# Information Sciences: Signal Processing

## Lecture 1: Signals and Systems

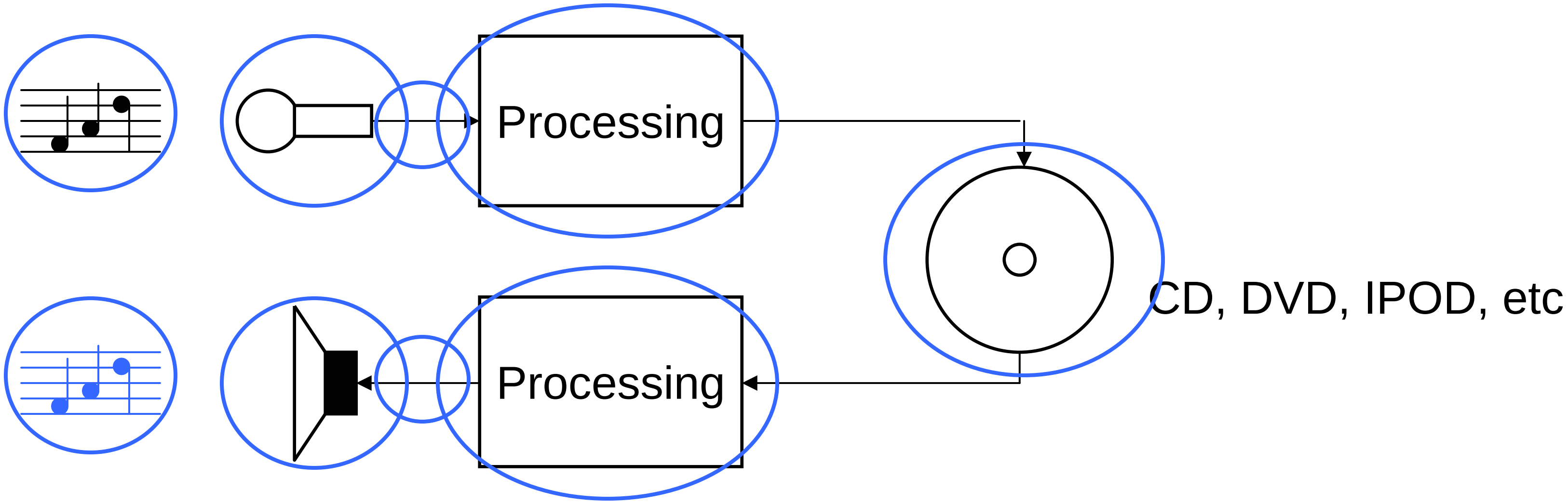
- Goal of the course:

  Record audio (music, voice) on a CD.. and reproduce it!

- The first step… We need some “hardware” to do the job (e.g. microphones, loudspeakers, and other “stuff”) We will call a “system” each piece of hardware.
- Each system transforms “objects” that we call “signals.” We will need a mathematical representation for signals and the transformations introduced by the systems

In this lecture we will describe what signal and systems are and the mathematical representation that we use to work with them

- A first rough idea of the hardware…




- This is what we want to record/reproduce, i.e. sound. We call this a signal
- Microphones and loudspeakers transform sound to electric quantities. They are systems
- These electric quantities are also signals. Systems transform signals into other signals!
- We need some more processing, i.e. other systems to transform the electrical signals into “holes” that we can drill on the CD…
- We can see a CD as an object that takes at the input what we record and gives as an output what we read. This is also a system

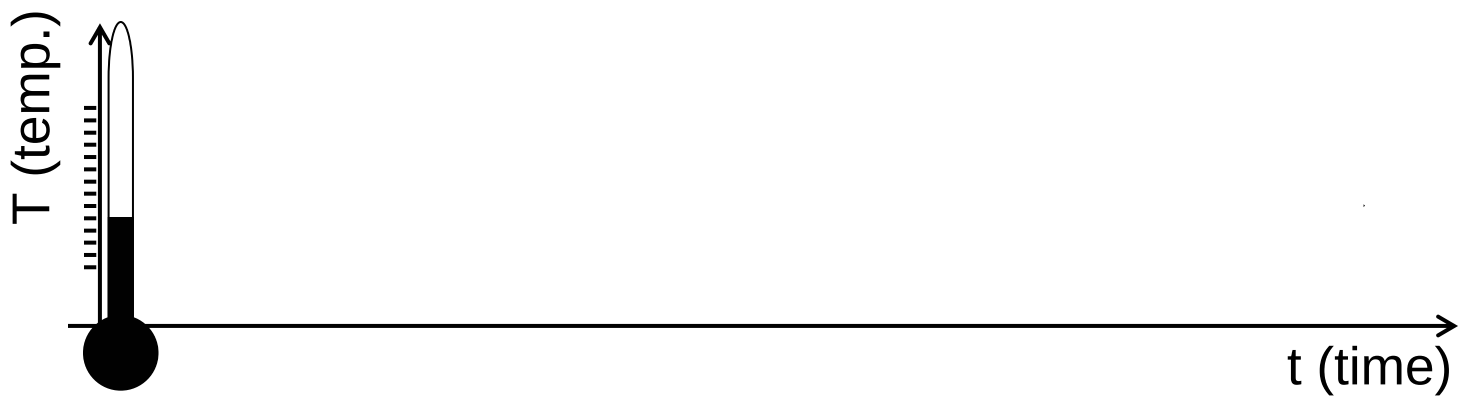
- Let's see more in detail what a signal is:
  - We use signals to represent sounds, electric signals, holes on a disk. We can say that…

  I. A signal represents the measure of a certain physical quantity
- Hence, we can use signals to represent any measure we want (the speed of a car, the temperature of the lake, the amount of remaining snow in a glacier, etc.
- In order to acquire signals we need transducers. A transducer transforms the signal of interest into another one that we can process (typically an electrical signal). Loudspeakers and microphones are transducers, but also thermometers, tachymeters, anemometers, barometers, antennas, tape heads, etc.

