segmentació sigui un nombre reduït de regions, ja que no té sentit que una estructura/objecte del món que s'analitza, com ara el camí, els marges o els obstacles, es vegi representada a la imatge per un gran nombre de regions. Per tant, com el nombre de parts de l'escena es reduït (no hi ha moltes àrees que estiguin definint la vora o el camí i en una escena es suposa que hi ha pocs obstacles) i per definir aquests elements no és necessari generar un gran nombre de regions en la imatge segmentada, es convenient limitar el nombre màxim de regions que formen la imatge a un valor petit. Per aquesta raó i amb l'objectiu de simplificar el procés de segmentació, es fixa un nombre màxim de regions que existiran durant la segmentació. Aquest nombre ha de tenir en compte quina és la complexitat de la imatge.

Aquest objectiu entra en conflicte amb el fet que el nombre d'objectes a l'entorn és elevat i la seva il·luminació heterogènia. Per obtenir una imatge segmentada, en el tipus d'escenes considerades, amb un nombre de regions màxim prefixat, és necessari forçar si cal, la unió de regions petites a una regió veïna de valor (color) pròxim i mida gran. Aquest procés implica l'eliminació de les regions petites que apareixen a la imatge, fet que no resulta preocupant per que o bé són objectes petits que es troben a prop del vehicle i no són significatius en l'entorn, o bé són objectes que encara que siguin de mida considerable, es troben a gran distància del mòbil i per tant la seva no inclusió a la descripció de l'entorn no resulta transcendent. A mesura que el mòbil s'apropi als objectes grans que es troben lluny, la seva mida a la imatge anirà creixent, generant una regió a la segmentació i d'aquesta manera s'actualitzarà la descripció de l'entorn amb aquest objecte.

El mètode de segmentació proposat ha de resoldre situacions més complexes que els utilitzats en altres sistemes de navegació [Wallance 87] [Thorpe 88] [Turk 88] que es basen en la tècnica d'agrupament per segmentar la carretera en les imatges -veure secció 2.2-, ja que requereix d'un nombre inferior de restriccions -no assumint que el color de totes les matèries que formen el camí i el color del marge són diferents- i un temps de procés, per imatges d'igual resolució, més curt per generar una segmentació coherent amb l'escena.

