DRIVES

Development and Research on Innovative Vocational Education Skills

COMPUTER VISION FUNDAMENTALS

U2.E4. IMAGE PROCESSING AND MANIPULATION

Computer Vision Expert

May 2021, Version 1

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

The student is able to …

CVE.U2.E4.PC1	understand and distinguish the different types of image processing methods, namely, analogue and digital image processing.
CVE.U2.E4.PC2	comprehends what sampling, quantization, and image interpolation are.
CVE.U2.E4.PC3	understands the differences between image resize and remapping.
CVE.U2.E4.PC4	understands what aliasing is and is able to distinguish between spatial aliasing and anti-aliasing.
CVE.U2.E4.PC5	understands and defines spatial domain methods and frequency domain methods.
CVE.U2.E4.PC6	knows important concepts for image enhancement in spatial domain like point processing and mask processing
CVE.U2.E4.PC7	understands what contrast stretching is.
CVE.U2.E4.PC8	knows some commonly used transformation functions used in contrast enhancement.

LEARNING OBJECTIVES

The student is able to …

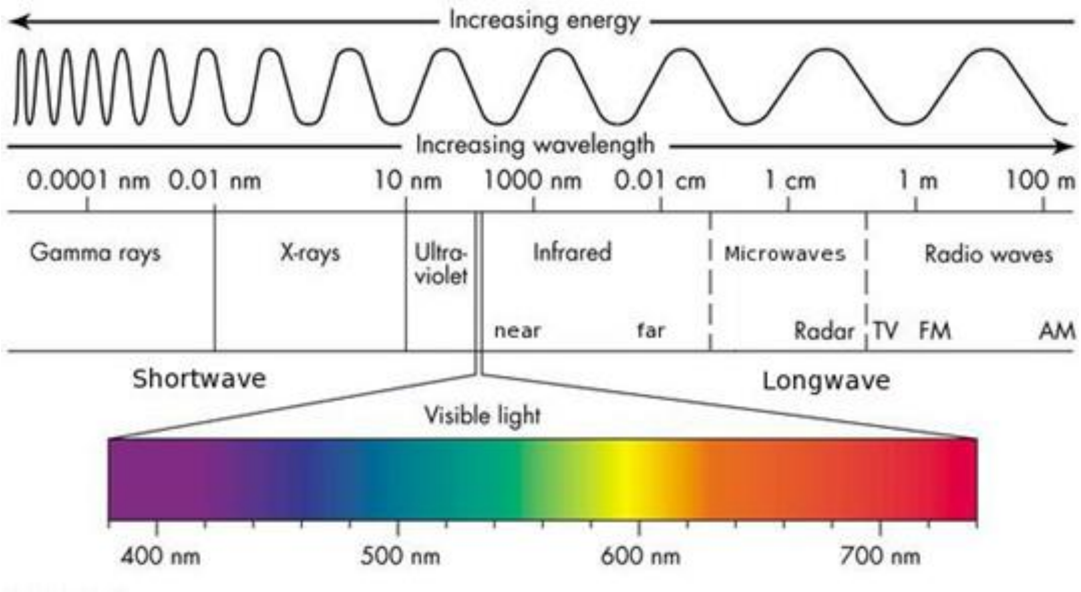
CVE.U2.E4.PC9	understands what histogram and histogram processing is and knows how histogram equalization works.
CVE.U2.E4.PC10	understands how Singular Value Decomposition (SVD) and Singular Value Equalization (SVE) works.
CVE.U2.E4.PC11	knows arithmetic and logic operations and understands their differences
CVE.U2.E4.PC12	understands image subtraction and image averaging process.
CVE.U2.E4.PC13	understands the differences between linear and nonlinear spatial filtering.
CVE.U2.E4.PC14	define and differentiate convolution and correlation.
CVE.U2.E4.PC15	understands the differences between linear and nonlinear smoothing filters as well as statistical filters.
CVE.U2.E4.PC16	to define and understand Laplacian, gradient and sobel operator.

An image is a two-dimensional function ***f(x,y)***, where *x* and *y* are the **spatial** coordinates, and the amplitude of ***f*** at any pair of coordinates (*x*,*y*) is called the intensity of the image at that level.

If ***x,y*** and the **amplitude** values of ***f*** are **finite** and **discrete quantities**, the image is called a **digital image**.

A digital image is composed of a finite number of elements called **pixels**, each of that has a particular location and value.

The principal source for the images is the electromagnetic (EM) energy spectrum.



An example of Imaging in all of the bands

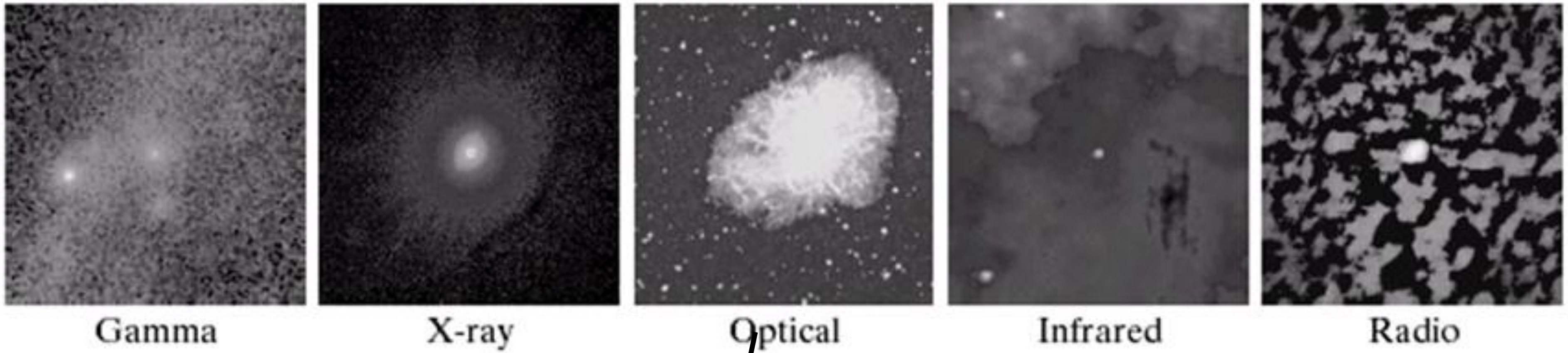
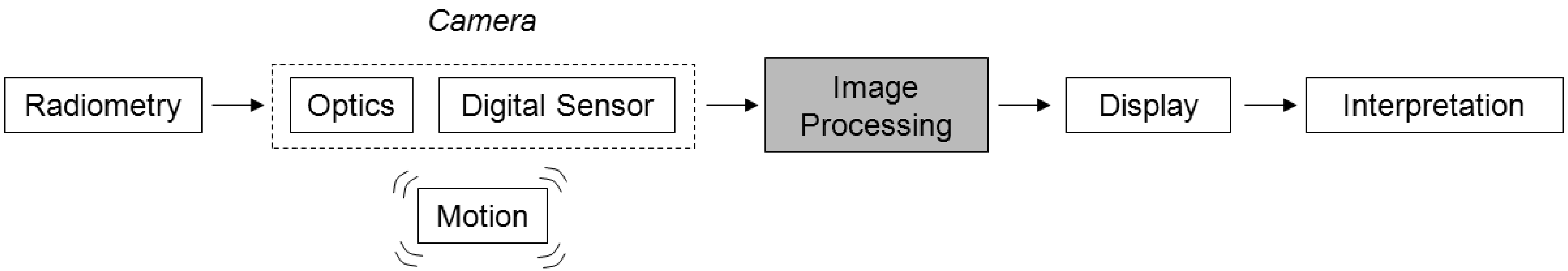


