

Contrast, equalització, quantització.

QUIM LLIMONA TORRAS - 146662

3.1 Canvis de contrast en una imatge

3.1.1 Canvi de contrast lineal

En aquests exercicis, realitzaré canvis de contrast lineals en imatges. És a dir, agafaré cert rang de valor de la imatge i faré que la imatge de sortida només tingui píxels continguts en un altre rang. La funció de transferència serà per interpolació lineal.

El primer exercici consisteix en augmentar el contrast d'una imatge de baix contrast. Per fer això, primer de tot cal mirar quins són els valors mínim i màxim de la imatge, i aleshores fer una transformació en què aquests valors esdevinguin 0 i 1 i la resta s'interpolin adequadament. En comptes d'agafar el màxim i mínim absoluts, he triat uns marges on ja hi haguessin bastant píxels. Així, n'hi haurà alguns que quedaran saturats (clipping), però el contrast serà més gran. Els valors per la imatge que mostro són (0.45,0.9).



Imatge original



Imatge amb més contrast

A partir d'ara, prendré la imatge amb el contrast millorat com a base per a efectuar noves transformacions.

La primera serà disminuir el contrast; en comptes de tenir una imatge que va de 0 a 1, faré que vagi de 0.2 a 0.8 escalant els seus valors (matemàticament, seria equivalent a multiplicar cada píxel per 0.6 i sumar-li 0.2).

La segona, expandir els tons foscos. És a dir, he agafat la imatge, he saturat tots els valors per sobre de 0.4 a 1 i he expandit el rang (0,0.4) perquè ocupi tot el marge disponible. Així, els detalls foscos que abans no es distingien bé ara es veuen amb claredat. No obstant això, la imatge resultant sembla molt més clara que la original. Això és perquè l'histograma de la imatge té certa semblança a una gaussiana, i per tant hi ha més valors a prop del 0.4 que del 0. Per tant, al fer l'escalat hi haurà més valors a prop de l'1 que del 0.

La tercera, expandir els tons clars. És anàloga a l'anterior; tots els valors per sota de 0.6 es saturen a 0, i el rang (0.6, 1) s'expandeix per veure millor els detalls molt clars. Passa el mateix d'abans: la imatge queda més fosca perquè donada la tendència gaussiana de l'histograma, hi ha més valors prop del 0.6 que de l'1.

La quarta, invertir el contrast de la imatge (efecte negatiu). És un efecte molt típic, i molt fàcil d'obtenir; només cal assignar a cada píxel de la imatge amb valor x el valor $(1-x)$.

La cinquena, fer un clipping, és a dir, posar tots els nivells per sobre d'un llindar (0.7) a aquest llindar. Aquest efecte pot ocórrer de forma natural per la limitació del rang dinàmic de la càmera o es pot fer artificialment per sobresaturar la imatge, simulant, per exemple, que hi ha una font de llum potent propera a l'objecte. En una implementació més bona, però, no es faria un llindar dur sinó que hi hauria una transició més suau, ajuntant un primer tram lineal amb un de logarítmic després, per exemple, o fent servir funcions sigmoïdals com $\text{atan}(x)$.

La sisena, binaritzar la imatge a 0.5. És a dir, fer que la imatge resultant només tingui 2 nivells possibles: 0 o 1. És el típic efecte de "blanc i negre", que fa una transformació molt bàsica de fotos a dibuixos. Degut a la implementació que he fet servir, han aparegut uns artefactes de color. Caldria passar la imatge a blanc i negre, o fer servir una altra implementació. Ho he deixat així perquè estèticament m'agrada com queda.

Aquests són els resultats:



