

ESCUELA INTERNACIONAL DE FÍSICA MÉDICA PROCESAMIENTO DE IMÁGENES MÉDICAS

Mauro Valente^{†*}

† Instituto de Física E. Gaviola CONICET & Universidad Nacional de Córdoba; Argentina.
Particle Physics FLUKA Group LHC @ CERN, *Genève*, Switzerland.
Universidad de La Frontera, Chile.

July 21th - August 1st 2015



UNIVERSIDAD
NACIONAL
DE COLOMBIA

ESCUELA INTERNACIONAL FÍSICA MÉDICA

Universidad Nacional de Colombia

*E-mail contact: valente@famaf.unc.edu.ar

Prefacio

El presente trabajo *Procesamiento de Imágenes Médicas* es un compendio original de notas elaborado por Mauro Valente en 2015, para ser utilizado como material de estudio y referencia para el curso superior en el marco de la *ESCUELA INTERNACIONAL FÍSICA MÉDICA* organizada por la Universidad Nacional de Colombia entre los días 21 de Julio y 1 de Agosto de 2015.

El contenido de estas notas está dedicado a formalismos y metodologías para el área específica de Procesamiento de Imágenes en el ámbito de la física médica. Se pretende lograr un primer acercamiento a las temáticas de las diferentes áreas del procesamiento de imágenes en física médica. Este trabajo se presenta como una exposición de fundamentos teóricos básicos acompañando con ejemplos prácticos de aplicación destinados a ejercitar los contenidos formales en situaciones prácticas de interés.

El contenido incluye breves repasos sobre elementos básicos interacción de radiación ionizante con la materia, focalizado en los procesos físicos de interacción con material biológico que dan lugar a la formación de imágenes radiológicas.

El entero contenido del presente trabajo está bajo gerenciamiento por parte del autor y la institución organizadora, Universidad Nacional de Colombia, para disponer de carácter exclusivo de su uso y distribución.

PROGRAMA Y CONTENIDO

MÓDULO I: Introducción al procesamiento de imágenes

1. Qué es una imagen?
2. Información contenida en una imagen

MÓDULO II: Representación matemática y operaciones algebraicas

1. Almacenamiento y transferencia computacional
2. Visualización - *display*
3. Representación matricial y operaciones a nivel de pixel.
4. Escaleo de Intensidad
5. Ecuilización de Histograma
6. Resaltado de bordes

Índice

1.. Introducción al procesamiento de imágenes	6
1.1.. Qué es un imagen ?	6
1.2..Cuál es la información que contiene una imagen ?	8
2.. Representación matemática y operaciones algebraicas	11
2.1.. Almacenamiento y transferencia digital	11
2.2.. Visualización de imágenes	13
2.2.1.. Ejemplo: Imágenes Gray Scale	14
2.2.2.. Ejemplo: Imágenes color RGB	14
2.2.3.. Ejemplo: Suma y diferencia de Imágenes	16
2.3.. Contraste: Relación entre intensidad y visualización	16
2.4.. Histograma de una imagen	16
2.5.. Gradientes en una imagen: Detección de bordes	17
2.5.1.. Ejemplo: Detección de bordes	17

MÓDULO I

Introducción al procesamiento de imágenes

1.. Introducción al procesamiento de imágenes

El capítulo 1. presenta una muy breve descripción introductoria a las imágenes digitales. La exposición inicia definiendo el concepto de Imagen, en relación al ámbito analógico y digital. Luego, se explica resumidamente el significado del contenido de información dentro de una Imagen vista como un objeto de manipulación.

1.1.. Qué es un imagen ?

Antes de dar una respuesta formal a la pregunta “*qué es una imagen?*” conviene comenzar por introducir ciertos conceptos y nociones básicas que permitirán establecer el contexto en el cual habrá de considerarse esta cuestión, a modo de fijar algunas ideas claras sobre definiciones y posterior tratado (procesamiento) de imágenes digitales.

Conceptualmente, el ámbito en el que debe considerarse el tratamiento de imágenes pueden describirse en términos de tres diferentes componentes principales:

Adquisición de la imagen: La adquisición de la imagen se realiza por medio de sistemas de transducción, o serie de transductores, los que por medio de la manipulación de la luz (visible o no), *i.e.* u alguna otra forma de radiación, producen una representación del objeto de interés, dando así lugar a la imagen. Ejemplos básicos de este tipo de sistemas, son: ojo humano, sensores de cámara fotográficas, sensores de video, tomógrafos de rayos X, ultrasonido, resonancia magnética, telescopios, radares. Los sistemas de transducción, inevitablemente, implican el agregado de ruido a la imagen. Además, poseen una resolución limitada y por tanto se ve afectada la apreciación de dicha imagen

Procesamiento de la imagen: El procesamiento digital de la imagen es el método, o conjunto de métodos, orientados a eliminar la mayor cantidad posible de interferencias indeseables, como el ruido, que son agregados durante la adquisición de la imagen. También están orientados al mejoramiento de las características de la imagen como, tales como: brillo, contraste, definición de contornos, color, etc. por medio de herramientas matemáticas. El procesamiento es complementado por técnicas de codificación y almacenamiento, así como la transmisión.

