Guion V: Cuanficación Escalar

Información sobre la entrega de la práctica

Las prácticas se entregarán en un único fichero comprimido Practica05ApellidoNombre.zip. El fichero contendrá:

* Las funciones de Matlab a realizar en ficheros .m con los nombres de las funciones que se indiquen en el guion.
* Los trozos de código a realizar, que se entregarán todos en los pasos correspondientes de un único fichero .m llamado Practica05ApellidoNombre.m . Este fichero lo crearás modificando el fichero .m Practica05MolinaRafael.m en el servidor.
* Las discusiones y respuestas solicitadas en el guion se entregarán en un único fichero pdf. El nombre del fichero será Practica05ApellidoNombre.pdf. Lo construirás editando Practica05MolinaRafael.doc y salvándolo en formato pdf.

# Cuantificación

**Paso 1**

El primer paso de este guion es un ejercicio muy sencillo. Vamos a diseñar un cuantificador muy simple para datos en [0:255] y se lo aplicaremos a varias imágenes.

X=[0:255];

for i=0:8

factor=2^i;

Q\_X=uint8(floor(factor\*(floor(X/factor)+0.5)));

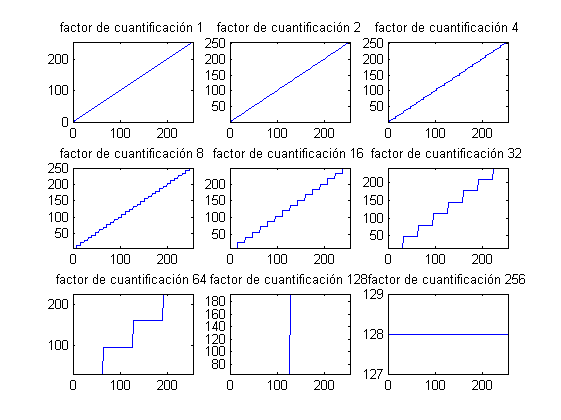
subplot(3,3,i+1)

plot(X,Q\_X); axis('tight');

title(['factor de cuantificación ',num2str(factor)])

end

Entiende que hace el cuantificador y las figuras obtenidas.



**Paso 2**

A continuación aplicaremos estos cuantificadores a la imagen bridge.pgm. Escribe en el paso 2 de Practica05ApellidoNombre.m código de Matlab que:

1. lea la imagen bridge.pgm en la matriz a, la convierta a double y la almacene en adouble
2. calcule y guarde en un vector de 9 componentes la entropía de las imágenes cuantificadas utilizando

uint8(floor(factor\*(floor(adouble(:)/factor)+0.5)));

con los 9 factores, factor=2^i, i=0:8.

1. calcule y guarde, en un vector de 9 componentes, el error cuadrático medio de cuantificación entre la imagen original y la cuantificada para los 9 factores y
2. dibuje en una gráfica con dos figuras las entropías obtenidas y la curva (factor, error cuadrático medio de cuantificación).

**Paso 3**

En el paso 3 de Practica05ApellidoNombre.pdf

1. incluye las entropías y los errores cuadráticos medios calculados en la tabla adjunta. Explica el contenido de la tabla
2. incluye las gráficas de la entropía y la curva (factor, error cuadrático medio de cuantificación).

**Soluciones**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=8 | N=7 | N=6 | N=5 | N=4 | N=3 | N=2 | N=1 | N=0 |
| Entropía | 0 | 0.9433 | 1.7796 | 2.7161 | 3.6950 | 4.6870 | 5.6805 | 6.6753 | 7.6686 |
| Error cuadrático | 3.87e+05 | 1.53e+05 | 4.30e+04 | 1.12e+04 | 2.76e+03 | 702.47 | 193.40 | 64.39 | 0 |

Entropía

Gráfico

Descripción generada automáticamente con confianza media

Error

Gráfico, Gráfico de líneas

Descripción generada automáticamente

**Paso 4**

Vamos ahora a aplicar diferentes cuantificadores a una imagen usando las funciones definidas en Matlab. Comenzaremos cargando la imagen camera.pgm y dibujando su histograma

clear all;close all;

A=imread('camera.pgm');

imhist(A)

Observa que para que conozcas otra función de Matlab hemos utilizado la función imhist. No obstante es mejor que sigas utilizando hist o histc. Ya debes entender bien cómo funcionan estas funciones. El histograma de la imagen es

C:\Users\Rms\Documents\rms\cursos\CRIM\curso 2014-2015\08 Cuantización Escalar\Guiones\untitled.tif

Este histograma no es, evidentemente, muy uniforme.