FIGURE 1.18 Images of the Crab Pulsar (in the center of images) covering the electromagnetic spectrum. (Courtesy of NASA.)

Visible light

The output of the digital sensor is a “raw” digital image. This output consists of an array of digital count values with each value representing the brightness, or gray level, of a pixel in the image. Image processing is usually employed in the imaging chain to improve the efficacy of the image data.



SAMPLING VS QUANTIZATION

A continuous model is appropriated for some situations but for others, it is more convenient to work with digital signals (signals which have a discrete domain and range). **Sampling** refers to the process of digitizing the domain and the process of digitizing the range is called **quantization.**

The **sampling** rate determines the spatial resolution of the digitized image, while the **quantization** level determines the number of grey levels in the digitized image.

Sampling is related to coordinates values.

Quantitization is related to intensity values.

SAMPLING

The sampling rate determines the spatial resolution of the digitized image.

1024

512

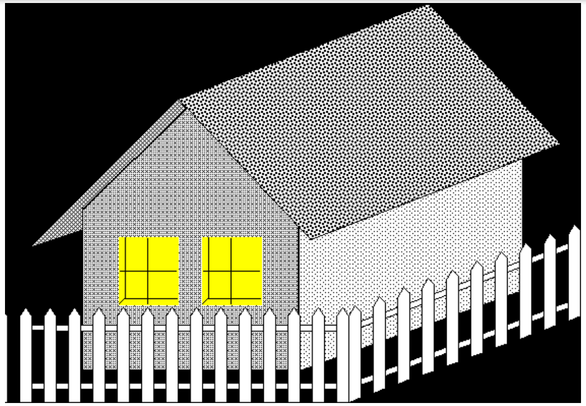
256

128

64

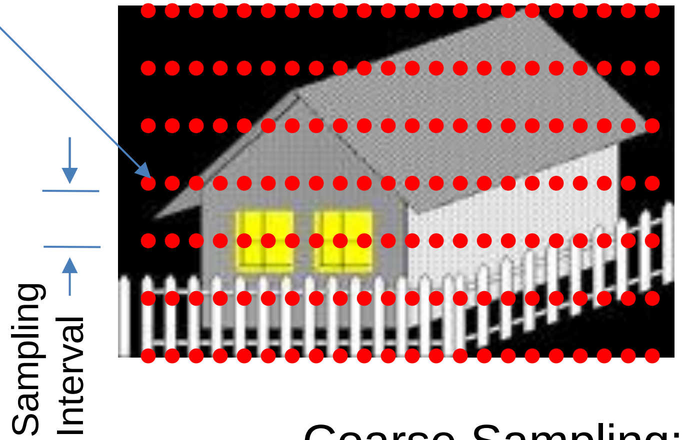
32

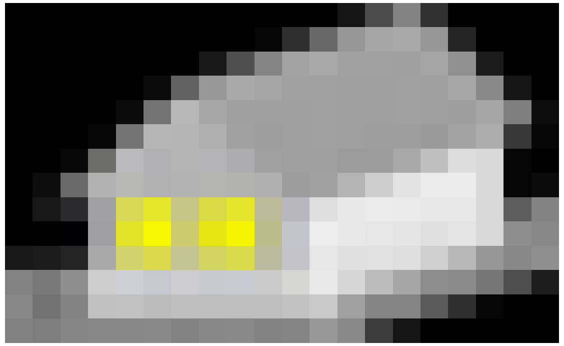
SAMPLING: SPATIAL RESOLUTION



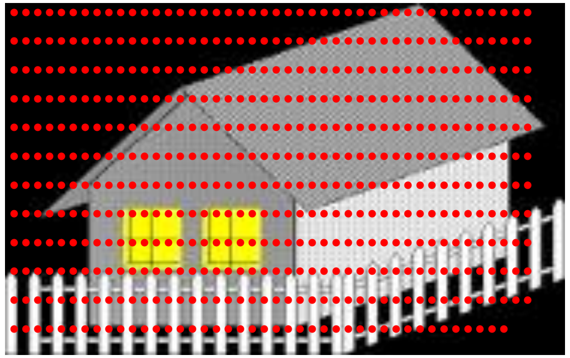
Digitize this image

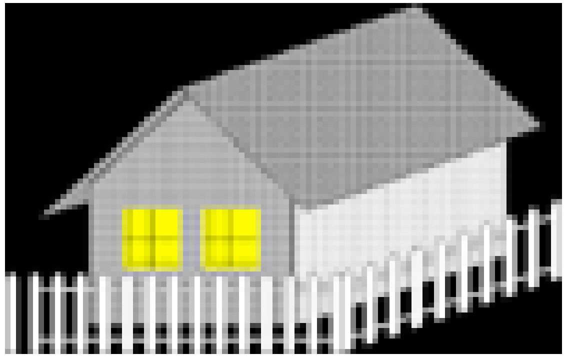
Sample picture at each red point





Coarse Sampling: 20 points per row by 14 rows





Finer Sampling: 100 points per row by 68 rows

QUANTIZATION

Quantization makes the range of a signal discrete, so that the quantized signal takes on only a discrete, usually finite, set of values. Quantization is generally irreversible and results in loss of information, unlike sampling (where under suitable conditions exact reconstruction is possible). Therefore, it introduces distortion into the quantized signal that cannot be eliminated. One of the basic choices in quantization is the number of discrete quantization levels to use.

With L levels, N = $\log_2$ L bits are needed to represent the different levels, or conversely, with N bits we can represent L = 2^N levels.

QUANTIZATION

Figure below shows the reconstruction and transition levels for a uniform quantizer.