Presentación al observador: Esta etapa representa la metodología empleada para exponer la imagen, tanto sea impresión (*hardcopy*) o por medios electrónicos. Deben considerarse ciertos aspectos de percepción humana al momento de diseñar la presentación de la imagen al observador así como las frecuencias de despliegue en el dispositivo.

Sobre las características de diseño de sistemas para la representación de señales:

- Los dispositivos sensoriales son capaces de realizar una cantidad limitada de mediciones de las señales de entrada (*input*) y estas mediciones deben ser suficientes y adecuadas para asegurar aproximaciones aceptables, ergo útiles, del objeto de estudio. La determinación *a priori* sobre cuáles mediciones realizar y, a su vez, cómo deben ser utilizados los *inputs* para que aproximen de la mejor manera a la señales “verdaderas” es **el problema** que debe ser resuelto.
- La selección del tipo de procesamiento, eventualmente en conjunto con la codificación, para una dada señal, requiere una precisa interpretación de las componentes de la señal, es decir un conocimiento acabado de los procesos físicos por los cuales las imágenes son adquiridas y representadas.

- El sistema de visión humano puede ser utilizado en ciertas etapas de procesamiento para dicha interpretación, mientras que los dispositivos de despliegue (*display*) sintetizan la información de una imagen limitando a una cantidad finita de respuestas básicas de despliegue, tal como sucede por ejemplo en los puntos de fósforo de un tubo de rayos catódicos (monitor). La determinación de tamaño y forma de las respuestas de despliegue, su configuración (intensidad o amplitud y posición espacial) pueden controlarse de procurando optimizar la calidad o fidelidad, de las imágenes.

El funcionamiento del sistema visual humano (*Human Visual System, HVS*) representa un punto de partida canónico para describir conceptos generales sobre los sistemas de adquisición de imágenes. La luz visible es radiación electromagnética cuya longitud se ubica dentro del espectro electromagnético, la retina del ojo humano percibe normalmente longitudes de onda de visión humana entre 390 nm y 780 nm, aproximadamente.

Las tres dimensiones básicas de la luz son:

1. **Intensidad** (brillantez o amplitud) asociada a la percepción visual del flujo luminoso (fluencia de fotones) que despierta o refleja el objeto de interés.
2. **Frecuencia** (o su inversa, longitud de onda) percibida por el ojo humano como el color de la luz.
3. **Polarización** (o ángulo-fase de oscilación), el cual no es perceptible por los humanos bajo circunstancias ordinarias, pero puede ponerse en evidencia utilizando elementos ópticos muy sencillos (polarímetros).

Existe un grupo de particular relevancia dentro de los sistemas de visión denominado “*Sistema lineal e invariante en el tiempo (SLI)*”, que cuentan con las dos siguientes características:

Función de respuesta al impulso: Es la función que describe el comportamiento evolutivo temporalmente de un sistema. En el caso de la visión humana este sistema es el funcionamiento orgánico de la vista. Una vez obtenida la respuesta al impulso, se realiza la convolución de la función obtenida con otra función a fin de observar y poder así conocer la respuesta del sistema a esa nueva función.

Función de modulación de transferencia: Es una función que describe al sistema de visión en el dominio de la frecuencia, *i.e.* es una transformación (de Fourier en este caso) de la respuesta a impulso. El ojo puede modelarse como un sistema que modifica funciones dependiente de tres variables, una temporal y dos espaciales. La función respuesta también puede ser representada como una función de tres variables.

A partir del análisis del sistema visión humano, pudiendo omitir las variables espaciales y estudiando sólo la evaluación temporal, se encuentra que el ojo es incapaz de distinguir una sucesión rápida de imágenes. Y, al estudiar su respuesta en frecuencia, se determina que su comportamiento corresponde a un “filtro de pasa-bajo”, con frecuencia de corte alrededor de 30 imágenes por segundo. Este fenómeno puede apreciarse en la observación de televisores antiguos que utilizan frecuencias de barrido verticales menores a la rapidez necesaria para que el ojo humano logre ver una imagen continua. Este efecto se denomina *flicker*. Por supuesto, existe la sensibilidad del sistema de adquisición (visión) y por tanto la respuesta-frecuencia del ojo puede variar, según la intensidad de la luz. Se conoce que ante imágenes poco brillantes la frecuencia de corte es menor, mientras que para imágenes altamente brillantes la frecuencia de corte aumenta.

