Multimedia systems

Prof.dr.sc. Davor Petrinović
Prof.dr.sc. Sonja Grgić
Prof.dr.sc. Daniel Hofman

Introduction to multimedia

prof.dr.sc. Davor Petrinović

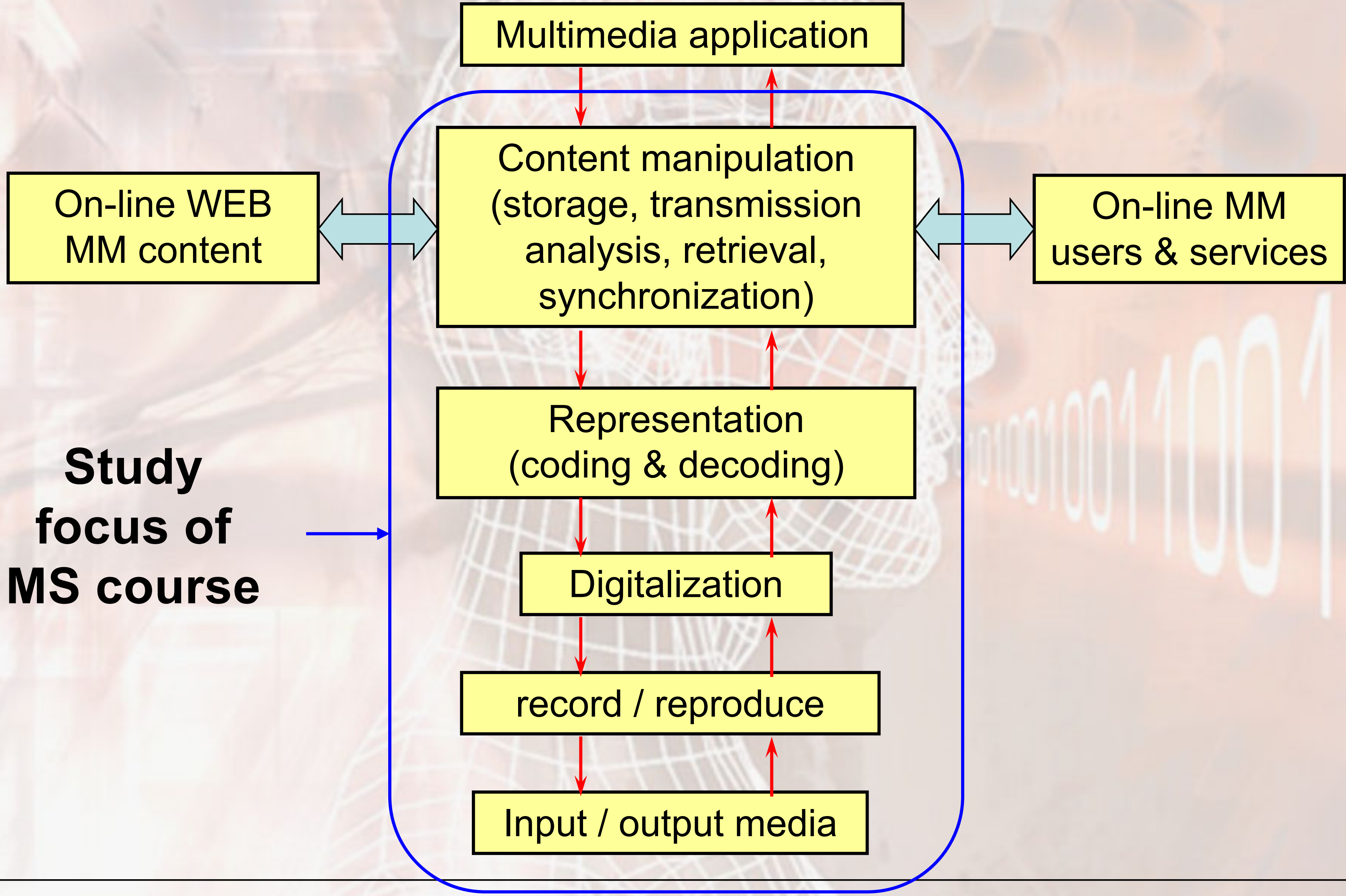
Multimedia

- What is multimedia
  - the term has many meanings ...
- The basic definition would be
  - A medium that uses a combination of several related types of media content for the purpose of better presentation of information.
  - The opposite of multimedia are "classic" content carriers: books, newspapers, etc. where information is presented exclusively in text, images, graphs or photographs.

Multimedia

- The key goal of multimedia is to expand such "dry" content with additional media:
  - audio recording (mono, stereo, surround sound),
  - narration (speech record),
  - video,
  - animations,
  - stereoscopy / stereography ,
  - virtual reality.
- ... resulting in much more attractive user-friendly content.

Multimedia system



Multimedia systems

- The course focuses on "lower" levels of multimedia systems that are directly related to input and output multimedia signals.
- Great attention will be paid to **procedures and algorithms for efficient digital representation** of these signals.
- Aspects of the **implementation and performance** of the lower layers of the MM system on computers will also be considered, given the specific requirements set by these algorithms.
- In the last cycle, intermediate levels related to **media manipulation** will also be considered.

Multimedia systems – course structure

- The course is conceived through three parts:
- First unit:
  - multimedia technologies and systems,
  - their architecture and application,
  - overview of media and data sources,
  - basics of coding and compression,
  - speech signal, modeling and analysis, parametric representation and coding,
  - speech coding standards, basics of synthesis and recognition,
  - audio signal, psychoacoustic model, coding procedures and standards.

Multimedia systems – course structure

- Second unit:
  - basic features of the human visual system,
  - color experience, chromaticity diagram, color rendering spaces,
  - creation and characteristics of images and video signals,
  - video signal forms: analog with components, analog composite, digital with components,
  - picture formats for standard quality television (SDTV) and high-definition television (HDTV),
  - principles of image compression; lossy and lossless compression. transformation coding, quantization, entropy coding,
  - overview of image compression standards and areas of application, image storage formats,
  - principles of video signal compression, removal of time redundancy by prediction between frames,
  - an overview of video compression standards and areas of application.

Multimedia systems – course structure

- Third unit:
  - storage, transmission and processing of multimedia data,
  - integration of multimedia content ,
  - synchronization,
  - multimedia systems and software tools,
  - overview of the complexity of image compression algorithms (JPEG) and video (MPEG) and problems with actual implementation,
  - examples of fast algorithms for execution on computer systems (DCT and Huffman),
  - processor and computer architecture for performing MM algorithms - based on hardware and software implementations,
  - ways to protect MM data, DRM, Example of system implementation (e.g. MP3 player).

Applications of multimedia

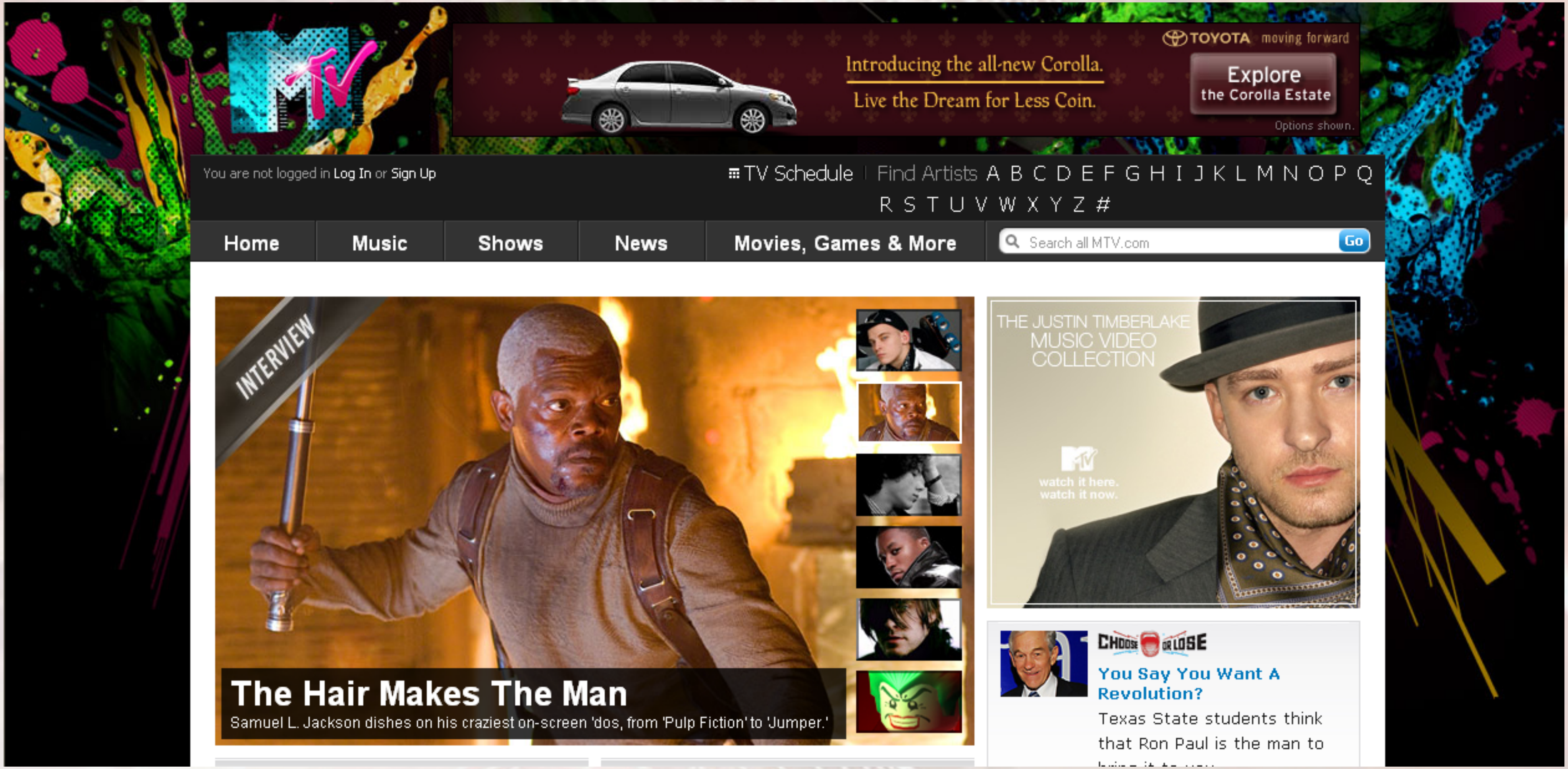
- In online information media :
  - newspaper houses : Večernji, Jutarnji, Vjesnik, ..
  - portals: Index, Monitor, T-portal, ...
  - RTV companies : HRT, Nova, Dnevnik, RTL ...
  - Radio Stations: 101, Plavi, Otvoreni, Cibona …
- In the entertainment industry:
  - music industry : MTV, Sony BMG, EMI, Universal, Warner
  - film industry