- A second fact concerning signals..
  - A physical quantity constant over time is not very interesting (there is not much processing to do)

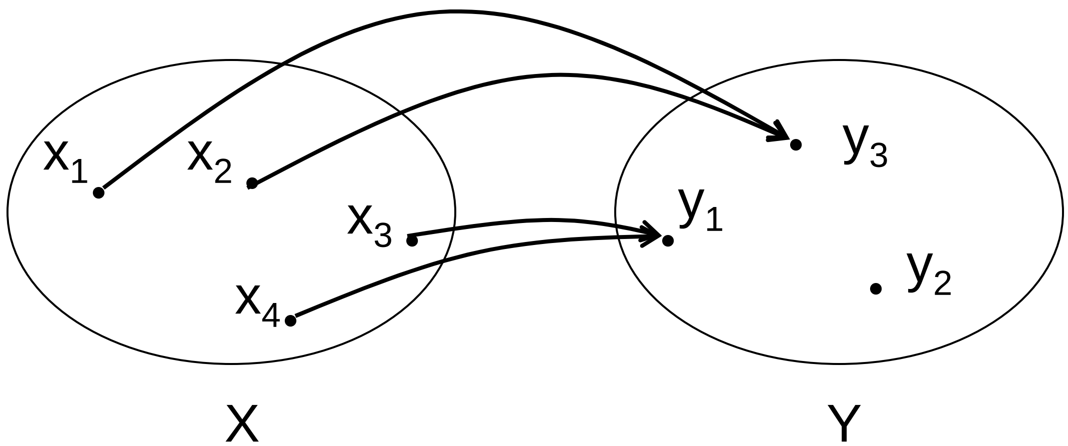
  II. We are interested in the evolution of a signal as a function of time
- Example, the temperature of the air:

We can use a thermometer and record temperature over time




We can represent a signal as a function of time!

- Remember what a function is…
  - It is a map that associate elements of a set to elements of a second set:

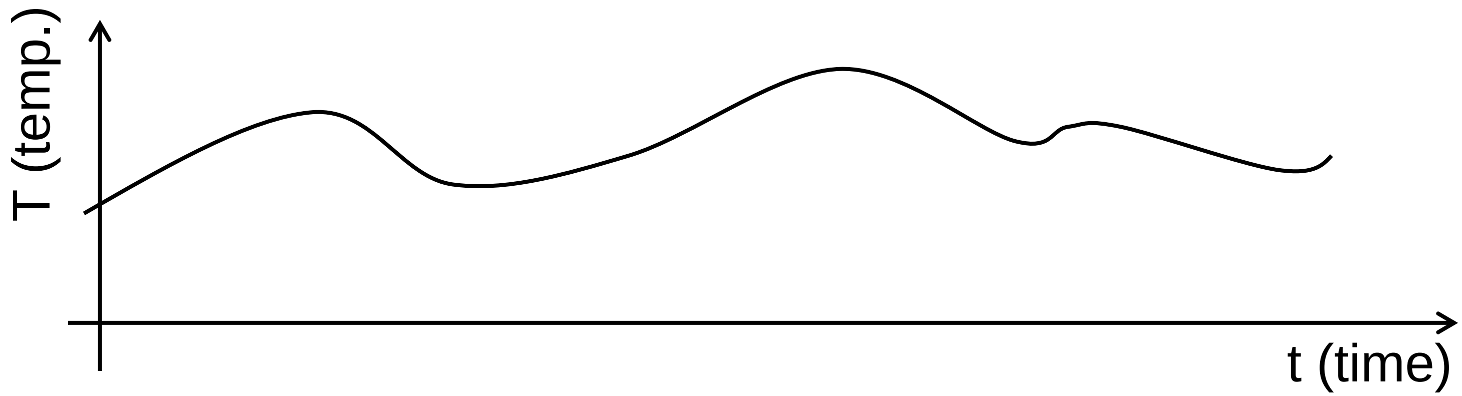



  - We write

$$\begin{aligned} f: \quad X &\to Y \\ x &\mapsto y \end{aligned}$$

  - X and Y are called the domain and the codomain of the function

- In the case of temperature:

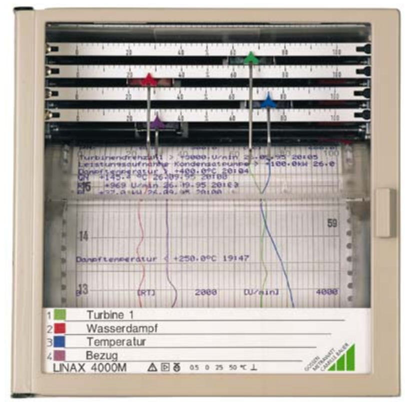



  - The temperature T is a function of time t. What are the domain and the codomain? Time and temperature are real quantity, we can define the “temperature” signal, T(t) as

$$T: \quad \mathbb{R} \to \mathbb{R}$$
$$t \mapsto T(t)$$

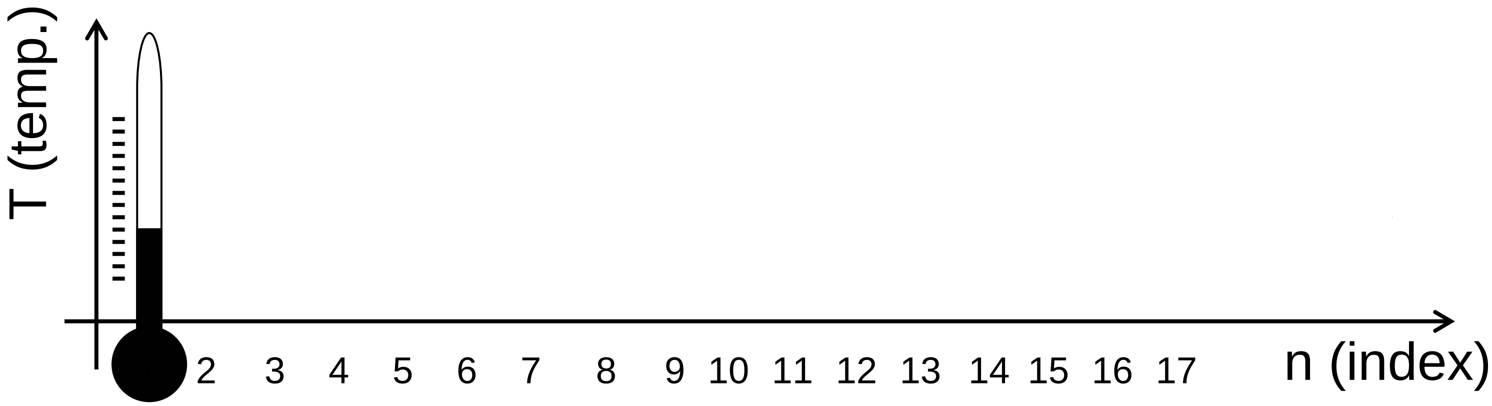
  - Remark that the function does not need to cover entirely the codomain, e.g. temperature cannot be any real number (it cannot be less than -273.15 C)

