Digital Image Processing
Image Transform (Fourier) / Quantization / Quality

O. Le Meur
olemeur@irisa.fr

Univ. of Rennes 1
http://www.irisa.fr/temics/staff/lemeur/

December 6, 2010

Introduction

Introduction

What is a signal?

This is a function of one or more variables (time, distance, temperature...).

A signal carries information!!

Signal processing is required to obtain a more desirable form.

Types of signals:

- Natural (lighting) or synthetic (from a lab);
- one dimension or more;
- Continuous/discrete.

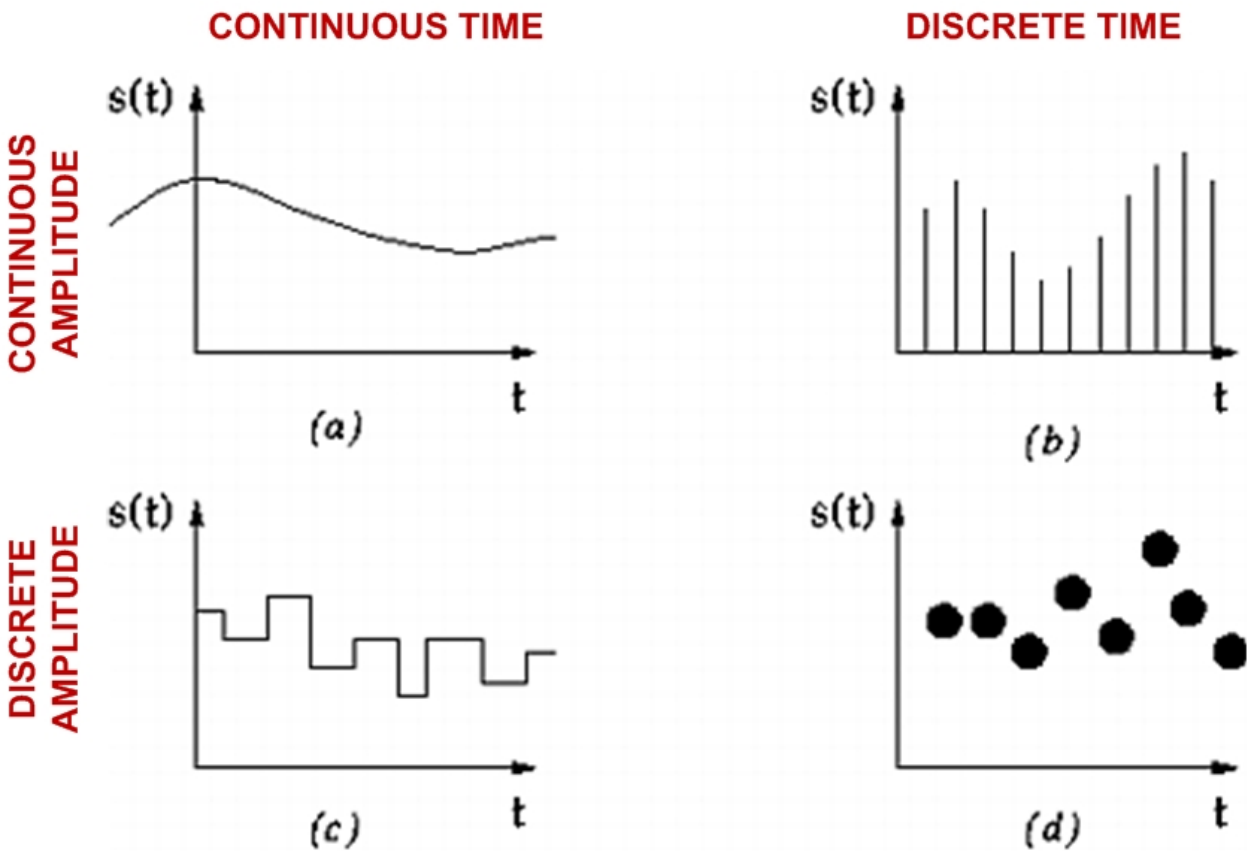
Introduction

Classification of signals

- Continuous-time vs Discrete-time signals:
  - → A continuous-time signal contains a value for all real numbers along the time axis;
  - → A discrete-time signal only has values at equally spaced intervals along the time axis.
- Analog vs Digital:
  - → Analog signal corresponds to a continuous set of possible function values;
  - → Digital signal corresponds to a discrete set of possible function values.
- Deterministic vs Random Signals:
  - → The past, present and future of a deterministic signal are known with certainty (its values can be determined by a mathematical expression, rule or table). Otherwise, it is called Random and its properties are explained by statistical techniques.
- Periodic vs Aperiodic: periodic signals repeat with some period T;
- Finite vs Infinite length.

Introduction

Classification of signals



Sampling = discrete time (or space)

Quantization = discrete amplitude

Introduction

There are many different ways to describe or represent a signal mathematically.

Most of the time a linear combination of elementary time functions. These functions are called basis functions.

Let the set of basis functions be defined by: $\{\phi_p\}$, (different posssibilities for the range of p)

$$x(t) = \sum_n a_n \phi_n(t)$$

Notations and reminder

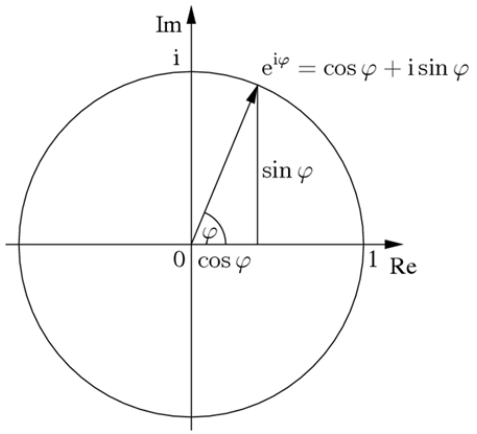
- a continuous signal $x(.)$;
- a discrete signal $x[.]$;
- a matrix $\mathbf{A}$.

Notations and reminder

- f is the frequency (Hz); $w = 2\pi f$ is the pulsation (radian/s); $T = \frac{1}{f}$ is the period (s);
- $x \in L^p(\mathcal{R})$ means that $\int_{-\infty}^{+\infty} |x(t)|^p \, dt < +\infty$. For instance, the space $L^2(\mathcal{R})$ is composed of the functions x with a finite energy $\int_{-\infty}^{+\infty} |x(t)|^2 \, dt < +\infty$;
- Euler's formulas: $exp(j\theta) = \cos\theta + j\sin\theta$;

$$exp(j\theta) = \cos\theta + j\sin\theta$$
$$exp(-j\theta) = \cos\theta - j\sin\theta$$

$$exp(j\theta) + exp(-j\theta) = 2\cos\theta$$
$$exp(j\theta) - exp(-j\theta) = 2j\sin\theta$$



From Wikepedia.

Reminder

- $sinc(x) = \frac{sin \pi x}{\pi x}$;
- Inner product:

$$\begin{aligned} \langle x, y \rangle &= \int_{-\infty}^{+\infty} x(t) \times y^*(t) dt \\ \langle x, y \rangle &= \sum_{n=-\infty}^{+\infty} x(n) \times y^*(n) \end{aligned}$$

- Condition of orthogonality for real basis function:

$$\int_{t_1}^{t_2} \phi_n(t) \phi_m(t) dt = \begin{cases} 0 & n \neq m \\ \lambda_m & n = m \end{cases}$$

where λ_m is a constant ($\lambda_m = 1$, orthonormal).

- Condition of orthogonality for complex basis function:

$$\int_{t_1}^{t_2} \phi_n(t) \phi_m^*(t) dt = \begin{cases} 0 & n \neq m \\ \lambda_m & n = m \end{cases}$$

where λ_m is a constant ($\lambda_m = 1$, orthonormal).

Image transformation

Image transformation

What is a transformation?

$$im[x, y] \xrightarrow{T} IM[u, v]$$

- im is the original image;
- IM is the transformed image;
- x, y (or u, v) represents the spatial coordinates of a pixel.

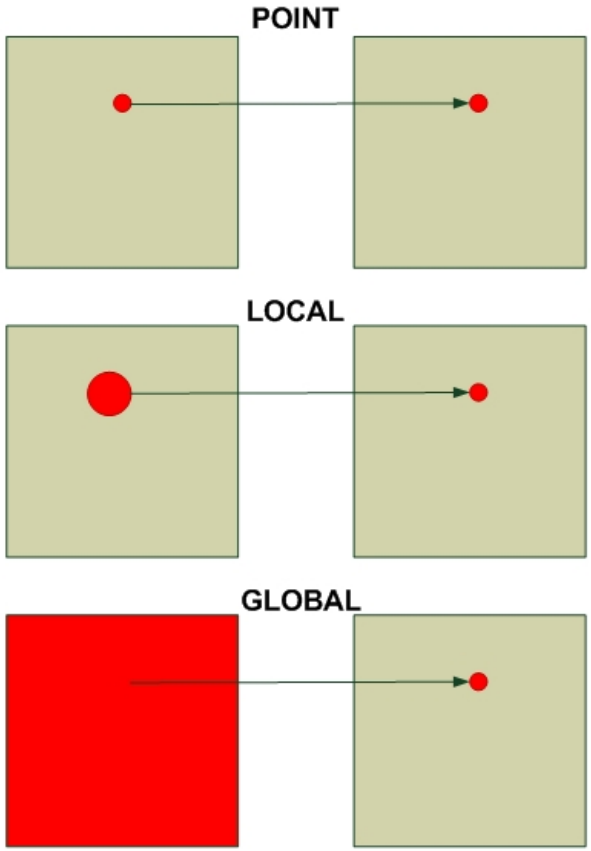
Goal of a transformation

The goal of a transformation is to get a new representation of the incoming picture. This new representation can be more convenient for a particular application or can ease the extraction of particular properties of the picture.

Image transformation

There exist 3 types of transformation:

- Point to point transformation:
  The output value at a specific coordinate is dependent only on one input value but not necessarily at the same coordinate;
- Local to point transformation:
  The output value at a specific coordinate is dependent on the input values in the neighborhood of that same coordinate;
- Global to point transformation:
  The output value at a specific coordinate is dependent on all the values in the input image.

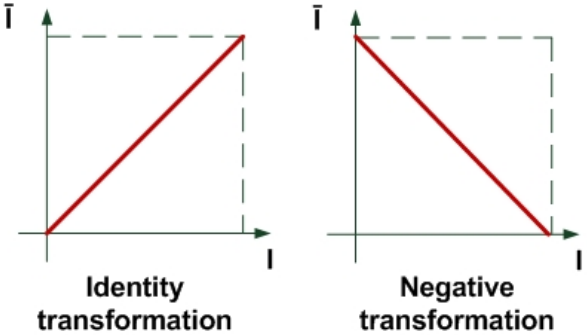


Note that the complexity increases with the size of the considered neighborhood...

Point to point transformation

- Geometric transformation such as rotation, scaling...
  Example of rotation:

- Scalar quantization
- LUT (Look-Up-Table)



Remark: a point to point transformation doesn't require extra memory...

Example of local to point transformation

Note that local to point transformation is also called neighborhood operators.

Examples of local to point transformation:

- Linear transformation:
  For instance, a tranverse filter (FIR) $IM[x,y] = \sum_i \sum_j h(i,j) im[x-i, y-j]$, where h is the 2D impulse response.
- Non-linear transformation such as adaptive linear filtering and rank filtering (median)...
- Morphological filtering (erosion, dilatation...) based on the definition of a structuring element.

Remark: Non linear filters are not reversible in general.

Global to point transformation

One of the most important transformation is the Fourier transform that gives a frequential representation of the signal.

Fourier series (development in real term):

Every signal x ($x \in L^2(t_1, t_1 + T)$) can be decomposed into a linear combination of sinusoidal and cosinusoidal component functions.

$$x(t) = b_0 + \sum_{n=1}^{+\infty} a_n cos(2\pi nt) + \sum_{n=1}^{+\infty} b_n sin(2\pi nt)$$

a_n and b_n are used to weight the influence of each frequency component.