Applications of multimedia

Applications of multimedia

Applications of multimedia

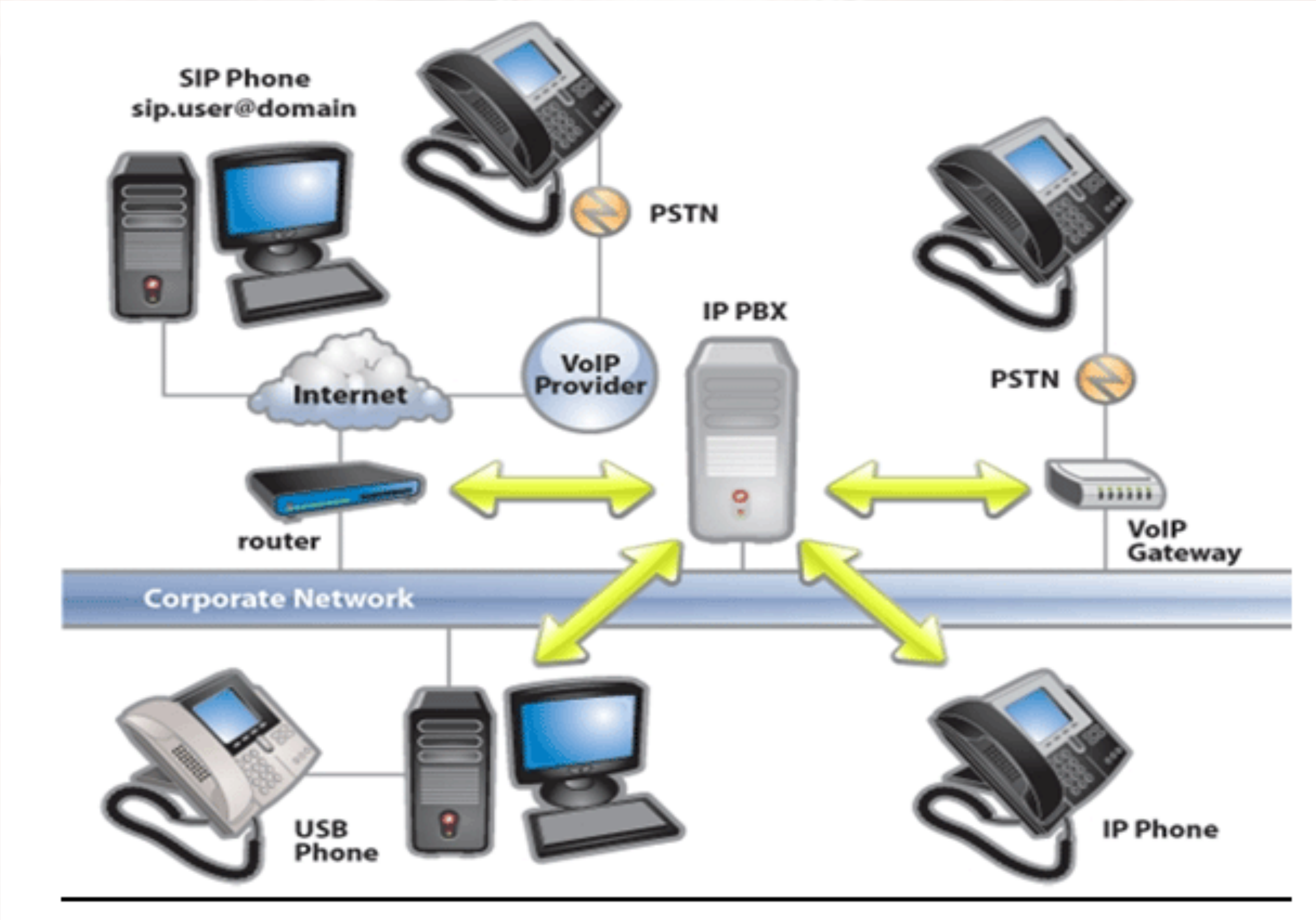
Applications of multimedia



Applications of multimedia

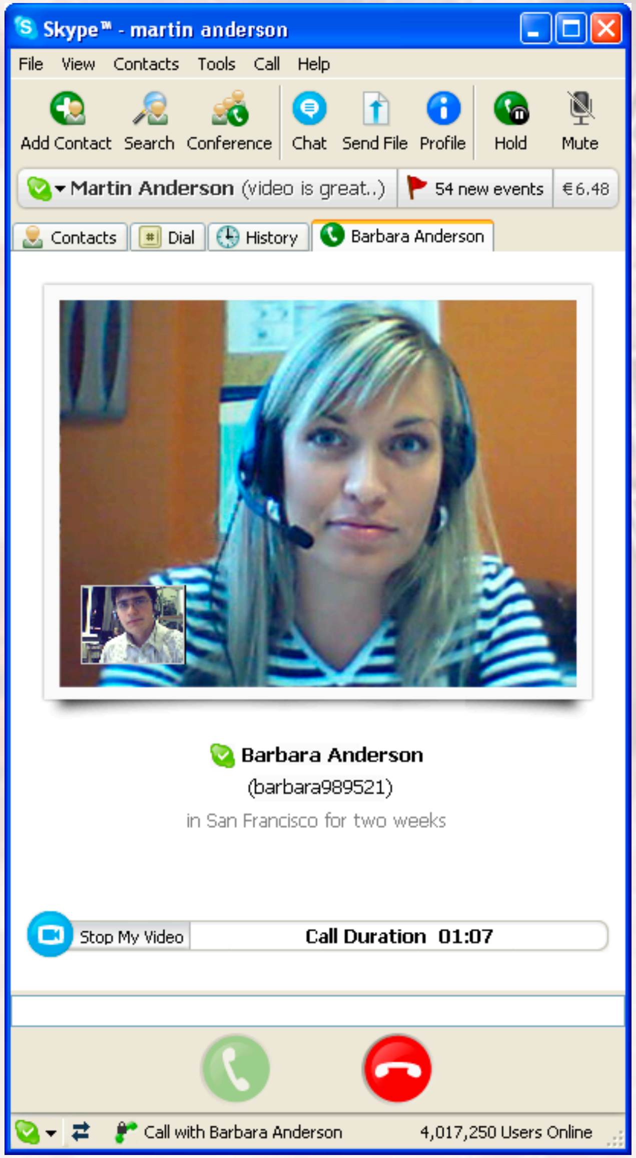
- In communications:
  - VOiP: MSN Messenger, Skype, SIP, AMR
  - video telephony , 3G UMTS,
  - video sharing , YouTube.
- For the purpose of creating interactive presentations and multimedia content:
  - PowerPoint,
  - Adobe (Macromedia) Flash,
  - MovieMaker, Adobe Premiere,
  - DVD authoring.