La dependencia según las variables espaciales del comportamiento del ojo también ha sido caracterizada, mejorando la comprensión de *por qué y cómo vemos*. Por ejemplo, en las pantallas de cristal líquido (*Liquid Crystal Display - LCD*), los cristales de un mismo color se colocaban en líneas verticales, dando lugar a resolución “desagradable” a la vista del observador. Este hecho

motivó investigaciones para buscar soluciones para mejorar la calidad de la imagen, identificando dos posibilidades: disminuir el tamaño de los cristales de tal forma que la resolución aumentara, o cambiar la disposición de los cristales de forma que en lugar de formar líneas verticales formaran líneas diagonales. En cualquiera de estas posibilidades la apariencia de la imagen mejora significativamente a la vista del observador humano. Tecnológicamente, se implementó la segunda opción porque no implica un aumento en los costos de producción.

Un experimento clásico empleado para determinar la habilidad del sistema de visión humano, en cuanto a la capacidad para discriminar la luminosidad consiste en la observación de un plano uniformemente iluminado en una región extensa, suficiente para ocupar el campo entero de la visión. Esta área es usada como difusor (como vidrio opaco, por ejemplo) que es iluminado por debajo por una fuente de luz cuya intensidad I , puede ser variada. Luego, se superpone a este campo un incremento de iluminación ΔI en la forma de un *flash* de corta duración que se presenta como un círculo concéntrico al centro del campo iluminado uniformemente, como lo muestra la figura 1. Para valores bajos de ΔI , *i.e.* no suficientemente brillante, el observador humano no consigue identificar la presencia del *flash* de luz circular, ya que no resulta perceptible. Contrariamente para valores de ΔI suficientes para poner de manifiesto un aumento apreciable de la intensidad durante el disparo del *flash*, resulta que el observador humano puede dar cuenta de la presencia de éste. La cantidad correspondiente a un incremento de iluminación discriminable de 50% del tiempo con iluminación de fondo I , es denominada *razón o promedio de Weber*. Valor pequeños de esta cantidad implican cambio pequeños en los porcentajes discriminables, determinando así una “buena” discriminación de luminosidad. Por tanto, de modo opuesto, valores grandes de esta cantidad hacen referencia a que significan que un cambio grande de porcentaje en intensidad es requerido para lograr la percepción. Esto representa “baja” discriminación de luminosidad.

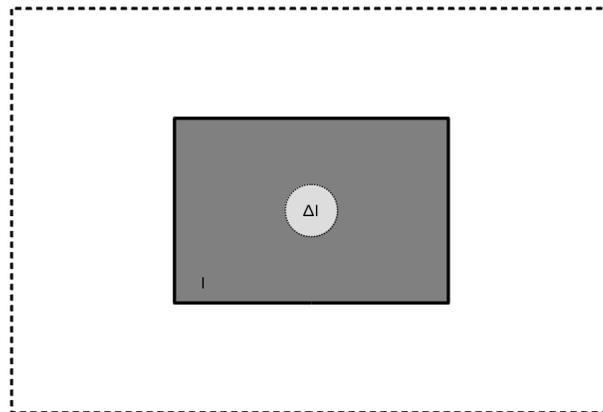


Figura 1. Esquema de un montaje experimental básico para discriminación de luminosidad: El recuadro gris oscuro, que representa una región que cubre todo el campo visual del observador, es iluminado con intensidad I . La región circular, en gris claro, es iluminada con intensidad ΔI en forma de destellos tipo *flash*.

1.2.. Cuál es la información que contiene una imagen ?

La utilidad del elemento “imagen” como objeto de manipulación de información es es muy amplia y abarca muchos campos. Ejemplos, entre la infinidad de los mismos, pueden ser imágenes obtenidas con fines de diagnóstico médico, satelitales, de radar, meteorológicas, artísticas, etc. La información contenida refiere, en términos generales, a distribuciones témporo-espaciales de observables (magnitudes) físicas cuando se trata de imágenes “reales”, o bien simplemente representaciones matemáticas de arreglos numéricos, en caso de tratarse de imágenes “abstractas”.

Básicamente, existen dos tipos de imágenes digitales:

Mapa de Bits: Se trata de imágenes generadas a partir de un arreglo (rejilla/cuadrícula) almacenada en un archivo. Cada uno de los “cuadritos de la grilla” se denomina *pixel* y es el elemento utilizado para guardar información de los valores de la función o magnitud. Por su parte, la grilla refiere a la distribución espacial así como el universo de la malla y resolución de muestreo.