256 levels

32 levels

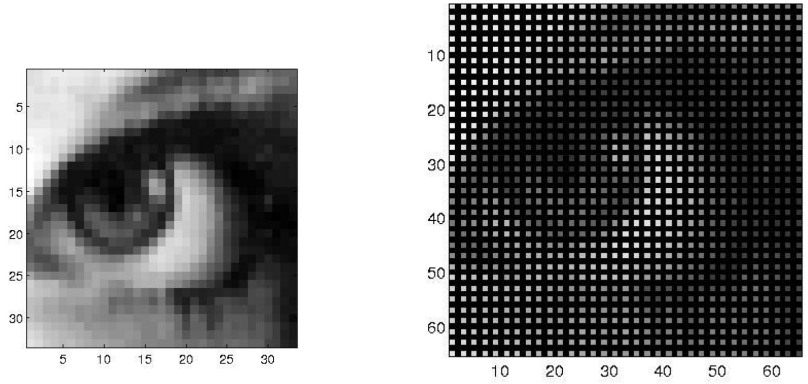
16 levels

8 levels

4 levels

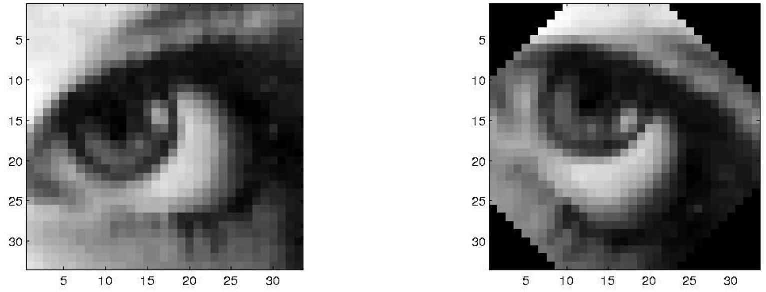
2 levels

Image interpolation occurs in all digital photos at some stage. It happens anytime you resize or remap (distort) your image from one pixel grid to another.



Resizing (resampling)

Inpainting (restauration of holes)



Remapping (geometrical tansformations like rotation, change of perspective, etc)

Morphing, nonlinear transformations

Aliasing comes in several forms:

SPATIAL ALIASING, IN PICTURES

moire patterns arise in

image warping & texture mapping

jaggies arise in rendering

TEMPORAL ALIASING, IN AUDIO

when resampling an audio signal at a lower sampling frequency, e.g. 50KHz (50,000 samples per second) to 10KHz.

TEMPORAL ALIASING, IN FILM/VIDEO

strobing and the “wagon wheel effect”

ALIASING: **SPATIAL ALIASING**

Is insufficient sampling of data along the space axis, that occurs because of the insufficient spatial resolution of the acquired image.

During image synthesis:

when sampling a continuous (geometric) model to create a raster image, e.g. scan converting a line or polygon.

Sampling: changing a continuous signal to a discrete signal.

During image processing and image synthesis:

when resampling a picture, as in image warping or texture mapping.

Resampling: sampling a discrete signal at a different sampling rate.

Example:

“zooming” a picture from n_x by n_y pixels to sn_x by sn_y pixels

s>1: called **upsampling** or **interpolation**

can lead to blocky appearance if point sampling is utilized

s<1: called **downsampling** or **decimation**

can lead to moire patterns and jaggies

FREQUENCY DOMAIN

It is possible to visualize and analyze a signal or a filter in the spatial domain or in the frequency domain.

Spatial domain: x, distance (generally in pixels).

Frequency domain: can be measured with:

- *w*, angular frequency in radians per unit distance;
- *f*, rotational frequency in cycles per unit distance. $\omega = 2\pi f.$

The period of a signal, $T = 1/f = 2\pi/\omega.$

Examples:

The signal [0 1 0 1 0 1 ...] has frequency f=.5 (.5 cycles per sample).

The signal [0 0 1 1 0 0 1 1 ...] has frequency f=.25 .

ANTI-ALIASING

Aliasing problem of a

downsampled image

Aliasing problem avoided

by using band limiting

Low Pass Filter

DRIVES
Development and Research on Innovative Vocational Education Skills

The Fourier transform is used to transform between the spatial domain and the frequency domain. A transform pair is symbolized with "$\leftrightarrow$", e.g. $f \leftrightarrow F$.

Spatial Domain

Signal: $f(x)$

Frequency Domain

Spectrum: $F(\omega)$

Fourier Transform:

$$F(\omega) = \int_{-\infty}^{+\infty} f(x) e^{-i\omega x} dx$$

Inverse Fourier Transform:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{i\omega x} d\omega$$

Where, $i = \sqrt{-1}$. *F* will be complex, in general.

Spatial Domain	Frequency Domain
impulse train, period T	impulse train, period 2 π/T
discrete with sample spacing T	periodic with period 2 π/T
convolution of signals: $f(x) \otimes g(x)$	multiplication of spectra: $F(w)G(w)$
multiplication of signals: $f(x)g(x)$	convolution of spectra: $F(w) \otimes G(w)/2\pi$
$(\omega_c/\pi)\operatorname{sinc}((\omega_c/\pi)x)$ where $\operatorname{sinc}(x)=\dfrac{\sin \pi x}{\pi x}$	Box $(w/2w_c)$ where box(x) = {1 if |x|<1/2, 0 otherwise and w_c is cutoff frequency

Point Processing: enhancement of any point in an image depends on only the gray-level at that point.

Mask Processing/Filtering: where the values of the mask coefficients control the nature of the process.