**Paso 5**

Vamos a aplicarle un cuantificador uniforme con dos intervalos (necesitaríamos, por tanto, como mucho un bit por píxel para codificar la imagen cuantificada), mostraremos la imagen cuantificada y calcularemos el error cuadrático medio de cuantificación. Entiende bien la función quantiz, la usaremos con cierta frecuencia, y el cálculo del error. Observa también como funciona reshape

dA=double(A);

frontera\_particion =[127];

valores\_cuantizados=[63 , 192];

[index,quants]=quantiz(dA(:),frontera\_particion,valores\_cuantizados);

dqA=reshape(quants,size(dA));

qA=uint8(dqA);

subplot(1,2,1), imshow(A); title('Imagen Original')

subplot(1,2,2), imshow(qA); title('Cuantificación uniforme con dos niveles')

error=(dA-dqA).\*(dA-dqA);

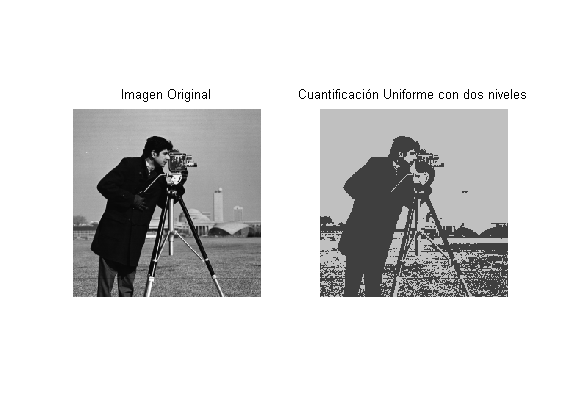
qerror=sum(error(:))/numel(error);

fprintf('Error cuadrático medio de cuantificación %e\n',qerror);

La salida es

Error cuadrático medio de cuantificación 1.580957e+03

y las imágenes original y cuantificadas son



Observa que cambiando los dos valores de valores\_cuantizados puede cambiar el error de cuantificación

**Paso 6**

Vamos ahora a aplicar el cuantificador de Max-Lloyd a esta imagen con dos intervalos de cuantificación

[particion, vcuantizada,qerror] = lloyds(dA(:),2);

[index,quants]=quantiz(dA(:),particion,vcuantizada);

qA=uint8(reshape(quants,size(A)));

subplot(1,2,1), imshow(A); title('Imagen Original')

subplot(1,2,2), imshow(qA); title('Cuantización Max-Lloyd con dos niveles')

formatspc='Partición= %4.2f; V. Cuantificadas =%4.2f, %4.2f; Error =%4.2f.\n';

fprintf(formatspc,particion,vcuantizada(1),vcuantizada(2),qerror)

C:\Users\Rms\Documents\rms\cursos\CRIM\curso 2014-2015\08 Cuantización Escalar\Guiones\untitled.tif

Partición= 88.54; V. Cuanticadas =23.73, 153.35; Error =597.31.

Observa que el error de cuantificación es mucho menor

Entiende muy bien que contienen las variables particion, vcuantizadas. Observa que en partición hay un elemento menos que en las variables cuantificadas.

**Paso 7**

En este guion estamos viendo hasta ahora unos ejercicios muy sencillos de cuantificación. Normalmente la cuantificación no va a ser aplicada a la imagen o señal original, se aplicará a predicciones o a datos transformados. Las predicciones y transformación de datos las veremos en temas siguientes. Vamos ahora a practicar con diferentes cuantificadores.