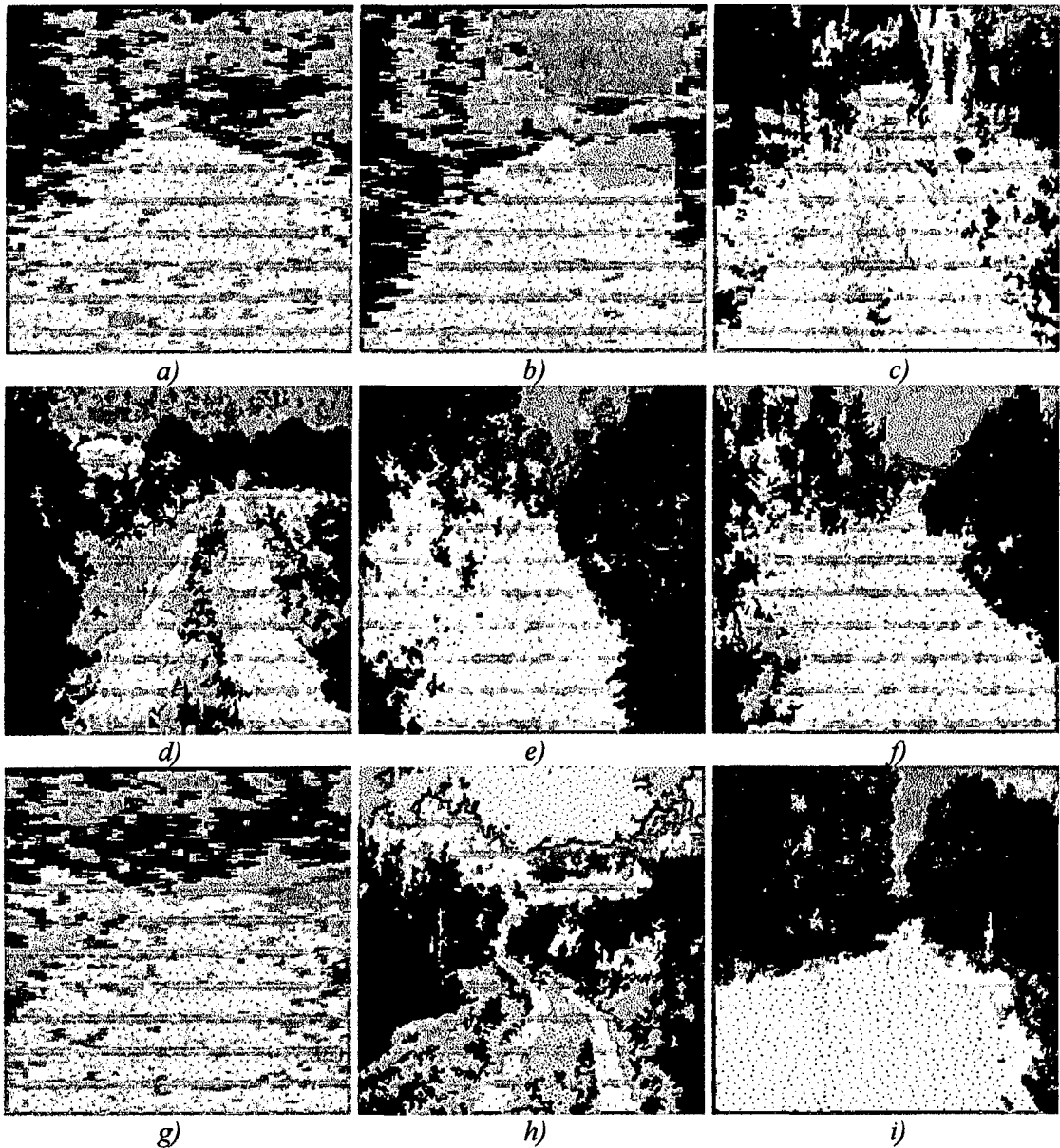


Figura 4.18. Segmentació de les imatges H/I de la figura 4.6 utilitzant la tècnica de creixement de regions presentada.

La figura 4.18 mostra el resultat de la segmentació de les 9 imatges H/I de la figura 4.6. El nombre dels diferents tipus de regions manipulades en la segmentació -imatges de 256x184 pixels⁶-, queden recollits a la taula 4.4. L'elevat nombre de regions que es creen durant la segmentació de la imatge és degut a la complexitat de les escenes i a que la distància permesa entre un pixel i una regió per que aquest sigui agregat es fixa

⁶ En la configuració de la càmera utilitzada, l'àrea de l'entorn on es troba l'àrea del camí que es vol analitzar, es projecta en les 184 files inferiors de la imatge.

a un valor petit⁷ per afavorir que la forma de les regions respecti la forma dels objectes de l'escena. Les regions fusionades ho són per ser veïnes d'una regió de color similar. Les regions absorbides són regions petites, que s'han fusionat a la regió veïna de color més pròxim, encara que no sigui de color similar. A l'annex B es presenten els detalls de la implementació del procés de segmentació de les imatges.

	Nombre regions creades	Nombre regions fusionades	Nombre regions absorbides	Nombre final regions
a.	2180	698	1145	37
b.	746	221	504	21
c.	1070	615	433	22
d.	1460	854	560	46
e.	828	400	416	12
f.	871	511	341	19
g.	2851	1006	1750	95
h.	2198	1304	833	61
i.	1067	522	532	13

Taula 4.4. Nombre de regions manipulades durant la segmentació de les imatges de la figura 4.18.

4.2.4. Estudi Comparatiu de les Tècniques de Segmentació d'Imatges Color

A les seccions 4.1.1 i 4.2.2 s'han presentat alguns dels espais de representació del color i dels mètodes de segmentació d'imatges color de més interès. Així mateix, s'ha proposat un nou espai de representació de la informació color (secció 4.1.2) i una adaptació, tenint en compte les característiques de les imatges analitzades, de l'algorisme de segmentació de creixement de regions (secció 4.2.3). En aquesta secció es discutirà sobre els avantatges i inconvenients del mètode de segmentació proposat respecte els altres mètodes de segmentació d'imatges color presentats.

Els avantatges que presenta el mètode de segmentació utilitzat són:

- La decisió de si el hue és una característica representativa del color dels objectes, i per tant útil per a la segmentació d'imatges, es fa analitzant l'expressió utilitzada per la transformació RGB a Hue, i els valors RGB de la imatge. La decisió no es pren en funció de la saturació i/o la intensitat lluminosa del pixel com es proposa en els altres mètodes, i per tant no és necessari en cap moment calcular el valor de la saturació.

⁷ Per la segmentació de les imatges de la figura 4.18 s'han considerat que si un pixel defereix en més de 4.5 en el valor H/I respecte les regions veïnes, aquest pixel es susceptible de crear una nova regió. En la secció 4.1.3 s'indica quin ha estat el procediment per determinar aquest valor.