Reducció de contrast



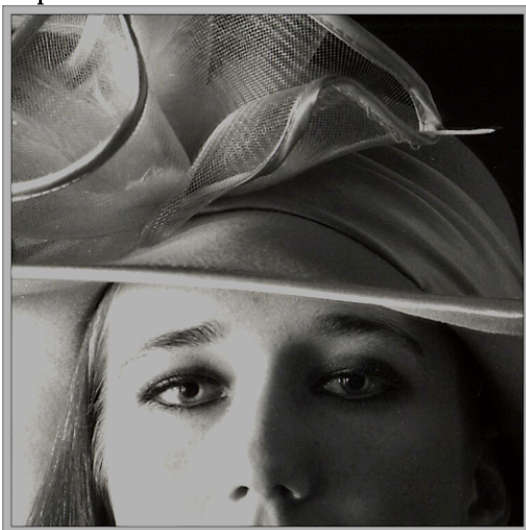
Expansió de tons foscos



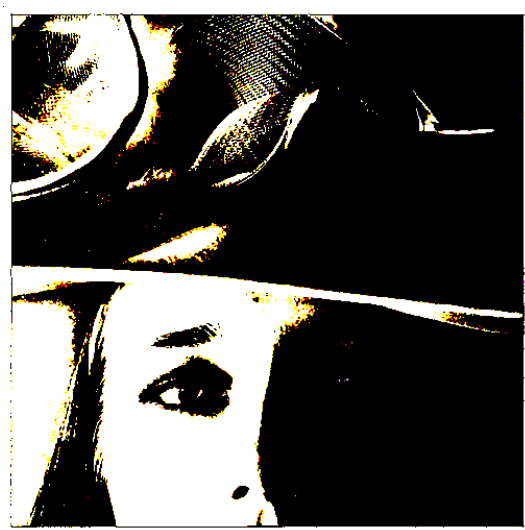
Expansió de tons clars



Inversió de contrast



Clipping



Binarització

3.1.2 Canvi de contrast no lineal

Els canvis de contrast no lineal funcionen de manera similar als lineals, amb la peculiaritat que tenen un factor gamma que controla la curvatura de la línia. Per un factor gamma = 1, són equivalents. Si gamma és inferior a 1, la funció es torna còncava. Si és superior, es torna convexa.

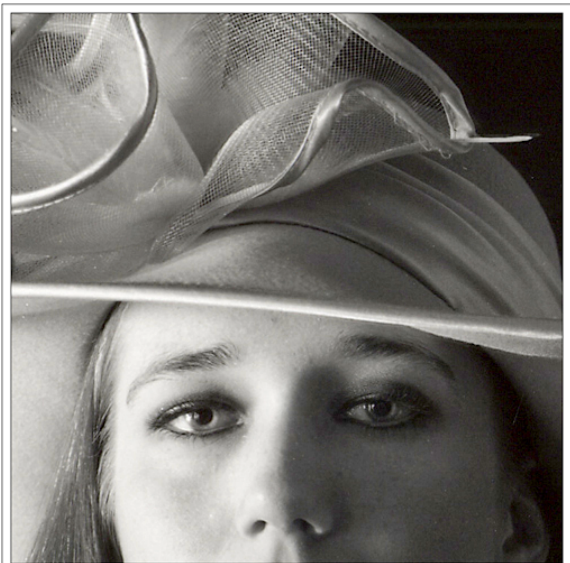
Aquest tipus de transformacions són molt habituals com a transformació general de la imatge dels monitors dels ordinadors, i és molt útil estar-hi familiaritzat per poder-los calibrar correctament i gaudir d'una imatge molt més vívida en termes de color i contrast.

Com afecta la gamma a la imatge? Si la gamma es fa més petita, la funció es torna més còncava, per tant es potencien més els tons alts. La imatge, per tant, serà més clara, i viceversa.

Per veure millor els detalls foscos de la imatge presentada abans, per tant, caldrà posar una gamma inferior a 1. Ho he provat amb 0.7, i efectivament s'aprecia molt millor el que hi ha a les regions en penombra:



Imatge amb correcció lineal



Imatge amb correcció no lineal

3.2 Equalització de l'histograma de la imatge

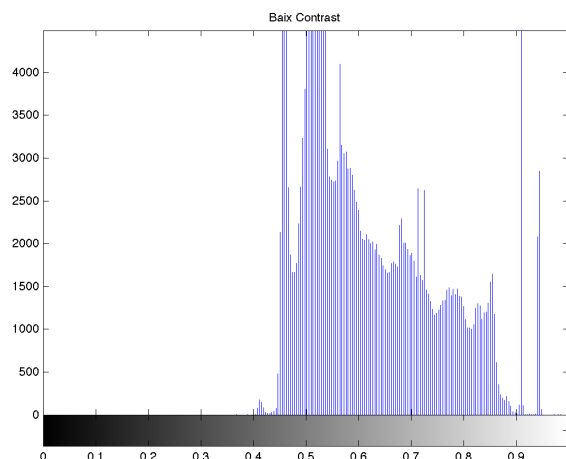
Nota: he treballat sempre amb imatges en el rang (0,1) però quantitzades a 256 nivells.