- Can we record and process the temperature signal?
  - We can use some electromechanical systems, e.g.:


  - This is not very sophisticate! It is very difficult to compute simple quantities like the average temperature on a certain time interval
  - A computer can only perform a finite number of operations per second. It is not possible to process an arbitrary temperature signal

- To simplify acquisition and processing, we can measure temperature every $T_S$ seconds:




- Now the temperature is defined on a discrete set of points. We number these points and call n the index of the measurement.
- We represent the measurements by defining the function:

$$\begin{aligned} \bar{T}: \quad & \mathbb{Z} \to \mathbb{R} \\ & n \mapsto \bar{T}(n) \end{aligned}$$

Where $\mathbb{Z}$ is the set of the integer numbers

- To summarize, we have seen two ways to represent the evolution of temperature with time:

$$\begin{aligned} T: \quad & \mathbb{R} \to \mathbb{R} \\ & t \mapsto T(t) \end{aligned}$$

Temperature is given for each time $t \in \mathbb{R}$

We call this a continuous-time signal

$$\begin{aligned} \bar{T}: \quad & \mathbb{Z} \to \mathbb{R} \\ & n \mapsto \bar{T}(n) \end{aligned}$$

Temperature is given for a discrete set of time instants, indexed by $n \in \mathbb{Z}$

We call this a discrete-time signal

- Remark: in the analysis course you will learn the definition of continuous function. Be careful! Do not confuse that definition with that of continuous-time signal! Here, continuous-time refers only to the domain of the signal, that is the set of real numbers, and not to the "graph" of the signal, that can be arbitrary, even not continuous

- What is the relation between a continuous time signal and a discrete time signal associated to the same physical quantity? If we give the index n=0 to the measurement corresponding to t=0, then

$$n \to t = nT_S$$

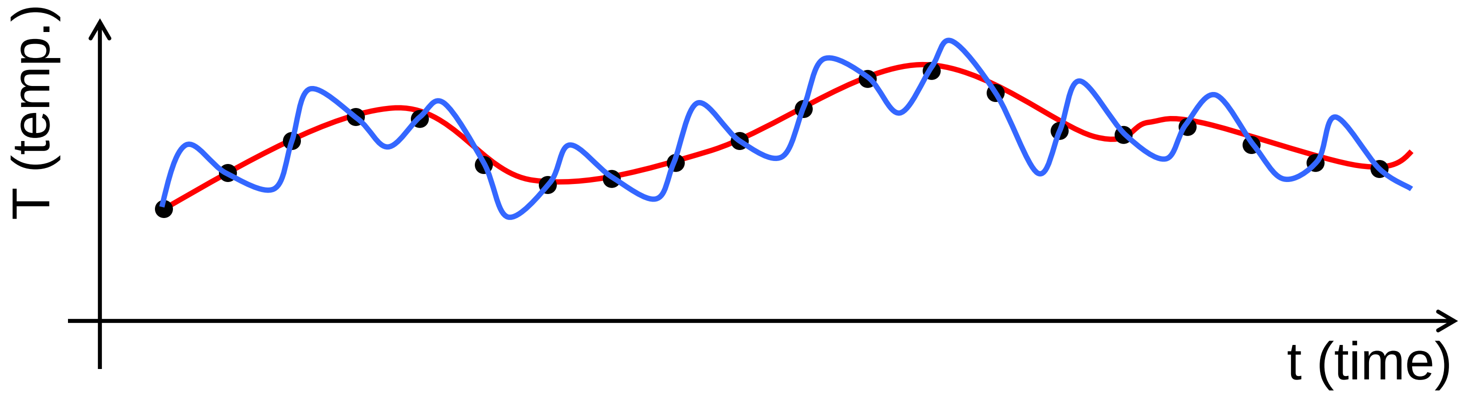
and

$$\bar{T}(n) = T(nT_S)$$

- We call this operation sampling of the continuous time signal $T(t)$, i.e. we say that $\bar{T}(n)$ is a sampled version of $T(t)$

- The quantity $T_S$ determines how often we take a measurement, we call it the sampling period
- The quantity $f_S = 1/T_S$ is the number of measurements per second, we call it the sampling frequency (we measure it in Hertz)

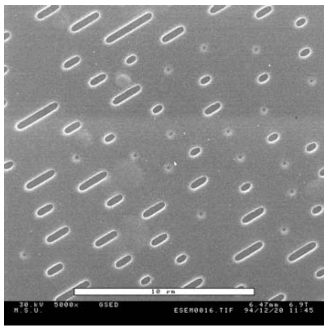
- Is the sampled signal equivalent to the continuous time signal? i.e., is there a one to one correspondence between the sampled signal and the continuous time signal?



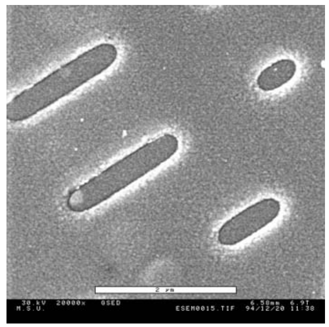

No! There are many continuous-time signals that give the same sampled signal. Discrete-time signals seem “less powerful” than continuous-time signals

- However, intuitively, if the sampling period is small enough the discrete-time signal captures most of the trend of the continuous time signal (more on this later)

- Both continuous time and discrete-time signals take (in general) real values. Can we record real values on a CD?