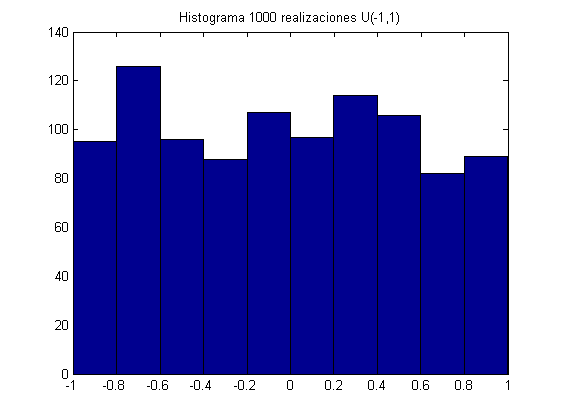
Generamos 1000 datos de una distribución U(-1,1) y mostramos su histograma. Para que todos obtengamos las mismas realizaciones usaremos la función rng('default')

close all; clear all;

rng('default');

X=2\*rand(1000,1)-1; %entiende este paso

hist(X); title('Histograma 1000 realizaciones U(-1,1)')



**Paso 8**

Antes de analizar el histograma, como curiosidad: como sabes la variancia de una distribución U(a,b) es (b-a)^2/12 que vale (1-(-1))^2/12=1/3 en nuestro caso. Si calculamos la varianza muestral tenemos

var=sum((X-mean(X)).\*(X-mean(X)))/numel(X)

que en el ejemplo produce 0.3206 que no está mal, lo cual es normal porque tenemos muchos datos.

**Paso 9**

Volvamos al histograma. Éste nos indica que nuestras observaciones son candidatas a ser cuantificadas utilizando un cuantificador uniforme. Escribe en el paso 9 de Practica05ApellidoNombre.m código Matlab para:

1. Construir un cuantificador uniforme con 2**n** valores cuantificados con n=1,2,3,4,5,6,7,8.
2. Calcular en función de n, el error cuadrático medio de cuantificación sobre los datos en X y también su error cuadrático medio de cuantificación teórico.
3. Dibujar, en función de n, las dos curvas de errores obtenidas. Dos subfiguras distintas en una misma figura.

La función linspace de Matlab, con valor inicial -1 y final 1, puede ser útil para este ejercicio.

**Paso 10**

En el paso 10 de Practica05ApellidoNombre.pdf:

1. Completa la tabla adjunta que contiene el error cuadrático medio calculado y el teórico. Utiliza 5 decimales
2. Incluye las dos curvas obtenidas correspondientes al error cuadrático medio calculado y al teórico.
3. Discute y explica además qué ocurre cuando aumentamos n en el paso anterior.
4. Por último, indica cómo deberíamos codificar los índices de cuantificación: Huffman, aritmética o longitud fija.

**Solución**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático medio | 0.07998 | 0.02083 | 0.00502 | 0.00129 | 0.00031 | 0.00008 | 0.00002 | 0.00000 |
| Error cuadrático medio teórico | 0.08333 | 0.02083 | 0.00520 | 0.00130 | 0.00032 | 0.00008 | 0.00002 | 0.00000 |

**Interfaz de usuario gráfica, Gráfico, Gráfico de líneas

Descripción generada automáticamente**

Conforme aumentamos n la cantidad de bloques en que distribuímos los datos aumenta, y, por tanto, las etiquetas asignadas a dichos bloques se acercan a los valores reales de nuestros datos. Si tuvieramos infinitos bloques, habría un bloque para cada número cuya etiqueta sería exactamente la que corresponde con dicho número y el error sería 0.

Los índices de cuantización formarán un alfabeto de 2^n elementos en nuestro caso con probabilidades muy cercanas a 1/(2^n) ya que la distribución es uniforme y sabemos que usando Huffman con probabilidades de 1/(2^n) el resultado es óptimo. Por tanto yo optaría por codificar usando Huffman que debería dar un resultado cercano al óptimo para este ejercicio.

**Paso 11**

Generamos 2000 realizaciones de una distribución de Laplace de media cero y varianza 1/30 Observa cómo generamos las realizaciones, ejecuta la orden help laprnd. Observa también que esta distribución tiene mucho menos varianza que una U(-1,1), ¿lo entiendes?. Por simplicidad nos quedaremos sólo con las observaciones que están en el intervalo (-1,1), todas. Dibujaremos el histograma de las realizaciones. Mira la diferencia con el histograma de las realizaciones de la distribución uniforme que obtuvimos con anterioridad.

close all; clear all; rng('default');

X=laprnd(1000,1,0,sqrt(1/30));

X=X(X>=-1 & X<=1);

hist(X); title('Histograma 2000 realizaciones Laplace(0,1/30)')

****

**Paso 12**

Como puedes observar el histograma no es muy uniforme. Esto nos indica que nuestras observaciones no son candidatas a ser cuantificadas utilizando un cuantificador uniforme. No obstante, escribe en el paso 12 de Practica05ApellidoNombre.m código Matlab para:

1. construir un cuantificador uniforme con 2**n** , n=1,2,3,4,5,6,7,8, niveles de cuantificación en el intervalo (-1,1) y aplicarselo a estos datos en X,
2. calcular, en función de n, el error cuadrático medio de cuantificación sobre los datos en X,
3. dibujar, en función de n, la curva de errores obtenida.