L'histograma d'una imatge és la freqüència relativa dels seus nivells de gris. En el cas d'imatges amb valors continus, cal especificar unes regions en les quals s'ajuntaran els valors. També es pot definir l'histograma continu, que no és més que la funció de densitat de probabilitat (pdf) de la imatge. En el nostre cas, però, les imatges tenen valors sencers discrets d'entre un rang de 256. Per tant, podem calcular l'histograma especificant 256 regions que corresponen a cadascun dels nivells de gris disponibles.

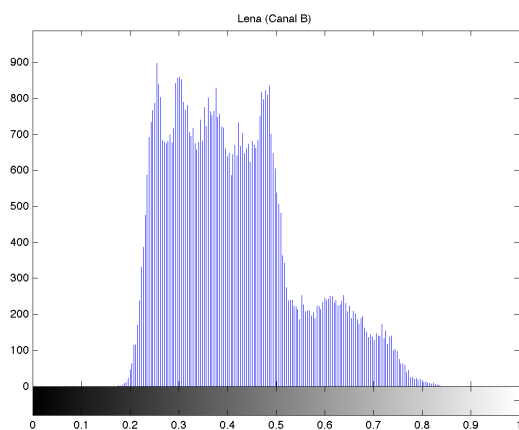
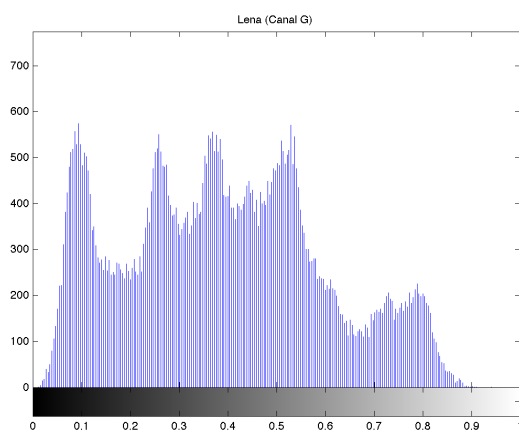
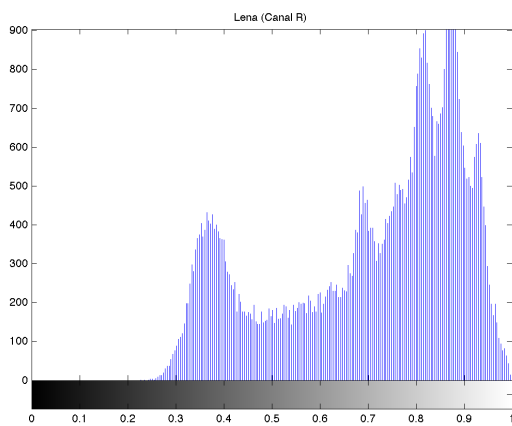
Per il·lustrar el concepte, he calculat l'histograma de la imatge feta servir abans (versió de baix contrast), on es veuen clarament les regions que he fet servir per expandir el contrast, i el de cadascun dels canals de color de la Lena:



Imatge de baix contrast



Histograma de la imatge



Lena en color, amb els seus histogrames (Vermell, Verd i Blau, en ordre).

Podem observar com el vermell té valors molt alts a la part més a prop de l'1 i els altres no, i això indica que la imatge tira cap al vermell.

Abans hem vist com "millorar" una imatge augmentant-ne el seu rang mitjançant transformacions lineals o paramètriques del contrast. Una generalització d'aquest sistema és l'equalització d'histograma: donat l'histograma de la imatge, realitzem una transformació sobre aquesta de manera que l'histograma resultant ompli tots els "buits" que el primer tenia, i eixampli les regions d'alta probabilitat i aixafi les de baixa. Amb això aconseguim que tots els nivells de gris tinguin una representació semblant, i per tant, que no quedin regions massa fosques ni massa clares.

L'algorisme es presenta als apunts de teoria, i es basa en assignar a cada píxel de la imatge el valor de la funció de probabilitat acumulada (cdf) en aquell nivell de gris. És a dir, quina proporció de píxels de la imatge està per sota d'aquell nivell. La seva implementació en MATLAB es troba a la carpeta adjunta de codi, al fitxer equalitza.m.