Applications of multimedia



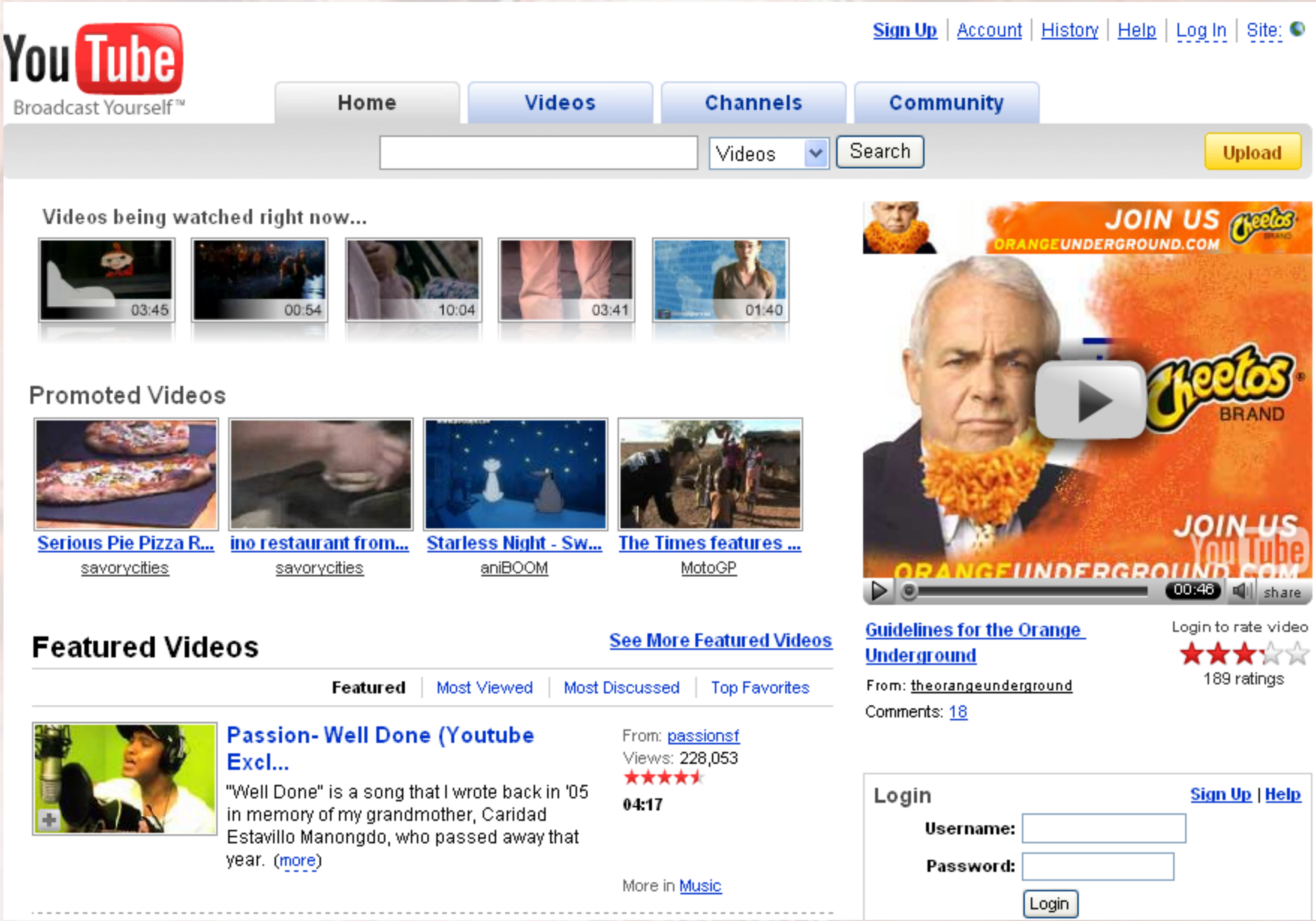
Applications of multimedia

Applications of multimedia



Applications of multimedia

Applications of multimedia

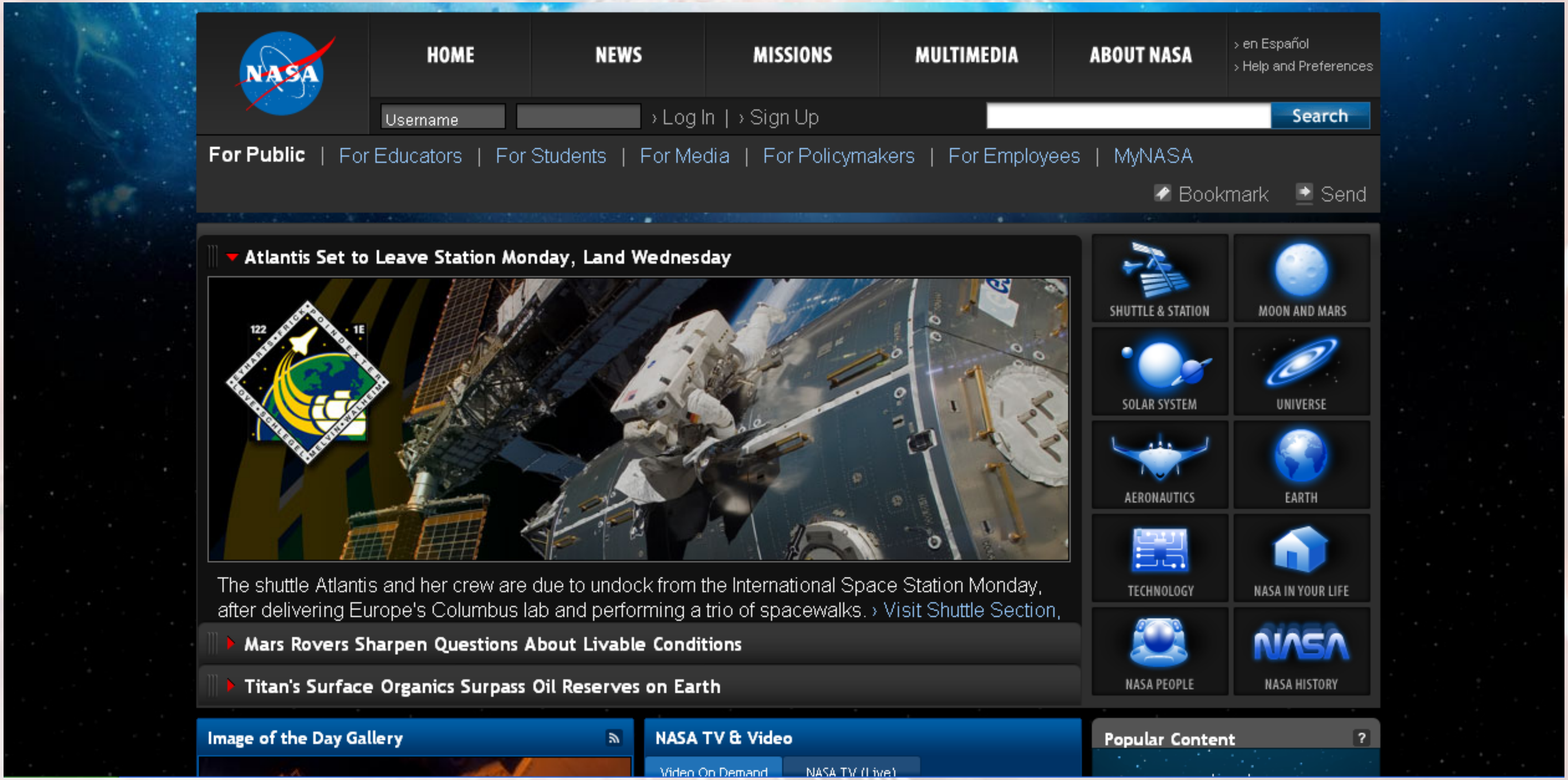


Applications of multimedia

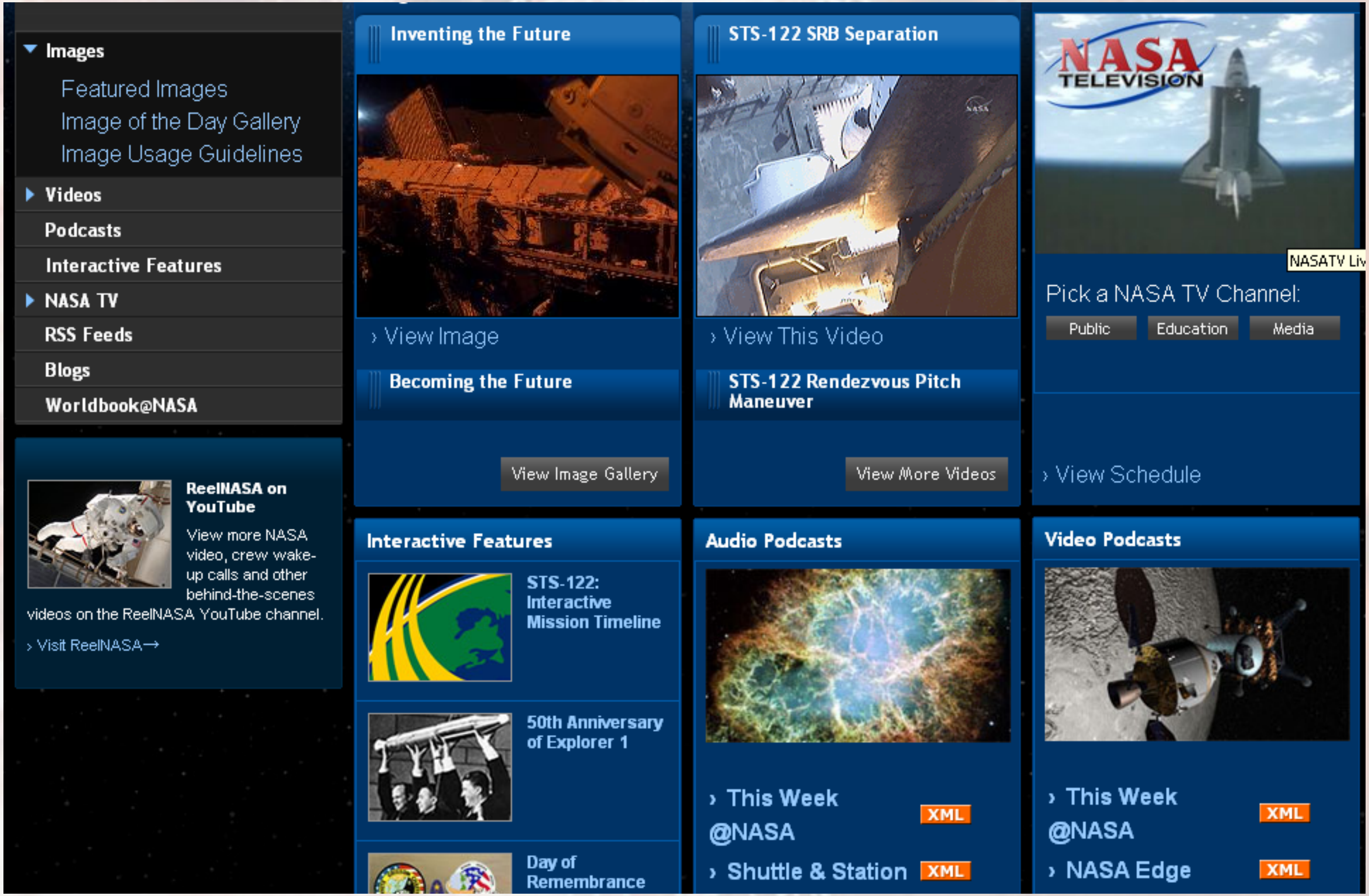
- In education and science:
  - academic and scientific contents ,
  - school educational contents,
  - Google Earth, NASA, Wikipedia, MathWorld,
- In e-commerce :
  - eBay, Amazon, ...
  - wholesale chains: Konzum, Mercator, ...
  - computer stores: HGspot, Vemil, M-San,...