- La característica coloromètrica del píxel a la imatge H/I esta representada en un únic valor, i no en dos o tres com és el cas dels altres espais de representació del color. Aquest fet fa que la complexitat i el nombre d'operacions necessaris per segmentar la imatge es redueix considerablement.
- Els mètodes de segmentació presentats a la secció 4.2.2, tenen en compte la intensitat dels píxels, i per tant la presència d'ombres a l'escena pot fer aparèixer regions que en estiguin associades. En tot cas, aquests algorismes s'haurien de modificar per segmentar de forma adequada el tipus d'imatges considerades.
- El nombre de regions que apareixen en una imatge segmentada depèn, principalment, de com de restrictiva és la propietat que compleixen tots els píxels que pertanyen a una regió (com per exemple la màxima diferència de color permesa entre píxels de la mateixa regió). Una propietat restrictiva permet diferenciar en regions diferents objectes que tenen una característica semblant, però també pot produir que un mateix objecte estigui representat en varies regions de la imatge segmentada degut a les variacions de la característica per una il·luminació no homogènia o pel soroll inherent al procés de captació de la imatge. El mètode de segmentació proposat és un compromís entre aquests dos efectes. La propietat utilitzada a l'hora de definir les regions pot ser restrictiva, el que permet segmentar en regions diferents objectes amb una petita variació de la característica representada en la imatge, però al mateix temps, es generen un gran nombre de petites regions. Al llarg de la segmentació de la imatge, quan una regió ha finalitzat de créixer s'analitza la seva mida. El fet que les regions de mida petita siguin absorbides per una regió veïna de mida gran fa que el nombre de regions generades al finalitzar la segmentació sigui baix, i en tot cas, que es pugui limitar a un valor màxim.
- La segmentació de la imatge es realitza analitzant una única vegada tots els píxels de la imatge, i per determinar a quina regió s'afegeix un píxel p únicament s'analitza el veïnatge de p . Això fa que el temps necessari per segmentar la imatge sigui inferior al de les altres tècniques de segmentació ja que moltes d'elles necessiten analitzar globalment la imatges (les que es basen en l'histograma d'una o més característiques de la imatge), i/o han de recórrer varies vegades la imatge per segmentar-la.

Si bé, donades les propietats del mètode de segmentació proposat, aquest permet segmentar adequadament les imatges captades de camins mal o feblement estructurats, també presenta limitacions com ara:

- La resolució coloromètrica de l'espai H/I pot resultar reduïda per certes aplicacions on cal diferenciar objectes de color molt semblants, és a dir, tenen un hue molt pròxim o únicament es diferencien en el seu nivell de saturació. Si bé en el primer cas el problema pot ser resolt augmentant la resolució de l'espai H/I (utilitzar més de 8 bits per representar el valor H/I d'un pixel), la solució a la segona situació passa necessàriament per considerar i analitzar la saturació durant la segmentació, com ho fan algunes de les tècniques presentades a la secció 4.2.2. Aquest fet que es dóna en entorns interiors i/o artificial (construïts per l'acció de l'home), es pot freqüent en els entorns naturals.
- Malgrat que el mètode de segmentació proposat és deriva d'una tècnica genèrica com és la de creixement de regions, ha estat adaptat per segmentar imatges captades en entorns naturals, i per tant no es pot considerar de propòsit general. El fet que no es detectin en molts casos els canvis d'il·luminació de l'escena o es suprimeixin la regions petites de la imatge segmentada, són efectes no desitjat per certes aplicacions.

4.3. Classificació de les Regions de la Imatge

Com s'ha indicat anteriorment, en el model de camí transitable s'ha establert que un camí és una àrea connexa, que pot tenir una superfície heterogènia, una direcció i una amplada localment constants (veure secció 3.2). Conseqüentment, tota agrupació de regions que formin la projecció del camí en la imatge, ha de complir aquestes restriccions. Per tant, en aquest últim pas de la segmentació, es seleccionaran un conjunt de regions amb colors similars als colors-camí que compleixen les restriccions geomètriques de camí.

Per a la realització d'aquest procés és necessari la imatge generada per la segmentació de baix nivell (basada únicament en la informació color), una estimació del desplaçament relatiu del vehicle entre les dues últimes imatges captades, i la descripció de l'entorn (obtinguda a partir de l'anàlisi de les imatges anteriors de la seqüència). La classificació de les regions és el procés que indicarà quines són les regions de la imatge que estan associades al camí, als marges, i, si n'hi ha, als obstacles.

En la descripció de l'entorn s'indica quina és la posició del camí al món. A partir d'aquesta informació, coneixent el desplaçament relatiu del vehicle, des de l'última

vegada que s'ha actualitzat la descripció de l'entorn fins s'ha capturat la imatge que s'analitza actualment, i utilitzant el model de formació de la imatge, es pot determinar en quina zona de la imatge que s'analitza, a la que s'anomena **zona d'interès**, ha d'aparèixer el camí. Les regions de la imatge són classificades en camí, marge o obstacle potencial, en funció del seu color (respecte els colors-camí) i de la seva posició a la imatge (respecte la zona d'interès).

4.3.1. Definició de les Zones d'Interès

A partir de l'anàlisi d'una seqüència d'imatges, s'ha construït una descripció de l'entorn, on s'indica la posició, amplada i direcció del camí que es segueix actualment. D'un altre banda, si es coneix el desplaçament relatiu del vehicle (ΔP) entre la posició en que s'ha captat l'última imatge utilitzada per generar la descripció de l'entorn (P_{t-1}) i la imatge que s'analitza, es pot deduir quina és la posició al món del vehicle ($P_t = P_{t-1} + \Delta P$) en l'instant t en la que s'ha captat la imatge que s'analitza actualment. En la descripció de l'entorn es defineixen dues àrees, que corresponen a l'àrea del camí més propera a la posició P_t (àrea 1 en la figura 4.19.a) i al camí (àrea 2 en la figura 4.19.a).