En el cas d'imatges en color, es pot equalitzar cada canal per separat. Tanmateix, si es treballa en un espai de color RGB pot ser que la tonalitat de color general de la imatge quedi molt alterada; això passarà si els canals tenen distribucions molt diferents. Una millora de l'algorisme és passar la imatge a l'espai HSV o algun amb equivalent perceptual, i equalitzar-ne només alguns canals (per exemple, deixant el hue intacte). Aquesta millora, però, la deixo per futures experimentacions.

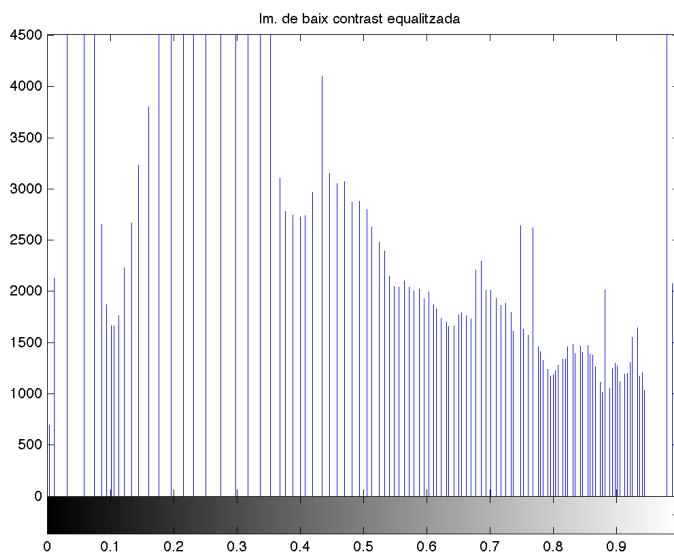
A continuació en mostro els resultats, tant per la imatge del barret com per la Lena.



Imatge original



Imatge equalitzada



Histograma de la imatge. Ara ocupa tot l'ample disponible, i és més regular que abans. En algunes regions els nivells estan molt espaiats (expansió), i en d'altres molt junts (compressió).

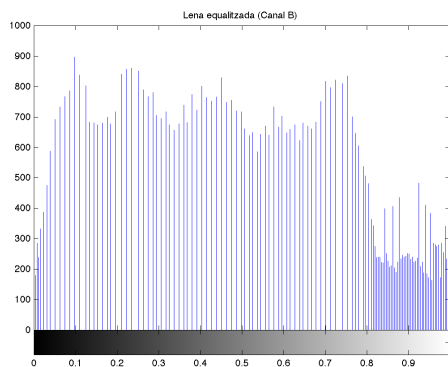
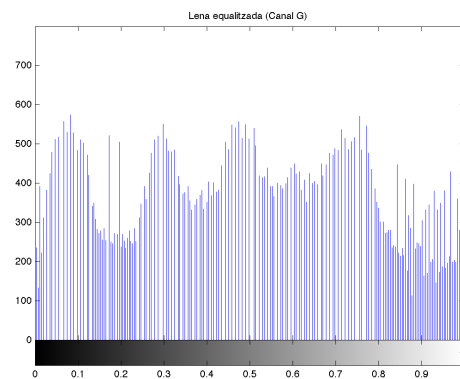
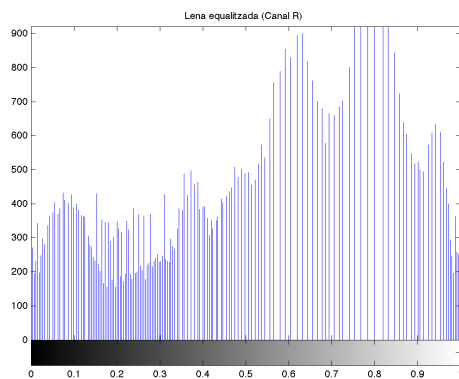


Lena original



Lena equalitzada

L'equalització ha eliminat la tendència al vermell de la imatge, com he comentat abans que passa quan s'equalitzen tots els canals indiscriminadament.



Histograma equalitzat. Efectivament, ara tots els canals ocupen tot l'ample disponible, i les regions que abans tenien probabilitat baixa tenen els valors molt propers.

3.2.1 Entropia d'una imatge

L'entropia d'una imatge és una mesura del seu desordre, i està relacionada amb la predictibilitat de la mateixa donat el seu histograma. Per exemple, si tenim un histograma on un nivell negre tingui un valor molt alt i la resta de valors siguin baixos, la imatge tindrà poca entropia perquè ja podem predir que serà majoritàriament negra. En canvi, si l'histograma és constant (equivalent a una imatge completament sorollosa), l'entropia és màxima ja que no podem predir res concret sobre el contingut; donat un píxel, podria prendre qualsevol valor amb igual probabilitat.