Applications of multimedia

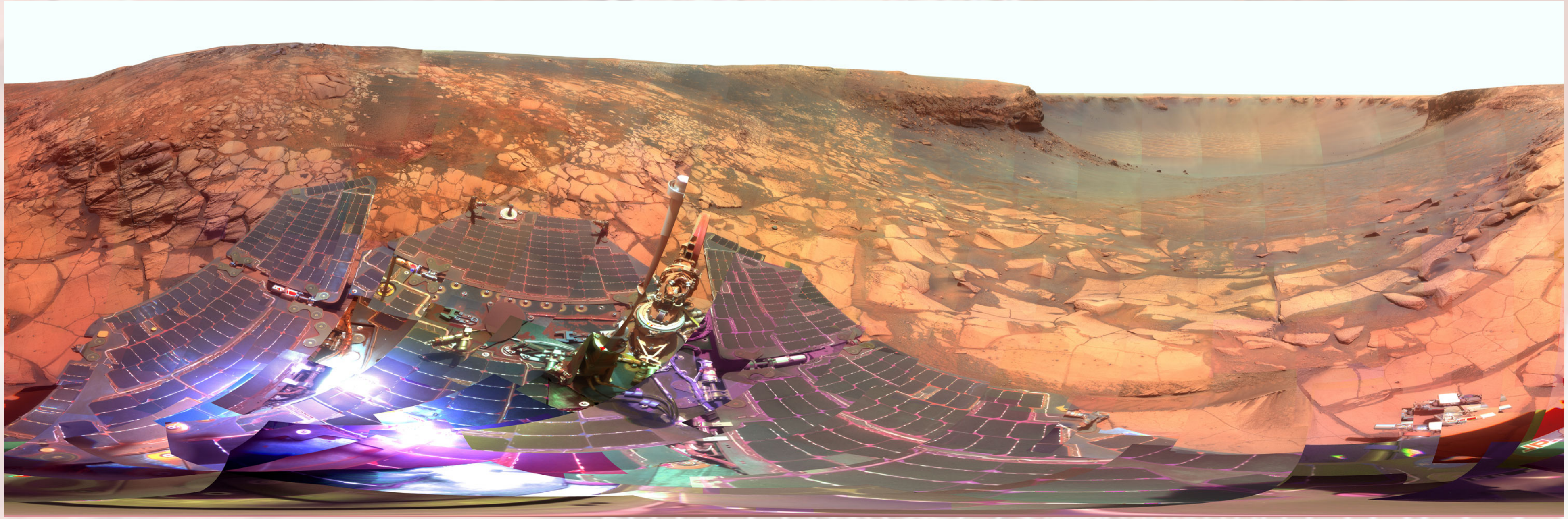
Applications of multimedia



Applications of multimedia



Applications of multimedia



Applications of multimedia

What we learned

- Definition of multimedia
- The importance of multimedia technologies
- The structure of the multimedia system
- Applications of multimedia

Media in multimedia

prof.dr.sc. Davor Petrinović

Media

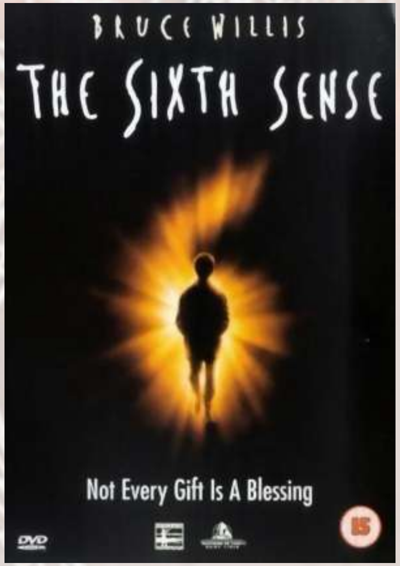
- What are the media that are the subject of the study of MM systems?
- ... any form of signal or phenomenon that represents certain cognitively useful or interesting information for a person.
- The media involved in multimedia are therefore always **related to human**, either as a source or user of this media information.

Media

- Direction of the flow of media is important, because a person can be:
  - signal source, or
  - its user or signal sink.
- In addition to human, the source of the media can be:
  - the material world that surrounds us with sources of various natural signals, or else
  - artificially created signals and phenomena that can create the impression of an apparent or virtual world.

The signals we perceive

- Signals can be divided according to the way they are perceived, which is closely related to our senses.
- The main types of stimuli we feel are:
  - vision,
  - hearing,
  - taste,
  - smell and
  - touch.

\+  ☺

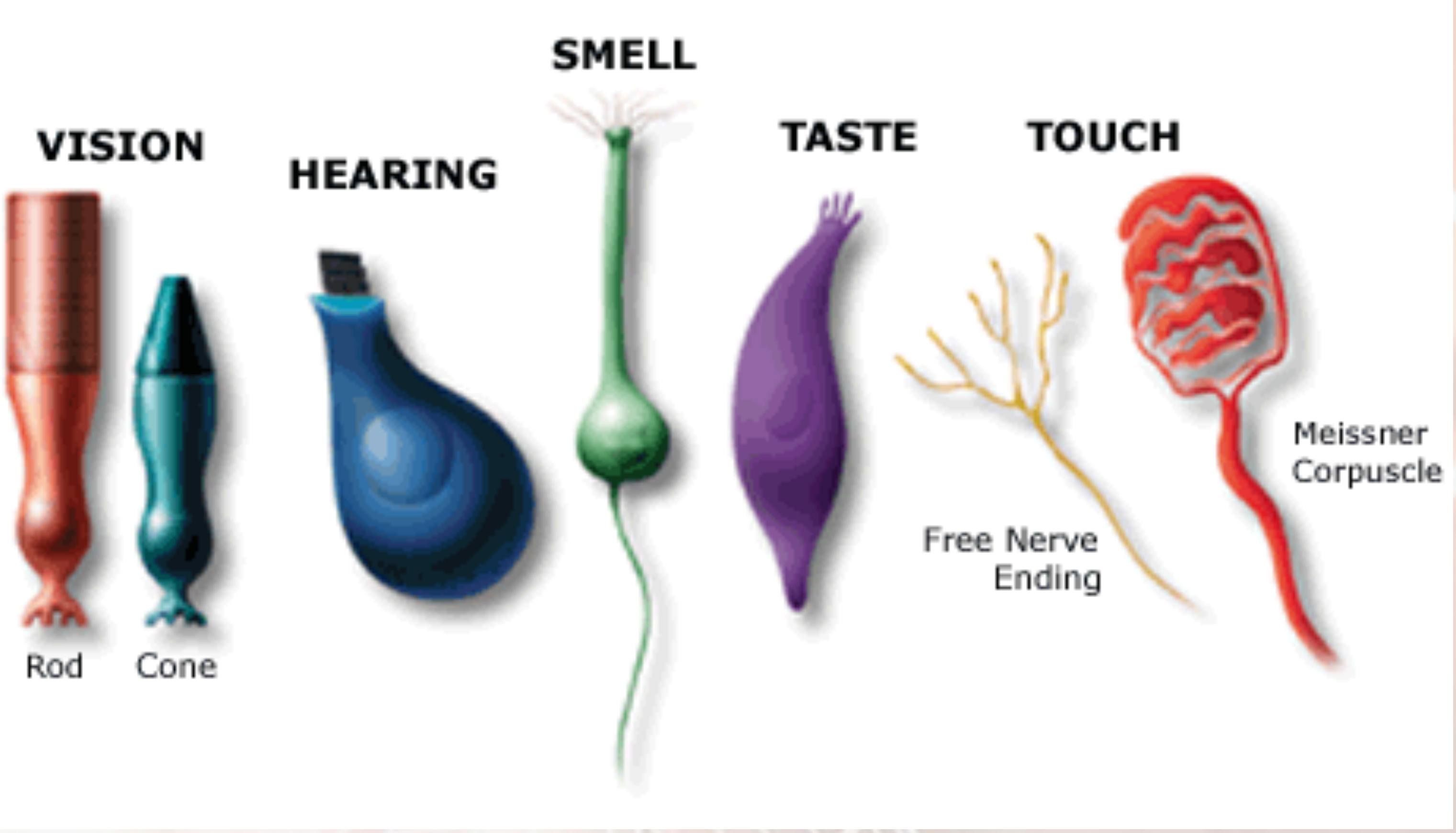
Human senses

The signals we perceive

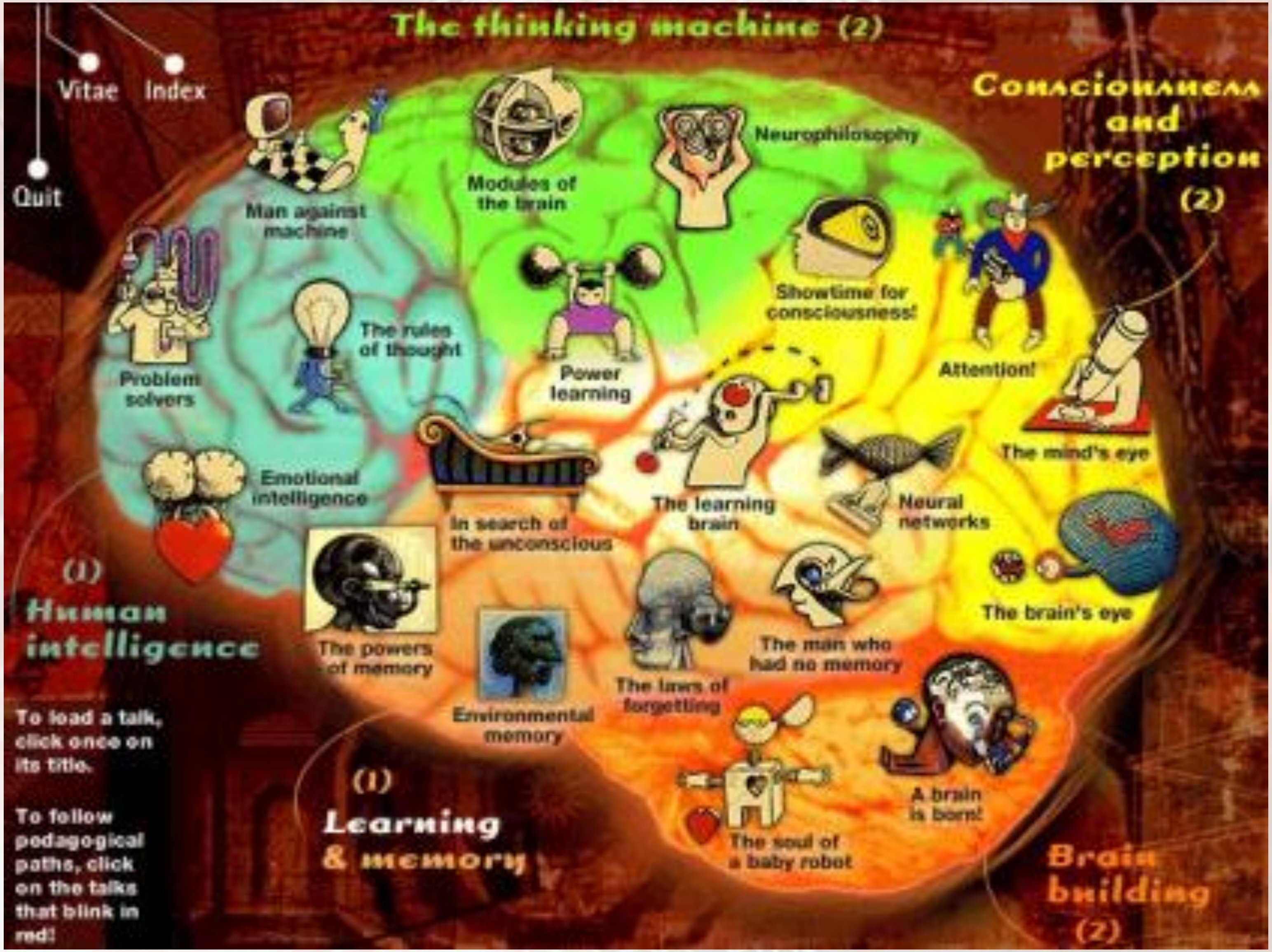
- Receptors can be divided into the following groups (according to Wikipedia):
  - Mechanoreceptors - for the sense of touch, pressure, hearing and balance,
  - Thermoreceptors - for the feeling of cold and heat,
  - Nociceptors - for the sensation of pain,
  - Electromagnetic receptors - for the sense of vision,
  - Chemoreceptors - for the sense of taste, smell, oxygen, glucose and carbon dioxide.