Global to point transformation

Fourier series (development in complex term):

$$x(t) = \sum_{n=-\infty}^{+\infty} C_n exp(j2\pi n \frac{t}{T})$$

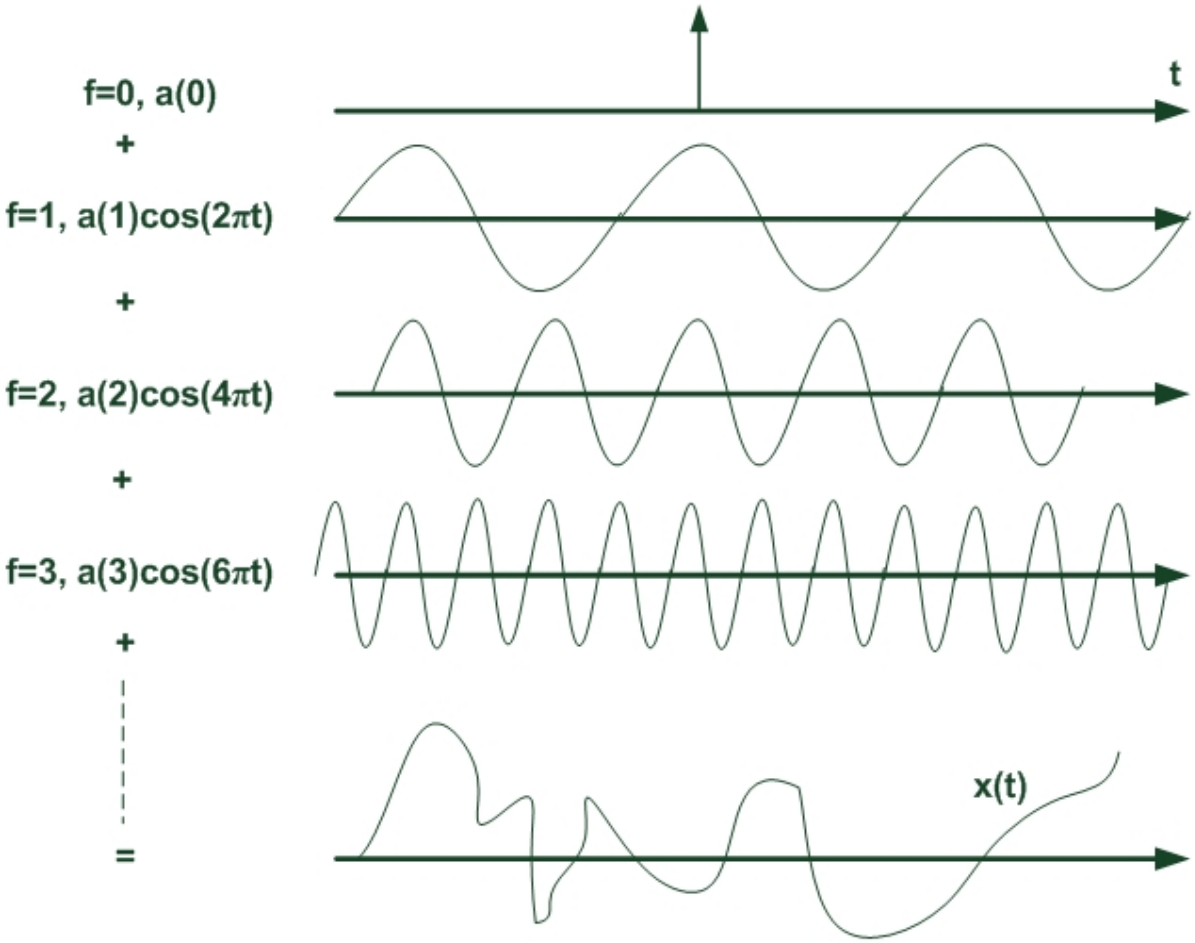
with,

$$C_n = \frac{1}{T} \int_{t_1}^{t_1+T} x(t) exp(-j2\pi n \frac{t}{T}) dt$$

Remark:

- $C_n = \frac{a_n - jb_n}{2}$, $n \geq 1$; $C_n = \frac{a_{-n} + jb_{-n}}{2}$, $n \leq 1$; $C_{-n} = C_n^*$;
- Note that if the function to be represented is also T-periodic, then t_1 is an arbitrary choice. Two popular choices are $t_1 = 0$, and $t_1 = -T/2$.

Global to point transformation



Fourier transformation

Fourier transformation

Fourier transformation in $L^2(\mathcal{R})$

$$x(t) \xrightarrow{\mathcal{F}} X(f)$$

Inverse Fourier Transform:

$$x(t) = \int_{-\infty}^{+\infty} X(f) exp(j2\pi ft) df$$

Fourier Transform:

$$X(f) = \int_{-\infty}^{+\infty} x(t) exp(-j2\pi ft) dt$$

provided that $x \in L^2(\mathcal{R})$ and $X \in L^2(\mathcal{R})$.

Fourier transformation

From the output of the Fourier transform, we define:

- The frequency spectrum: $Real\,(X(f)) + jImg\,((X(f)))$
  The Fourier transform of a function produces a frequency spectrum which contains all of the information about the original signal, but in a different form.
- Magnitude spectrum: $|Real\,(X(f)) + jImg\,(X(f))|$
- Phase spectrum: $Arctg\left(\frac{Img(X(f))}{Real(X(f))}\right)$
- Power spectrum: $Real\,(X(f))^2 + Img\,(X(f))^2$

Fourier transformation

- Linearity:

$$\begin{aligned} ax(t) &\xrightarrow{\mathcal{F}} aX(f) \\ ax_1(t) + bx_2(t) &\xrightarrow{\mathcal{F}} aX_1(f) + bX_2(f) \end{aligned}$$

- Complex conjugate: $x^*(t) \xrightarrow{\mathcal{F}} X^*(-f)$

$$\begin{aligned} x^*(t) &= \left(\int_{-\infty}^{+\infty} X(f) exp(j2\pi ft) df \right)^* \\ x^*(t) &= \int_{-\infty}^{+\infty} X^*(f) exp(-j2\pi ft) df \end{aligned}$$

In the same way, $x^*(-t) \xrightarrow{\mathcal{F}} X^*(f)$ and $x(-t) \xrightarrow{\mathcal{F}} X(-f)$.

Fourier transformation

- Hermitian symmetry: if $x(t) \in \mathcal{R}$, we deduce $X(-f) = X^*(f)$

$$
\begin{aligned}
X(f) &= \int_{-\infty}^{+\infty} x(t) exp(-j2\pi ft) dt \\
X^*(f) &= \left(\int_{-\infty}^{+\infty} x(t) exp(-j2\pi ft) dt \right)^* \\
X^*(f) &= \int_{-\infty}^{+\infty} x^*(t) exp(j2\pi ft) dt
\end{aligned}
$$

Given that $x(t) \in \mathcal{R}$, we have $x^*(t) = x(t)$ that implies $X^*(f) = X(-f)$.

Fourier transformation

Note that, as we use the space $L^2(\mathcal{R})$, the inner product of $x \in L^2(\mathcal{R})$ and $y \in L^2(\mathcal{R})$ is given by:

$$\langle x, y \rangle = \int_{-\infty}^{+\infty} x(t) y^*(t) dt$$

The norm is then given by:

$$\|x\|^2 = \langle x, x \rangle = \int_{-\infty}^{+\infty} |x(t)|^2 \, dt$$

$$\begin{aligned}
\int_{-\infty}^{+\infty} |x(t)|^2 \, dt &= \int_{-\infty}^{+\infty} x(t) x^*(t) dt \\
&= \int_{-\infty}^{+\infty} x(t) \left[\int_{-\infty}^{+\infty} X^*(f) exp(-j2\pi ft) df \right] dt \\
&= \int_{-\infty}^{+\infty} X^*(f) \left[\int_{-\infty}^{+\infty} x(t) exp(-j2\pi ft) dt \right] df \\
&= \int_{-\infty}^{+\infty} X^*(f) X(f) df \\
\int_{-\infty}^{+\infty} |x(t)|^2 \, dt &= \int_{-\infty}^{+\infty} |X(f)|^2 \, df
\end{aligned}$$

Fourier transformation

$$\int_{-\infty}^{+\infty} |x(t)|^2 \, dt = \int_{-\infty}^{+\infty} |X(f)|^2 \, df$$

This formula (Parseval's theorem or energy conservation) proves that the energy is conserved by the Fourier transform.

Loosely, the sum (or integral) of the square of a function is equal to the sum (or integral) of the square of its transform.

Fourier transformation

- Translation in time/space domain: $x(t-t_0) \xrightarrow{\mathcal{F}} exp(-j2\pi t_0 f)X(f)$

$$x(t-t_0) \xrightarrow{\mathcal{F}} \int_{-\infty}^{+\infty} x(t-t_0)exp(-j2\pi tf)dt$$

$$k = t - t_0$$

$$x(t-t_0) \xrightarrow{\mathcal{F}} \int_{-\infty}^{+\infty} x(k)exp(-j2\pi(k+t_0)f)dk$$

$$x(t-t_0) \xrightarrow{\mathcal{F}} exp(-j2\pi t_0 f)\left[\int_{-\infty}^{+\infty} x(k)exp(-j2\pi fk)dk\right]$$

Displacement in time or space induces a phase shift proportional to frequency and to the amount of displacement.

- Frequency shift: $x(t)exp(\pm j2\pi f_0 t) \xrightarrow{\mathcal{F}} X(f \pm f_0)$

Displacement in frequency multiplies the time/space function by a unit phasor which has angle proportional to time/space and to the amount of displacement. Amplitude modulation.

Fourier transformation

- Convolution in time/space domain: $x_1(t) * x_2(t) \xrightarrow{\mathcal{F}} X_1(f) \times X_2(f)$

$$x_1(t) * x_2(t) \xrightarrow{\mathcal{F}} \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} x_1(\tau) x_2(t-\tau) d\tau \right] exp(-j2\pi ft) dt$$

$$x_1(t) * x_2(t) \xrightarrow{\mathcal{F}} \int_{-\infty}^{+\infty} x_1(\tau) \left[\int_{-\infty}^{+\infty} x_2(t-\tau) exp(-j2\pi ft) dt \right] d\tau$$

Translation

$$x_1(t) * x_2(t) \xrightarrow{\mathcal{F}} \left[\int_{-\infty}^{+\infty} x_1(\tau) exp(-j2\pi f\tau) d\tau \right] \times X_2(f)$$

- Convolution in frequency domain: $x_1(t) \times x_2(t) \xrightarrow{\mathcal{F}} X_1(f) * X_2(f)$

These properties mean that if two functions are multiplied in one domain, then their Fourier transforms are convolved in the other domain.

Fourier transformation

- Time and frequency scaling: $x(\lambda t) \xrightarrow{\mathcal{F}} \frac{1}{|\lambda|} X(\frac{f}{\lambda})$

$$x(\lambda t) \xrightarrow{\mathcal{F}} \int_{-\infty}^{+\infty} x(\lambda t) exp(-j2\pi tf) dt$$

$$k = \lambda t$$

$$x(\lambda t) \xrightarrow{\mathcal{F}} \int_{-\infty}^{+\infty} x(k) exp(-j2\pi k \frac{f}{\lambda}) \frac{dk}{|\lambda|}$$

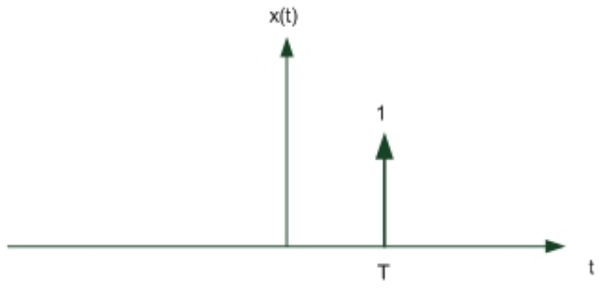
$$x(\lambda t) \xrightarrow{\mathcal{F}} \frac{1}{|\lambda|} \left[\int_{-\infty}^{+\infty} x(k) exp(-j2\pi k \frac{f}{\lambda}) dk \right]$$

Multiplication of the scale of the time/space reference frame changes by the factor λ scales the frequency axis of the spectrum of the function.

All the previous properties are what makes Fourier transforms so useful and practical.

Fourier transformation

- Translated Dirac, $x(t) = \delta(t - T)$

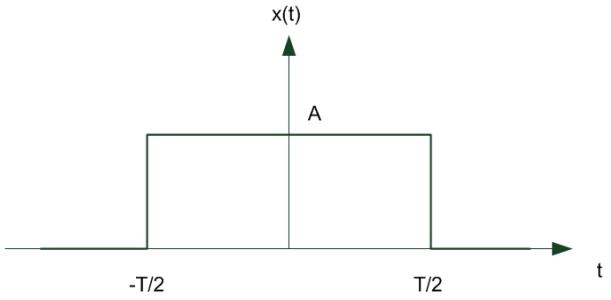


$$X(f) = \int_{-\infty}^{+\infty} x(t) exp(-j2\pi tf) dt$$
$$X(f) = exp(-j2\pi Tf)$$

If $T = 0$, $X(f) = 1$. The Fourier transform is constant whatever the frequency.

- $x(t) = 1$, $X(f) = \delta(f)$

- $x(t) = Arect(t|T)$



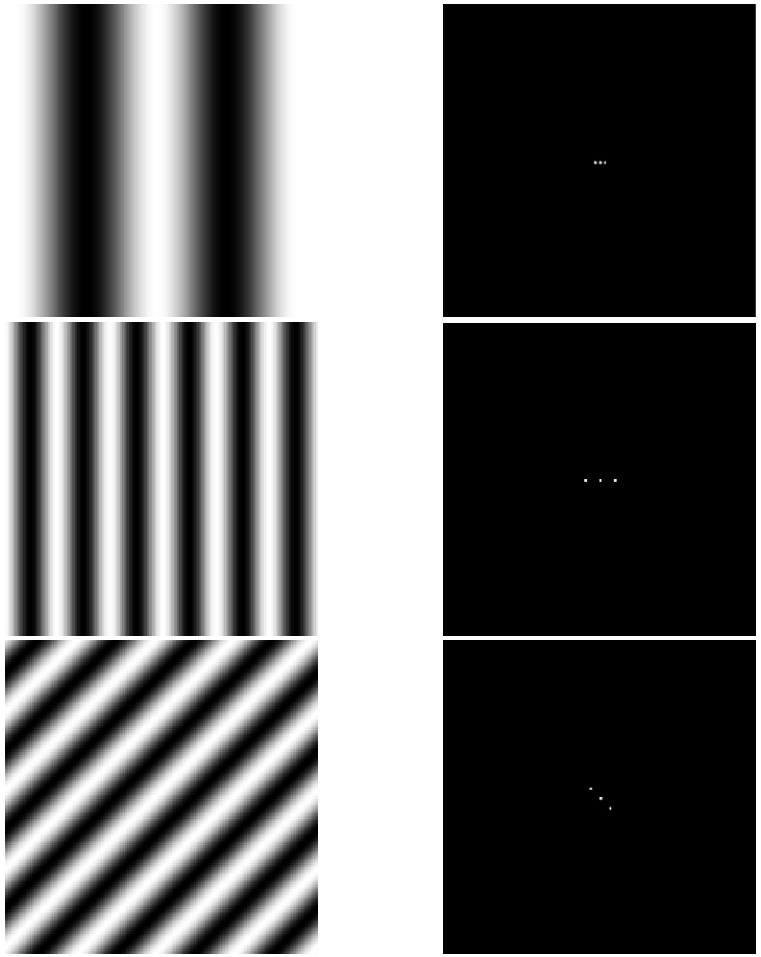
$$\begin{aligned} X(f) &= \int_{-\infty}^{+\infty} x(t) exp(-j2\pi tf) dt \\ &= A \int_{-\frac{T}{2}}^{+\frac{T}{2}} exp(-j2\pi tf) dt \\ &= -\frac{A}{j2\pi f} [exp(-jfT\pi) - exp(jfT\pi)] \\ X(f) &= ATsinc(fT) \end{aligned}$$

Fourier transformation

- Ideal low-pass filter in frequency domain: $X(f) = rect(f|\frac{F}{2})$ selects low frequencies over $\left[-\frac{F}{4}, \frac{F}{4}\right]$. The impluse response of the filter is calculated with the inverse Fourier integral:

$$
\begin{aligned}
x(t) &= \int_{-\infty}^{+\infty} X(f) exp(j2\pi ft) df \\
&= \int_{-\frac{F}{4}}^{+\frac{F}{4}} exp(j2\pi ft) df \\
&= \frac{1}{j2\pi t}\left[exp(j2\pi \frac{F}{4} t) - exp(-j2\pi \frac{F}{4} t)\right] \\
sin\theta &= \frac{exp(j\theta) - exp(-j\theta)}{2j} \\
x(t) &= \frac{F}{2} sinc(\frac{\pi}{2} tF)
\end{aligned}
$$

Fourier transformation



Time sampling

Time sampling

To be more flexible and to gain in accuracy...