CONTRAST STRETCHING

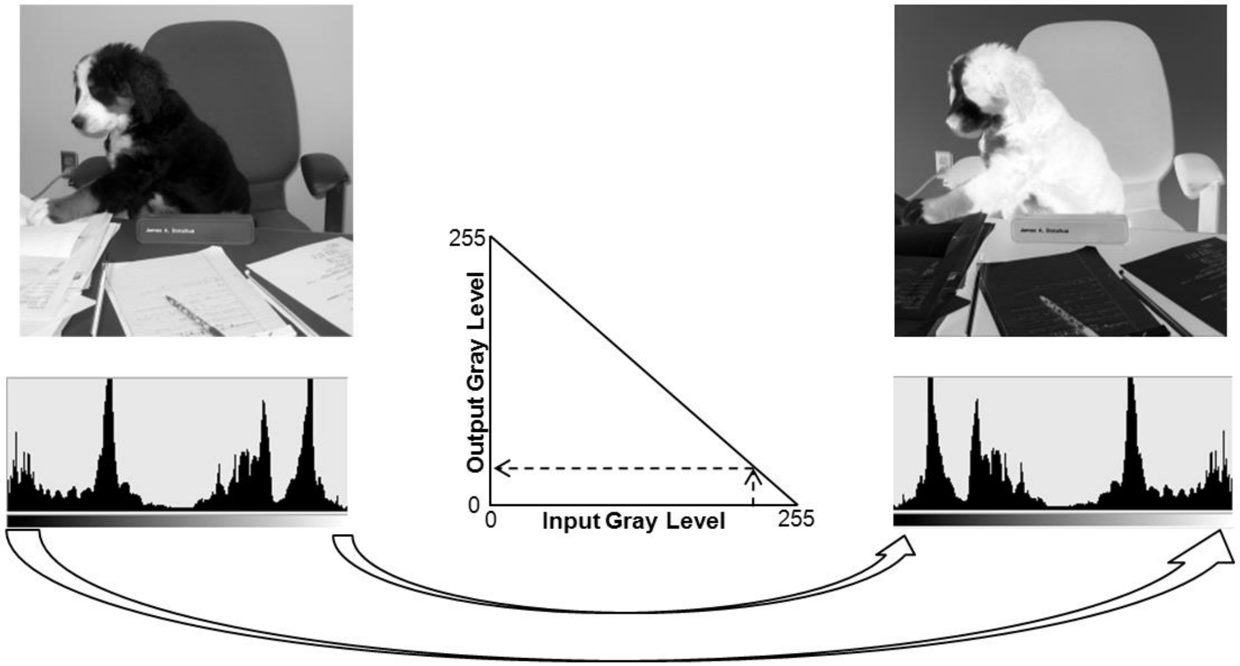
Contrast: is the difference in visual properties making an object/image distinguishable from other objects and the background.

Contrast Stretching: enhance the contrast in an image by stretching the range of intensity values to span a desired range of values.

CONTRAST STRETCH

A low-contrast image can be transformed into a high-contrast image by remapping or stretching the gray-level values such that the histogram spans the full range. The simplest contrast stretch is a linear transform, and this is given by:

$$g'(x,y) = \mathrm{INT}\left\{\frac{255}{GL_{max} - GL_{min}}\left[g(x,y) - GL_{min}\right]\right\},$$



The contrast reversal transform

Contrast enhancements improve the perceptibility of objects in the scene by increasing the brightness difference between objects and their backgrounds.

Contrast enhancements are typically performed as a contrast stretch followed by a tonal enhancement.

A contrast stretch improves the brightness differences accross the image and tonal enhancements improve the brightness differences in the shadow (dark), midtone (grays), or highlight (bright) regions.

Suprathreshold Contrast Sensitivity

$$1 \leq \frac{\frac{\partial I'}{\partial x}}{\frac{\partial I}{\partial x}} \leq (1+\tau) \quad \text{where,} \quad \tau \geq \lambda.$$

$I'(x, y)$ is the contrast enhanced image.

$I(x, y)$ is the image.

τ single parameter to express the contrast enhancement of an image.

The lower bound guarantee that contrast reduction does not occur at any point in the image and the upper bound assures that the contrast enhancement is bounded.

DRIVES
Development and Research on Innovative Vocational Education Skills

The Method for Gray Images

$$f(\Omega) = \frac{1}{4|\Omega|} \sum_{p \in \Omega} \sum_{q \in N_4(p)} \frac{I'(p) - I'(q)}{I(p) - I(q)}$$

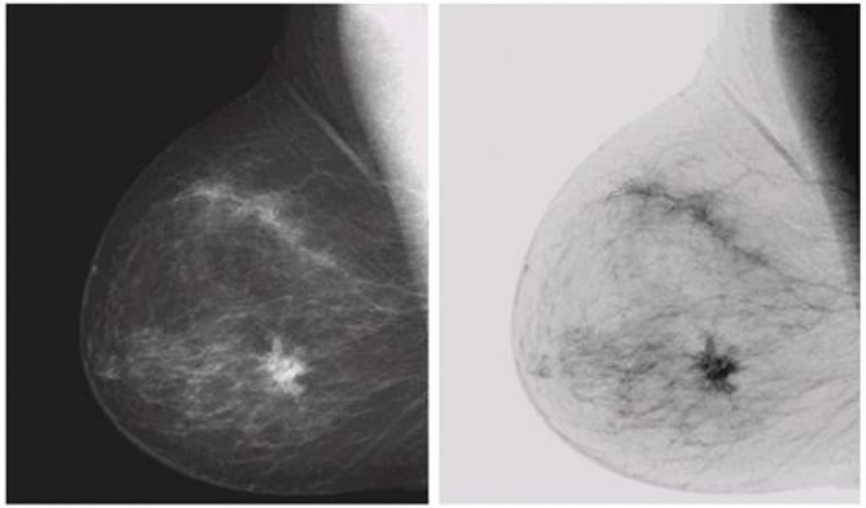
Scalar functions I(*p*) and I'(*p*) represent the gray values at pixel *p* of the input and output images, Ω denotes sets of pixels that makes up the image.

The objective function, defined above, can be directly used as a metric to evaluate the amount of average contrast enhancement (ACE) achieved across the whole image. The maximum average contrast that can be achieved without respecting saturation constraints is given by 1+ π. .

Images Negatives

$$s = (L-1) - r$$

s corresponds to the pixel value of the output image and r is the pixel value of the input image.



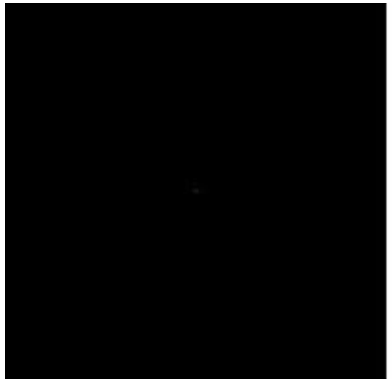
The picture on the left corresponds to an Original digital mammogram.

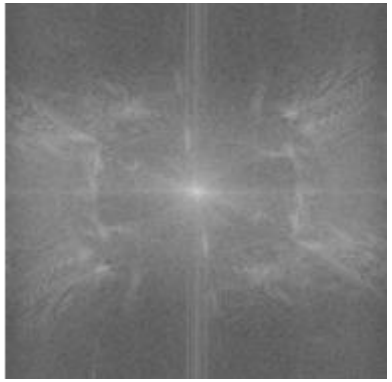
The right picture is a negative image obtained using the negative transformation

Logarithmic Transformations

$$s= c \log(1+r)$$

s corresponds to the pixel value of the output image and r is the pixel value of the input image.