The signals we perceive

- Shape of individual receptors



What we do with these signals?



The signals we emit

- The main signals we emit are:
  - our visual appearance in space,
  - sounds we produce such as: voice, singing, sighing or sounds caused by movement or playing music,
  - our tactile action on the environment,
  - thermal radiation,
  - chemical impact on the environment, ... primarily the smell we spread.

Multimedia systems (MS)

- The main role of MS is to manipulate the described media for the purpose of their:
  - recording or acquisition,
  - reproduction or presentation,
  - transmission or distribution,
  - efficient storage,
  - analysis and content search.
- The most common application of MS is for the purpose of:
  - connecting people, informing, presenting, learning or having fun.

Multimedia signals

- Signals carrying media information are defined by certain physical or chemical quantities, such as e.g.
  - EM waves in the visible or IR part of the spectrum,
  - sound waves described by sound pressure variations,
  - mechanical signals, measurable through the displacements, velocities or accelerations of objects,
  - signals measurable as concentrations of certain chemical compounds (molecules) in the air or water.

The acquisition task of media information

- Multimedia systems, as a rule, include the task of media acquisition performed by :
  - sensors that convert physical quantities that carry media information into an equivalent more easily measurable form.
  - The conversion to an analog electrical signal is most often performed :
    - some of the measurable parameters of this signal (e.g. voltage, current, frequency, etc.) are proportional to the original physical quantity.

The acquisition task of media information

- Commonly used sensors in MM systems are:
  - electroacoustic transducers; various types of microphones (dynamic, condensor, ribbon),
  - two-dimensional image or video sensors (CCD, CMOS, IR sensor),
  - various computer input units.
- Sensors used in the broader context of virtual reality but also for the purpose of managing MM systems:
  - displacement and position detectors,
  - accelerometers,
  - gesture detectors (electronic glove).

Reproduction of MM content

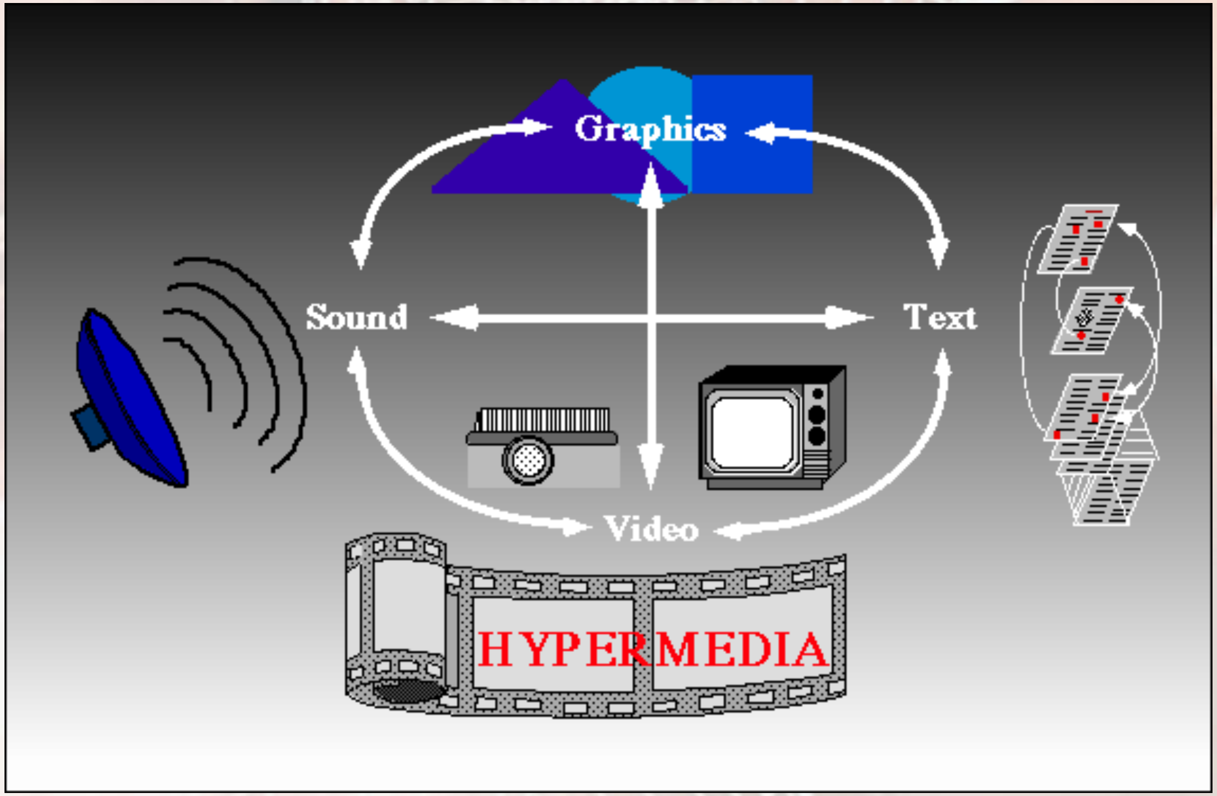
- Reproduction .... a key task of the MM system!
- This is where **multimodality** comes to the fore, because when playing a certain content, we try to present as many parallel media as possible in a **connected and interactive** form :
  – text,
  – audio recordings,
  – photos,
  – video,
  – computer graphics and animations.

Interactivity of reproduction

- What is interactivity?
- During playback, the user has the ability to easily and quickly manage the presentation of content by selecting individual related media
- ... “**content navigation** ”.
- It is this interactivity that distinguishes:
  - linear media content and
  - hypermedia content (multimedia analogy of hypertext)

Hypermedia content

- Different types of media that are intertwined with a series of hyperlinks



- Naturalness of interaction
  - gesture based management

What have we learned?

- Media in MM
- Perception of MM signals, receptors
- Human as s source of MM signals
- Importance of MM technologies
- MM signal
- Acquisition and Reproduction of MM content
- Interactivity
- Hypermedia content

Historical development of multimedia

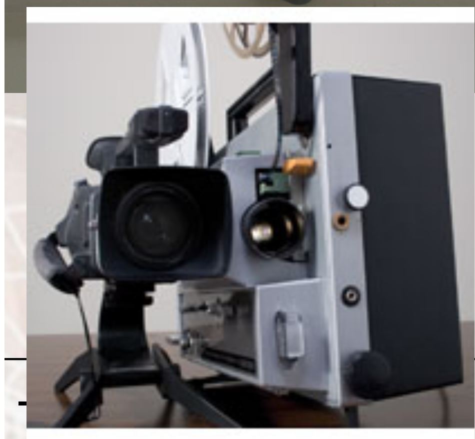
prof.dr.sc. Davor Petrinović

Multimedia development

- Until the advent of computers, presentations of certain media content were related to separate dedicated playback devices :
  - overhead projector (display of graphics and text),
  - audio system for playing audio content (audio tapes, gramophone records),
  - cinema / video projections for video content,
  - slide projector for displaying photos or slides for presentation.

Media technology from the beginning of the last century

Media Technology from 1970'



Multimedia development

- **Computers** are a key driver of MM development.
- The ultimate goal of development:
  - **networked computer as a universal and integral system for playing any MM content !**
- Today's computers are already very close to this goal with their capabilities or even exceeding it.
- With the development of technology, the expected standard of reproduction quality is constantly being raised and stricter requirements are being placed on the MM system !
- .... there will surely be enough work for all of you ☺

Multimedia computer

Multimedia development

- The focus of today's development:
  - the ability to play and even record MM content **anywhere and anytime** !
- Mobile platforms for content playback :
  - media players (iPod, iPod-Video, “MP4” players, ...),
  - modern cell phones with the ability to play and record audio and video content (iPhone, Android),
  - handheld computers (Palm, IPAQ, ..),
  - mobile video and audio streaming (DAB, DVBT, pocket TV, Internet radio, ... over networks such as GSM, GPRS, UMTS, Wi-Fi, WiMax, …).

Example of mobile MM platforms from early 2000

Multimedia development

- Exceptional variety of different MM content players :
  - differences in processing capabilities,
  - differences in resolution, size, refresh rate and quality of display units,
  - differences in power consumption limits (battery power limitation of mobile platforms),
  - differences in the type and capacity of available memory for content storage,
  - differences in the existence and bandwidth of the communication channel for the transmission of MM content (LAN, Wi-Fi, GPRS, UMTS),
  - differences in the operating systems used.

Multimedia development

- The diversity of MM platforms also places specific demands on the organization of MM content
- ... the need for a scalable media format!
  - all devices must be able to display the same content, but in accordance with the capabilities of each platform,
  - the content must be organized in layers, where the lowest layer allows for basic quality, with the least amount of information to describe and the least complexity of implementation;
  - each subsequent layer increases quality, at the expense of increasing information and computational complexity.
- The problem of efficient distribution of MM content is extremely complex!

What have we learned?

- Multimedia development
- Multimedia 50 years ago
- The importance of computers in multimedia
- Mobile MM platforms
- Specifics and diversity of MM platforms

Digital signal processing

prof.dr.sc. Davor Petrinović

Digital signal processing

- Another extremely important driver of the evolution of MM technologies is the development of the theory of **Digital Signal Processing** (DSP).
  - A more accurate name would be signal processing in the time discrete domain, i.e., over the time samples of the signal.
  - The name DSP derives from the fact that signal samples are represented by numbers over which the necessary mathematical operations are performed.
- Processing is performed by either software and/or hardware solutions.