x5000

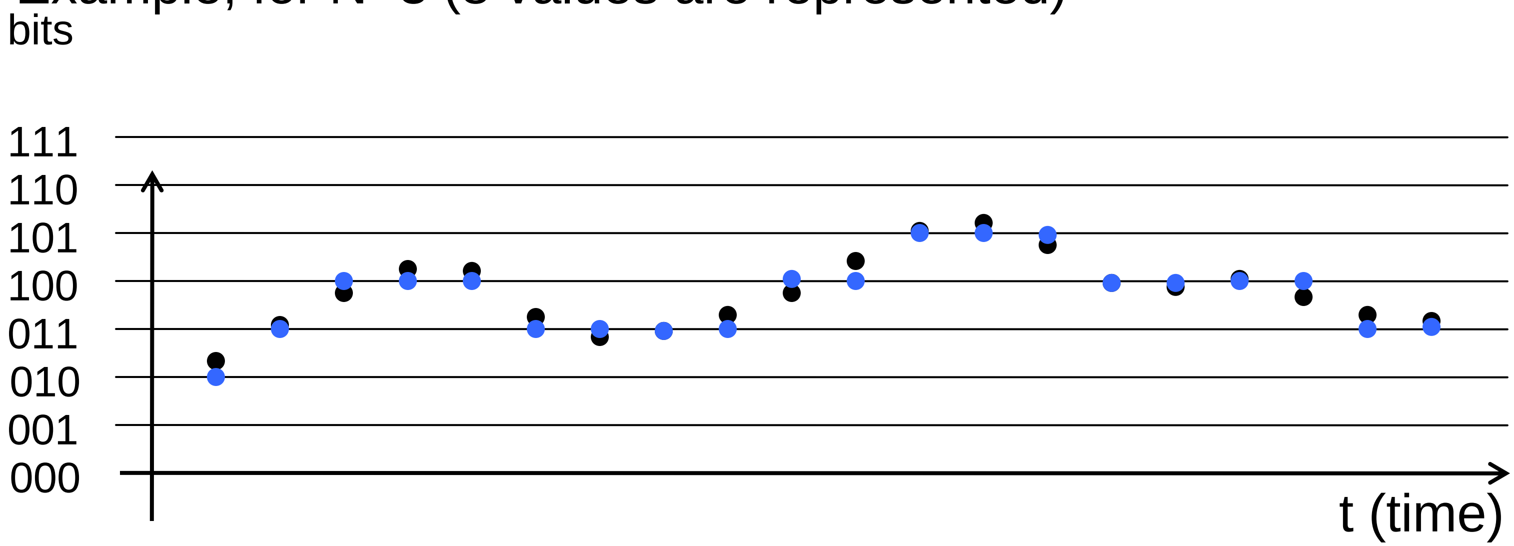


x20000

- A hole on the CD can take only two states, we say that it represents one bit of information (we call the two states “0” and “1”)

- If we combine several bits we can represent more than two values. For example, with N bits we can represent $2^N$ different values.
- We have the freedom to choose the $2^N$ values, but we only represent a finite set of values
- We have to approximate the amplitude of the signals with the values of the set. This operation is called quantization

- If the number of bits N is large, the quantized signal is a good approximation of the original signal
- Example, for N=3 (8 values are represented)

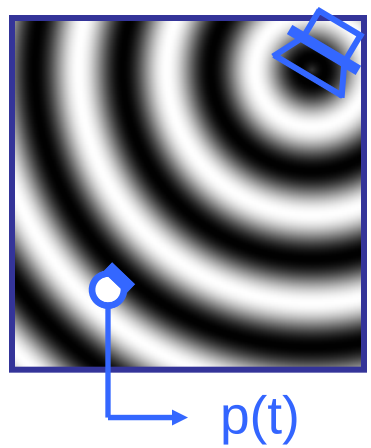



- Original samples
- Samples after quantization

The original signal is represented by the sequence of bits

010, 011, 100, 100, 100, 011, …

- Let's go back to the problem of recording sound. What is sound exactly? Sound is the perception of the variation of air pressure.
- A source of sound, like the membrane of a loudspeaker, produces a variation of air pressure that propagates in space



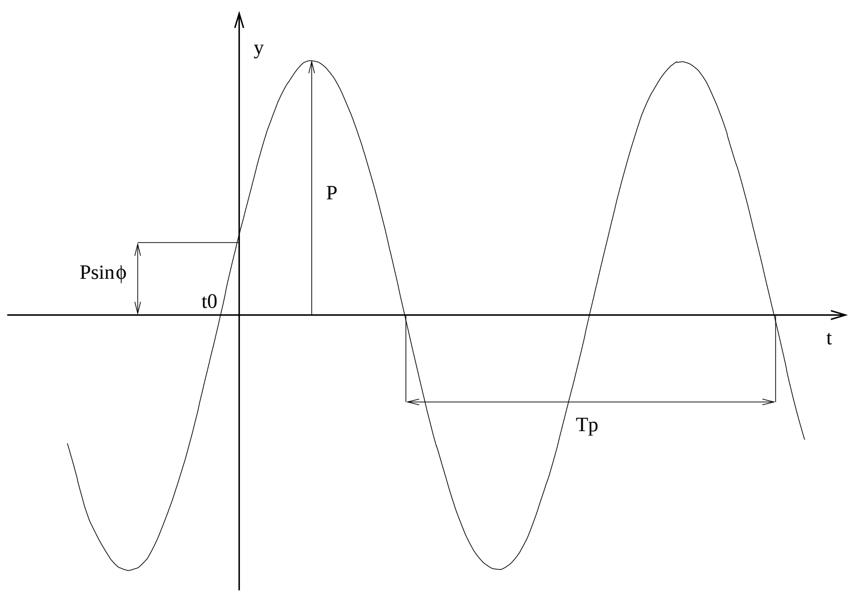

- A microphone placed at a certain position produces an electric signal proportional to the pressure p(t). It works exactly like the thermometer of the previous example. p(t) is a continuous time signal

- Let see a special signal that you will encounter often in the future. This is called the sinusoid.
- You can obtain something very close to a sinusoid by using some musical instruments or a tuning fork:

Tuning fork on resonance box

- The generated sound pressure has the form

$$p(t) = P \sin(2\pi f t + \phi)$$




$P$ is the amplitude
$f$ is the frequency
$\phi$ is the phase
$T_P = 1/f$ is the period
$\omega = 2\pi f$ is the frequency in radians/s

- Example

$$p(t) = \cos(2\pi 440 t)$$

this is a sinusoid with P=1, f=440 Hz (this perceived as an "A"/"La") and $\phi = \pi/2 = 90^\circ$

f=440 Hz means that during 1 second there are 440 oscillations.