Conegudes la posició P_t del vehicle (càmera) a l'entorn i els paràmetres intrínsecs i extrínsecs del sistema d'adquisició d'imatges, es pot deduir en quina zona de la imatge es projecta el camí pel que circula el mòbil. Les zones 1 i 2 en la figura 4.19.b, corresponen a la projecció de les àrees del món associades a les àrees 1 i 2 de la descripció. La zona 1 de la imatge és aquella en la que s'analitzen els colors de les matèries que formen el camí per definir els colors-camí (veure secció 4.2.2).

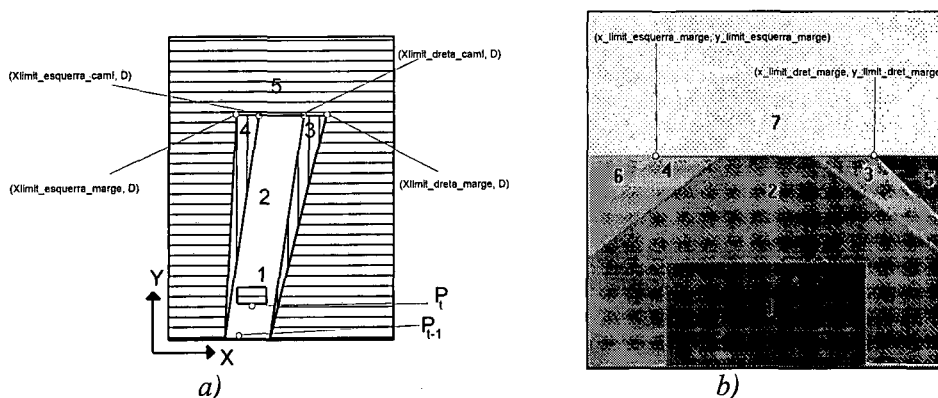


Figura 4.19. Àrees que es defineixen en el procés de classificació de regions.

a) A l'entorn. 1- Àrea inicial camí a la posició P_t , 2- Àrea camí, 3,4- Àrea variació de direcció a la dreta o a l'esquerra, 5- Marge.

b) A la imatge 1-Àrea inicial camí, 2-Àrea camí, 3, 4- Àrea camí amb variació de direcció de $+\alpha$ graus o de $-\alpha$ graus, 5,6- Marges dret i esquerra del camí, 7- Part superior.

Que les zones 1 i 2 de la imatge coincideixen amb la situació real del camí a la imatge depèn de que la descripció de l'entorn sigui coherent amb el món que representa, i que l'estimació del desplaçament relatiu del vehicle (ΔP) sigui correcte. Si la primera condició ha de ser garantida per la pròpia metodologia, la segona depèn de les característiques del vehicle i especialment de l'entorn, ja que les rrelliscades dels pneumàtics o una resposta incorrecte dels actuadors del vehicle fan que el desplaçament i orientació real del vehicle no coincideixi amb els indicats pels sensors.

Tenint en compte la possible existència d'aquests errors, es defineixen 7 zones a la imatge (figura 4.19.b) que corresponen a les posicions esperades a la imatge de la zona central del camí més propera al mòbil (1); del camí (2); del camí si hi ha una diferència de $+\alpha$ graus entre l'orientació real del vehicle i la indicada pels sensors (3); del camí si hi ha una diferència de $-\alpha$ entre l'orientació real del vehicle i la indicada pels sensors (4); i finalment, la resta de l'entorn que es considerat com marge (5, 6 i 7). Les zones 3 i 4 indiquen quina és la incertesa a considerar a l'hora de definir la zona que ocupa el camí a la imatge, utilitzant la descripció de l'entorn i l'estimació del desplaçament relatiu del vehicle. Com es mostra a la figura 4.19, les 7 zones definides a la imatge tenen associades 5 àrees a la descripció de l'entorn.

Per determinar la relació entre les 5 àrees de l'entorn i les posició a la imatge de les 7 zones, cal utilitzar les equacions A.6 i A.7 assumint que el pendent del terra és constant. En funció de la seva posició a la imatge, les regions generades per la segmentació, s'etiqueten amb:

1. Regió dintre de la zona inicial del camí
2. Regió dintre del camí (però fora de la zona inicial)
3. Regió dintre del camí si hi ha un error en la direcció menor o igual de $+\alpha$ graus
4. Regió dintre del camí si hi ha un error en la direcció menor o igual de $-\alpha$ graus
5. Regió dintre del marge dret
6. Regió dintre del marge esquerra
7. Regió dintre de la part superior

Si una regió de la imatge està inclosa simultàniament en varies de les zones esmentades, s'indica totes les zones en les que apareix.