Imágenes vectoriales: Son objetos geométricos independientes (segmentos, polígonos, arcos, etc.), cada uno de ellos definido por distintas propiedades matemáticas de morfología, posicionamiento, intensidad, etc. Por ejemplo, un círculo de rojo quedaría definido por la posición de su centro, el radio, el grosor de línea y el color.

La categorización recientemente introducida resulta canónica, del momento que se analiza según las características dimensionales de los objetos que serán “colocados/almacenados” en las posiciones determinadas por la grilla. El objeto más sencillo es un escalar (x), luego le siguen vectores (\vec{x}), pasando a matrices ($X(m, n)$); y en sentido general, tensores $\mathbf{X}_{\alpha_1, \dots, \alpha_m}^{\beta_1, \dots, \beta_n}$.

A modo de ejemplificar, la modalidad más básica de información contenida en una imagen es el caso en que a cada *pixel* (i, j) del arreglo de la grilla (en el contexto de imágenes 2D) le corresponde un único valor numérico, es decir un escalar (o cardinal 1 del conjunto). Este es el caso de fotografías llamadas “blanco y negro” (en términos formales se trata de “escala de grises”), ya que la información contenida en cada posición dentro de la grilla de la imagen es un tono/tonalidad de gris dentro de la escala que recorre del blanco (prefijado como el máximo de iluminación posible) al negro (considerado el nivel de ausencia absoluta de intensidad de luz); en el contexto de imágenes fotográficas¹. Puede demostrarse que cualquier magnitud física “monocanal”, *i.e.* monovaluada puede ser representada en su distribución témporo-espacial por medio de imágenes con valores escalares en los *pixels*.

Pasando ahora al caso de imágenes a color, en este caso ya no es suficiente la información de un único valor para determinar unívocamente un color, ya que la representación requiere de la “mezcla”, formalmente combinación (lineal) de diferentes “colores base”, que en términos formales representan la *Eigen*- base, mientras que los elementos del vector son los pesos estadísticos de la combinación (lineal). En el caso de imágenes fotográficas a color, lo convencional es el uso de tres canales para la multibanda, típicamente rojo, verde y azul; lo que recibe la denominación *RGB - Red Green Blue*. La utilización de información multibanda se extiende a una amplísima y muy variada diversidad de situaciones, desde imágenes de captura de radar, en diferentes frecuencias (para mejorar y complementar la información radiométrica, respecto del caso monobanda) o cualquier magnitud física que requiera representación vectorial. En este contexto, a cada posición espacial (i, j) le corresponde un vector; que en el ejemplo de la multibanda RGB es la combinación de los pesos estadísticos de “la proporción de cada color base” necesaria para la mezcla que da lugar al color de interés. Por lo tanto, es evidente que el color rojo se representa por $(1, 0, 0)$, el verde $(0, 1, 0)$ y el azul $(0, 0, 1)$. Cualquier color será representado aprovechando la propiedad matemática de que los vectores anteriores constituyen una *Eigen*- base, es decir ortonormal² y completa, que permite expresar cualquier vector arbitrario (a, b, c) en términos de éstos, *i.e.* $(a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)$.

¹No así para otro tipo de imágenes, como radiografías por ejemplo, en las que las tonalidades se refieren a propiedades del material y no a intensidad lumínica detectada.

²Es decir, ortogonal y normalizada.

MÓDULO II

Representación matemática y operaciones algebraicas

2.. Representación matemática y operaciones algebraicas

El capítulo 2. introduce muy brevemente los rasgos fundamentales sobre la representación matemática de las imágenes. Se describen características básicas de algunas propiedades como histograma, contraste, base de representación; entre otros. Se provee de ejemplos sencillos acompañando la exposición.

2.1.. Almacenamiento y transferencia digital

Los sistemas de computación fueron originalmente desarrollados para realizar cálculos matemáticos de manera rápida. Desde entonces, se los utiliza en una amplia, y creciente, variedad de aplicaciones, que incluye presentación, visualización y almacenamiento de información. Existen dispositivos, además, que permiten optimizar los procesos de almacenamiento y transferencia, por medio de redes y sistemas de comunicación, fijos y portátiles.

La utilización de sistemas computacionales en diagnóstico médico (*medical imaging*) comenzó a principios de la década de 1970. Y, desde entonces, cobró cada vez más importancia. Los primeros usos fueron en el campo de medicina nuclear con la finalidad de obtener imágenes de la cinética de radiofármacos en diferentes órganos. Actualmente, los sistemas computacionales son vitales en *medical imaging*, incluyendo tomografía computada, resonancia magnética, y métodos de *imaging* funcional; sólo por mencionar algunos.

Es bien conocido que el sistema de numeración típicamente utilizado está basado en la cantidad de dedos. Los sistemas de cómputo, contrariamente, disponen de dos opciones³ y por tanto se utiliza un sistema binario para el almacenamiento y manipulación de números, información en definitiva.