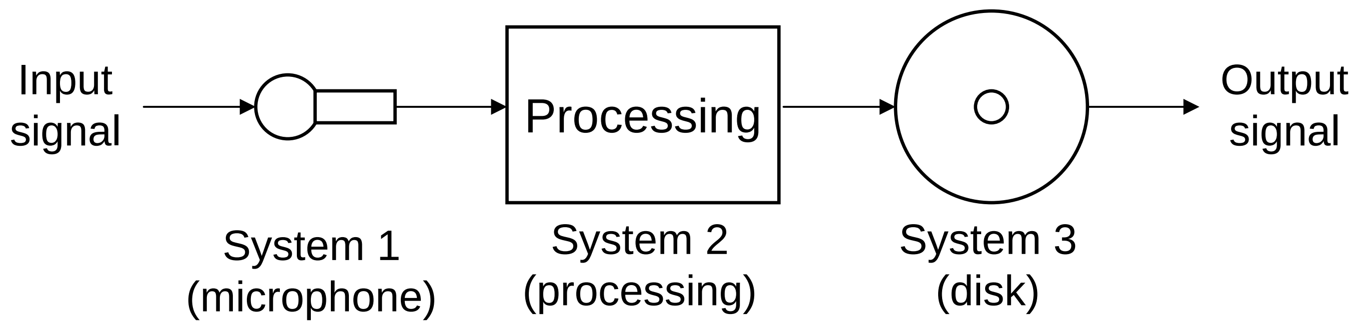
Equivalently, we can say that the period has a duration $T_P$=1/440 s

- What happens if we sample the signal p(t)? A common sampling frequency used for audio signals is $f_s$=44.1 KHz. We obtain

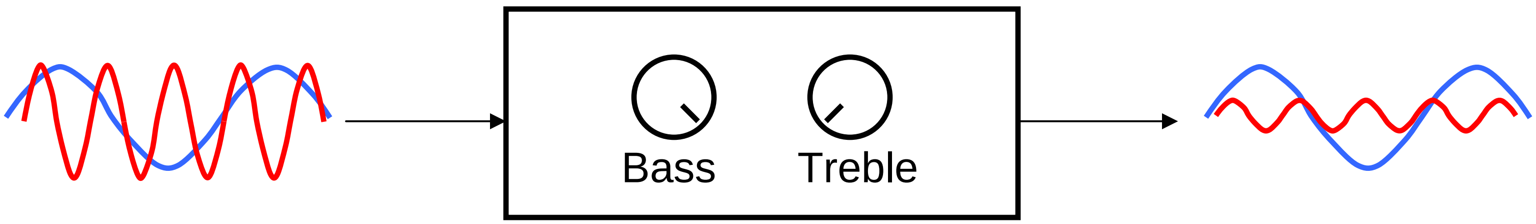
$$\bar{p}(n) = p(nT_S) = P\sin(2\pi f_D n + \phi)$$

with $f_D = \frac{f}{f_S}$

- We said that a system is an element that transforms a signal into another signal. Why do we need systems?
  - To convert the signal to a different medium. These systems are the transducers (e.g. the microphone)
  - To modify a signal. E.g., the tone control of an Hi-Fi chain
  - To control a device. E.g. the heating systems is activated by a signal obtained by processing the temperature
- We often use a chain of systems, i.e. we build a bigger system by connecting a sequence of subsystems (e.g. the Hi-Fi chain)
- We often draw a block diagram of a chain of system. This is a graphical representation of the chain. For example,

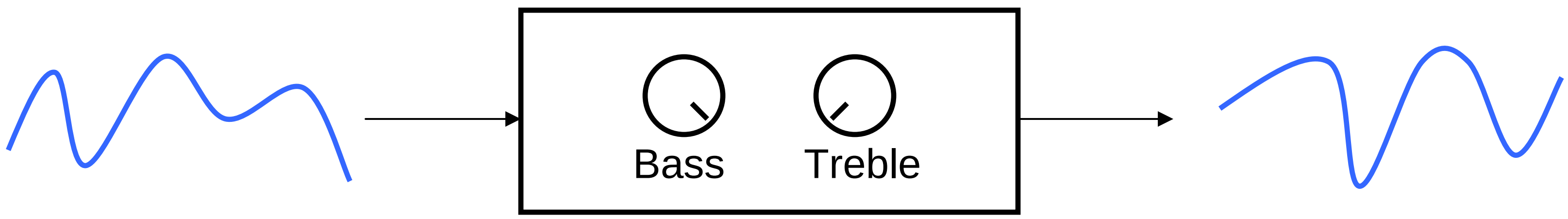


- A system transform signals into other signals. Since signals are functions, a system transform functions into other functions!
- Example: tone control of an Hi-Fi chain




- A sinusoid at low frequency passes unchanged through the tone control
- A sinusoid at high frequency is attenuated by the tone control

- In general, to describe completely a system, we need to give the output signal corresponding to each possible input signal. This is very complicate



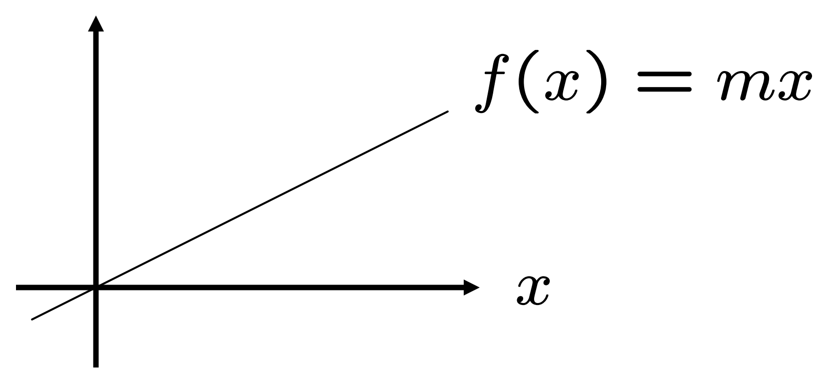

- There is a class of systems that are easier to describe and study. These are called linear systems
- Let's see first what a linear function is…

- Let's take a function f that maps real numbers into real numbers. We say that f is linear if

$$\forall x_1, x_2 \in \mathbb{R}, \quad f(x_1 + x_2) = f(x_1) + f(x_2) \quad \text{Additivity}$$

$$\forall x, a \in \mathbb{R}, \quad f(ax) = af(x) \quad \text{Homogeneity}$$

- In this case, the graph of the function is always a line:




- The slope m identifies completely a linear function defined on real numbers!