Digital signal processing

- DSP theory matured as early as the 1960s and 1970s, but it gained momentum with the advent of computers as ideal platforms for its implementation.
- Until then, all processing is performed exclusively by analog circuits (networks or circuits) that perform the desired operations on time-continuous signals.
  - The maximum complexity of analog processing is directly determined by the complexity of the network or circuit that implements it!

Analog signal processing

- an example of a 6th order filter design using operational amplifiers and passive R and C components

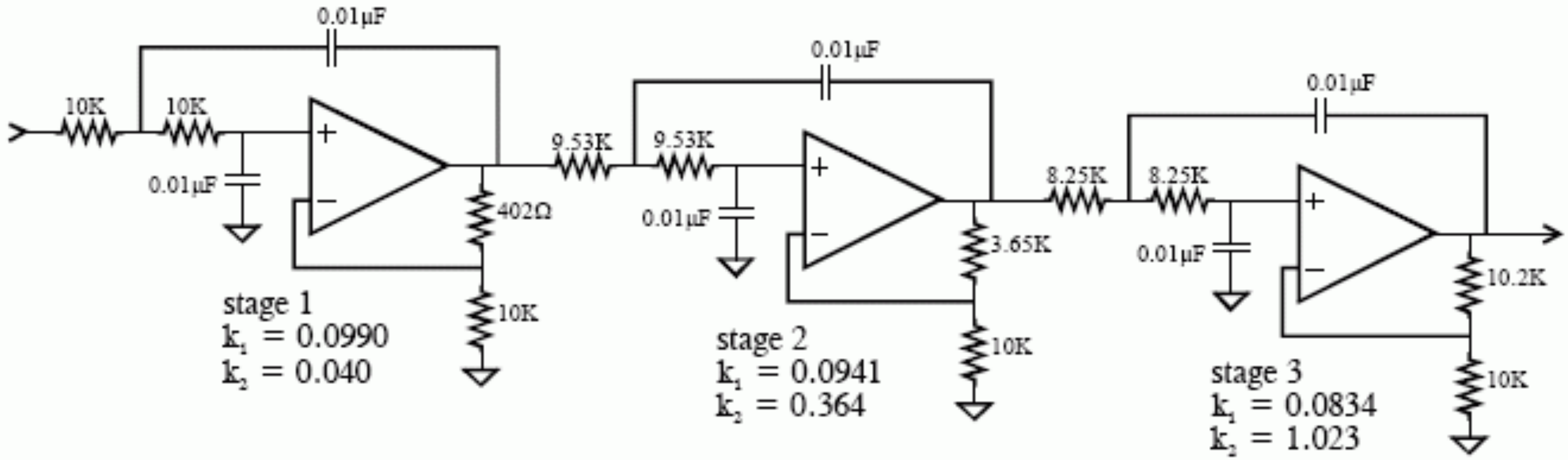
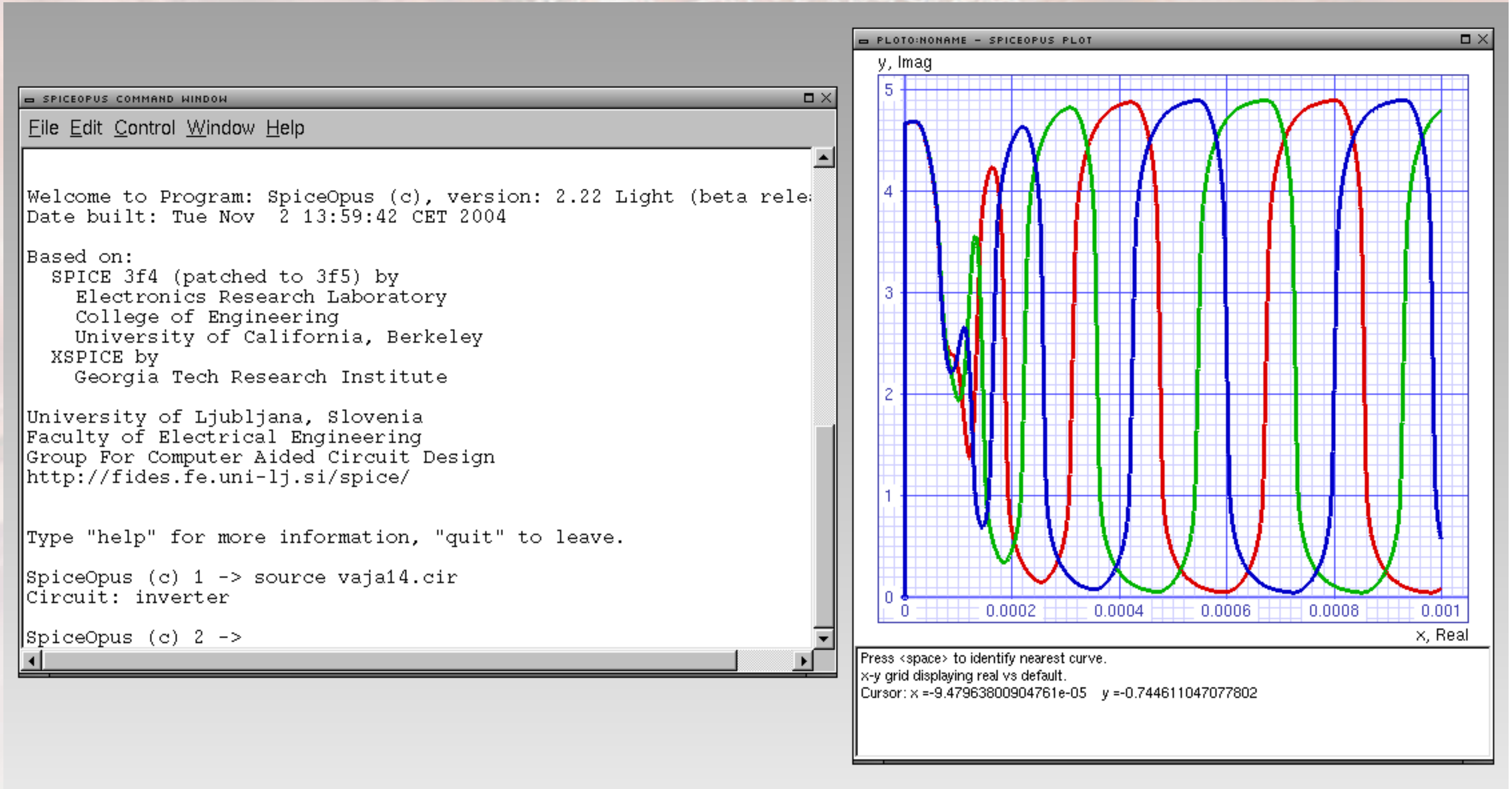


FIGURE 3-9
A six pole Bessel filter formed by cascading three Sallen-Key circuits. This is a low-pass filter with a cutoff frequency of 1 kHz.

Analog signal processing

- Initially, DOS theory is used exclusively to simulate the operation of analog circuits :
  - to simplify the circuit design and to
  - investigate the influence of certain parameters on the operation of the circuit.
- The behavior of the electrical network is described by systems of differential equations,
  - the solution (network response) is found using numerical integration procedures.