Digital algorithms are now ... everywhere from television, CD, embedded electronic (car...).

Whether sound recordings or images, most discrete signals are obtained by sampling an analog signal. But, be careful, we have to follow some constraints in order to be able to reconstruct the analog signal.

The simplest way to discretize an analog signal x is to record its sample values at intervals T: $x(t)$ can be approximated by $\{x(nT)\}_{n\in\mathcal{Z}}$.

Definition

Diracs are essential to make the transition from functions of real variable to discrete sequences.

A Dirac can be defined as:

$$\int_{-\infty}^{+\infty} \delta(t)x(t)dt = x(0)$$

$$\delta(t) = \begin{cases} 0 & \text{if } t \neq 0 \\ +\infty & \text{for } t = 0 \end{cases}$$

$$\int_{-\infty}^{+\infty} \delta(t)dt = 1.$$

Remark: $\{\delta_\tau : \delta(t-\tau)\}_{\tau \in \mathcal{Z}}$ is an orthonormal basis:

$$\langle \delta(t-n), \delta(t-m) \rangle = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{for } m = n \end{cases}.$$

Properties of a Dirac

- Scaling and symmetry: $\delta(\lambda t) = \frac{\delta(t)}{|\lambda|}$

$$\begin{aligned} \int_{-\infty}^{+\infty} \delta(\lambda t) dt &= \int_{-\infty}^{+\infty} \delta(k) \frac{dk}{|\lambda|} \\ \delta(\lambda t) &= \frac{\delta(t)}{|\lambda|} \end{aligned}$$

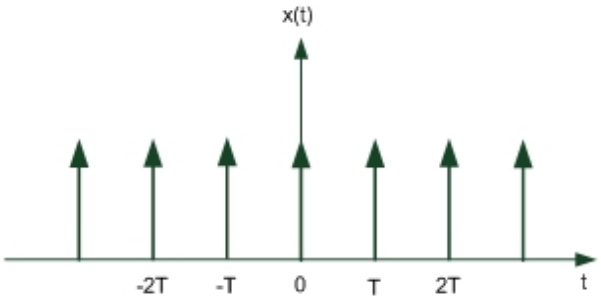
Note that if $\lambda = -1$, $\delta(-t) = \delta(t)$.

- Product: $x(t) \times \delta(t - T) = x(T) \times \delta(t - T)$
- Convolution: $x(t) * \delta(t - T)$

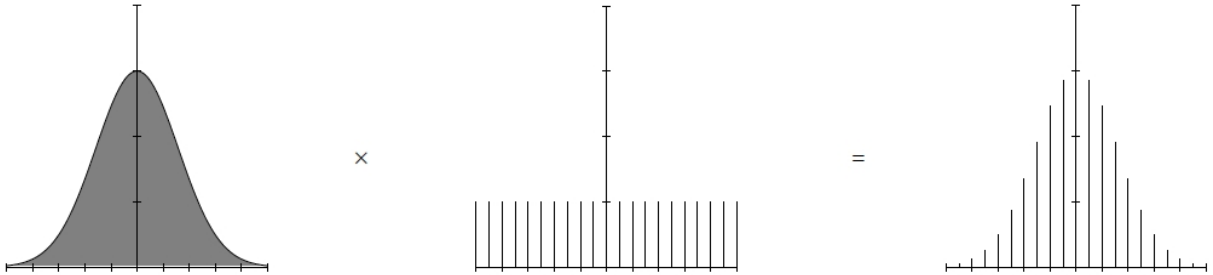
$$\begin{aligned} x(t) * \delta(t - T) &= \int_{-\infty}^{+\infty} x(\tau)\delta(t - T - \tau) d\tau \\ &= \int_{-\infty}^{+\infty} x(\tau)\delta(\tau - (t - T)) d\tau \\ &= x(t - T) \end{aligned}$$

The effect of convolving with a translated Dirac is to time-delay $x(t)$ by the same amount (copy of the signal $x(t)$ at the position T).

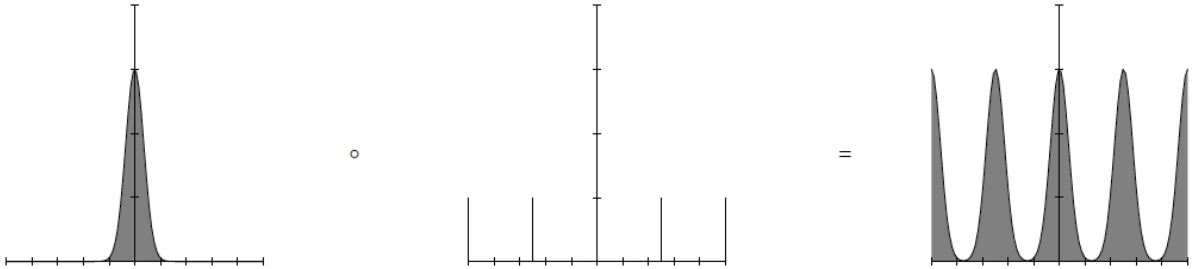
- Dirac comb: $x(t) = \sum_{n \in \mathcal{Z}} \delta(t - nT)$. Notation: $\Pi_T(t)$



- The product of a comb with a signal $x(t)$ conducts to a sampling:



- The convolution of a comb with a signal $x(t)$ conducts to a periodization of $x(t)$:



FT of a Dirac

- Dirac: $x(t) = \delta(t)$, $\delta(t) \xrightarrow{\mathcal{F}} 1$, constant function whatever the frequency

$$\begin{aligned} X(f) &= \int_{-\infty}^{+\infty} \delta(t) exp(-j2\pi tf) dt \\ X(f) &= exp(-j2\pi 0f) = 1 \end{aligned}$$

- Translated Dirac: $x(t) = \delta(t-T)$, $\delta(t-T) \xrightarrow{\mathcal{F}} exp(-j2\pi Tf)$

$$\begin{aligned} X(f) &= \int_{-\infty}^{+\infty} \delta(t-T) exp(-j2\pi tf) dt \\ X(f) &= exp(-j2\pi Tf) \end{aligned}$$

- Constant signal: $x(t) = 1$, $1 \xrightarrow{\mathcal{F}} \delta(f)$. The FT of a constant signal is a weight at the null frequency.

- Dirac comb: $x(t) = \sum_{n \in \mathcal{Z}} \delta(t - nT)$
  This is a periodic function, since $x(t+T) = x(t)$, $\forall t$. Its Fourier serie is given by:

$$
\begin{aligned}
x(t) &= \sum_{n=-\infty}^{+\infty} C_n exp(j2\pi n \frac{t}{T}) \\
C_n &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \delta(t) exp(-j2\pi n \frac{nt}{T}) dt = \frac{1}{T} \\
x(t) &= \frac{1}{T} \sum_{n=-\infty}^{+\infty} exp(j2\pi n \frac{t}{T})
\end{aligned}
$$

By using the Fourier Trasnform,

$$
\begin{aligned}
\mathcal{F}(x(t)) &= \mathcal{F}\left(\frac{1}{T} \sum_{n=-\infty}^{+\infty} exp(j2\pi n \frac{t}{T}) \right) \\
\mathcal{F}(x(t)) &= \frac{1}{T} \sum_{n=-\infty}^{+\infty} \mathcal{F}\left(exp(j2\pi n \frac{t}{T}) \right) \\
\mathcal{F}(x(t)) &= \frac{1}{T} \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} exp(-j2\pi t(f - \frac{n}{T})) dt \\
\sum_{n \in \mathcal{Z}} \delta(t - nT) &\xrightarrow{\mathcal{F}} \frac{1}{T} \sum_{n \in \mathcal{Z}} \delta(f - \frac{n}{T})
\end{aligned}
$$

Sampling

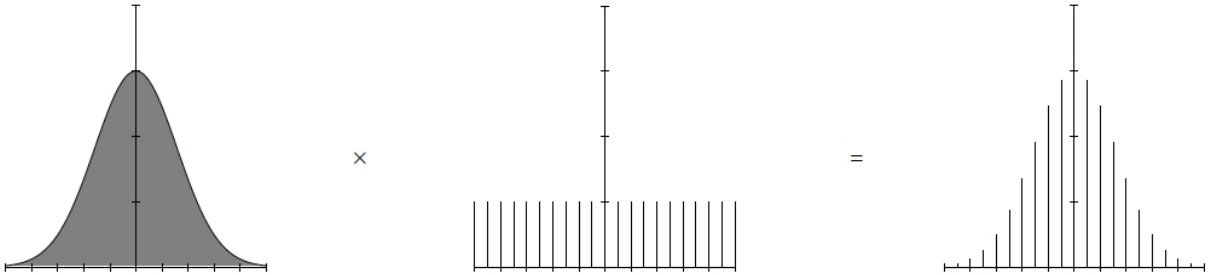
Sampling a signal is equivalent to multiplying it by a grid of impulses.

Sampling

Let $x(t)$ be a continous function, its sampled version $y(t)$ is given by:

$$\begin{aligned} y(t) &= x(t) \times \Pi_T(t) \\ &= x(t) \times \sum_{n \in \mathcal{Z}} \delta(t - nT) \\ &= \sum_{n \in \mathcal{Z}} x(t)\delta(t - nT) \\ y(t) &= \sum_{n \in \mathcal{Z}} x(nT)\delta(t - nT) \end{aligned}$$

where $\Pi_T(t)$ is the sampling function (sampling frequency $\frac{1}{T}$).



Fourier Transform of a sampled signal

Let $x(t)$ be a continous function, its sampled version $y(t)$, $y(t) = x(t) \times \Pi_T(t)$

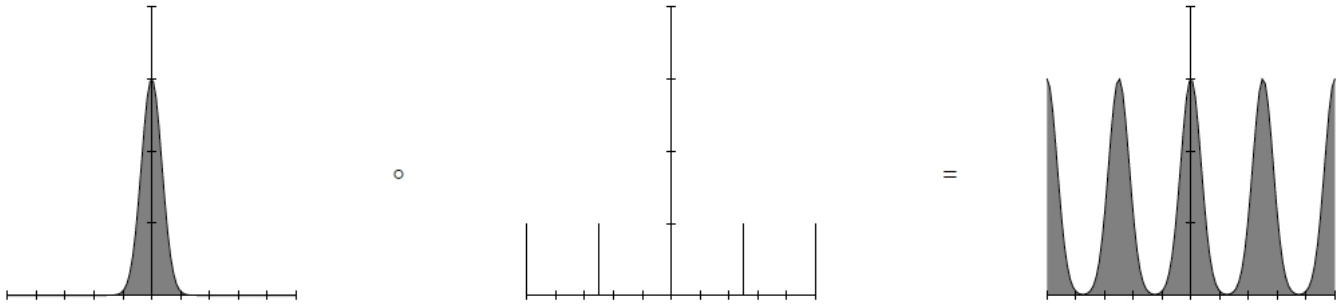
$$
\begin{aligned}
y(t) &= x(t) \times \sum_{n \in \mathcal{Z}} \delta(t - nT) \\
\mathcal{F}(y(t)) &= \mathcal{F}\left(x(t) \times \sum_{n \in \mathcal{Z}} \delta(t - nT) \right) \\
\mathcal{F}(y(t)) &= X(f) * \mathcal{F}\left(\sum_{n \in \mathcal{Z}} \delta(t - nT) \right) \\
\mathcal{F}(y(t)) &= X(f) * \left[\frac{1}{T} \sum_{n \in \mathcal{Z}} \delta(f - \frac{n}{T}) \right] \\
\mathcal{F}(y(t)) &= \frac{1}{T} \sum_{n \in \mathcal{Z}} X(f) * \delta(f - \frac{n}{T}) \\
\mathcal{F}(y(t)) &= \frac{1}{T} \sum_{n \in \mathcal{Z}} X(f - \frac{n}{T})
\end{aligned}
$$

The Fourier representation of the sampled function is a superposition of the Fourier transform $X(f)$ and all its replicas at positions $\frac{n}{T}$. For $n = 0$, the spectrum is the primary spectrum and for $n \neq 0$ we have copies of the spectrum.