Les zones 1 i 2 de la imatge corresponen a la posició esperada del camí a la imatge, que es deduïda a partir de la descripció de l'entorn i el coneixement, gràcies als

sensors odomètrics, del desplaçament relatiu (ΔP) del mòbil entre la generació de la descripció actual i la captació de la imatge que s'analitza. Com s'ha de preveure la possibilitat de que el vehicle no estigui orientat correctament respecte al camí (ja sigui per una maniobra d'evitació d'un obstacle, per una derrapada de les rodes del vehicle o per un control de la trajectòria erroni) o un possible canvi de la direcció o de l'amplada del camí en l'àrea de l'entorn observada per primera vegada, es fa necessari la definició de les zones 3 i 4 a la imatge, que corresponen a zones on també es pot trobar el camí en aquests casos. La mida de les zones (àrees) 3 i 4 queden definides pel valor α de la variació de direcció màxima considerada, entre la direcció definida pel camí a la imatge, i la que s'espera que havia de definir.

Com més gran sigui el valor de α , més gran podrà ser la diferència entre la direcció/posició esperada del camí i la direcció/posició real del camí vista des del mòbil (càmera). Per contra, valors grans de α fan que una part important del marge del camí sigui considerada com una zona on es pot trobar el camí, i per tant si el color d'una regió del marge és semblant a un dels colors-camí, fet bastant habitual en els camins mal o feblement estructurats, serà classificada com camí ampliat. Per tant l'elecció del valor de α , és un compromís entre permetre variacions importants entre la direcció/posició esperada i real del camí, el temps de procés, i minimitzar el risc que una regió del marge amb color similar al camí també sigui considerada com camí. En tot cas, el valor adequat de α és l'error en l'estimació de la direcció del vehicle, paràmetre que depèn de les característiques del vehicle i de l'entorn.

Per determinar a la descripció de l'entorn, la mida i posició de les zones 3 i 4, es calcula prèviament el rectangle envoltant del camí (zona 2) a la descripció. Donats els punt $(X_{\text{limit_esquerra_camí}}, D)$ i $(X_{\text{limit_dret_camí}}, D)$ -posició dels límits esquerra i dret del camí a la descripció actual de l'entorn a una distància D , on D és la distància màxima a la que s'ha detectat el camí, figura 4.19- s'obtenen els punts $(X_{\text{limit_esquerra_marge}}, D)$ i $(X_{\text{limit_dret_marge}}, D)$ mitjançant les equacions:

$$X_{\text{limit_esquerra_marge}} = X_{\text{limit_esquerra_camí}} - (D \cdot \sin \alpha) \quad (4.4)$$

$$X_{\text{limit_dret_marge}} = X_{\text{limit_dret_camí}} + (D \cdot \sin \alpha) \quad (4.5)$$

Amb les equacions A.4 i A.5, es determinen la posició dels píxels $(x_{\text{limit_esquerra_marge}}, y_{\text{limit_esquerra_marge}})$ i $(x_{\text{limit_dret_marge}}, y_{\text{limit_dret_marge}})$ a la imatge.

A la taula 4.5 es pot veure la incidència del valor α en el nombre de píxels a analitzar en la imatge -nombre de píxels de les zones 1, 2, 3 i 4-, i per tant en el temps de procés. En aquest exemple, el camí considerat té una amplada de 3 metres, i les

característiques del sistema d'adquisició d'imatges utilitzat són les que es proposen a l'annex B.1. El nombre de píxels associats a cada valor de α , s'obté utilitzant les equacions 4.4, 4.5, A.4 i A.5. En la tercera columna s'indica l'increment en el nombre de píxels a analitzar, i en la quarta apareix la mida màxima, en metres, de les zones 3 i 4 a la descripció de l'entorn (màxima diferència prevista entre la posició indicada del camí a la descripció de l'entorn i la posició real del camí), en funció de l'angle α .

α	Nº Píxels	Increment	Variació màx.
0°	20554		0
2°	21536	4.7%	0.7
4°	22564	4.7%	1.4
6°	23532	4.2%	2.09
8°	24841	5.5%	2.78
10°	25456	2.4%	3.47
12°	26306	3.3%	4.16
14°	27378	4.0%	4.84
16°	27862	1.7%	5.51
18°	28432	2.0%	6.18
20°	28588	0.5%	6.84
>21°	28672	0.2%	>7.49

Taula 4.5. Nombre de píxels que es troben en les zones 1 a 4 en funció de α .

4.3.2. Procés de Classificació

La classificació de les regions de la imatge segmentada es fa tenint en compte el seu color i la seva posició a la imatge, de la següent forma:

- Totes les regions que estan dintre, total o parcialment, de la zona inicial del camí (zona 1) a la imatge i tenen un color igual a un dels colors-camí, es classifiquen com a **camí**.
- Totes les regions que estan totalment dintre de les zones 1 o 2, i tenen un color similar a un dels colors-camí, són també classificades com **camí**.
- Totes les regions que es troben parcialment dintre de les zones 3 o 4 i fora de les 5, 6 i 7, i tenen un color similar a un dels colors-camí, són classificades com **camí ampliat**.
- Totes les regions sense etiquetar i que queden envoltades per regions marcades com camí o camí ampliat, es marquen com **obstacles potencials**. Aquestes regions es caracteritzen per estar dintre de la zona que es considera camí, però tenen un color diferent als colors-camí. Per tant, en funció de l'alçada de l'objecte al que està associada cada una de les regions classificades com obstacle-potencial, l'objecte serà considerat obstacle o

camí. Si l'alçada de l'objecte associat a la regió classificada com obstacle-potencial és superior a un llindar, es considerarà que és un obstacle. En cas contrari, la regió classificada com obstacle-potencial està associada a un material diferent (gespa, grava, etc.) que també forma la superfície del camí. El valor del llindar que ens determina a partir de quina alçada els objectes són considerats obstacles, depèn de les característiques del mòbil i fa referència al màxim desnivell del terreny que aquest pot superar.