A continuación se presenta un ejemplo, a modo de brevísimo refresco sobre bases numéricas:

El número 4862 en base decimal representa:

$$4 \times 10^3 + 8 \times 10^2 + 6 \times 10^1 + 2 \times 10^0 = 4000 + 800 + 60 + 2$$

Mientras que el número 1101 en base binaria es:

$$1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 8 + 4 + 0 + 1$$

La memoria y almacenamiento en computadores se basan en objetos llamados “bits”, en referencia a *Binary digITS*, que equivale a un *switch*, puede estar encendido (*ON*) o apagado (*OFF*). En algunas arquitecturas tecnológicas cada *bit* es un transistor en *ON* u *OFF*. En dispositivos magnéticos, cada *bit* es una porción de cinta o disco, que puede ser magnetizado.

Los *bits* se agrupan en *bytes*, cada uno consiste en 8 *bits*. La capacidad de los computadores en cuanto a memoria o la capacidad de almacenamiento de un dispositivo se describe típicamente en término de *bytes* o múltiplos como kb, Mb, Gb, Tb, etc. Los *bits* también se agrupan para formar *words*, “palabras”. La cantidad de *bits* en una *word* depende del sistema computacional, sea 32-b o 64-b, por ejemplo.

Una imagen digital es un arreglo rectangular, típicamente cuadrado, de números u objetos. La unidad de porción de imagen, *pixel* en referencia a *picture element*, contiene información sobre la posición de éste en el arreglo (grilla) así como el objeto allí almacenado, valor numérico.

El efecto de las propiedades de *pixel* se observa en la figura 2, donde se aprecia la estrecha y directa relación entre el pixelado y la resolución espacial del muestreo.

³En definitiva, se trata de si “pasa” o “no pasa” corriente. En el caso de métodos modernos, como *high performance quantum computing* puede ser el estado de *spin*.

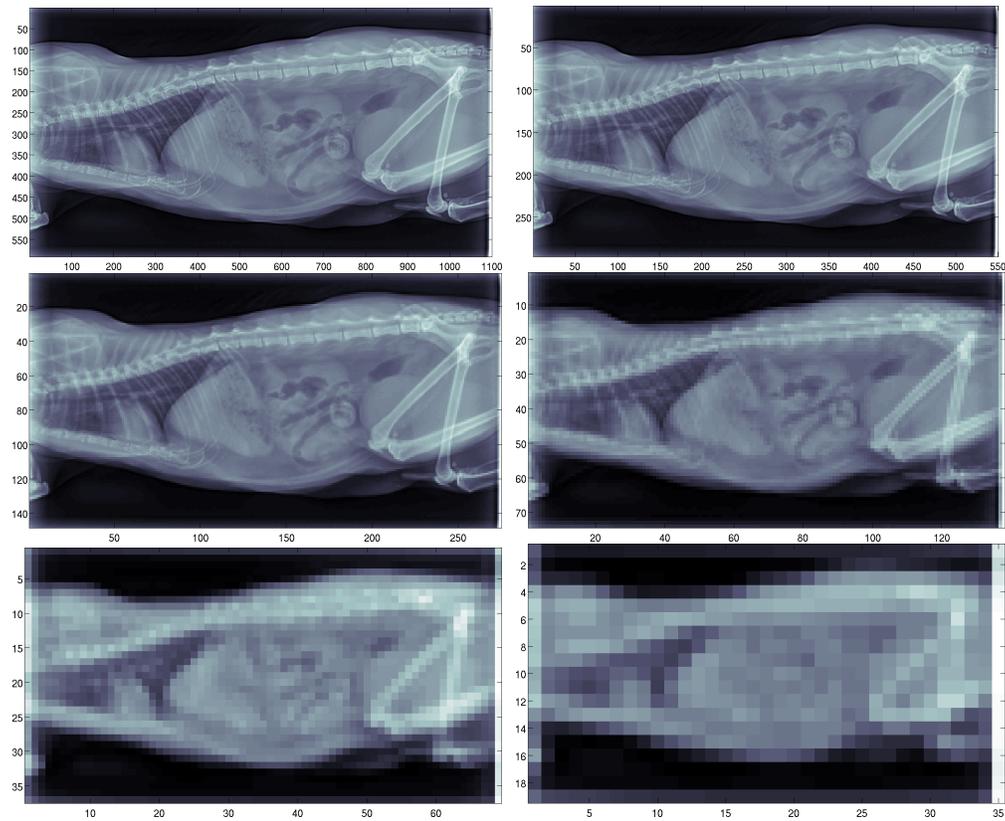


Figura 2. Efecto del pixelado, resolución espacial: de arriba hacia abajo y de izquierda a derecha se reporta la radiografía por contraste de absorción de rayos X blandos de un animal pequeño, variando la resolución espacial (aumentando el tamaño físico) del *pixel* en potencias de 2. Imagen de micro-tomografía por rayos X (μ) obtenida en la *facility* de Laboratorio de Investigación e Instrumentación en Física Aplicada a la Medicina e Imágenes por Rayos X (LIIFAMIR^x).