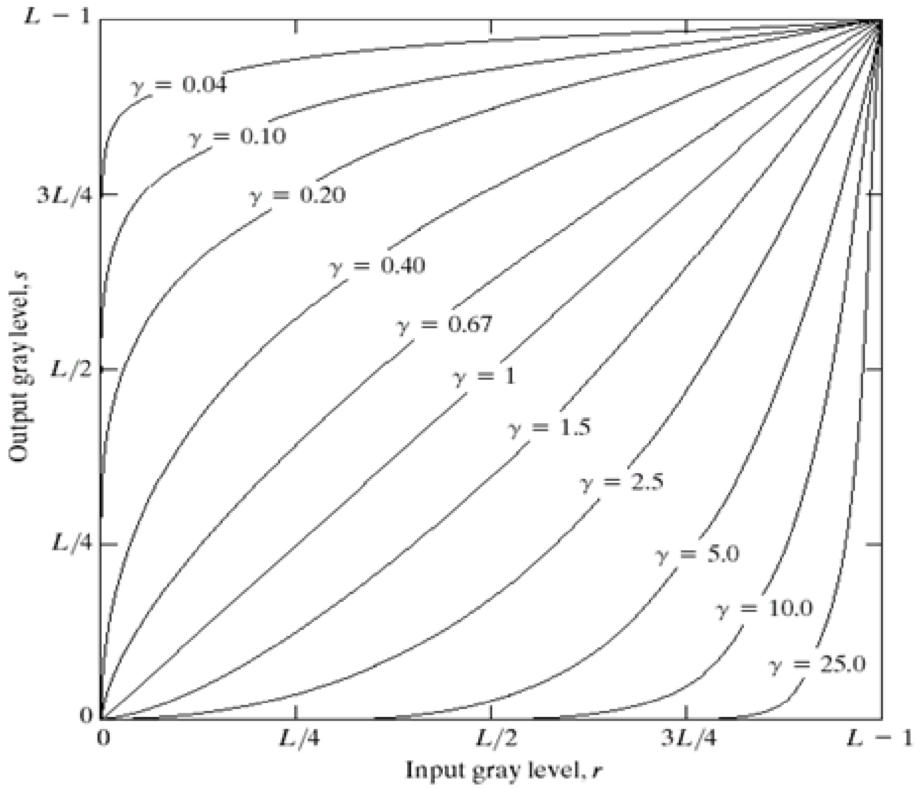


The picture on the left corresponds to Fourier spectrum of Barbara's image. The right picture is the result of applying the log transformation.

Logarithmic Transformations

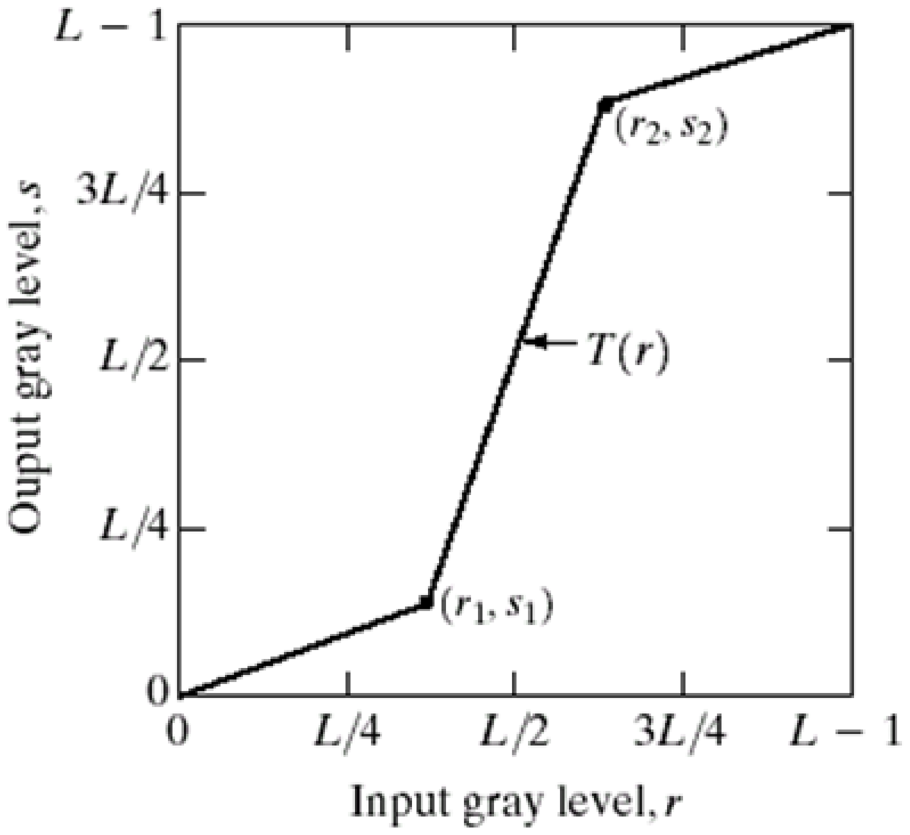
$$s = c\, r^{\gamma}$$

s corresponds to the pixel value of the output image and r is the pixel value of the input image. ($\gamma \geq 0$ and $0 \leq r \leq 1$)



Plots for various values of γ (c=1).

Piecewise-Linear Transformations



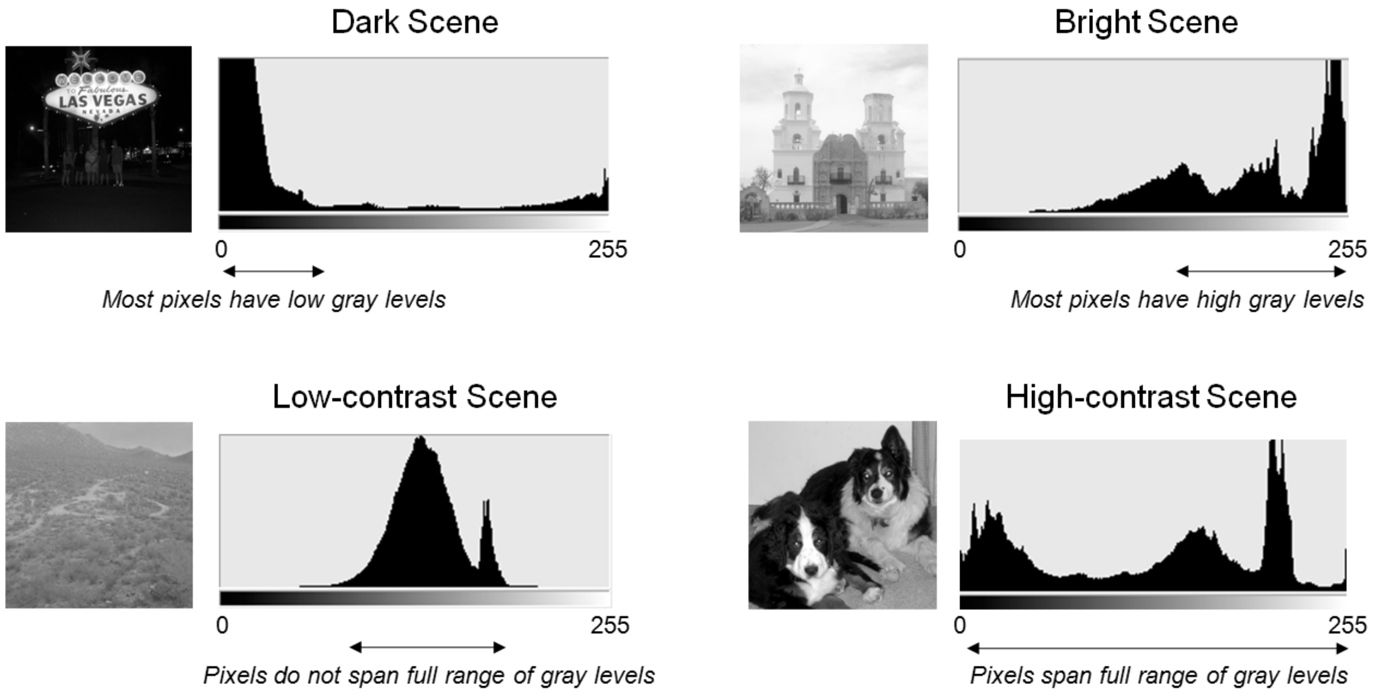
An example of piecewise linear transformation function