En el cas de les imatges, hi ha un detall que cal tenir en compte; si un valor no està present a la imatge, cal descartar-lo abans de calcular l'entropia, ja que si el considerem com a part de l'histograma l'entropia seria infinita: si un nivell mai apareix, no podem predir res sobre ell (a part de la seva no-existència, evidentment).

La fórmula que he fet servir està als apunts, i la seva implementació al fitxer adjunt entropia.m. Calculant l'entropia de la imatge en blanc i negre del barret abans i després d'equalitzar-la he obtingut el següent:

	Abans d'equalitzar	Després d'equalitzar
Entropia (bits/píxel)	6.6218	6.5819

Per tant, l'entropia creix a l'equalitzar les imatges. És lògic, ja que els valors que abans no estaven presents en la imatge o tenien una probabilitat baixa ara estan presents, i per tant hi ha més incertesa.

3.3 Quantització

3.3.1 Quantitzador uniforme de N nivells

Per entendre com es pot implementar un quantitzador uniforme de manera molt senzilla, he analitzat el següent fragment de codi:

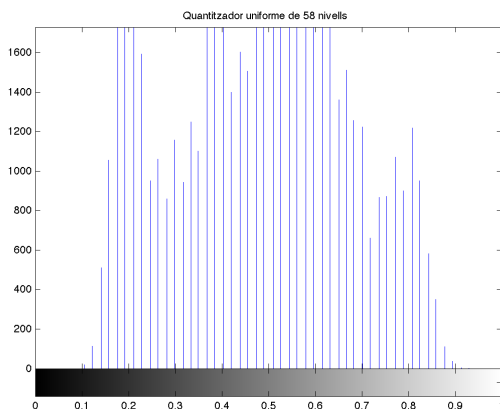
```
minim=min(min(lenna));  maxim=max(max(lenna));  M=maxim-minim;  g1=57/M;  g2=255/57;
lenna_quantitz1=round(g1*(lenna-minim))*g2;      g1=10/M;      g2=255/10;
lenna_quantitz2=round(g1*(lenna-minim))*g2;      g1=2/M;      g2=255/2;
lenna_quantitz3=round(g1*(lenna-minim))*g2;
```

En aquest cas, la quantització es pot descomposar en els següents passos:

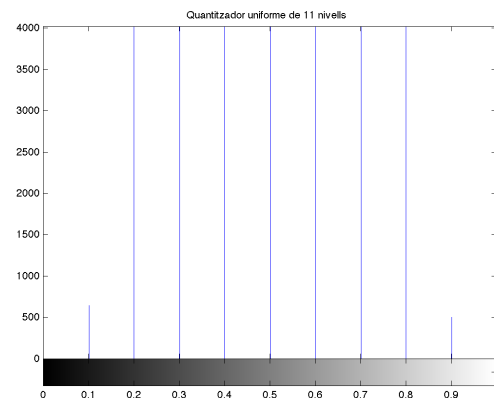
- Normalitzar la imatge en el rang (0,1). Això es fa restant-li el mínim i dividint-la pel seu rang, màxim - mínim.
- Multiplicar la imatge pel nombre de nivells - 1. Així, la imatge anirà entre 0 i aquest nombre. Es resta 1 perquè el 0 també és un nivell.
- Arrodonir la imatge perquè agafi valors discrets.
- Dividir-la pel nombre de nivells - 1 per tornar-la a tenir en el rang (0,1).
- Multiplicar per 255 per tenir-la en el rang (0,255).

En realitat, normalitzar-la fent servir el màxim i mínim de la imatge equival a fer una transformació lineal de contrast com la que he fet al principi de tot de la pràctica per augmentar el rang dinàmic de la imatge. He omès aquest pas a la meua funció per tenir un quantitzador més pur i, per tant, més sensible a certs errors, com ara a passar-li imatges de baix contrast. El que he fet és assumir que la imatge ja entra dins el rang (0,1). També he omès l'últim pas, ja que m'agrada més treballar amb imatges normalitzades.