Fourier Transform of a sampled signal

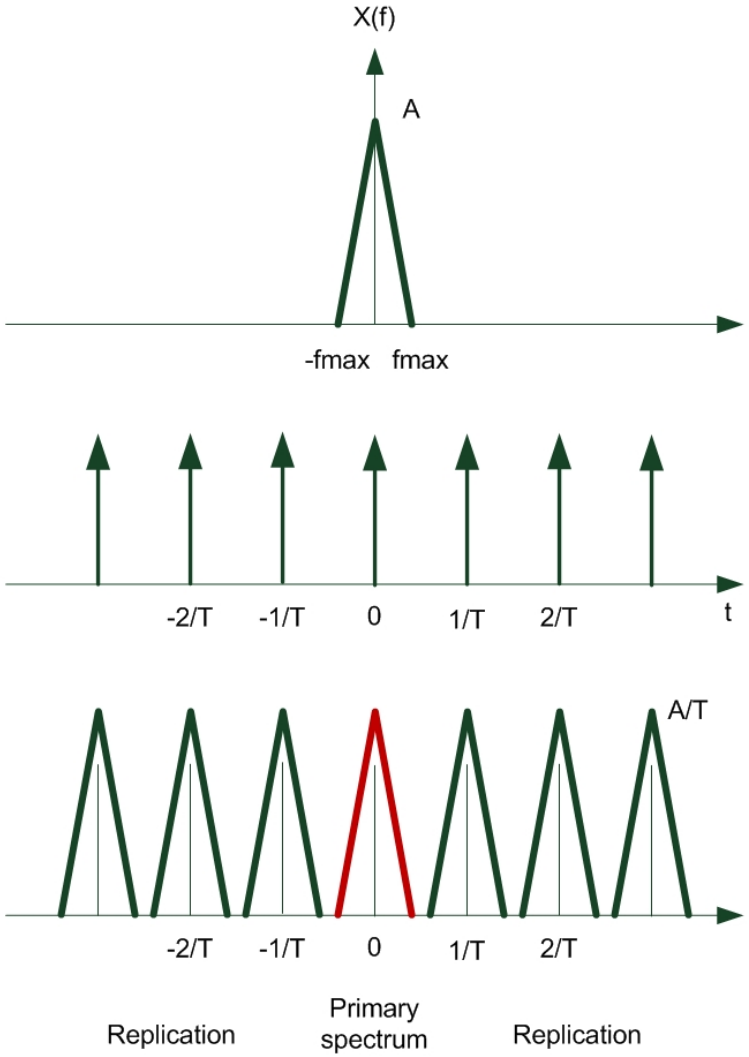
Let $x(t)$ be a continous function, its sampled version $y(t)$, $y(t) = x(t) \times \Pi_T(t)$
Let $X(f)$ be the Fourier transform of $x(t)$ and $Y(f)$ the Fourier transform of its sampled version $y(t)$.

$$Y(f) = \frac{1}{T} \sum_{n \in \mathcal{Z}} X(f - \frac{n}{T})$$

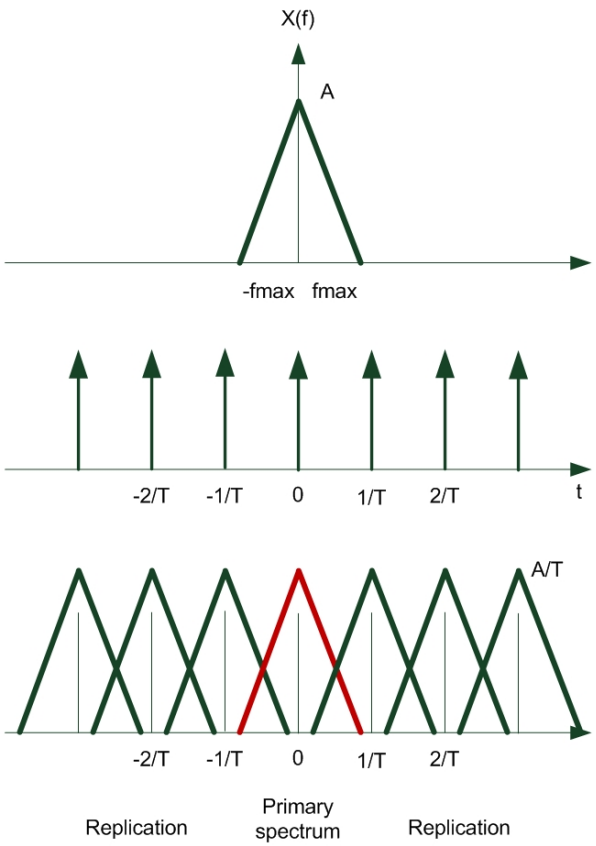


The sampling operation replicates the primary spectrum periodically in the Fourier domain.

Fourier Transform of a sampled signal



Fourier Transform of a sampled signal



Aliasing effect...

We won't be able to recover the original signal with a good quality. The transformation is no more inversible!

The sampling frequency is too small.

Shannon-Nyquist theorem

To ensure a perfect reconstruction of a signal $x(t)$ (to avoid aliasing effect), the sampling frequency f_e must verify:

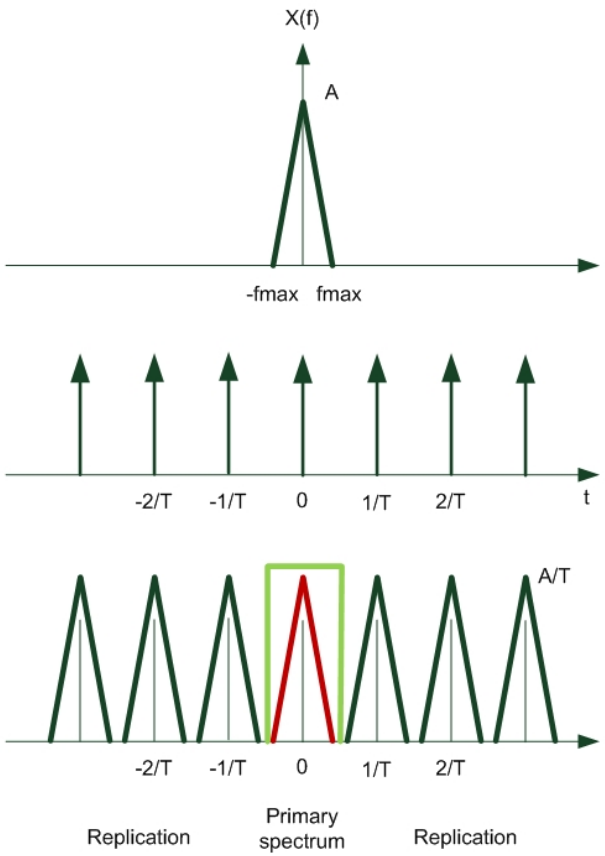
$$f_e > 2 \times f_{max}$$

where f_{max} is the maximal frequency of the spectrum $X(f)$.

The periodization doesn't provoke an aliasing effet.

Whittaker-Shannon theorem

If $x(t)$ is bandlimited to a frequency range $[-f_{max}, f_{max}]$, that is, its Fourier transform $X(f)$ is zero for $|f| > f_{max}$, then $x(t)$ can be completely specified by samples taken at the Nyquist sample rate of $2f_{max}$.

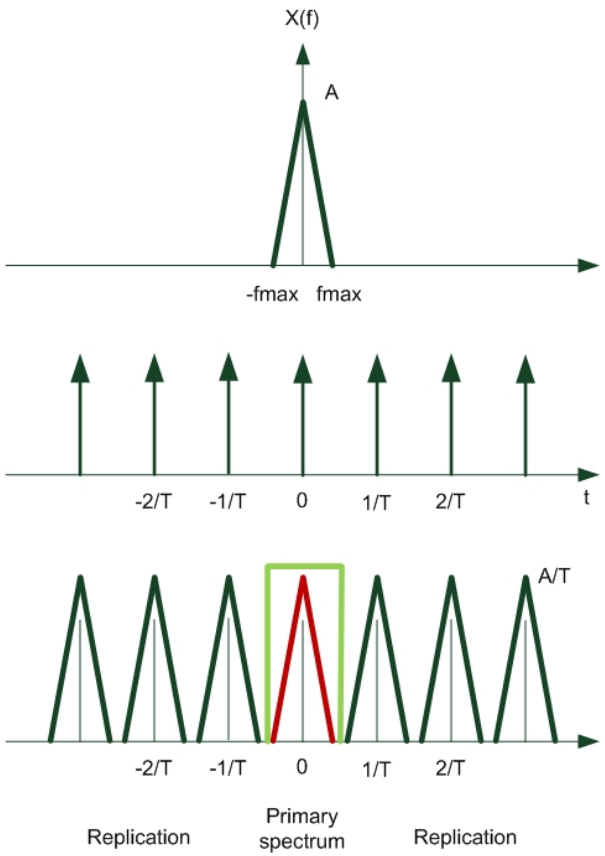


Whittaker-Shannon theorem

To recover the signal $x(t)$, we can use a low-pass filter in frequency domain. For instance, the ideal low-pass filter in the frequency domain is given by
$G(f) = T_e rect(f/f_e)$

$$X(f) = Y(f) \times rect(f/f_e)$$

$$x(t) = \sum_{k=-\infty}^{+\infty} x(kT_e) sinc((t - kTe)f_e)$$



The sampling of a physical signal always induces an aliasing effect, even with an ideal sampler. Paley-Wiener theorem indeed showed that there is no bandlimited signal having a finite energy...

In many cases, the spectrum of the signal is not perfectly known and can be corrupted by noise. It is thus required to filter the analog signal before the sampling operation.

1. Pre-filtering to obtain a bandlimited signal;
2. Sampling by checking the value of the sampling frequency;
3. Low-pass filtering to retrieve the signal.

Discrete Fourier Transform

Definition of 1D DFT

The discrete Fourier transform (DFT) is the transform of a finite sequence. Since the time-domain is discrete the spectrum is periodic

The DFT of a sequence of N complex numbers $x_0, \ldots, x_{N-1}$ is given by:

$$X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] exp(-j\frac{2\pi}{N}nk),\ k = 0, \ldots, N-1$$

$$\Longleftrightarrow$$

$$X[k] = \langle x, e_k \rangle, \text{ where } e_k(n) = exp(j\frac{2\pi}{N}nk).$$

The family $\left\{e_k(n) = exp(j\frac{2\pi}{N}nk)\right\}_{0 \leq k < N}$ is an orthogonal basis of the space of signals of period N.

The IDFT of a sequence of N complex numbers $X_0, \ldots, X_{N-1}$ is given by:

$$x[n] = \sum_{n=0}^{N-1} X[k] exp(j\frac{2\pi}{N}nk),\ n = 0, \ldots, N-1$$

The DFT computes the $X[k]$ from the $x[n]$, while the IDFT shows how to compute the $x[n]$ as a sum of sinusoidal components $X[k]exp(\frac{j2\pi}{N}kn)$ with frequency k/N cycles per sample.

Definition of 1D DFT

$$X[k] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]exp(-j\frac{2\pi}{N}nk)$$

The weighting coefficient $\frac{1}{N}$ can be removed or replaced by $\frac{1}{\sqrt{N}}$. In the first case, the weighting coefficient of the IDFT is equal to $\frac{1}{N}$ whereas in the second case it is equal to $\frac{1}{\sqrt{N}}$.

What is important is that $DFT \circ IDFT = Id$.

Discrete bi-dimensional Fourier transformation

$im[k,l]$ is a discrete picture having a size $N \times M$.

Inverse Fourier Transform:

$$im[k,l] = \sum_{u=0}^{N-1}\sum_{v=0}^{M-1} IM[u,v]exp(j2\pi(\frac{k}{N}u + \frac{l}{M}v))$$

Fourier Transform:

$$IM[u,v] = \frac{1}{\sqrt{NM}}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} im[k,l]exp(-j2\pi(\frac{k}{N}u + \frac{l}{M}v))$$

- $u = 0, \ldots, N-1$, $\upsilon_k = \frac{u}{N}$, $\upsilon_k \in [0,1]$;
- $v = 0, \ldots, M-1$, $\upsilon_l = \frac{v}{M}$, $\upsilon_l \in [0,1]$.
  The unit is cycles per unit length (pel^{-1}).

Discrete bi-dimensional Fourier transformation

A frequential decomposition in terms of exponential basis images ...

$$
\begin{aligned}
im[k,l] &= IM[0,0] \times exp(j2\pi(0 \times k + 0 \times l)) + \\
&\quad IM[0,1] \times exp(j2\pi(0 \times k + l)) + \\
&\quad IM[1,0] \times exp(j2\pi(k + 0 \times l)) + \ldots \\
im[k,l] &= IM[0,0] + \\
&\quad IM[0,1]\,(cos(2\pi l) + jsin(2\pi l)) + \\
&\quad IM[1,0]\,(cos(2\pi k) + jsin(2\pi k)) + \\
&\quad IM[1,1]\,(cos(2\pi(k+l)) + jsin(2\pi(k+l))) + \ldots
\end{aligned}
$$

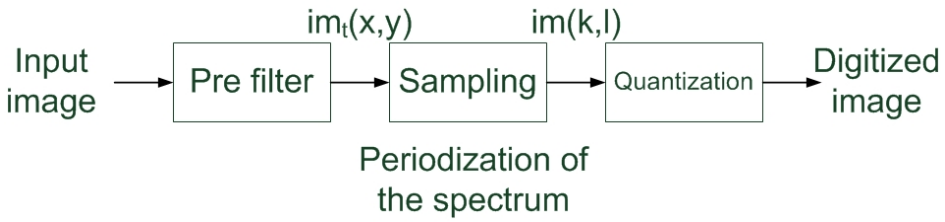
$IM[u,v]$ are the weighting coefficients of the spatial frequencies present in the signal.