- La resta de regions, les que no han estat classificades com camí, camí ampliat o no són obstacles potencials, es marquen com **marge**. Les regions etiquetades com marge estan normalment situades a la part dreta, esquerra i superior de la imatge, parcial o totalment dintre de les zones 5, 6 o 7.

Així doncs, aquesta classificació de les regions de la imatge, etiqueta les regions en quatre tipus (camí, camí ampliat, obstacle potencials i marge) segons el color de la regió i la seva posició a la imatge (on les zones de la imatge s'ha definit a partir de la posició del camí a la descripció actual de l'entorn, i el possible error en l'orientació del vehicle). A partir de la seqüència de resultats generats per la classificació s'actualitzarà la descripció de l'entorn. Per generar la descripció de l'entorn, cal verificar la coherència de la classificació realitzada, calcular l'alçada de les regions marcades com obstacle potencial, i realitzar la projecció inversa per obtenir la posició en coordenades del món de les regions que apareixen a la imatge. En l'annex B es presenta en detall la implementació del procés de classificació de les regions.

La figura 4.20 mostra el resultat de la classificació de les imatges de la figura 4.18. Les línies negres que apareixen en les imatges corresponen als límits de les 7 zones de la imatge definides a partir de la descripció actual de l'entorn⁸ (que ha estat generada a partir de l'anàlisi de les imatges anteriors de la seqüència). Les regions de color verd són les que ha estat classificades com camí i el seu color és el color-camí més freqüent (color 1). Les regions de color groc són les que han estat classificades com camí ampliat i el seu color és semblant al color 1. Les regions de color blau cel i blau marí són les que ha estat classificades com camí i el seu color és el segon color-camí (color 2), i les que han estat classificades com camí ampliat i el seu color és semblant al segon color-camí, respectivament. Les regions de color vermell corresponen a les que

⁸ El fet que les línies que delimiten les 7 zones de la imatge, no tinguin la mateixa situació en les 9 imatges de la figura 4.19, es degut a que són camins d'amplades diferents i que han estat captats amb diferents sistemes de visió i amb diferents configuracions (distància focal, posició i orientació de la càmera respecte el pla de terra).

que han estat classificades com obstacles potencials i les de color taronja són les que han estat classificades com marge.

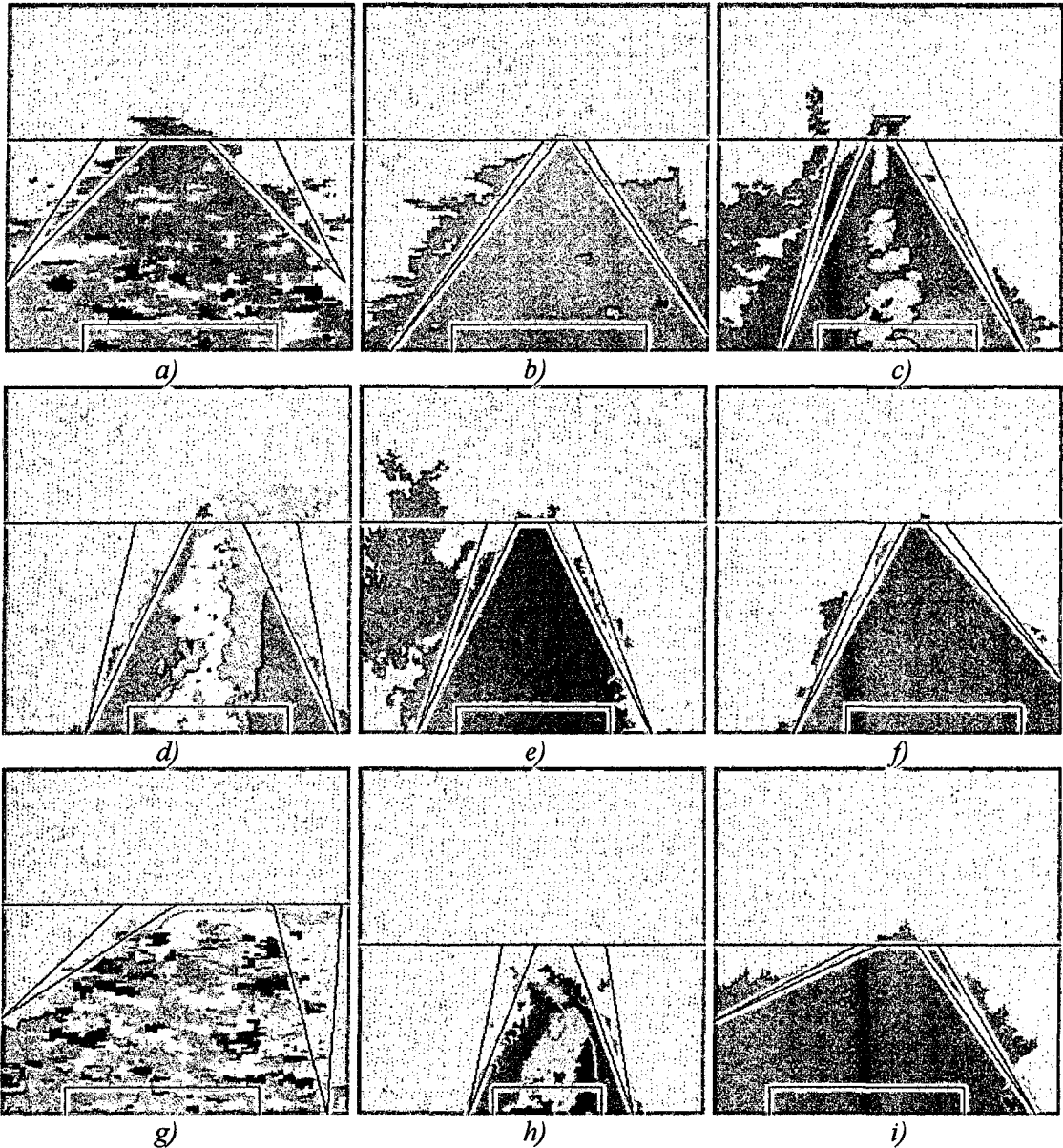
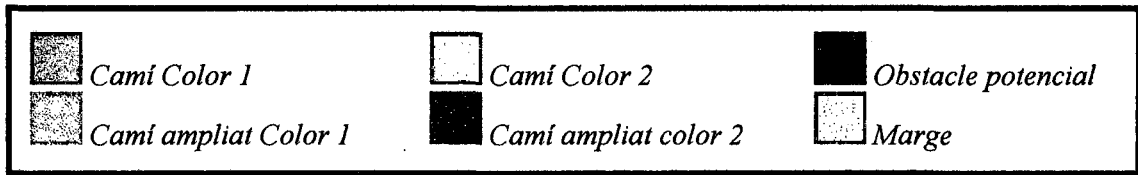