Histogram : corresponds to a discrete function $h(r_k)=n_k$, where r_k is the k^{th} gray level in the range of [0, L-1] and n_k is the number of pixels having gray level r_k.

Normalized histogram : is $p(r_k)=n_k/n$, for $k=0,1,\ldots,L-1$ and $p(r_k)$ can be considered to give an estimate of the probability of occurrence of ray level r_k.

CONTRAST ENHANCEMENTS: GRAY-LEVEL HISTOGRAM

Most of the contrast enhancement methods use the gray-level histogram, created by counting the number of times each gray-level value occurs in the image, then dividing by the total number of pixels in the image to create a distribution of the percentage of each gray level in the image. The gray-level histogram describes the statistical distribution of the gray levels in the image but contains no spatial information about the image.



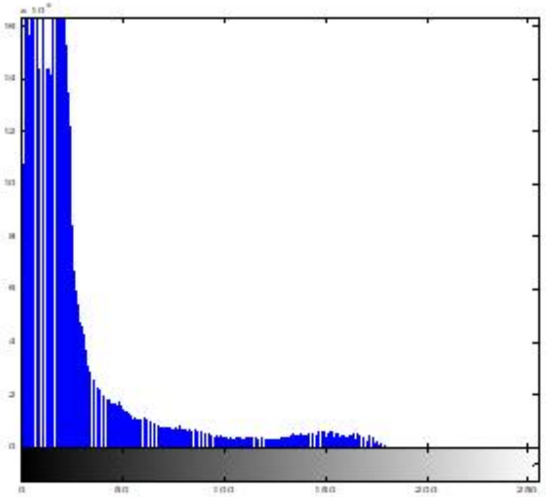
The gray-level histogram for different brightness and contrast scenes

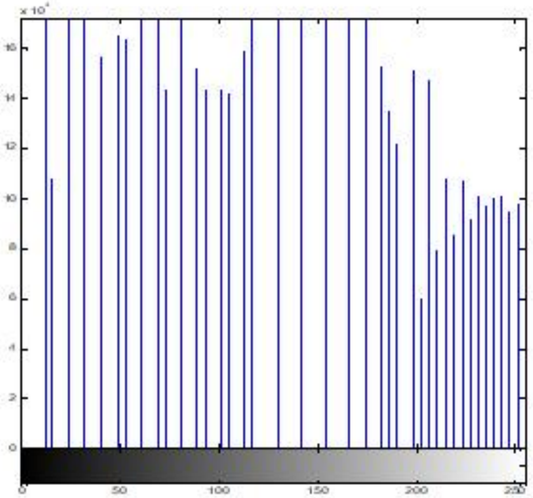
Histogram equalization : can be defined as a method that increases the dynamic range of the gray-levels in a low-contrast image to cover full range of gray-levels. It is achieved by having a transformation function that is the Cumulative Distribution Function (CDF) of a given PDF of gray-levels in a given image.

The new intensity value of pixel x is calculated by:

$$I(x) = \text{round}\left(\frac{cdf(x) - \min cdf}{1 - \min cdf} \times (L-1) \right)$$

The probability function of the output levels is uniform.





Histogram Equalization

SINGULAR VALUE DECOMPOSITION (SVD)

The multiplication of two orthogonal square matrices, U and V can be written in any matrix,

A. The matrix containing the sorted singular values on its main diagonal, Σ.

$$\mathbf{A=U\Sigma V^T}$$

DRIVES
Development and Research on Innovative Vocational Education Skills

Note: σ_1 is much bigger than other σs, so changing this, it will affect on the reconstructed image, for example changing σ_1 will directly change the luminance of the image.

G(0.5, 1) : corresponds to a synthetic intensity matrix whose pixel values have Gaussian distribution with mean of 0.5 and variance of 1 with the same size of the original image.

ξ: is ratio of the largest singular value of the generated normalized matrix over a normalized image.

$$\boxed{\xi = \frac{\max\left(\Sigma_{G(\mu=0.5,\sigma=1)}\right)}{\max\left(\Sigma_A\right)}}$$