Analog signal processing – SPICE simulation tool



Pioneers of DSP

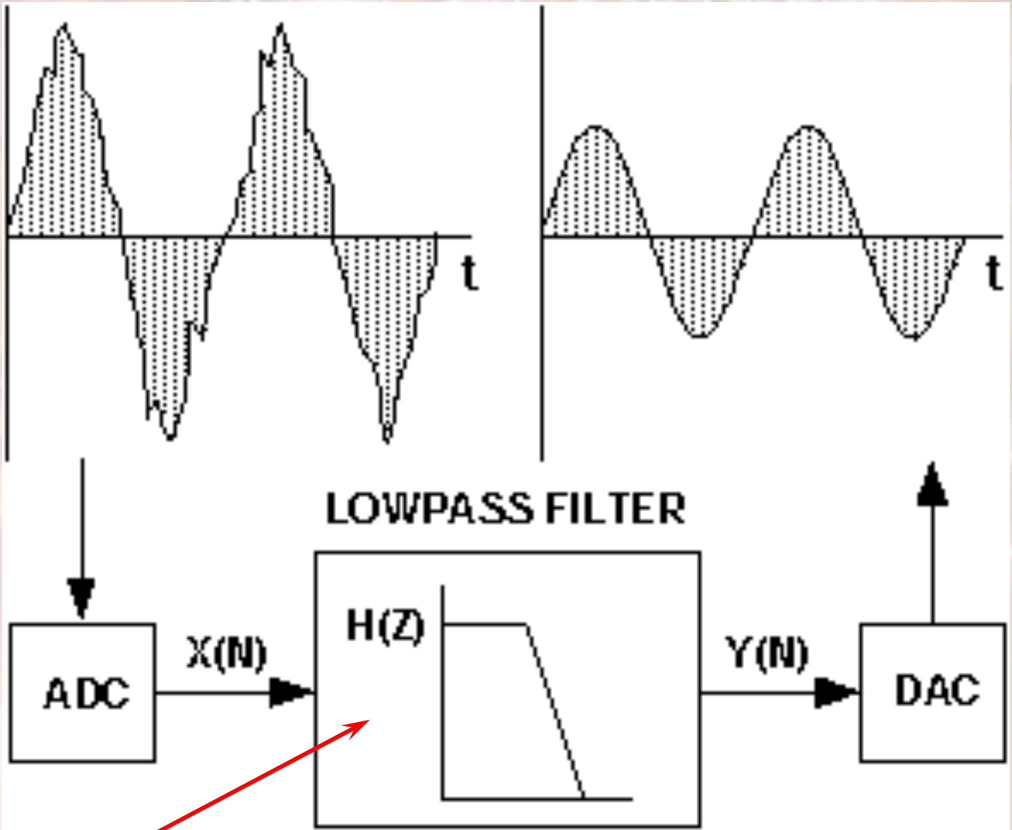
Alan V. Oppenheim

Ronald W. Schafer

Lawrence R. Rabiner

Digital signal processing

- Basic structure for digital signal processing



- processing of input signal samples - an example of a low-pass filter

Digital signal processing

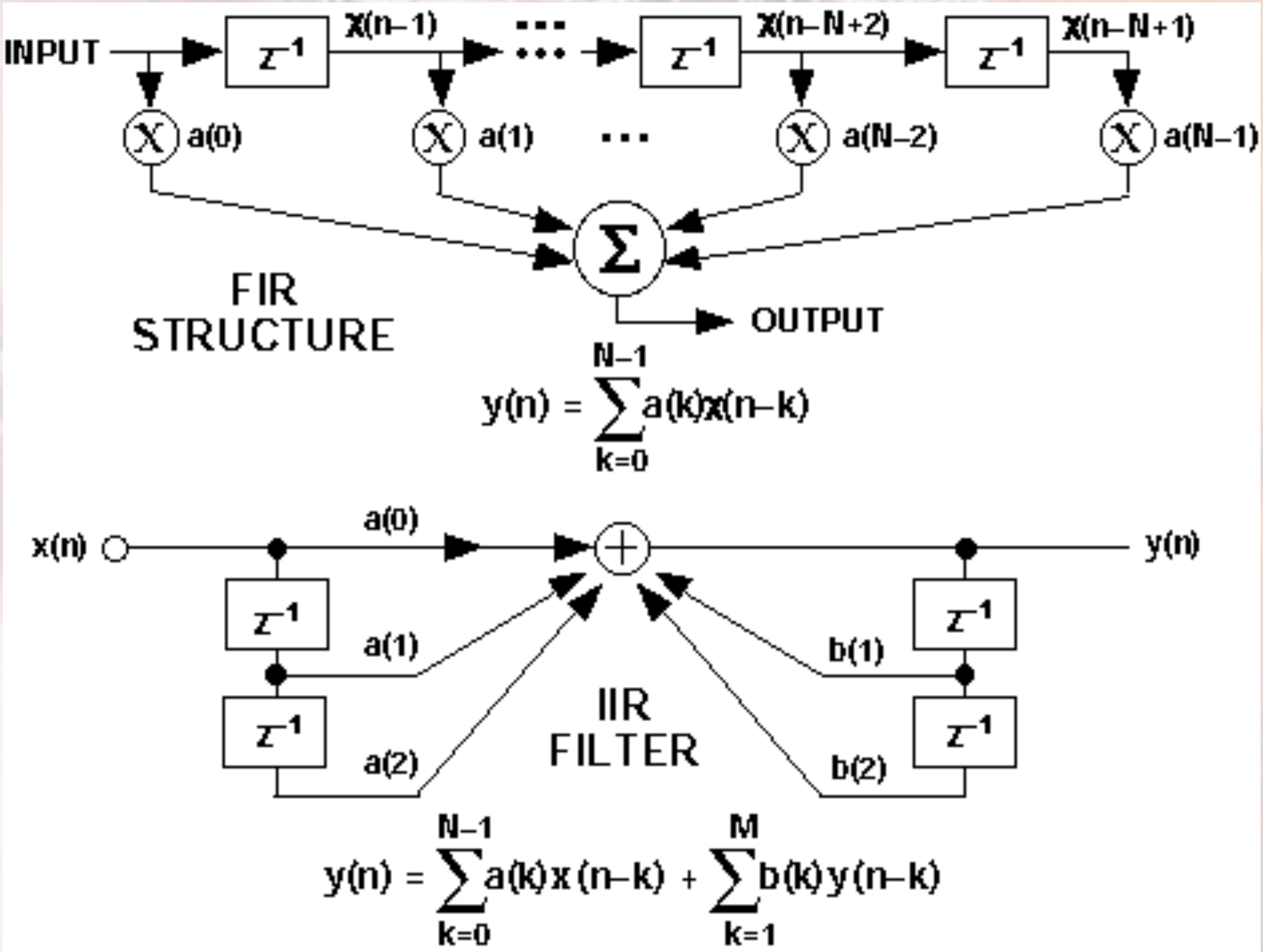
- Important advantages of DSP :
  - the accuracy of operations can be adjusted to the desired amount by system design,
  - the device is not ageing, ... the accuracy is always the same,
  - insensitivity to analog noise and cross-talk,
  - it is possible to perform processing that does not have an equivalent circuit in the analog domain at all,
  - software-only based implementation of processing operations is sufficient for many applications,

Digital signal processing

- advantages of DSP, ... continued
  - software implementation allows easy adaptation of the previously developed system to completely new tasks (Firmware flashing),
  - extremely complex signal processing is possible,
  - possibility of signal processing in blocks,
  - possibility of signal processing in transformation domains (DFT, DCT, ...),
  - digital signal representation is much more suitable for storage or transmission.

Digital signal processing

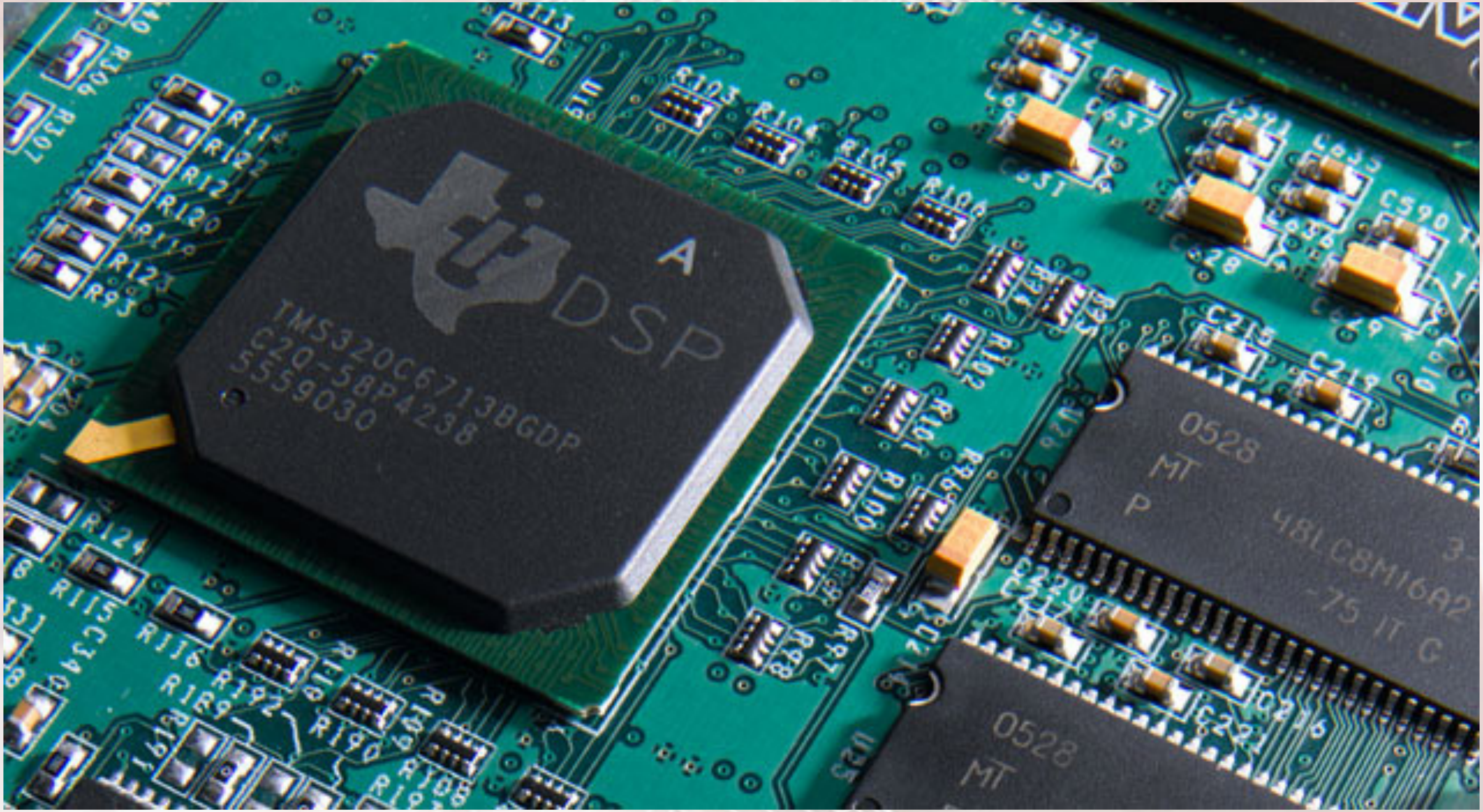
- an example of filter structures in a time-discrete domain



Dedicated processors for digital signal processing

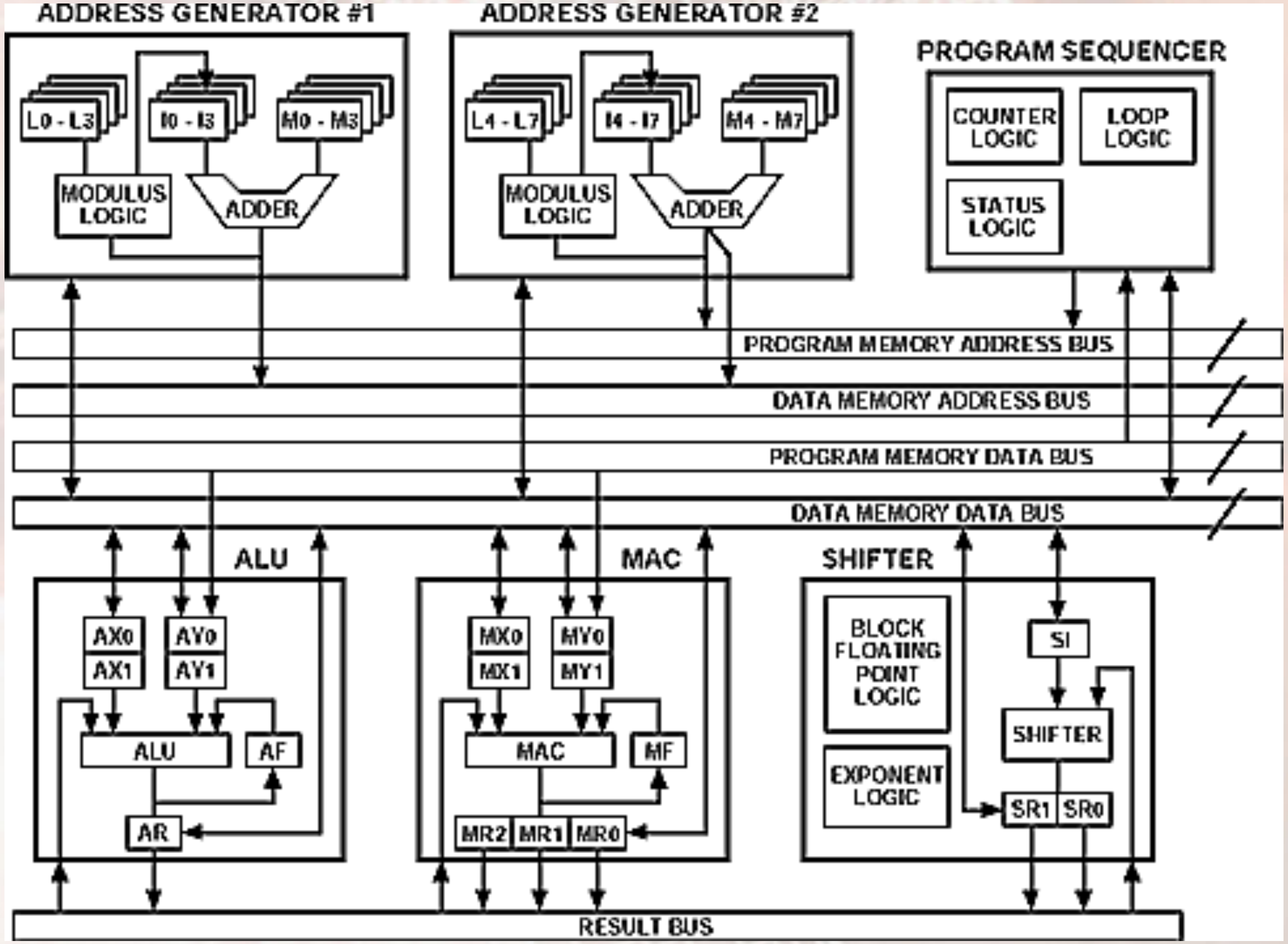
- For the needs of performing digital signal processing algorithms, specialized processors with an architecture adapted for such tasks have been developed, the so called
  - **Digital Signal Processors** (**DSP**)
- They have dedicated execution units for performing consecutive multiplication and addition operations.
- Support for time-parallel multiple operations performed within a single machine cycle.