Cuadro 1. Números en base decimal y binaria.

DECIMAL	BINARIO
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111
16	10000

2.2.. Visualización de imágenes

Como se mencionó, el *bitmap* es un modo elemental para representar imágenes digitales como información en el *hardware*, específicamente la memoria, de un computador. Consiste, básicamente, en formar arreglos de elementos (vectores, matrices, tensores) ordenados de modos específicos. En general, para el caso típico de imágenes 2D, se realiza un ordenamiento por filas de elementos de matriz (*pixels*) asignando a cada uno un valor que determina "el color" en esa posición de la imagen.

Ya se mencionó que en el caso de imágenes en tonalidades de grises, el valor del elemento de matriz es un escalar; mientras que para el caso de imágenes a color el valor de cada elemento de matriz es un vector de tres coordenadas, cada una de las cuales especifica "el grado de influencia" de los colores rojo (*Red "R"*), verde (*Green "G"*) y azul (*Blue "B"*), de modo que se denomina representación RGB). Existen otros modos de representación a color, como por ejemplo CMYK (cián, magenta, amarillo y negro).

Típicamente se emplean escalas (que determinan "rangos dinámicos") en 2^N bits, y se denomina N -bits. Es decir, para el caso más común de 8-bits, la escala es $[0, 255]$, ya que por costumbre se define el rango como $[0, 2^{N-1}]$.

El uso típico de 8-bits está basado, principalmente, en dos motivos. En primer lugar, estudios biométricos muestran que el ojo humano no es suficientemente sensible para diferenciar más de 256 niveles de intensidad para un dado color. Además, el rango de valores para los elementos de matriz determinan la necesidad en cuanto a la capacidad de almacenamiento en el computador.

Entonces, para imágenes en tonalidades de grises, conocidas como "de una banda" el rango para los valores de los elementos de matriz (escalares) es $[0, 255]$, mientras que para imágenes a color, los valores de elementos de matriz (vectores de 3 coordenadas) asumen valores en $([0, 255], [0, 255], [0, 255])$. Si embargo, también es frecuente encontrar representaciones norma--

lizadas ⁴ para imágenes a color, es decir, elementos de matriz en $([0, 1], [0, 1], [0, 1])$ para determinar los colores RGB.

Todos los colores en el rango visible pueden representarse como combinaciones RGB, variando desde el negro $(0, 0, 0)$ al blanco $(255, 255, 255)$. Por lo tanto, una imagen RGB es representada por un arreglo bidimensional de *pixels*, cada uno codificado en 3 *bytes* pudiendo asumir 256^3 diferentes valores de combinaciones vectoriales, es decir 16.8 millones de diferentes colores, aproximadamente.

Es posible cuantificar la diferencia entre dos colores (en representación digital, valores del trio vectorial RGB) calculando la distancia, según algún tipo de métrica, Euclídea por ejemplo, entre los vectores que los representan. Si el color C_1 representado por el vector (R_1, G_1, B_1) y el color C_2 representado por (R_2, G_2, B_2) . Entonces, en el espacio vectorial, la distancia $D(C_1, C_2)$ entre éstos está dada por:

$$D(C_1, C_2) = \sqrt{(R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2} \quad (1)$$

Para el caso particular de imágenes de una banda (tonalidades de grises) puede aplicarse la misma metodología descrita para imágenes RGB con la simplificación asociada al hecho de que en el espacio de colores, los vectores en la dirección del vector $(1, 1, 1)$ representan las diferentes tonalidades de gris.

Por tanto, existe la equivalencia de que para cualquier *pixel* de tipo RGB si se lo proyecta sobre $(1, 1, 1)$ se obtiene la contribución de cada tonalidad de gris. Entonces, se tiene:

$$\text{Proy} \equiv (R, G, B) \cdot (1, 1, 1) = R + G + B = |\vec{V}| |\hat{n}| \cos(\phi) \quad (2)$$

donde *Proy* es la proyección, \vec{V} es el vector que forma el punto (R, G, B) en el espacio de coordenadas del trío (representación vectorial), \hat{n} es el vector de proyección $(1, 1, 1)$ y ϕ es el ángulo que forma \vec{V} con \hat{n} . De aquí puede verse que $\text{Proy} = \frac{R+G+B}{\sqrt{3}}$ y debe atenderse de que este valor no exceda 255, de modo que es usual renormalizar para obtener $\text{Proy} = \frac{R+G+B}{3}$.