Pel que fa a l'histograma, quantitzant una imatge s'aconsegueix que hi hagi menys bandes actives ajuntant-ne diverses en una de sola:



Histograma amb quant. de 58 nivells



Histograma amb quant. de 11 nivells

Per tant, el nombre de nivells a les lènyes quantitzades és de 58, 11 i 3, respectivament. El resultat és el següent:



Lena original



Quantització amb 58 nivells



Quantització amb 11 nivells



Quantització amb 3 nivells

Quan es quantitza amb pocs nivells, a les zones on hi ha gradients de color hi apareixen unes bandes molt marcades, ja que el gradient queda segmentat en regions segons correspongui als valors quantitzats. Aquest efecte és especialment evident a la quantització d'11 nivells.

En algun casos es pot buscar com a efecte artístic (per simular un dibuix o quadre a partir d'una foto, per exemple), però en general ho considerarem com una distorsió no desitjada.

Després de fer diversos experiments amb la lena en escala de grisos, he vist que un nombre de nivells proper a 30 és acceptable per no apreciar clarament la quantització; encara hi ha banding, sobretot a la pell, però no es veu a primer cop d'ull. És un valor proper a 32, que equivaldria a 4 bits.



Lena original



Lena quantitzada amb 30 nivells de gris

Si es fan servir només dos nivells, el quantitzador també es pot fer servir com a binaritzador, obtenint un efecte similar al del primer exercici. En aquest cas no ha aparegut aquell artefacte de color:



Lena original



Lena binaritzada

3.3.2 Criteris quantitatius de fidelitat

Quan es desenvolupen algorismes de compressió d'imatges, convé definir una mesura de la qualitat de la imatge restaurada per veure quin compressor és millor. Una mesura molt bàsica és l'error quadràtic mitjà (MSE), que és la mitjana de la diferència al quadrat entre les dues imatges.

Aquesta mesura té un problema, i és que depèn del rang de valors de les imatges. Donat un compressor, una imatge normalitzada de 0 a 1 i una amb escala completa de 0 a 255 donaran valors completament diferents. El problema és major si es fa servir escala completa amb imatges de diferent resolució de bits. Per resoldre aquest problema i d'altres, la literatura acostuma a expressar les mesures amb la PSNR (peak signal to noise ratio), que és una mesura relativa en dB de la relació entre el rang de valors de la imatge i el rang de valors del soroll. Així els algorismes es poden comparar independentment del format de la imatge de test. A més, la PSNR és una mesura logarítmica (ja que està en dB), per tant permet comparar directament l'ordre de magnitud de l'error.

En el meu cas, les imatges estan normalitzades en el rang (0,1), per tant el MSE serà inferior a 1. He adaptat les funcions donades perquè funcionessin correctament. Aquests són els resultats de l'avaluació de la quantització uniforme:

	MSE	PSNR (dB)
Quant. unif. 58 nivells	2.5667e-05	45.9062
Quant. unif. 11 nivells	0.00075673	31.2106
Quant. unif. 3 nivells	0.021054	16.7666

Abans d'analitzar el resultat, vull destacar la utilitat de fer servir una escala en dB quan la magnitud varia molt. Mentre el MSE varia entre l'ordre de 10^{-2} i 10^{-5} , el PSNR va de 16 a 46. Per tant, és molt més fàcil d'escriure i d'interpretar.

Com menys nivells de quantització es fan servir, més gran és l'error, ja que els possibles valors estan més separats entre ells i més valors originals queden entremig. Això es correspon amb el resultat visual vist abans; a menys nivells, més distorsió es percep.

3.3.3 Quantitzador adaptat a la imatge a quantitzar

A la imatge quantitzada a 3 nivells que he mostrat abans, quasi tota la superfície és de color gris; només algunes regions són blanques o negres. Això passa perquè l'histograma segueix una distribució lleugerament gaussiana i no uniforme, però la quantificació s'ha fet uniformement partint el rang en parts iguals. Per tant, els valors extrems tenen menys probabilitat de sortir que els del mig.

Per minimitzar aquest efecte, es poden fer servir quantitzadors no uniformes amb els nivells adaptats a l'histograma de la imatge. Una manera senzilla però efectiva de fer això és quantitzar uniformement una transformació de la imatge que tingui una distribució uniforme, i després desfer aquesta transformació sobre la imatge quantitzada. L'equalització d'histograma és una d'aquestes transformacions, i ja l'he implementat abans.