Some remarks:

1. A picture having a size of $N \times M$ is a linear combination of $N \times M$ exponential basis images!!!!
2. The DFT of a picture yields a complex picture of the same size;

Some remarks:

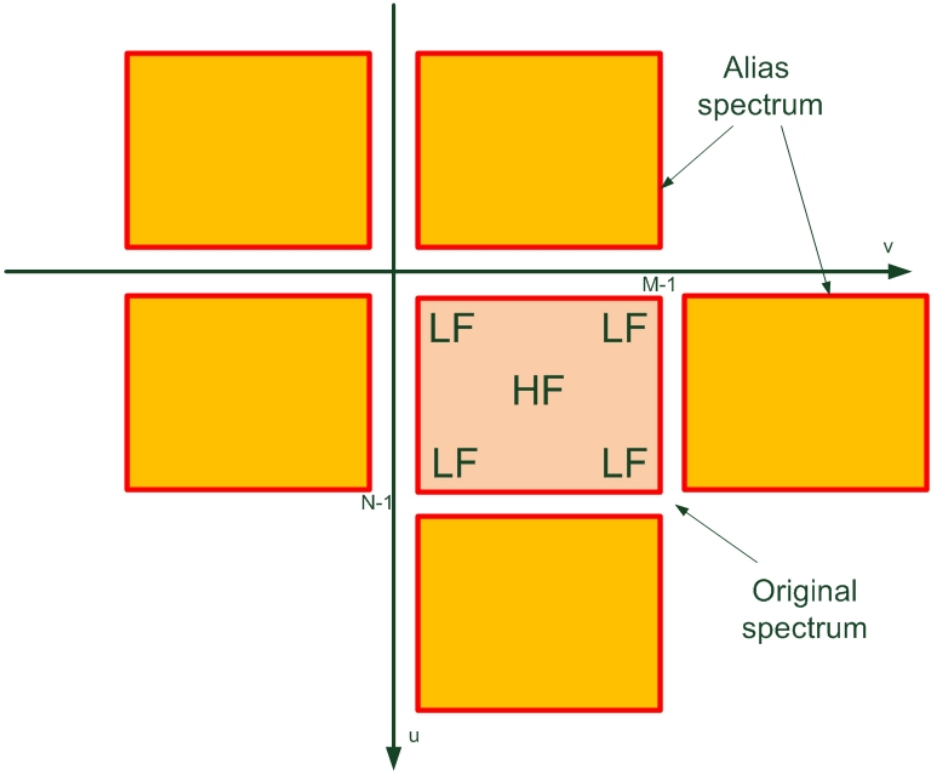
3 $IM[u, v]$ and $im[k, l]$ are periodic of infinite extent. Period = N for k, u. Period = M for l, v:

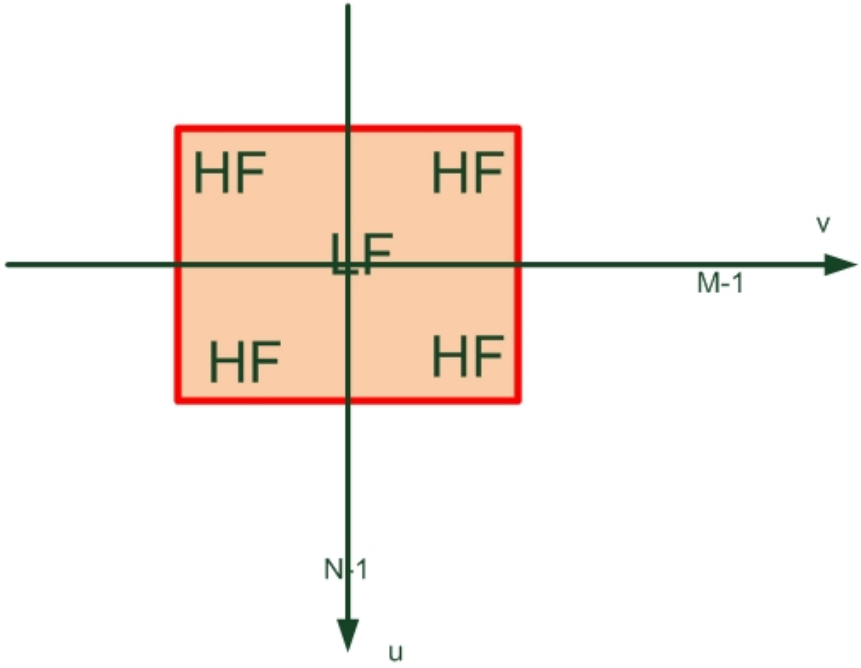


The image im_t is bandlimited and is spatially sampled, for example, with a rectangular array of $N \times M$ pels:

$$im[k, l] = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} im_t(k, l)\delta(k - iT_K)\delta(l - jT_L)$$

Some remarks:





Translation

Translation in the frequency domain

$$im[k,l]exp\left(j\frac{2\pi}{N}(u_0k+v_0l)\right) \xrightarrow{\mathcal{DFT}} IM[u-u_0,v-v_0]$$

Translation in the spatial domain

$$im[k-k_0,l-l_0] \xrightarrow{\mathcal{DFT}} IM[u,v]exp\left(-j\frac{2\pi}{N}(k_0u+l_0v)\right)$$

The multiplication by an exponential function in a representation modifies the origin in the other representation.

Average value

$$\overline{im} = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} im[k,l]$$

$$IM[u,v] = \frac{1}{\sqrt{NM}}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} im[k,l]exp(-j2\pi(\frac{k}{N}u+\frac{l}{M}v))$$

$$IM[0,0] = \frac{1}{\sqrt{NM}}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} im[k,l]$$

$$\overline{im} = \frac{1}{\sqrt{NM}}IM[0,0]$$

$IM[0,0]$ is proportional to the average value of the picture.

Separable transformation = separable basis decomposition

The 2D DFT can be computed as the separable product of two one-dimensional discrete Fourier transforms.

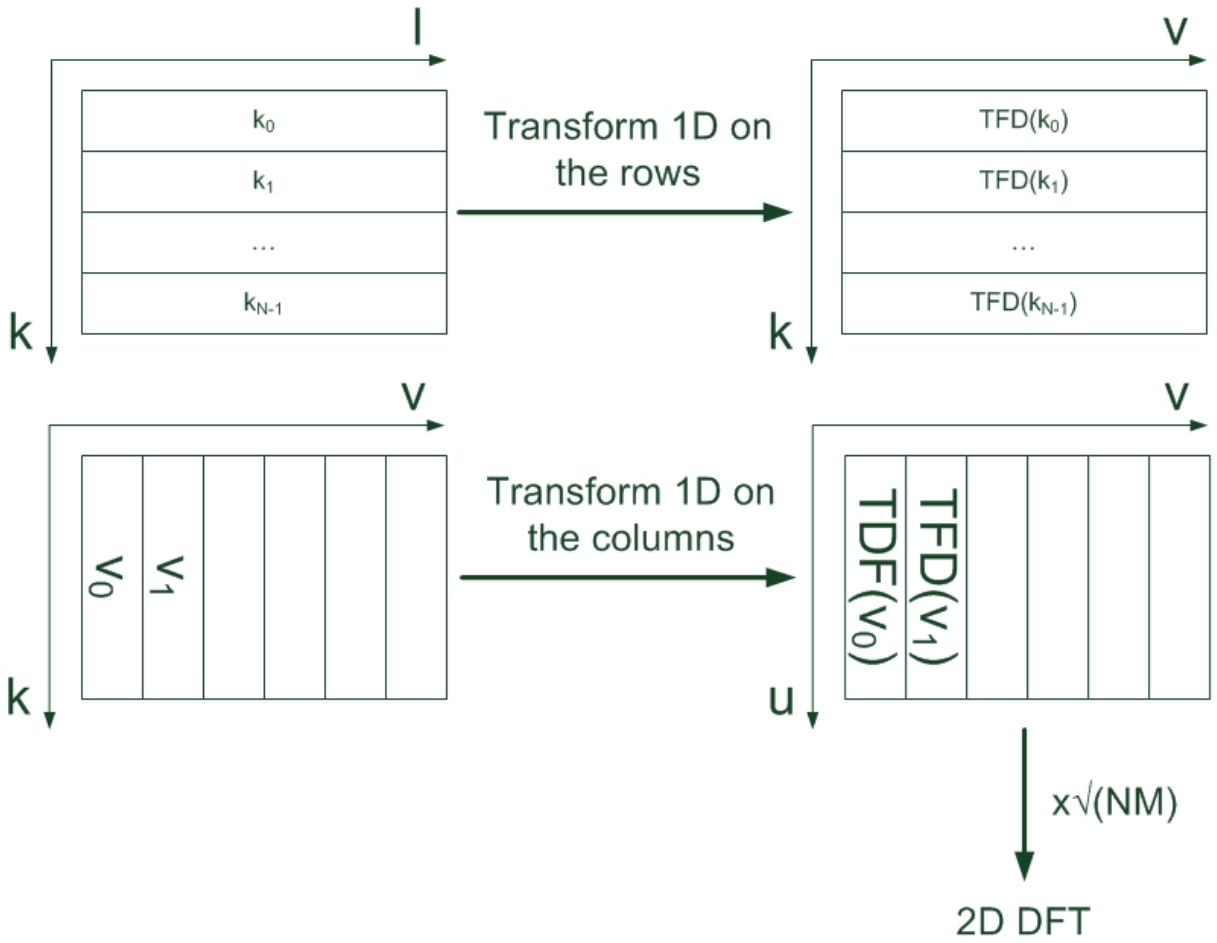
$$IM[u,v] = \frac{1}{\sqrt{NM}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} im[k,l] exp(-j2\pi(\frac{k}{N}u + \frac{l}{M}v))$$

Initialy the family $\left\{e_{k,l}[u,v] = exp(-j2\pi(\frac{k}{N}u + \frac{l}{M}v))\right\}_{0 \leq k < N, 0 \leq l < M}$ is used, but we can also use the family $\{e_k(u)e_l(v)\}_{0 \leq k < N, 0 \leq l < M}$.

$$\begin{aligned}
IM[u,v] &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} exp(-j2\pi \frac{k}{N}u) \left[\frac{1}{\sqrt{M}} \sum_{l=0}^{M-1} im[k,l] exp(-j2\pi \frac{l}{M}v)\right] \\
IM[u,v] &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} exp(-j2\pi \frac{k}{N}u) \left[\sqrt{M} DFT_{1D}(im[k,l])\right] \\
IM[u,v] &= \sqrt{N} DFT_{1D} \left[\sqrt{M} DFT_{1D}(im[k,l])\right]
\end{aligned}$$

We start by computing the Fourier coefficients on the image rows int the basis $\{e_l\}_{0 \leq l < M}$. A second transform is applied on the columns of the transformed images in the basis $\{e_k\}_{0 \leq k < N}$.

Separable transformation = separable basis decomposition



Separable transformation = separable basis decomposition

For a picture having a size $N \times N$, there are three ways to compute its DFT:

- DFT 1D: N^2 multiplications and $N(N-1)$ additions, $\mathcal{O}(N^2)$

$$X[k] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]exp(-j\frac{2\pi}{N}nk)$$

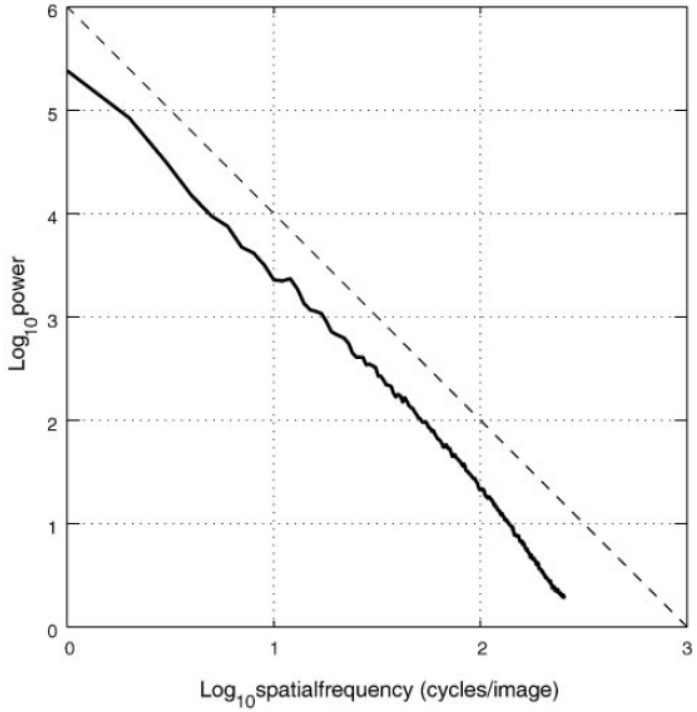
- DFT 2D: $N^2 \times N^2$ multiplications and $N^2(N-1)^2$ additions, $\mathcal{O}(N^4)$

$$IM[u,v] = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} im[k,l]exp(-j2\pi(\frac{k}{N}u+\frac{l}{N}v))$$

- DFT 2D separable transform: $2N \times N^2$ multiplications and $2N \times N(N-1)^2$ additions, $\mathcal{O}(N^3)$

Examples

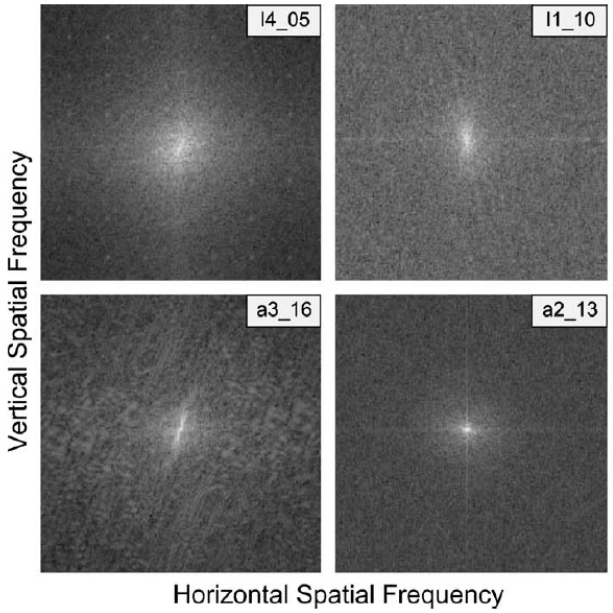
Remark: the spatial frequency amplitude spectra of natural images are well described by a power law $X(f) = \frac{k}{f^{\beta}}$. Estimated average values of β are ranged over 0.9 and 1.2. The power always falls with spatial frequency.



Power spectrum of a natural image (solid line) compared with $\frac{1}{f^2}$ (dashed line).

From E.P. Simoncelli and B.A.Olhausen, Natural Image Statistics and Neural Representation, Annun. Rev. Neuroscience, 2001.