- In the case of a systems, things are slightly more complicate… Let's write as h[s] the output signal of the system h, when s is the input signal
- The system h is linear if,
  - For any choice of two signals $s_1$, $s_2$

$$h[s_1 + s_2] = h[s_1] + h[s_2]$$

i.e. the output corresponding to the sum is the sum of the outputs. This is the property of additivity

  - For any choice of a signal s and a scalar a,

$$h[as] = ah[s]$$

i.e. if the input signal is scaled, the output signal is scaled by the same factor. This is the property of homogeneity

- It is not always easy to test system linearity, since we have to consider all the possible input signals
- Example: let's consider the system h:

  $h[s(n)]=s(n-1)$

  i.e. the output signal is obtained by delaying the input signal of one sample

  $h[s_1(n)+s_2(n)]=s_1(n-1)+s_2(n-1)=h[s_1(n)]+h[s_2(n)]$ additivity ok

  $h[as(n)]=as(n-1)=ah[s(n)]$ homogeneity ok

  The system is linear!
- Example: let's consider the system h:

  $h[s(n)]=s^2(n)$

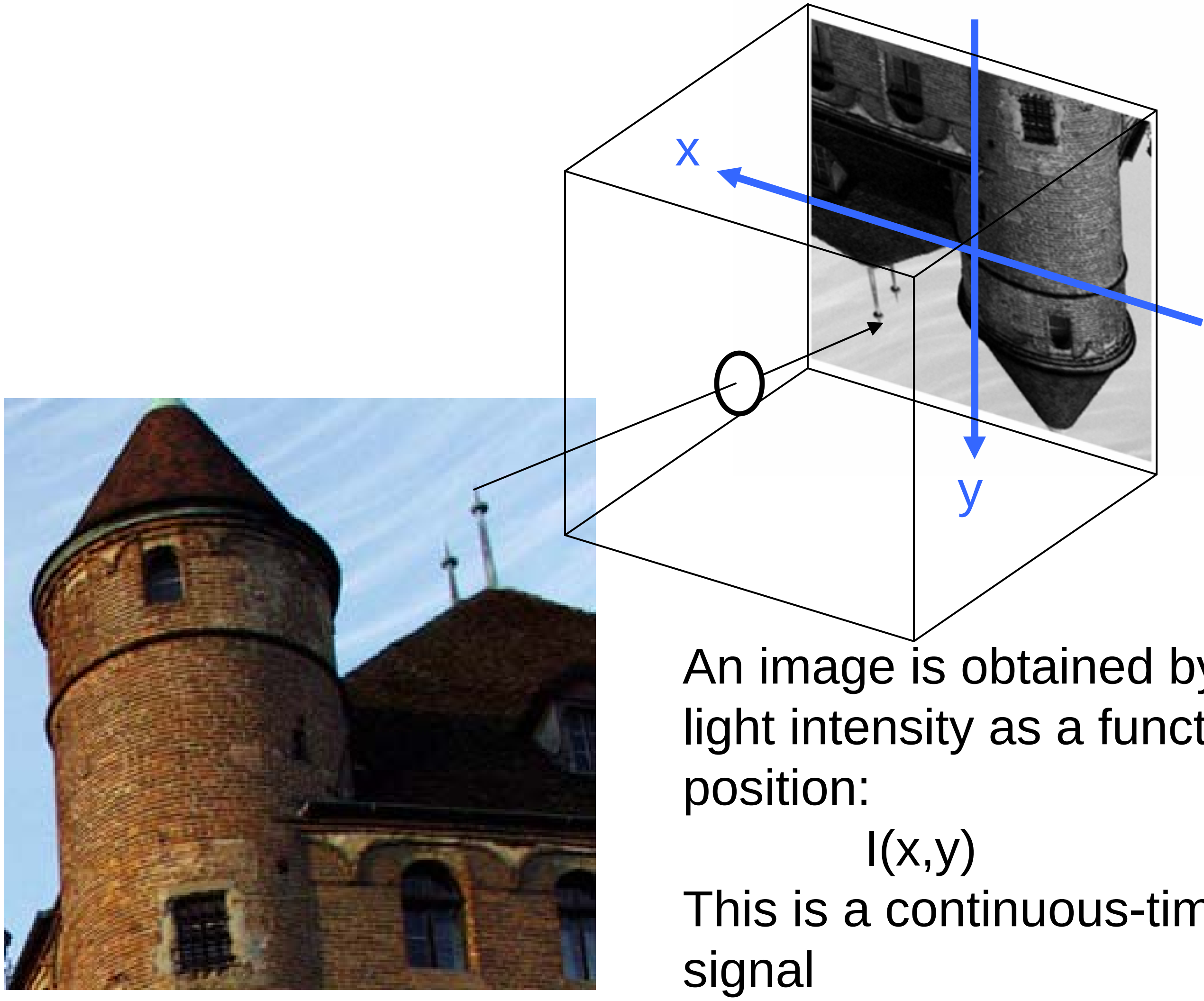
  i.e. the output signal is the square of the input signal

  $h[2s(n)]=4s^2(n)=4h[s(n)]$ homogeneity is not verified for a=2, the system is not linear (no need to check other cases!)

In this lecture we have seen that:

- We can use signals to represent the evolution of a measurement as a function of time
- Signals can be continuous-time, when the measure is performed continuously…
- … or discrete-time, when the measure is performed at certain instants regularly spaced
- Continuous time signal can be transformed into discrete time signal by sampling them
- Systems transform signals into other signals
- Among all the possible systems, the class of linear systems enjoy the properties of additivity and homogeneity

- There are signals that are function of space instead of time




An image is obtained by measuring light intensity as a function of the position:

$$I(x,y)$$

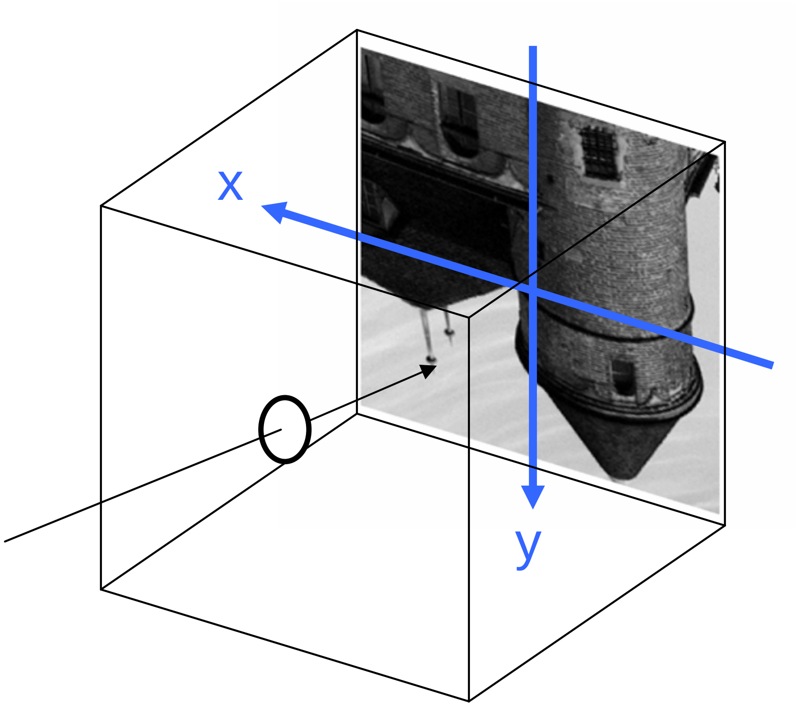
This is a continuous-time (space) signal