2.2.1.. Ejemplo: Imágenes Gray Scale

A continuación se realiza una simple demostración, a modo de ejemplo sobre cómo generar imágenes en escala de grises.

Se define el tamaño del arreglo (matriz) que dará lugar a la imagen, es decir cuya visualización gráfica corresponda a la imagen de interés. Se determina la base del *bitmap*, típicamente potencias de 2. Luego, debe asignarse valores a los *pixels* de este arreglo, es decir asignar valores a los elementos de matriz. Y por último, la visualización consiste en graficar colocando un tono de gris en la posición del *pixel* correspondiente.

2.2.2.. Ejemplo: Imágenes color RGB

A continuación se realiza una simple demostración, a modo de ejemplo sobre cómo generar imágenes en RGB.

Se define el tamaño del arreglo (matriz) que dará lugar a cada una de las tres (R, G, B) de la imagen. Se determina la base del *bitmap*, típicamente potencias de 2. Luego, debe asignarse valores a los *pixels* para cada uno de los tres arreglos, es decir asignar valores a los elementos de cada una de las tres matrices. Y por último, la visualización consiste en reunir las matrices en una única imagen, formando los vectores en la posición del *pixel* correspondiente.

⁴Al valor $2^N - 1$.

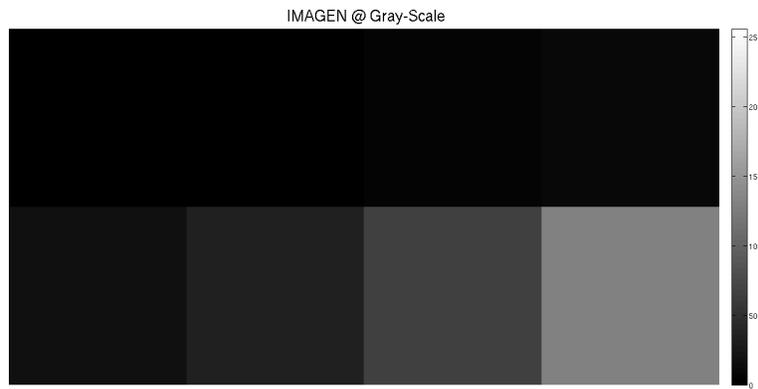


Figura 3. Imagen en tono de grises de dimensión 2×4 , en base 8-bits con elección arbitraria de las tonalidades de gris de cada *pixel*.



Figura 4. Imagen de tipo RGB de dimensión 5×10 , en base 8-bits con elección *random* de valores para bandas R, G y B.

2.2.3.. Ejemplo: Suma y diferencia de Imágenes

A modo de ejemplo de cómo aprovechar el álgebra matricial para la descripción de imágenes se propone realizar suma y diferencia entre dos imágenes arbitrarias. Para ello, se tiene en cuenta que coincidan las dimensiones de los objetos de modo que está definida la operación matemática. Luego, se procede a generar dos imágenes arbitrarias y realizar las operaciones suma y resta, de modo *pixel-a-pixel*, i.e.: $ImagenSuma(i, j) = Imagen1(i, j) + Imagen2(i, j)$, y también: $ImagenResta(i, j) = Imagen1(i, j) - Imagen2(i, j)$.

La figura 5 muestra dos imágenes con sus respectivas imágenes suma y resta.

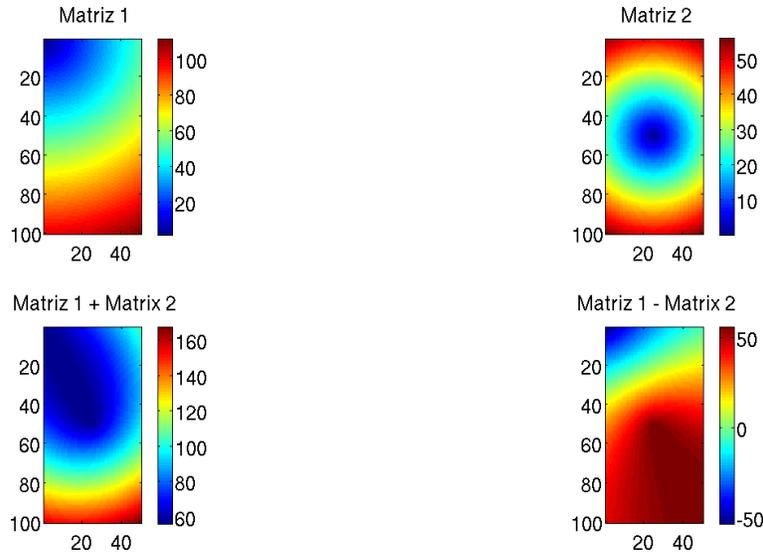


Figura 5. Imágenes arbitrarias (arriba izquierda y derecha) y correspondientes imágenes Suma (abajo izquierda) y diferencia (abajo derecha).