Examples

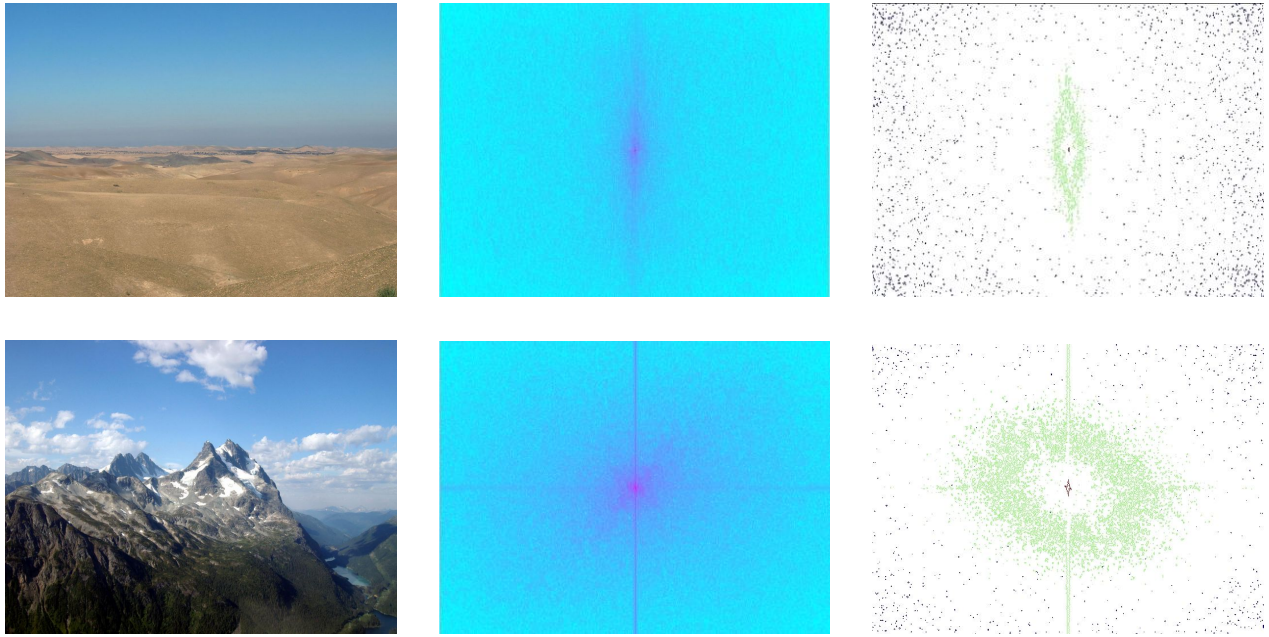


The spectra may be more or less anisotropic.

From R.M. Balboa, N.M. Grzywacz, Power spectra and distribution of contrasts of natural images from different habitats, Vision Research, 43, pp. 2527-2537, 2003.

Examples

From left hand-side to right: original picture, spectrum, contour lines.



Remark: For display purpose, a logarithm is applied on the spectrum.

Scalar Quantization

Principle

The quantization is a process to represent a large set of values with a smaller set.

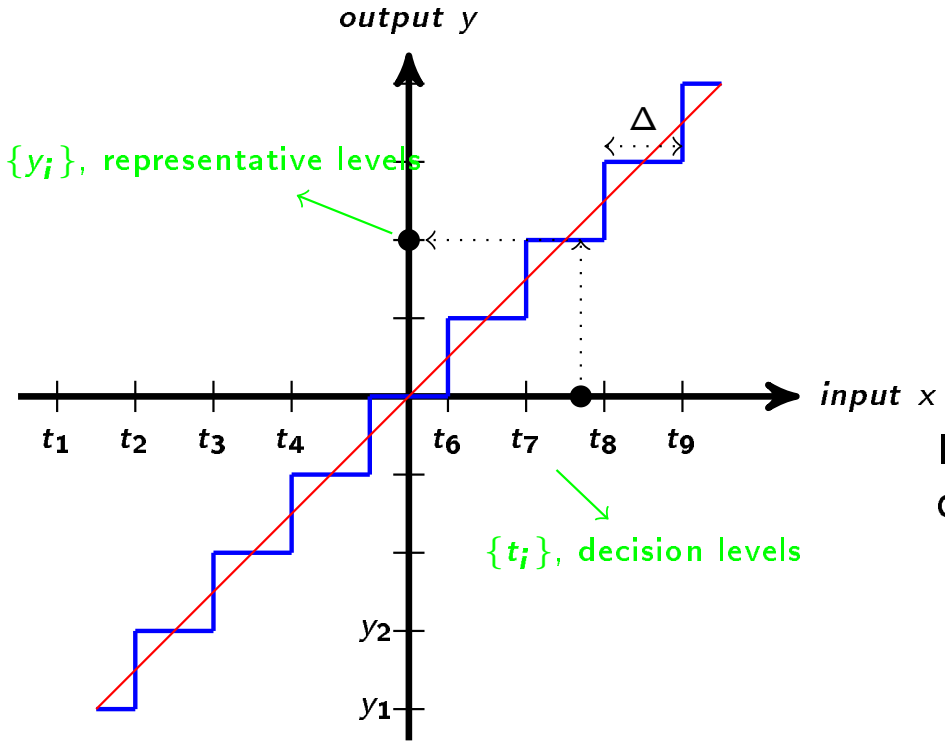
Scalar quantization:

$$\begin{array}{ccc} Q : \mathcal{X} & \longmapsto & \mathcal{C} = \{y_i, i = 1, 2, ...N\} \\ x & & Q(x) = y_i \end{array}$$

- N is the number of quantization level;
- $\mathcal{X}$ could be continue (ex: $\mathcal{R}$) or discret;
- $\mathcal{C}$ is always discret (codebook,dictionnary);
- $card(\mathcal{X}) > card(\mathcal{C})$;
- As $x \neq Q(x)$, we will lost some information (lossy compression).

Uniform quantization

In the uniform quantization, the quantization step size Δ is fixed, no matter what the signal amplitude is.



Definition:

- The quantization thresholds are uniformly distributed:
  $\forall i \in \{1, 2, ..., N\},\ t_i - t_{i+1} = \Delta$
- The output values are the center of the quantization interval:
  $\forall i \in \{1, 2, ..., N\},\ y_i = \frac{t_i + t_{i+1}}{2}$

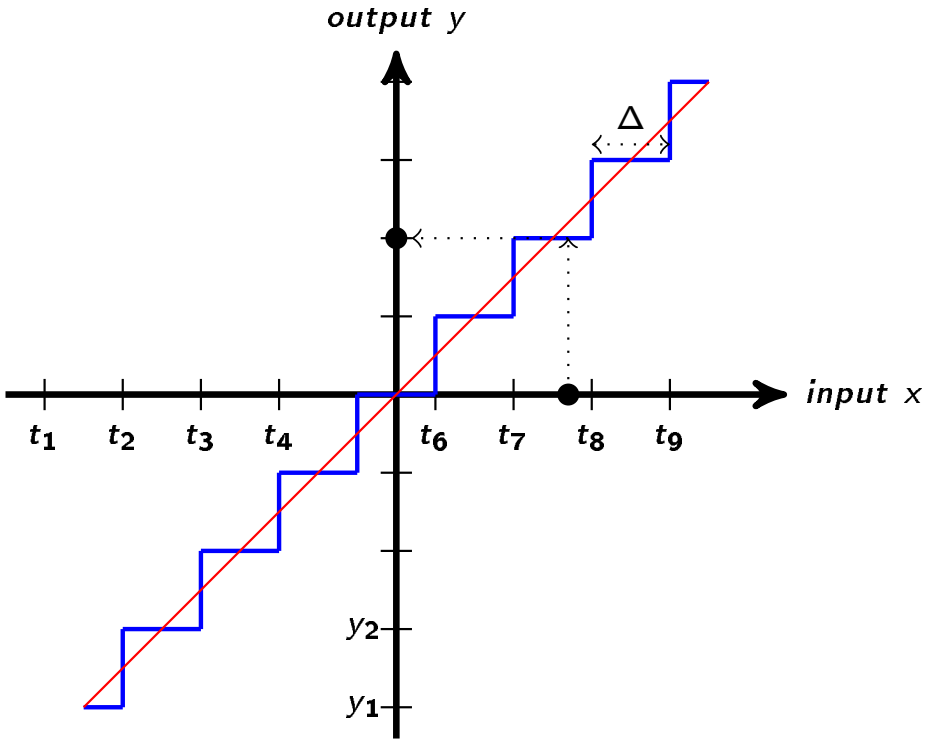
Example of the nearest neighborhood quantizer (see on the left):

$$Q(x) = \Delta \times \left\lfloor \frac{x}{\Delta} + 0.5 \right\rfloor.$$

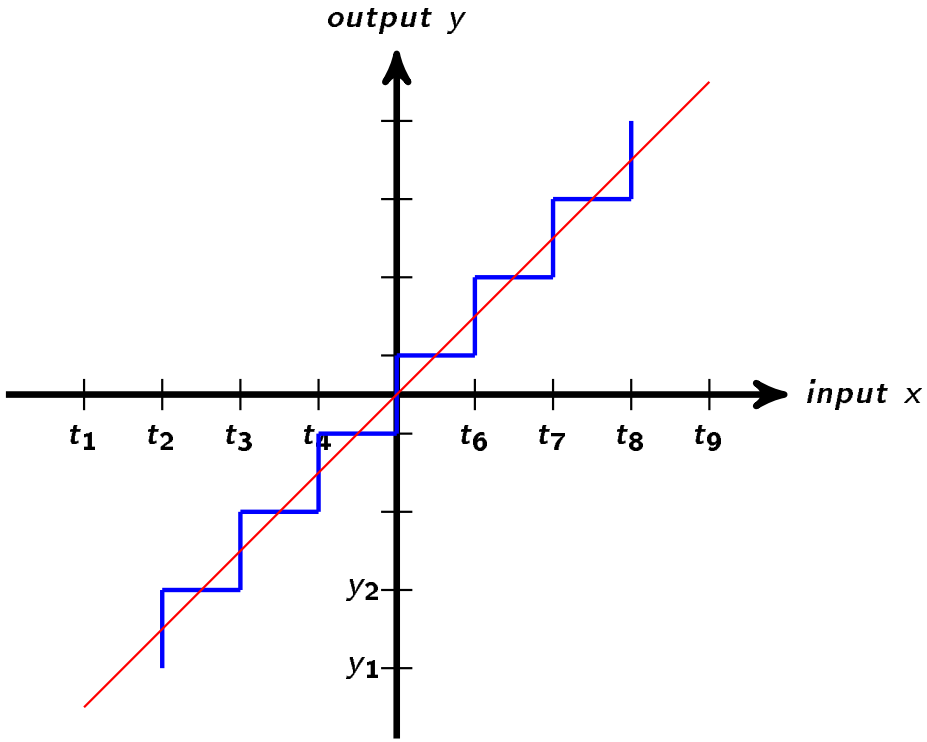
Uniform quantization

A uniform quantization is completely defined by the number of levels, the quantization step and if it is a mid-step or mid-riser quantizer.

A mid-step quantizer
Zero is one of the representative levels y_k

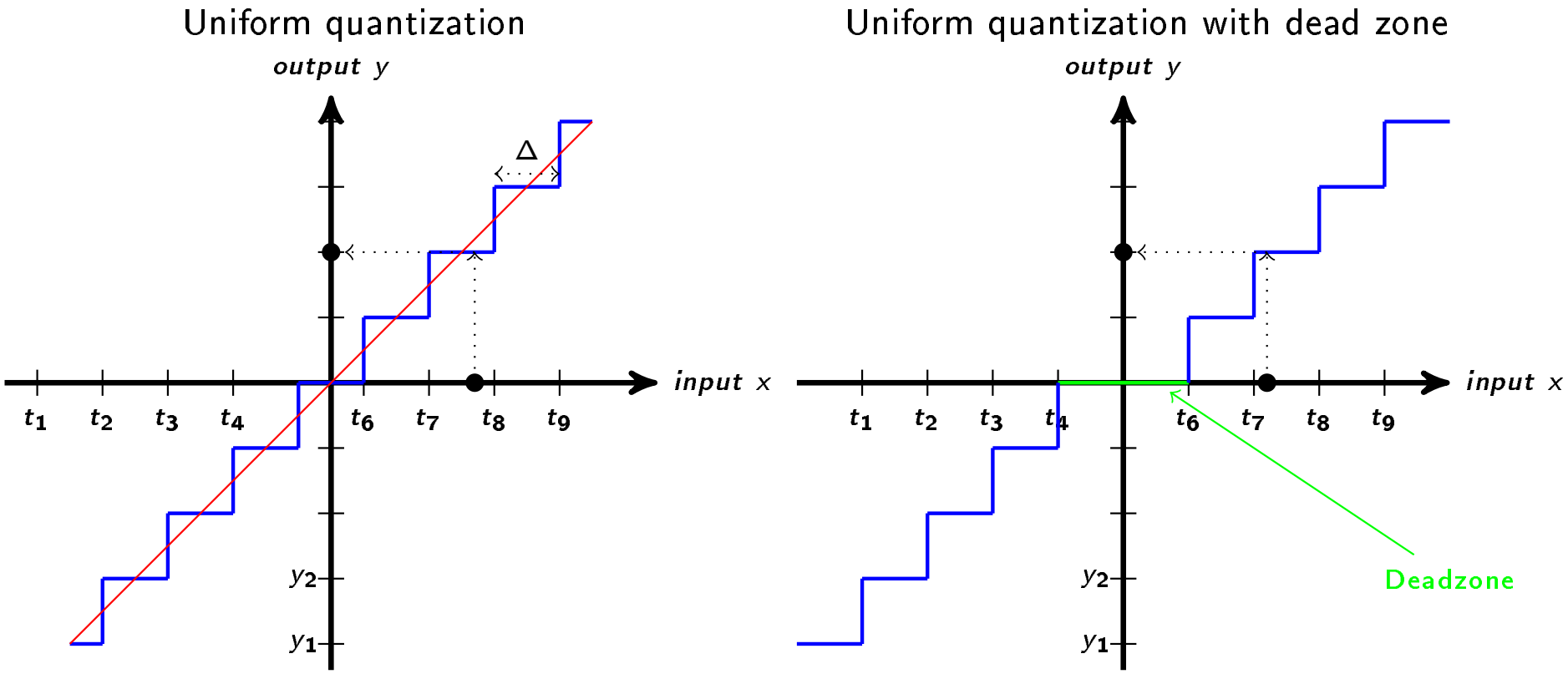


A mid-riser quantizer
Zero is one of the decision levels t_k



Usually, a mid-riser quantizer is used if the number of representative levels is even and a mid-step quantizer if the number of level is odd.