**Paso 13**

En el paso 13 de Practica05ApellidoNombre.pdf

1. Completa la tabla de errores cuadráticos medios. Utiliza cinco decimales
2. Incluye la curva de errores cuadráticos medios obtenidos
3. Por último indica cómo deberíamos codificar los índices de cuantificación: Huffman, aritmética o longitud fija.

**Respuesta**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático  medio | 0.15705 | 0.02854 | 0.00581 | 0.00134 | 0.00033 | 0.00008 | 0.00001 | 0.00000 |

**Gráfico, Gráfico de líneas

Descripción generada automáticamente**

En este caso los intervalos tendrán probabilidades muy dispares, por tanto una codificación que aproveche bien las probabilidades dispares y no utilice demasiada cabecera (Huffman no es una opción aquí ya que al no ser las probabilidades potencias negativas de dos, el tamaño de la cabecera lo hace peor que sus competidores), sería la codificación aritmética.

**Paso 14**

Discute y compara las tablas que has incluido en los pasos 10 y 13.

**Respuesta**

Tabla paso 10

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático medio | 0.07998 | 0.02083 | 0.00502 | 0.00129 | 0.00031 | 0.00008 | 0.00002 | 0.00000 |
| Error cuadrático medio teórico | 0.08333 | 0.02083 | 0.00520 | 0.00130 | 0.00032 | 0.00008 | 0.00002 | 0.00000 |

Tabla paso 13

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático  medio | 0.15705 | 0.02854 | 0.00581 | 0.00134 | 0.00033 | 0.00008 | 0.00001 | 0.00000 |

El error para pocos intervalos es muy superior en el caso de la distribución de Laplace como era de esperar, ya que dentro de cada intervalo las grandes poblaciones de datos se agrupan en los extremos derechos (en los intervalos de [-1,0)) o izquierdos (en los intervalos de (0,1]) lo que hace que el error sea mayor. Al aumentar mucho la cantidad de intervalos vemos como ese error desaparece, ¿pero cual es el precio a cambio?

Pongamos que N mayor implica mayor resolución como si fuera un monitor. En el caso de Laplace, para valores altos de N, la cantidad de segmentos es suficiente para suplir su falta de uniformidad, sin embargo estamos utilizando una resolución exageradamente alta para los valores extremos, que sabemos son mucho menos numerosos.

Por eso lo más idóneo sería utilizar cuantificaciones no-uniformes para el caso de distribuciones como la de Laplace.

También cabe destacar como el error real en la función uniforme supera las expectativas teóricas, mientras que la función de Laplace está muy por debajo.

**Paso 15**

Escribe en el paso 15 de Practica05ApellidoNombre.m código de Matlab para:

1. Construir a partir de los datos observados en X un cuantificador de Max-Lloyd con: 2n niveles de cuantificación usando la función lloyd de Matlab para n=1,2,3,4,5,6,7,8,
2. Calcular los errores de cuantificación
3. Dibujar la curva de errores en función de n.
4. Incluir en una figura con 8 subfiguras los límites de las particiones y los valores de cuantificación asignados a cada partición.

**Paso 16**

En el paso 16 de Practica05ApellidoNombre.pdf

1. Completa la tabla de errores cuadráticos medios. Utiliza 5 decimales.
2. Incluye la curva de errores cuadráticos medios obtenidos.
3. Incluye la figura con 8 subfiguras que dibuja los límites de las particiones con los valores de cuantificación asignados a cada partición.
4. ¿Qué conclusiones extraes de la figura?
5. Por último, indica cómo deberíamos codificar los índices de cuantificación: Huffman, aritmética o longitud fija.