$$Equalized\,\mathrm{Im}\,age = U_A\left(\xi\Sigma_A\right)V_A^T$$

DRIVES
Development and Research on Innovative Vocational Education Skills

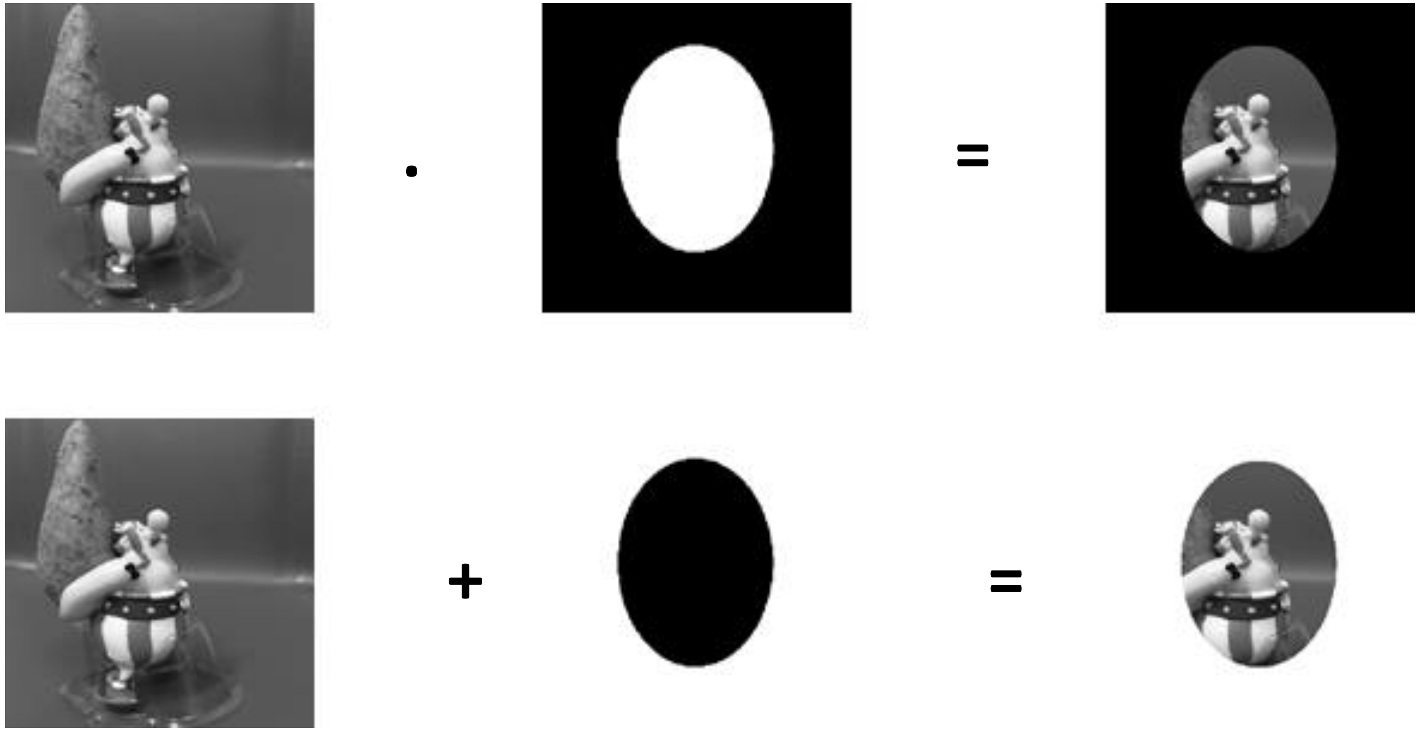
Basic Arithmetic Operations

Operation	Definition
ADD	c = a + b
SUB	c = a - b
MUL	c = a x b
DIV	c = a / b
LOG	c = log(a)
EXP	c = exp(a)
SQRT	c = sqrt(a)
TRIG.	c = sin/cos/tan(a)
INVERT	c = (2^B – 1) - a

Basic Logic Operations

Operation	Definition
NOT	$c = \bar{a}$
OR	$c = a + b$
AND	$c = a.b$
XOR	$c = a \oplus b = a.\bar{b} + \bar{a}.b$

ARITHMETIC AND LOGIC OPERATIONS



LINEAR AND NONLINEAR SPATIAL FILTERING

Filtering : spatial domain filtering

frequency domain filtering

Spatial domain filtering :

- linear spatial filtering (can be done with a convolution, is just the linear sum of values in a sliding window)
- nonlinear spatial filtering (A sliding median filter is a simple example of a non-linear filter)

Convolution

$$g(x,y)=\sum_{s=-a}^{a}\sum_{t=-b}^{b}\omega(s,t)f(x-s,y-t)$$

$$g=\omega * f$$

Convolution kernel, ω

1	-1	-1
1	2	-1
1	1	1

Rotate 180°

1	1	1
-1	2	1
-1	-1	1

Input Image, f

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2

Convolution

1	1	1
-1	2	1
-1	-1	1

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2

Final Output Image, g

5	4	4	-2
9	6	14	5
11	7	6	5
9	12	8	5

Input Image, *f*

1	1	1	
-2	4	2	3
-2	-1	3	3
2	2	1	2
1	3	2	2

Output Image, g

5	4		

$$g(x,y)=\sum_{s=-a}^{a}\sum_{t=-b}^{b}\omega(s,t)f(x+s,y+t)$$

$$g=\omega\circ f$$

Correlation kernel, ω

1	-1	-1
1	2	-1
1	1	1

Don't rotate, use directly

Input Image, *f*

2	2	2	3
2	1	3	3
2	2	1	2
1	3	2	2

Correlation

1	1	1
-1	2	1
-1	-1	1

Input Image, *f*

1	1	1	
-2	4	2	3
-2	-1	3	3
2	2	1	2
1	3	2	2

Output Image, g

5	10		

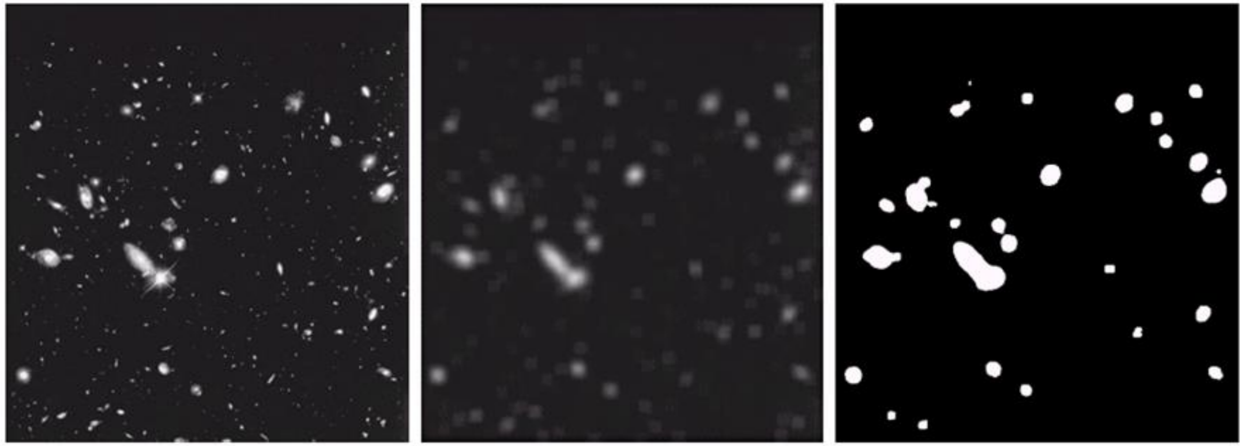
Final Output Image, g

5	10	10	15
3	4	6	11
7	11	4	9
-5	4	4	5

Linear Spatial Filtering

$$R = \omega(-1,-1)f(x-1,y-1)+\omega(-1,0)f(x-1,y)+\ldots+\omega(0,0)f(x,y)+\ldots+\omega(1,0)f(x+1,y)+\omega(1,1)f(x+1,y+1)$$