- Continuous images can be acquired by using an analog camera

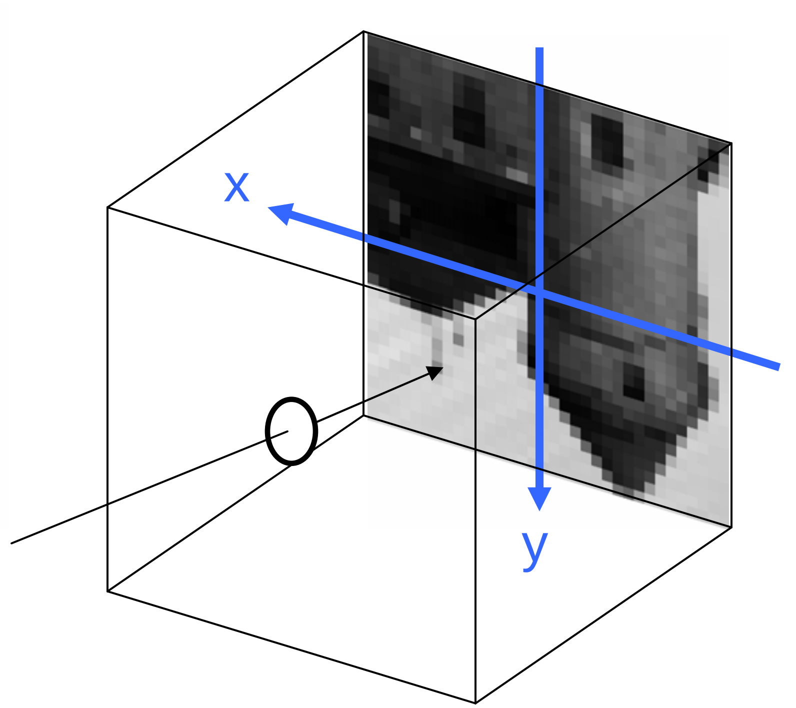
19th century studio camera

- However, today everyone buys a digital camera

- A digital camera produces a discrete image. As for audio signal, sampling allows to transform a continuous signal into a discrete one. For image acquisition, sampling is implicit in the sensor