Figura 4.20. Classificació de les regions de les imatges segmentades de la figura 4.18, basada en el color i posició de les regions.

En la figura 4.21, es pot comparar el resultat de la classificació de regions sobre la imatge RGB original. Els traços en color blanc, corresponen a la frontera entre regions classificades com camí o camí ampliati, i regions classificades com obstacle potencial o marge.

Com es pot veure a la figura 4.21, la determinació amb robustesa de la posició del camí a la imatge no és possible utilitzant únicament la informació continguda en una imatge. El problema que es presenten més sovint és la similitud de color entre el camí i part del seu marge. En aquests casos pot passar que una regió classificada com camí contingui part del marge (figures 4.12.b, c, d, e i f) o que una regió classificada com a marge contingui part del camí (figura 4.21.g). Per determinar la situació del camí amb fiabilitat, caldrà finalment imposar les restriccions del model genèric de camí (amplada i direcció del camí localment constants) i integrar la informació d'una seqüència d'imatges.

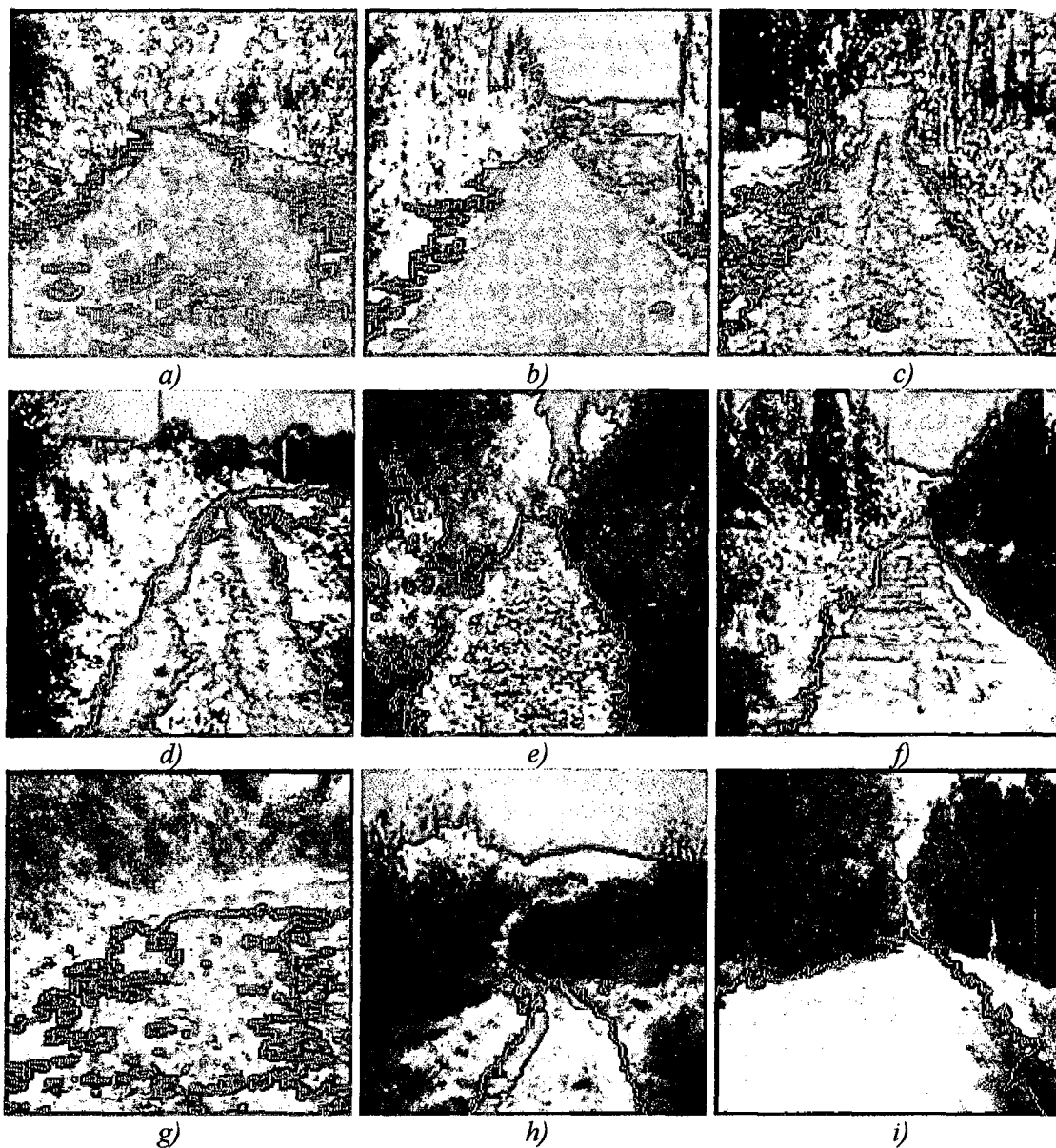


Figura 4.21. Superposició del resultat de la classificació de regions sobre la imatge original.

4.4. Conclusions

Al llarg d'aquest capítol s'ha presentat la part de la metodologia que es basa en el processament d'una única imatge per generar una primera aproximació de la posició que ocupa el camí a l'entorn, la imatge classificada. En la imatge classificada s'indica quines regions de la imatge estan associades al camí, al marge o a possibles obstacles. L'obtenció de la imatge classificada requereix de tres etapes:

1. Transformació d'espai de representació del color, del RGB en el que les càmeres color subministren les imatges a l'H/I.
2. Segmentació a baix nivell de la imatge mitjançant l'anàlisi del color
3. Classificació de les regions en funció dels colors de les matèries que formen el camí a l'àrea analitzada i de la posició a la imatge de les regions respecte la descripció de l'entorn