Smoothing Spatial Filters is known as averaging filter or lowpass filters, the types are linear and nonlinear.



original blurred Thresholded

The Laplacian

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y)$$

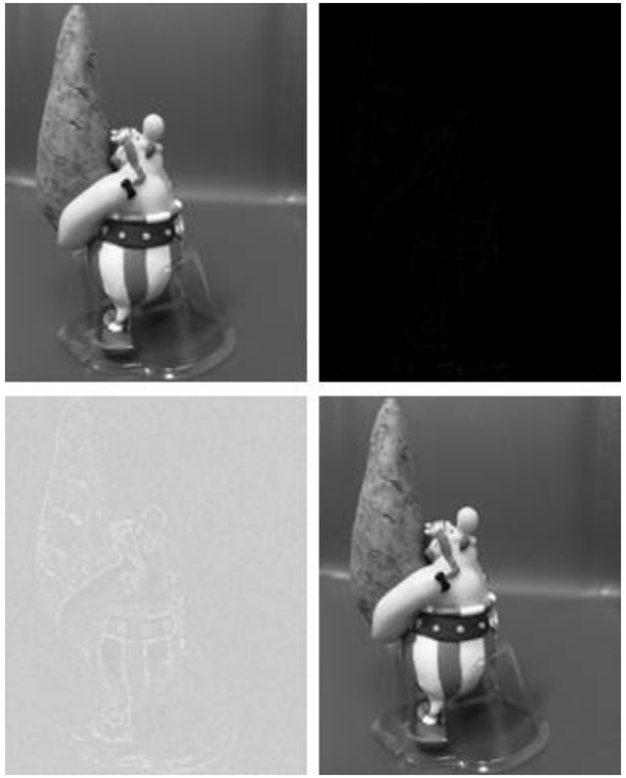
$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y)$$

The 2D Laplacian

$$\nabla^2 f = [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)] - 4f(x, y)$$

The 2D Laplacian

$$g(x,y)=\begin{cases} f(x,y)-\nabla^2 f(x,y) & \text{If the center coefficient of the Laplacian mask is negative} \\ f(x,y)+\nabla^2 f(x,y) & \text{If the center coefficient of the Laplacian mask is positive} \end{cases}$$



DRIVES
Development and Research on Innovative Vocational Education Skills

The Gradient

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

$$\nabla f = mag(\nabla \mathbf{f}) = \left[G_x{}^2 + G_y{}^2 \right]^{1/2} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

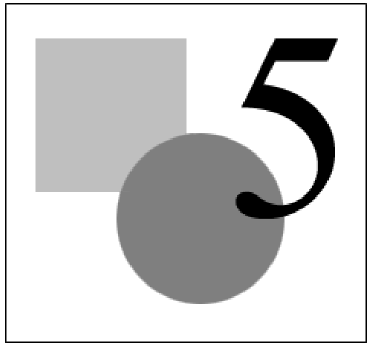
$$\nabla f \approx |G_x| + |G_y|$$

$$\nabla f \approx |G_x| + |G_y| = |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|$$

The Sobel Operator

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1



Horizontal Vertical

SUMMARY

- An image is a two dimensional function ***f(x,y)***, where *x* and *y* are the **spatial** coordinates, and the amplitude of ***f*** at any pair of coordinates (*x*,*y*) is called the intensity of the image at that level.
- The **sampling** rate determines the spatial resolution of the digitized image, while the **quantization** level determines the number of grey levels in the digitized image.
- **Quantization** is generally irreversible and results in loss of information, unlike sampling (where under suitable conditions exact reconstruction is possible).
- Image interpolation happens anytime you resize or remap (distort) your image from one pixel grid to Another.
- Aliasing has several forms: spatial in pictures, temporal in audio, and temporal in videos.
- Spatial domain methods: The Fourier transform is used to transform between the spatial domain and the frequency domain.

SUMMARY

- **Contrast Stretching:** enhance the contrast in an image by stretching the range of intensity values to span a desired range of values.
- Contrast enhancements functions: Images Negatives, Logarithmic Transformations and Piecewise Linear Transformations.
- **Histogram equalization :** can be defined as a method that increases the dynamic range of the gray-levels in a low-contrast image to cover full range of gray-levels.
- **Spatial domain filtering** can be linear spatial filtering, and nonlinear spatial filtering.

REFFERENCES

1. Introduction to image processing | Digital Image Processing. (n.d.). Retrieved July 4, 2021, from https://sisu.ut.ee/imageprocessing/book/1

Anbarjafari, G. (n.d.). *Vid_Lecture08Fin*. https://www.google.com.pk/search?q=1.+Digital+Image+processing%2C+Spatial+Filtering+Part+1+Gholamreza+Anbarjafari&oq=1.+Digital+Image+processing%2C+Spatial+Filtering+Part+1+Gholamreza+Anbarjafari&aqs=chrome..69i57.103j0j1&sourceid=chrome&ie=UTF-8

Fiete, R. D. (2010). Image Enhancement Processing. *Modeling the Imaging Chain of Digital Cameras*, 127–161. https://doi.org/10.1117/3.868276.ch9

Iftekharuddin, K. M., & Awwal, A. (2012). Sampling and Quantization. *Field Guide to Image Processing*, *i*, 5–5. https://doi.org/10.1117/3.923354.ch5

Ingeniería, E. De, & Oviedo, I. De. (n.d.). *Image interpolation Nearest neighbor Bilinear interpolation Bicubic*.

Transforms, F. (n.d.). *What is aliasing? What causes it?*

Pitas, P. I. (n.d.). *Digital Image Formation summary*.

REFERENCE TO AUTHORS

Regina Sousa

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0002-2988-196X

Diana Ferreira

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0003-2326-2153

Ana Luísa Sousa

- PhD student in Information System and Tecnologies
- Research Collaborator of the Algoritmi Research Center

0000-0001-5731-3583

António Abelha

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0001-6457-0756

José Machado

- Associate Professor with Habilitation at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-4121-6169

Victor Alves

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-1819-7051

REFERENCE TO AUTHORS

This Training Material has been certified according to the rules of **ECQA – European Certification and Qualification Association.**

The Training Material was developed within the international job role committee **“Computer Vision Expert”:**

UMINHO – University of Minho (https://www.uminho.pt/PT)

The development of the training material was partly funded by the EU under Blueprint Project DRIVES.

DRIVES

Development and Research on Innovative Vocational Education Skills

Thank you for your attention

DRIVES project is project under **The Blueprint for Sectoral Cooperation on Skills in Automotive Sector**, as part of New Skills Agenda.

The aim of the Blueprint is **to support an overall sectoral strategy and to develop concrete actions to address short and medium term skills needs.**

Follow DRIVES project at:

More information at:

www.project-drives.eu

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.