**Respuesta**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático  medio | 0.01387 | 0.00507 | 0.00174 | 0.00058 | 0.00021 | 0.00005 | 0.00001 | 0.00000 |

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Siguiendo un poco lo que ya comentamos en ejercicios anteriores, y ampliando las figuras últimas que están un poco saturadas de contenido con 256 cuantizaciones, podemos concluir que conforme crece la cantidad de particiones que creamos, más se aproximan los valores contenidos a la etiqueta seleccionada, y en consecuencia el error disminuye drásticamente como podemos ver en la curva de error.

En este caso las probabilidades de cada partición no serán en la mayor parte de los casos potencias negativas de 2, por lo que descartamos Huffman como una posible solución. Sí que sabemos que las probabilidades deberían estar compensadas, ya que los intervalos donde más datos tenemos son más estrechos, mientras que los otros son mucho más anchos, al menos en el caso de valores bajos de n.

Por tanto diría que para valores bajos de n la codificación de longitud fija debería ser la mejor, transicionando a aritmética conforme aumente el valor de n.

**Paso 17**

Discute y compara las tablas que has incluido en los pasos 10 , 13 y 16.

**Respuesta**

Tabla paso 10

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático medio | 0.07998 | 0.02083 | 0.00502 | 0.00129 | 0.00031 | 0.00008 | 0.00002 | 0.00000 |
| Error cuadrático medio teórico | 0.08333 | 0.02083 | 0.00520 | 0.00130 | 0.00032 | 0.00008 | 0.00002 | 0.00000 |

Tabla paso 13

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático  medio | 0.15705 | 0.02854 | 0.00581 | 0.00134 | 0.00033 | 0.00008 | 0.00001 | 0.00000 |

Tabla paso 16

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N=1 | N=2 | N=3 | N=4 | N=5 | N=6 | N=7 | N=8 |
| Error cuadrático  medio | 0.01387 | 0.00507 | 0.00174 | 0.00058 | 0.00021 | 0.00005 | 0.00001 | 0.00000 |

Antes de comentar los resultados un pequeño recordatorio de los contenidos en cada tabla:

Tabla 10 – cuantificador uniforme / distribución uniforme

Tabla 13 – cuantificador uniforme / distribución Laplace

Tabla 16 – cuantificador Max-Lloyd / distribución Laplace

Gran parte de los resultados están comentados con anterioridad así que nos centraremos en los contenidos añadidos de comparar las 3 gráficas simultáneamente. Podemos observar como los peores resultados se obtienen al utilizar cuantificadores que no son adecuados para el conjunto de datos de que disponemos. (Veremos como esos datos pueden procesarse para que sigan una distribución que nos interese para aplicar determinados cuantificadores). Y dentro de los dos que funcionan bien (Tabla10 y Tabla16) vemos como el cuantificador Max-Lloyd tiene considerablemente más éxito principalmente para valores inferiores de N.

Esto se debe a la alta varianza de la distribución uniforme en comparación. Al estar los datos uniformemente distribuídos, dar un valor que agrupe todos los de una clase es una tarea sumamente costosa que requiere valores altos de N, mientras que con la distribución de Laplace, al estar concentrados los datos entorno a una sola región, el cuantificador de Max-Lloyd nos permite realizar esas separaciones de forma menos costosa “ignorando” las regiones donde hay volúmenes de datos inferiores.

**Paso 18**

Por último, ¿Por qué crees que hemos utilizado la distribución de Laplace?

**Respuesta**

Hemos utilizado Laplace casi con certeza para dos objetivos:

1 – Mostrar la ineficiencia de usar un cuantificador uniforme en datos distribuídos de forma no-uniforme. Y la distribución de Laplace debido a su baja varianza es todo lo opuesto a lo que podría ser una distribución uniforme a la hora de agrupar valores según su histograma.

Resumen: propone una distribución con un histograma muy diferente al uniforme.

2 – Mostrar la fortaleza de los cuantificadores no-uniformes, que destacan especialmente cuando los volúmenes de datos están muy sesgados en torno a un mismo punto permitiendo tomar particiones muy pequeñas para las zonas con alta concentración de datos y particiones muy grandes para zonas con baja concentración de datos.

Resumen: los datos tienen una distribución que muestran precisamente para qué existen los cuantificadores no-uniformes.