Uniform quantization with dead zone



Interest:
To remove small coefficients by favouring the zero value. Increase the coding efficiency with a small visual impact.

Example of an uniform quantization

Example (Original picture quantized to 8, 7, 6, 4 bits/pixels)

(a) Original (8 bits per pixel)

(b) 7 bpp ($\Delta = 2$)

(c) 6 bpp ($\Delta = 4$)

(d) 4 bpp ($\Delta = 8$)

Optimal quantization

Optimal quantization

An optimal quantization of a random variable X having a probability distribution $p(x)$ is obtained by a quantizer that minimises a given metric:

- $\mathcal{L}^{\inf}$-norm: $D = \max |X - Q(X)|$;
- $\mathcal{L}^1$-norm: $D = E\left[|X - Q(X)|\right]$;
- $\mathcal{L}^2$-norm: $D = E\left[(X - Q(X))^2\right]$ called the Mean-Square Error (MSE). This is the most used.

Considering the MSE, we have:

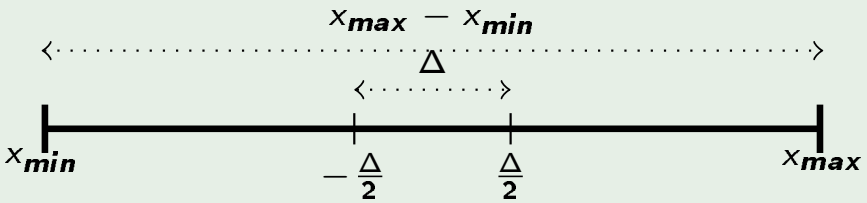
- if the random variable X is continue: $D = \int_{x_{min}}^{x_{max}} p(x)(x - Q(x))^2 dx$;
- if the random variable X is discret: $D = \sum_{k=1}^{n} p(x_k)(x_k - Q(x_k))^2$.

Example

Example (Quantization error)

Hypothesis:

- Uniform quantization with a quantization step Δ;
- $p(x)$ is a uniform probability distribution of a random variable X;
- N is the number of representative levels.



Quantization step: $\Delta = \frac{x_{max} - x_{min}}{N}$

Probability distribution: $p(x) = \frac{1}{x_{max} - x_{min}} = \frac{1}{N\Delta}$

Example

Example (Quantization error)

$$
\begin{aligned}
D &= \int_{x_{min}}^{x_{max}} p(x)(x - Q(x))^2 dx \\
&= N \times \int_{-\Delta/2}^{\Delta/2} p(x)(x - 0)^2 dx, \text{ 0 the mid-point} \\
&= N \times \int_{-\Delta/2}^{\Delta/2} \frac{1}{N\Delta} x^2 dx \\
&= \frac{\Delta^2}{12}
\end{aligned}
$$

Optimal quantizer in a nutshell

Uniform quantizer is not optimal if the source is not uniformly distributed.

Optimal quantizer

To find the decision levels $\{t_i\}$ and the representative levels $\{y_i\}$ to minimize the distortion D.

To reduce the MSE, the idea is to decrease the bin's size when the probability of occurrence is high and to increase the bin's size when the probability is low. For N representative levels and with a probability density $p(x)$, the distortion is given by:

$$\begin{aligned} D &= \int_{x_{min}}^{x_{max}} p(x)(x - Q(x))^2 dx \\ &= \sum_{k=1}^{N-1} \int_{t_k}^{t_{k+1}} p(x)(x - y_k)^2 dx \end{aligned}$$

Optimal quantizer in a nutshell

The optimal $\{t_i\}$ and $\{y_i\}$ satisfy:

$$\frac{\partial D}{\partial t_i} = 0 \text{ and } \frac{\partial D}{\partial y_i} = 0$$

Lloyd-Max quantizer

- $\frac{\partial D}{\partial t_i} = 0 \Rightarrow t_i = \frac{y_i + y_{i+1}}{2}$.
  t_i is the midpoint of y_i and y_{i+1}.
- $\frac{\partial D}{\partial y_i} = 0 \Rightarrow y_i = \frac{\int_{t_{i-1}}^{t_i} p(x)x dx}{\int_{t_{i-1}}^{t_i} p(x) dx}$.
  y_i is the centroid of the interval $[t_{i-1}, t_i]$.

$\Rightarrow$ given the $\{t_i\}$, we can find the corresponding optimal $\{y_i\}$.
$\Rightarrow$ given the $\{y_i\}$, we can find the corresponding optimal $\{t_i\}$.

How can we find the optimal $\{t_i\}$ and $\{y_i\}$ simultaneously?

Lloyd-Max algorithm in a nutshell

The Lloyd-Max algorithm is an algorithm for finding the representative levels $\{y_i\}$ and the decision levels $\{t_i\}$ to meet the previous conditions, with no prior knowledge.

Principle of the iterative process

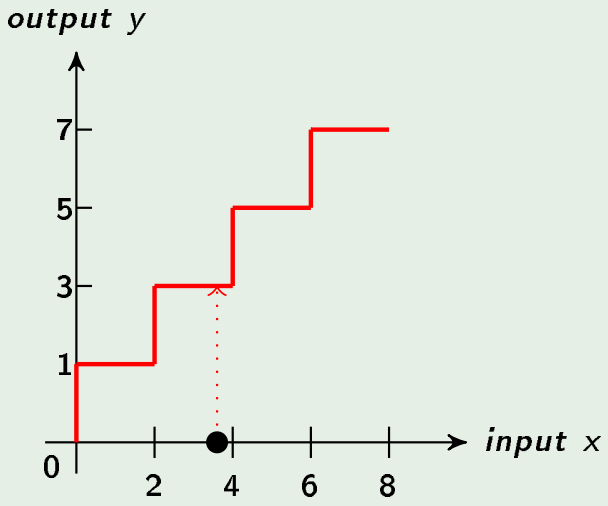
1. The iterative process starts for $k = 0$ with a set of initial values for the representative levels $\left\{y_1^{(0)}, \ldots, y_N^{(0)}\right\}$.
2. New values for decision levels are determined $t_i^{(k+1)} = \frac{y_i^{(k)} + y_{i+1}^{(k)}}{2}$
3. Compute the distortion $D^{(k)}$ and the relative errors $\delta^{(k)}$;
4. Depending on the stopping criteria ($\delta^{(k)} < \epsilon$), stop the process or update the representative levels $y_i^{(k+1)} = \frac{\int_{t_{i-1}^{(k+1)}}^{t_i^{(k+1)}} xp(x)dx}{\int_{t_{i-1}^{(k+1)}}^{t_i^{(k+1)}} p(x)dx}$ and go back to step 2.

Uniform, Semi-uniform and optimal quantizer

Suppose we have the following 1D discrete signal:

$$X = \{0, 0.01, 2.8, 3.4, 1.99, 3.6, 5, 3.2, 4.5, 7.1, 7.9\}$$

Example (Uniform quantizer: $N = 4$, Mid-riser and $\Delta = 2$)



$$t_i \in \mathcal{T} = \{t_0 = 0, t_1 = 2, t_2 = 4, t_3 = 6, t_4 = 8\}$$

$$r_i \in \mathcal{R} = \{r_0 = 1, r_1 = 3, r_2 = 5, r_3 = 7\}$$

Check the definition of a uniform quantizer...

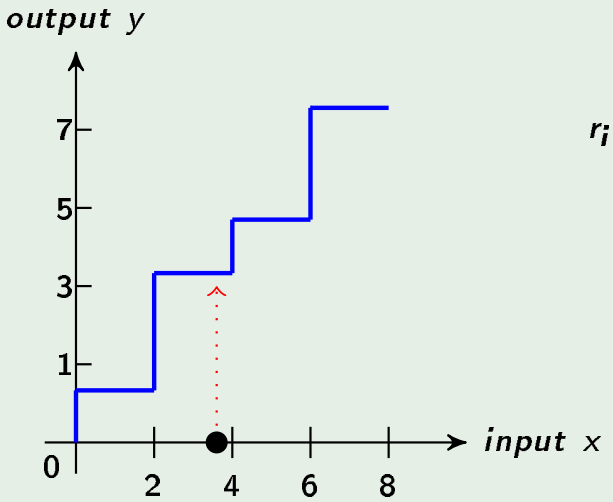
Quantized vector $X = \{1, 1, 3, 3, 1, 3, 5, 3, 5, 7, 7\}$ (MSE = 0.42)

Uniform, semi-uniform and optimal quantizer

Suppose we have the following 1D discrete signal:

$$X = \{0, 0.01, 2.8, 3.4, 1.99, 3.6, 5, 3.2, 4.5, 7.1, 7.9\}$$

Example (Semi-uniform quantizer: $N = 4$, Mid-riser and $\Delta = 2$)



$$t_i \in \mathcal{T} = \{t_0 = 0, t_1 = 2, t_2 = 4, t_3 = 6, t_4 = 8\}$$

$$r_i \in \mathcal{R} = \{r_0 = 2/3, r_1 = 3.25, r_2 = 4.75, r_3 = 7.5\}$$

Quantized vector $X = \{2/3, 2/3, 3.25, 3.25, 2/3, 3.25, 4.75, 3.25, 4.75, 7.5, 7.5\}$
(MSE = 0.31)

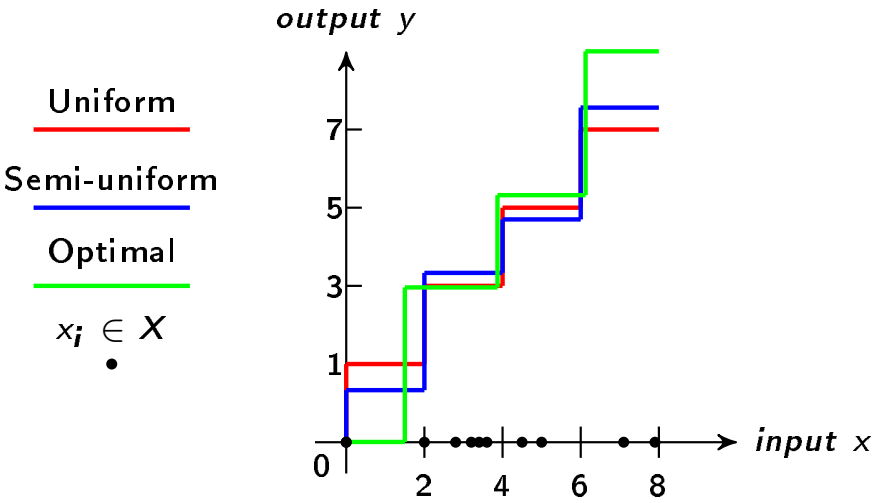
Uniform, semi-uniform and optimal quantizer

Suppose we have the following 1D discrete signal:

$$X = \{0, 0.01, 2.8, 3.4, 1.99, 3.6, 5, 3.2, 4.5, 7.1, 7.9\}$$

Example (Lloyd-Max Algorithm)

Given that $\mathcal{T} = \{0, 1.5, 3.87, 6.125, 8\}$ and $\mathcal{R} = \{0.005, 2.998, 4.75, 7.5\}$, the quantized vector is
$X = \{0.005, 0.005, 2.998, 2.998, 2.998, 2.998, 4.75, 2.998, 4.75, 7.5, 7.5\}$ (MSE=0.18).



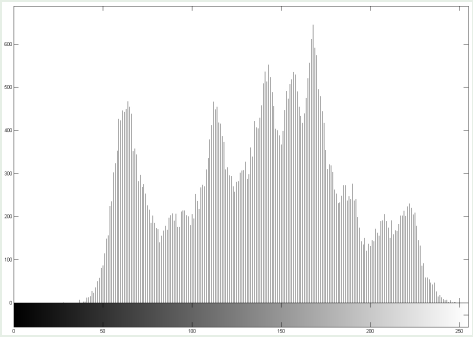
Uniform vs optimal quantizer

Example (N=5)

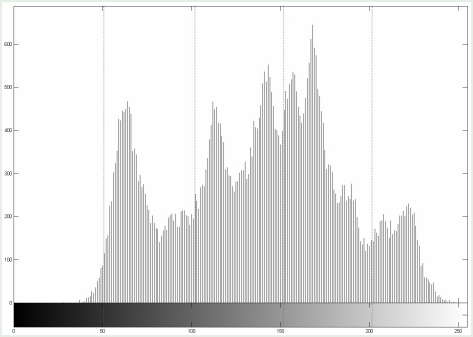
(a) Original

(b) Uniform

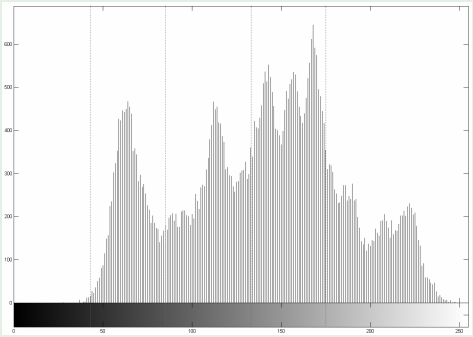
(c) Optimal



(d) Histo.



(e) Decision levels



(f) Decision levels

Distortion/quality assessment

Taxonomy

Distortion/quality metrics can be divided into 3 categories:

1. Full-Reference metrics (FR) for which both the original and the distorted images are required (benchmark, compression) ;
2. Reduced-Reference metrics (RR) for which a description of the original and the distorted image is required (network monitoring);
3. No-Reference (NR) metrics for which the original image is not required (network monitoring).

Each category can be divided into two subcategories: metrics based on signal fidelity and metrics based on properties on the human visual system.

Peak Signal to Noise Ratio (PSNR)

The PSNR is the most popular quality metric. This simple metric just calculates the mathematical difference between each pixel of the degraded image and the original image.

PSNR:

Let I and D the original and impaired images, respectively. These images having a size of M pixels are coded with n bits.

$$PSNR = 10 log_{10} \frac{((2^n - 1)^2)}{MSE} \text{ dB},$$

with the Mean Squared Error $MSE = \frac{\sum_{(x,y)} (I(x,y) - D(x,y))^2}{M}$.