El problema principal és com desfer la transformació; és a dir, com trobar la transformació inversa. El fet que la transformació sigui una funció discreta d'un conjunt de 256 valors a un altre de 256 valors i que sigui monòtona creixent, fa que no sigui quasi mai injectiva ni exhaustiva. Per exemple, en qualsevol expansió d'una regió es perd automàticament l'exhaustivitat, perquè un conjunt de punts es reparteix entre un conjunt més gran i sempre en quedaran de lliures, i en qualsevol contracció es perd l'injectivitat, perquè un conjunt de punts es mapeja a un conjunt més petit i sempre n'hi haurà que aniran a parar al mateix. Per tant, la seva inversa no està ben definida. Per solucionar aquest problema, alguns autors recomanen adaptar la funció a cada píxel que es vol reconstruir i fer que la funció tingui una imatge més gran interpolant els valors a partir dels seus veïns (http://www.eurasip.org/Proceedings/Eusipco/Eusipco1998/sessions/W%20M/WM%20P-7b/861/P05_2b_10_10.PDF).

Jo he adoptat una solució més simple, disponible al fitxer quantitzador.m. La no-injectivitat no es tracta; s'assigna el valor més alt dels disponibles. Per la no-exhaustivitat, s'interpolen els punts que no eren imatge, segons el comentat en aquest fòrum: http://www.mathworks.com/matlabcentral/newsreader/view_thread/253057

No tractar la no-injectivitat pot semblar un error, però en realitat les no-injectivitats més crítiques solen venir donades per zones de probabilitat zero, que resulten en una línia horitzontal a la funció de probabilitat acumulada. Aquestes zones no ens han de preocupar, ja que no formen part del domini útil (i, per tant, de les anti-imatges que voldrem obtenir). Les no-injectivitats puntuals que tenen lloc quan hi ha valors veïns que tenen imatges iguals per equilibrar un pendent fort de la funció tampoc són massa problemàtiques, ja que a l'estar tant properes si prenem una anti-imatge per l'altra l'error serà petit. En canvi, és important definir bé la inversa en el cas de no-exhaustivitat, perquè si fem transformacions a la imatge equalitzada i després volem recuperar l'histograma original pot ser que anem a parar a punts que no són del domini de la inversa.

En qualsevol aproximació, però, hi haurà un cert error, que es minimitzarà en major o menor mesura però que no es pot evitar completament donada la naturalesa destructiva de la transformació discreta durant l'equalització. En el cas de l'article que he citat, l'error té una PSNR d'entre 50 i 60 dB; en la implementació que he fet servir, els resultats són una mica pitjors. Per tant, no podem esperar que el quantitzador tingui un rendiment millor; si un quantitzador uniforme ja ens dóna aquest marge, segurament no val la pena provar-ho amb el no uniforme.

Aquest no és el cas pels exemples proposats, ja que el nombre de nivells és baix. Mostro a continuació el resultat de la quantització adaptada, comparada amb la uniforme amb el mateix nombre de nivells:



Quant. uniforme de 58 nivells



Quant. adaptada de 58 nivells



Quant. uniforme de 11 nivells



Quant. adaptada de 11 nivells



Quant. uniforme de 3 nivells



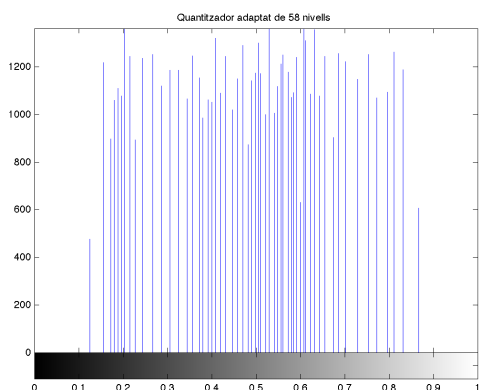
Quant. adaptada de 3 nivells

El cas més espectacular és el de 3 nivells; amb quantització uniforme quasi no es reconeixia la lena, mentre que amb l'adaptada se'n veu perfectament la silueta. En el cas d'11 nivells, el banding es manté però és molt menys molest (sobretot a la cara), i la imatge té un contrast més bo. A la de 58 és on menys s'aprecia l'efecte.