2.3.. Contraste: Relación entre intensidad y visualización

Conceptualmente, aumentar o disminuir el contraste en una imagen consiste, básicamente y visualmente, en aumentar o disminuir la pendiente de la línea recta con pendiente a 45 grados que representa los grises (con la precaución de no exceder los límites 0 – 255) entre *input* y *output*, como indica la figura 6.

La transformación correspondiente al cambio de contraste es:

$$output(m, n) = (input(m, n) - 2^{Y-1}) \tan(\phi) + 2^{Y-1} \quad (3)$$

donde Y es la escala en *bits* y el ángulo ϕ corresponde a las propiedades de la transformación lineal de contrastes, específicamente la pendiente (véase la figura).

2.4.. Histograma de una imagen

Dada la representación digital de una imagen por medio del arreglo de N filas por M columnas se determina una matriz $M \times N$, en la cual la representación digital de *bitmap* estará dada por la función distribución $f(m, n)$, para $n \in [0, N - 1]$ y $m \in [0, M - 1]$, típicamente N y M son potencias de 2, como ya se enunció.

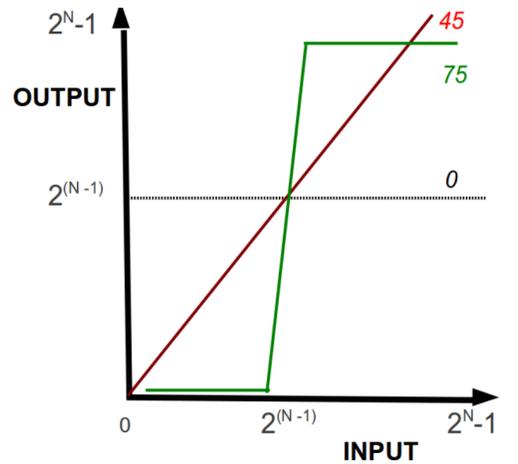


Figura 6. Representación del cambio de contraste entre *input* y *output*.

El histograma de una imagen $h(i)$, comúnmente denominado “*image enhancement*” o “*image characterization*” es un vector que da cuenta de la cantidad de *pixels* dentro de la imagen con un cierto valor de elemento. Es decir, para una imagen de α -bits, se tiene:

$$h(i) = \sum_{m=0}^{M+1} \sum_{N=0}^{N+1} \delta(f(n,m) - i) \quad \forall i \in [0, 2^{\alpha-1}] \quad (4)$$

Una de las técnicas genéricas, que luego se diversifica a una cantidad muy variada de metodologías específicas de procesamiento, es el método de convolución. Sea $w(k,l)$ un arreglo $2 \times K + 1, 2 \times L + 1$, centrado en el “origen” $(0,0)$ que coincide con el *pixel* central de la imagen. Puede considerarse a $w(k,l)$ como un *kernel* de convolución de modo que aplicado a la imagen $f(n,m)$ resulte:

$$g(m,n) = w(l,k) \otimes f(n,m) = \sum_{k=-K}^K \sum_{l=-L}^L w(k,l) \cdot f(m-k, n-l) \quad (5)$$

A partir de esta definición, pueden introducirse una gran cantidad de métodos específicos, entre los que se destacan las transformadas, como Fourier, Laplace, Radon, etc.

2.5.. Gradientes en una imagen: Detección de bordes

La detección de bordes, basada en métodos de aproximación numérica de los gradientes espaciales en las imágenes, es una de las herramientas más aprovechadas en el campo del *medical image processing*.

2.5.1.. Ejemplo: Detección de bordes

Si bien no se contempla un desarrollo detallado sobre el tema en esta exposición, traer a colación un ejemplo aporta a fijar las ideas respecto del potencial que representa aplicar álgebra matricial para el procesamiento de imágenes digitales.

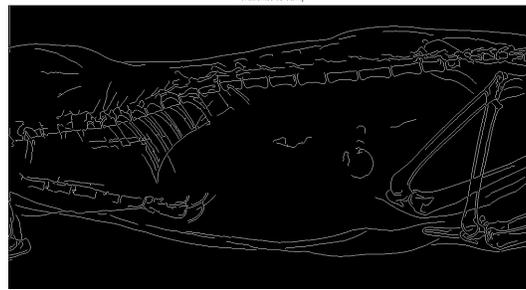


Figura 7. Detección de bordes (algoritmo Canny) en imagen radiológica. La performance depende de parámetros en el algoritmo de cómputo.