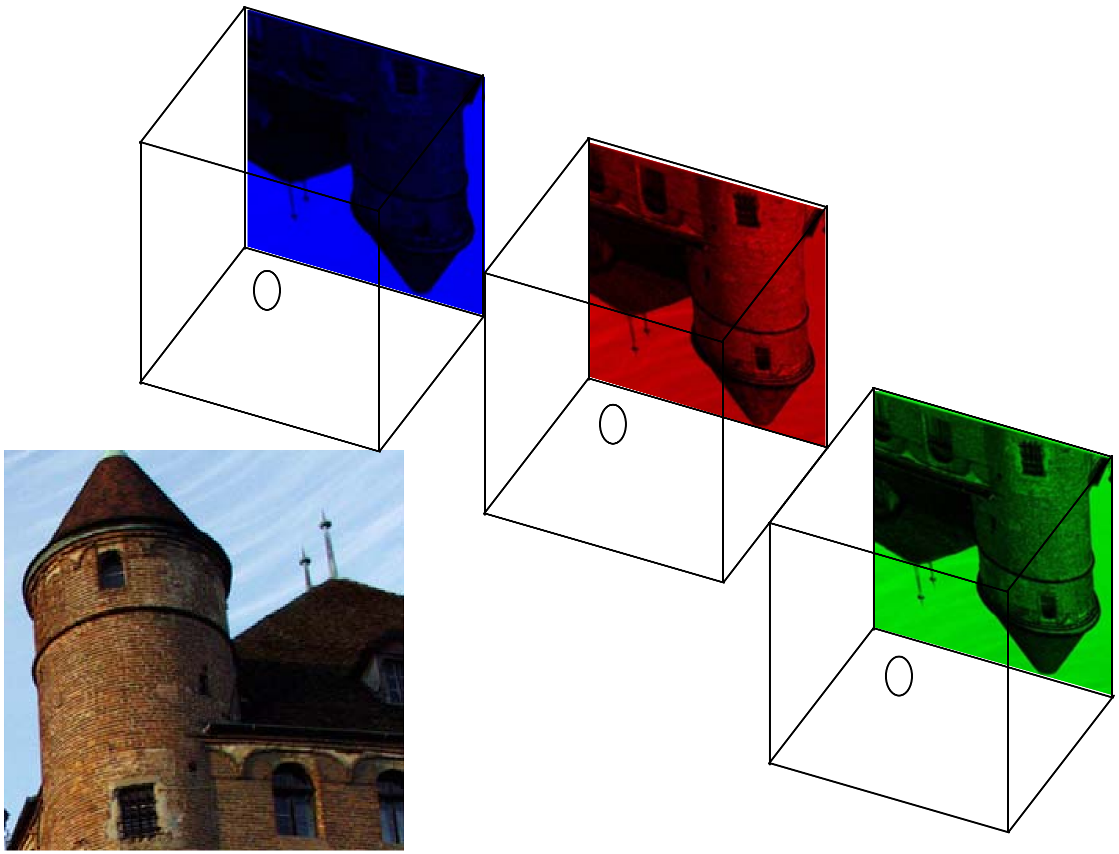



Analog camera

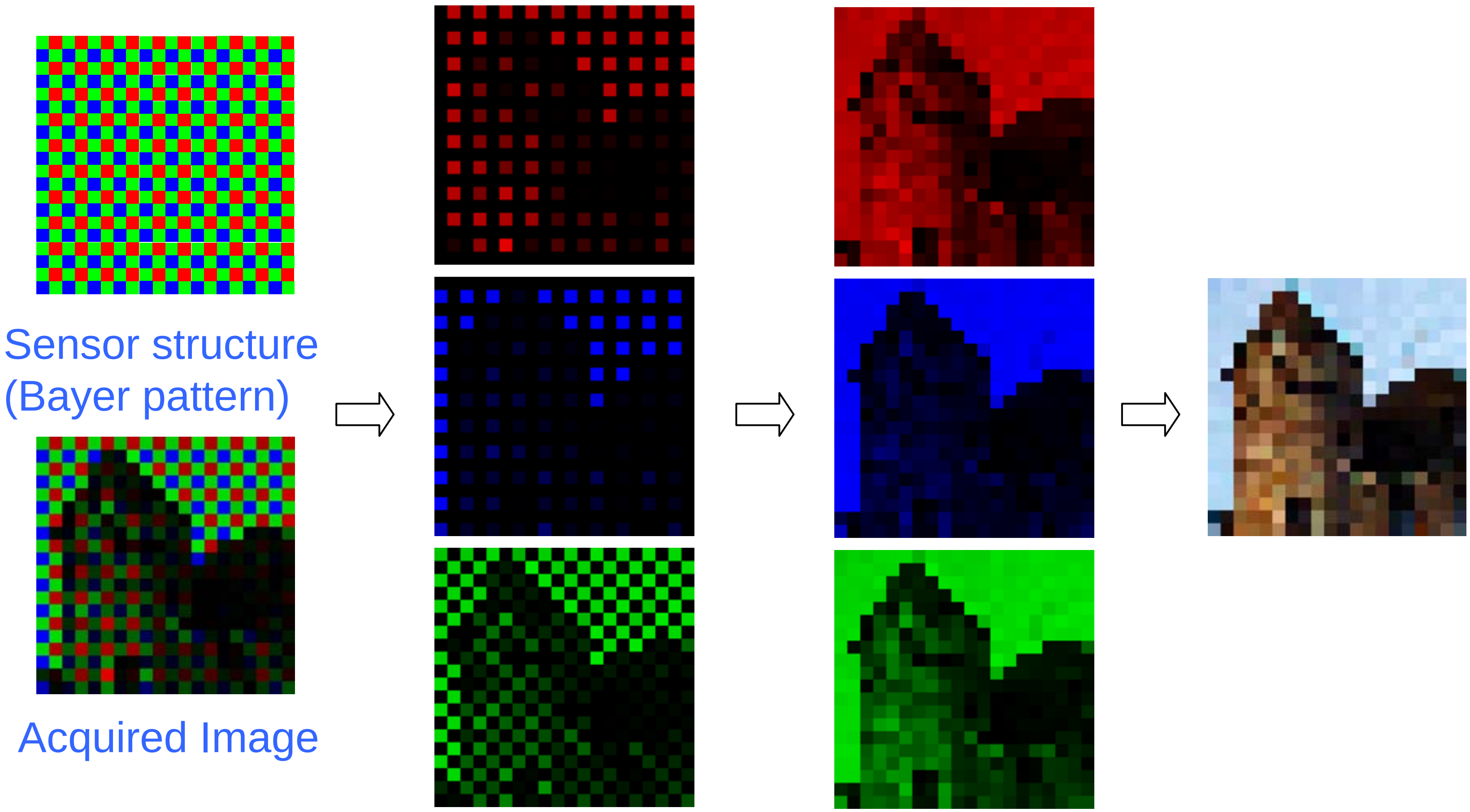



Digital camera

- What about color?
- Color is perceived by means of 3 types of photoreceptors: we need 3 sensors to acquire a color image

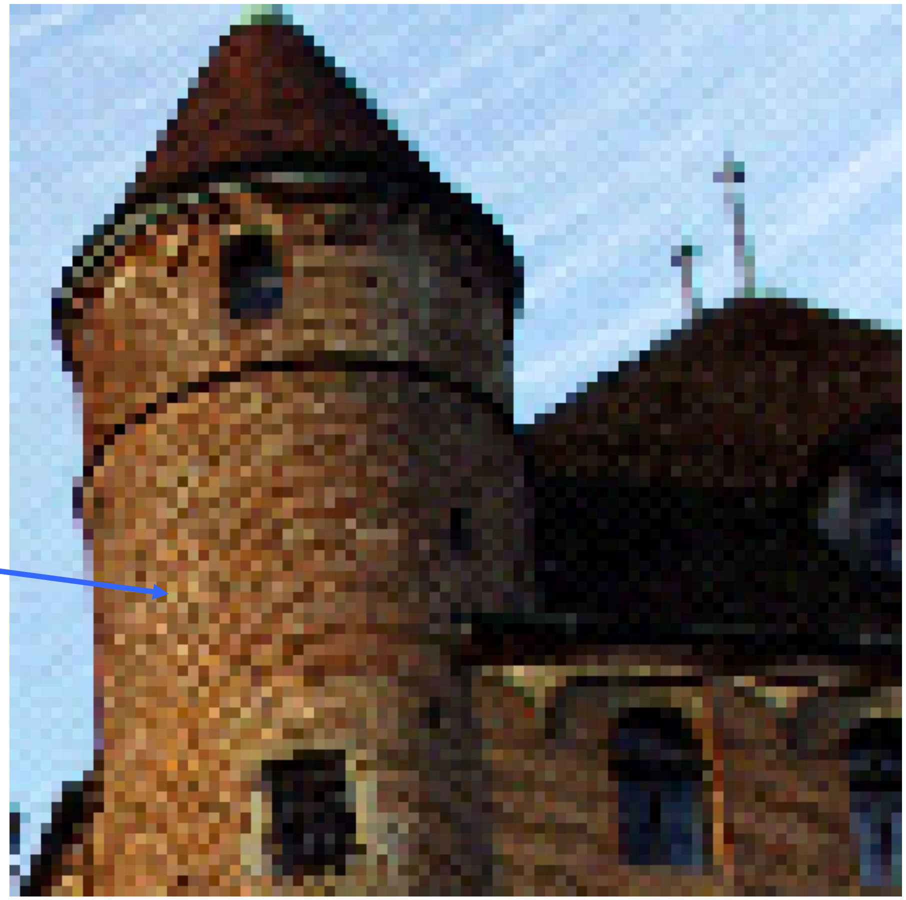


- A cheaper alternative is to interleave pixels sensitive to different colors on a single sensor



- How many Megapixels do you need?
  - The spacing of the pixels on the sensor corresponds to the sampling period of the sampling procedure
  - For a fixed image size, many “Megapixels” corresponds to high sampling frequency (In the third lecture, we will see why this is good)
  - Intuitively, many pixels allows to obtain an image with high level of detail, but there is more than that…
  - The procedure of image acquisition and interpolation may give strange artifacts when the resolution is low…

Artifacts!



8 Megapixels