Com es pot veure a la figura 4.21, aquest procés no és suficient per definir, de forma fiable, la zona de la imatge que està associada al camí. Com es presenta en el cinquè capítol, l'anàlisi de la seqüència d'imatges ens permet generar i actualitzar una descripció de l'entorn on s'analitza la coherència de la posició, a la descripció, del camí i dels obstacles. Així mateix, la imposició de les restriccions incloses en el model genèric de camí feblement estructurat permet definir la posició del camí en els casos en que part del marge i el camí presentin el mateix color.

Les principals idees i conceptes que s'han definit i utilitzat en la primera part de la metodologia presentada són:

- Definició de l'espai H/I, que resulta una característica força independent de les condicions d'il·luminació, molt discriminadora, i per tant interessant per segmentar imatges. La resolució coloromètrica utilitzada en l'espai H/I pot ser variable en funció de la informació que es vol obtenir de la imatge.
- Adequació del procés de segmentació de la imatge tenint en compte les característiques de les imatges analitzades: Identificació de les regions (posició i color) que defineixen el camí en la zona de la imatge més propera al vehicle (aquesta informació, els colors-camins, s'utilitza per guiar el procés de creixement de regions) i segmentació de la imatge en un nombre de regions que té en compte el nombre de zones rellevants que apareixen de la imatge (camí, marge i obstacles) més que el nombre d'elements presents a la imatge (que normalment és elevat donat la complexitat de les escenes

captades en entorns exteriors i la il·luminació no homogènia). Per assolir aquest objectiu, s'ha introduït en el procés de segmentació la reducció del nombre de regions mitjançant l'absorció de regions de mida petita, que no són rellevants per descriure l'entorn, per una regió veïna de mida gran.

- Classificació de les regions de la imatge segmentada en funció de si el seu color és semblant a un dels colors-camí i de la seva posició a la imatge respecte la posició esperada del camí que es segueix a la imatge. La posició en la que s'espera trobar el camí a la imatge queda definida per la descripció de l'entorn i el possible error en l'estimació del desplaçament/orientació relatiu del vehicle.