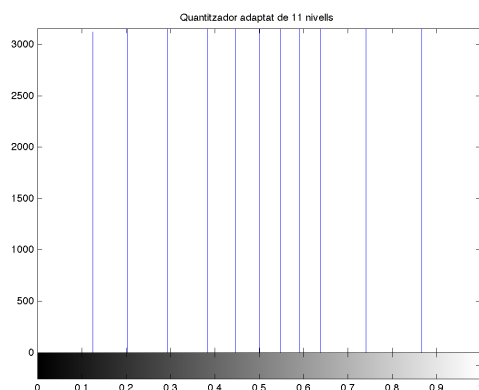
Tot això és coherent amb els resultats numèrics; en tots els casos hi ha hagut una millora d'entre 1 i 2 dB:

	MSE Q. uniforme	MSE Q. adaptat	PSNR Q. uniforme	PSNR Q.adaptat
58 nivells	2.5667e-05	1.9274e-05	45.9062	47.1503
11 nivells	0.00075673	0.00061248	31.2106	32.1291
3 nivells	0.021054	0.013634	16.7666	18.6538

L'histograma segueix tenint menys bandes però estan repartides de manera desigual; n'hi ha més a les regions de probabilitat alta, per cometre-hi menys error:



Quant. adaptada de 58 nivells



Quant. adaptada de 11 nivells

3.4 Halftoning

El dithering és una tècnica que aprofita la integració espacial amb filtre passabaixes del nostre sistema visual per aconseguir interpolar nivells de color o gris combinant trames d'un conjunt de nivells molt reduït. A més, elimina els problemes de banding associats amb la quantització. Aquesta tècnica és la que fan servir les impressores per imprimir amb tanta resolució, i en certa manera és també el fonament de les pantalles en color. També va ser molt popular en els inicis d'Internet mitjançant el format d'imatge comprimida GIF, que la fa servir per quantitzar extremadament les imatges.

Abans d'implementar la tècnica de difusió d'error descrita als apunts, he volgut provar un model de dither molt més simple, i més semblant al que es fa servir en altres camps del processat de senyal; sumar al senyal no quantitzat una component aleatòria que faci oscil·lar el valor quantitzat entre els dos llindars que hi ha al voltant de l'original. Aquesta component l'he calculat com una distribució uniforme d'amplitud parametritzada i mitjana nul·la, multiplicada per l'error de quantització sense dither. El paràmetre l'he posat a 2. Aquest és el resultat:



Dither aleatori, 3 nivells



Dither aleatori, 5 nivells



Dither aleatori, 11 nivells

Com es pot veure, no té res a veure amb la quantització uniforme o adaptada simple. Amb 3 nivells es reconeixen perfectament molts detalls de la imatge, i ja a la de 5 nivells el banding s'ha eliminat, a canvi de tenir un component sorollós. El resultat amb 11 nivells és encara millor.

La tècnica proposada de la difusió d'error dona uns resultats encara millors, fins i tot amb només 2 nivells, però és una mica més complexa. La implementació que n'he fet està al fitxer `error_diffusion.m`. La funció accepta com a paràmetre el nombre de nivells, que es distribueixen uniformement, o bé un vector amb la posició dels nivells desitjats.

Aprofitant això, he volgut comparar el rendiment de la tècnica amb quantització uniforme i adaptada. Per aquest últim cas, he passat els nivells de la quantització adaptada a l'algorisme. Com es pot comprovar a les figures següents, la millora és notable.



Quant. uniforme amb dither, 11 nivells



Quant. adaptada amb dither, 11 nivells



Quant. uniforme amb dither, 5 nivells



Quant. adaptada amb dither, 5 nivells



Quant. uniforme amb dither, 3 nivells



Quant. adaptada amb dither, 3 nivells



Quant. uniforme amb dither, 2 nivells



Quant. adaptada amb dither, 2 nivells

És especialment rellevant el cas de quantització adaptada amb dither de 2 nivells, ja que requereix enviar la mateixa quantitat d'informació (si no es fan servir altres tècniques de compressió) que les imatges binaritzades d'abans. En el cas no adaptat, en canvi, apareixen unes bandes verticals si es fa la imatge petita, segurament a causa de l'aliasing de la trama.

També he fet la prova amb imatges en color, tractant cada canal per separat. Els resultats són també excel·lents:



Quant. uniforme amb dither, 5 nivells



Quant. adaptada amb dither, 5 nivells



Quant. uniforme amb dither, 3 nivells



Quant. adaptada amb dither, 3 nivells



Quant. uniforma amb dither, 3 nivells



Quant. adaptada amb dither, 2 nivells

Cal destacar, novament, que la imatge adaptada de 2 nivells es pot transmetre amb només 3 bits per píxel (un per canal).