A high value indicates that the amount of impairment is small. A small value indicates that there is a strong degradation.

Peak Signal to Noise Ratio (PSNR)

The PSNR is not always well correlated with the human judgment (MOS Mean Opinion Score). The reason is simple: this metric does not take into account the properties of the human visual system.

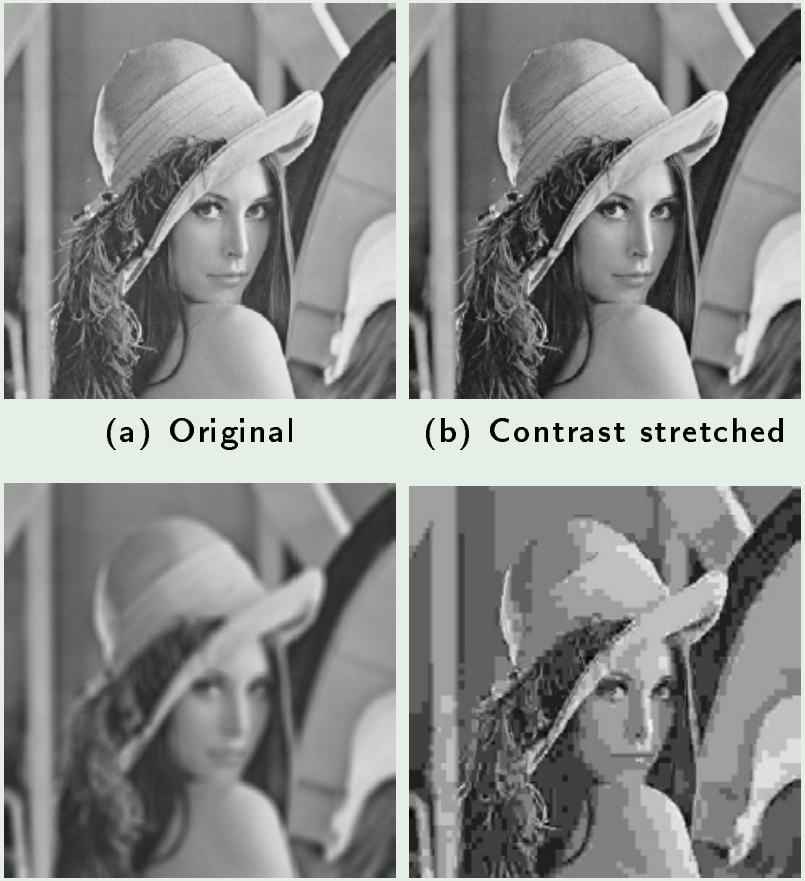
Example

(a) (b) (c) Original (d) Original+uniform noise

(a) and (b) from [?]: impact of Gabor patch on our perception.

Peak Signal to Noise Ratio (PSNR)

Example (These three pictures have the same PSNR...)

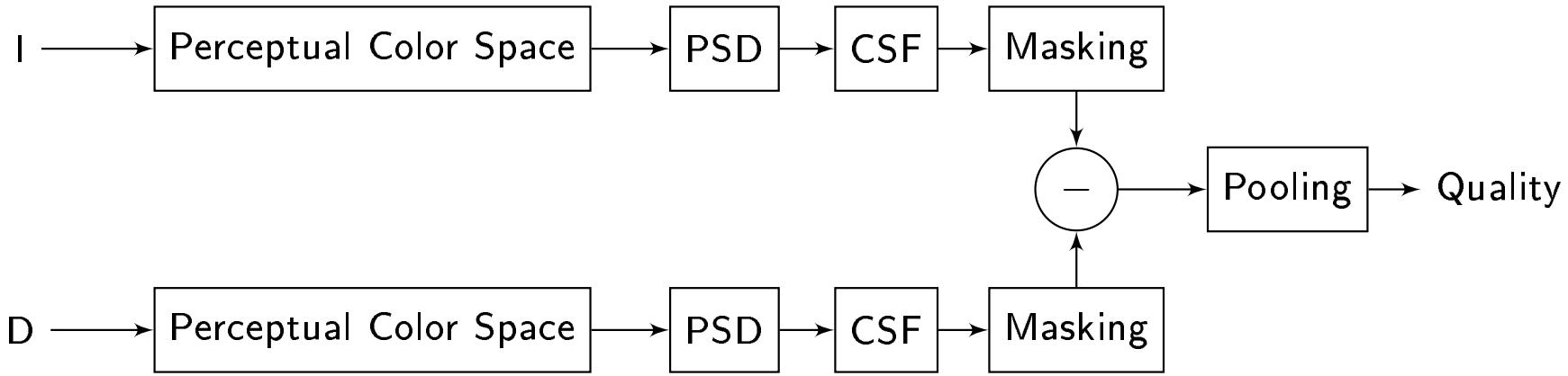


(a) Original (b) Contrast stretched

(c) Blur (d) JPEG

Metric based on the error visibility

For this type of metric, the behavior of the visual cells are simulated:

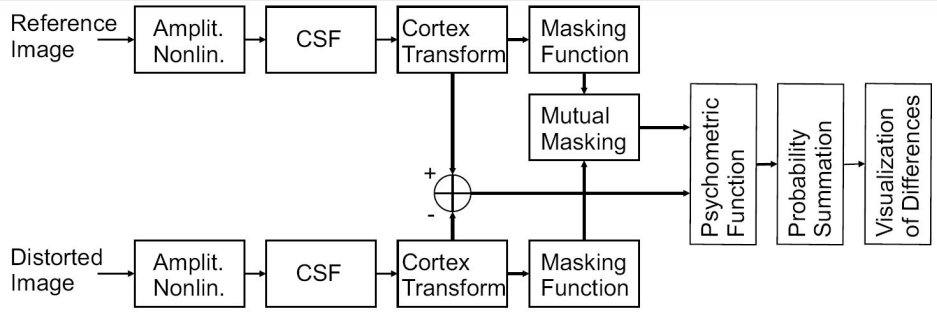


- PSD: Perceptual Subband Decomposition (Wavelet, Gabor, Fourier);
- CSF: Contrast Sensitivity Function.

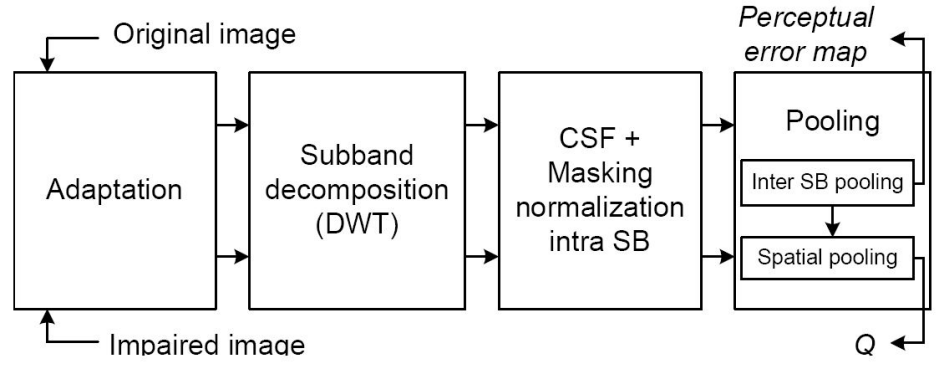
Metric based on the error visibility

Example

- VDP (Visible Differences Predictor) [Daly,93]:



- WQA (Wavelet-based Quality Assessment) [Ninassi et al.,08a]



- VQM (Video Quality Model) [Pinson et al.,04]...

Metric based on the structural similarity

SSIM standing for Structural Similarity index. Image degradations are considered here as perceived structural information loss instead of perceived errors [Wang et al.,04a].

Let I and D the original and the degraded images, respectively.

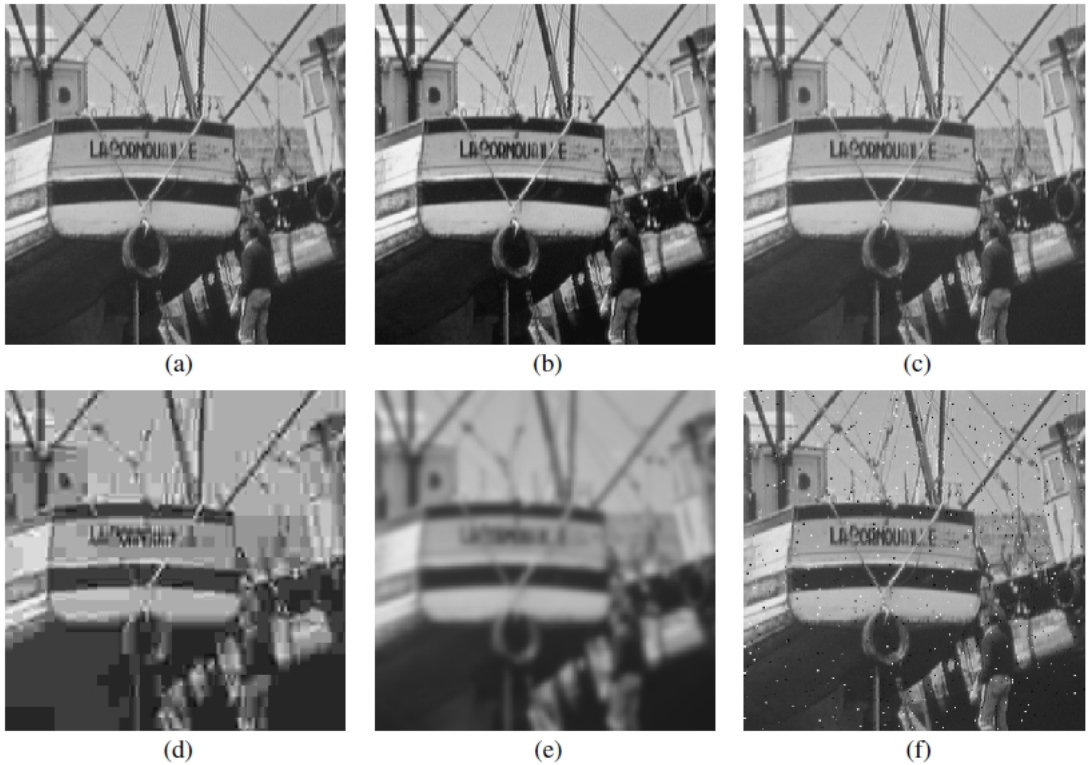
$$S(x,y) = l(x,y) \times c(x,y) \times s(x,y) \tag{1}$$

- The luminance comparison measure $l(x,y) = \dfrac{2\mu_x\mu_y}{\mu_x^2+\mu_y^2}$
- The contrast comparison measure $c(x,y) = \dfrac{2\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}$
- The structural comparison measure $s(x,y) = \dfrac{\sigma_{xy}}{\sigma_x\sigma_y}$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}$$

$\overline{SSIM} \to 1$, the best quality and $\overline{SSIM} \to 0$ indicates a poor quality.

Metric based on the structural similarity

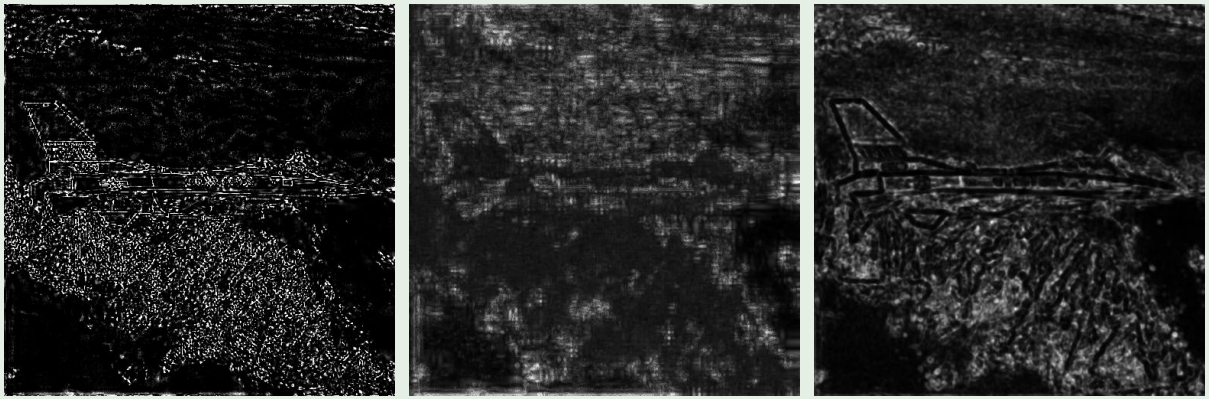


All pictures have the same MSE. (a) Original image; (b) Contrast-stretched image, $MSSIM = 0.9168$; (c) Mean-shifted image, $MSSIM = 0.99$; (d) JPEG, $MSSIM = 0.6949$; (e) Blurred image, $MSSIM = 0.7052$; (f) Impluse noise, $MSSIM = 0.7748$. Extracted from [Wang et al.,04a].

Example of distortion maps

Example

(a) Original (b) Degraded



(c) MSE (d) WQA (e) SSIM

Distortion/quality assessment

Conclusion

- Image Transformation: global to point;
- Fourier transformation;
- Time sampling with Shannon's theorem;
- Scalar quantization, a lossy process;
- Quality and distortion.

Suggestion for further reading...

- S. Daly. The visible differences predictor: An algorithm for the assessment of image fidelity. Digital Images and Human Vision, pp. 179-206, 1993, MIT Press.
- A. Ninassi, O. Le Meur, P. Le Callet, and D. Barba. On the performance of human visual system based image quality assessment metric using wavelet domain. Proc. SPIE Human Vision and Electronic Imaging XIII., Vol. 6806, pp. 680610-12, 2008.
- M. Pinson and S. Wolf. A new standardized method for objectively measuring video quality. IEEE Trans. Broadcasting, Vol. 50, N. 3, pp. 312-322, 2004.
- Z. Wang, A. C. Bovik, H.R. Sheikh and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE Trans. on Image Processing, Vol. 13, pp. 600-612, 2004.