An example of a TI DSP processor



Dedicated processors for digital signal processing

- Typical architecture of a simpler DSP processor



Digital signal processing

- example of digital filter implementation on ADSP21XX

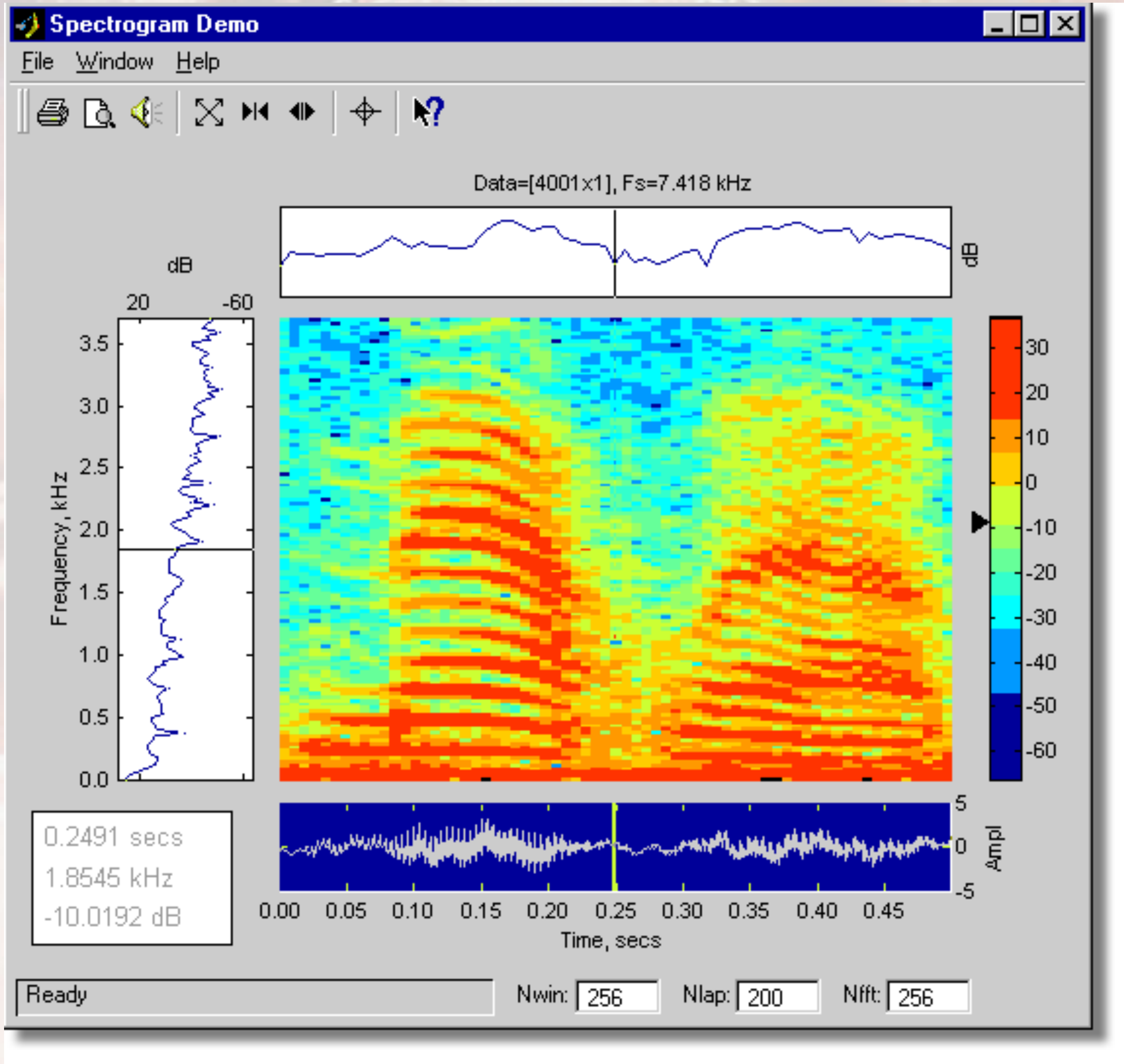
```
.ENTRY biquad;
biquad: CNTR = number_of_biquads
        DO sections UNTIL CE;
           SE=DM(I1,M2);
           MX0=DM(I0,M0), MY0=PM(I4,M4);
           MR=MX0*MY0(SS), MX1=DM(I0,M0), MY0=PM(I4,M4);
           MR=MR+MX1*MY0(SS), MY0=PM(I4,M4);
           MR=MR+SR1*MY0(SS), MX0=DM(I0,M0), MY0=PM(I4,M4);
           MR=MR+MX0*MY0(SS), MX0=DM(I0,M1), MY0=PM(I4,M4);
           DM(I0,M0)=MX1, MR=MR+MX0*MY0(RND);
sections:  DM(I0,M0)=SR1, SR=ASHIFT MR1 (HI);
         DM(I0,M0)=MX0;
         DM(I0,M3)=SR1;
         RTS;
.ENDMOD;
```

Digital signal processing

- Disadvantages of DSP:
  - the need for Analog/Digital (A/D) and Digital/Analog (D/A) converters for real-world connection .... to the media signals.
  - the need for a processor or dedicated hardware solutions even for the “dumbest” processing ... increased power consumption!
  - the maximum frequency of the input signal is limited by the sampling frequencies of the A/D and D/A converters and the processing capabilities of a DSP!

Example of digital audio signal processing

Example of digital speech processing



What have we learned?

- Definition of DSP
- Analog signal processing
- Benefits of DSP
- Examples of filter structures
- DSP processor
- Software implementation of filtration on DSP
- Disadvantages of DSP
- Illustration of typical applications

Signal digitization - sampling and reconstruction

prof.dr.sc. Davor Petrinović

Signal digitization

- Almost all modern MM systems immediately convert input signals into digital form.
- The key terms related to the digitization process are:
  - **sampling,** i.e., temporal or spatial discretization and
  - **quantization,** i.e., amplitude discretization.
- After digitization, all manipulations of the media signal are performed in the domain of signal samples.

Signal sampling

- Sampling is the process by which a time-continuous analog signal $x(t)$ is translated into a series of analog samples of the signal $x[n]$ at regular intervals:

$$x[n]=x(t_n), \quad \text{where } t_n=nT$$

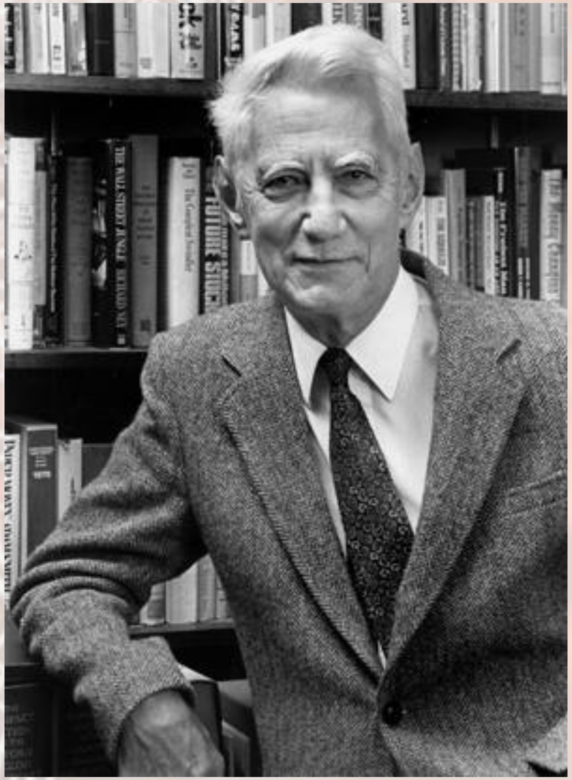
  - n is an integer index of the time sample, and
  - T is the sampling period, $f_s=1/T$
- Sampling of the signal is performed inside the A/D converter (and sometimes in front of it).

Unambiguity of time-discrete representation

- Are the input analog signal $x(t)$ and its samples $x[n]$ bound by a unique mapping in both directions?
  - are there two different signals $x(t)$ whose sampling gives the same samples $x[n]$?
  - is it at all possible to reconstruct the initial signal $x(t)$ from its samples $x[n]$ without error?
- The answer is given by **Nyquist-Shannon's sampling theorem** ...

Creators of the sampling theorem

Harry
Nyquist

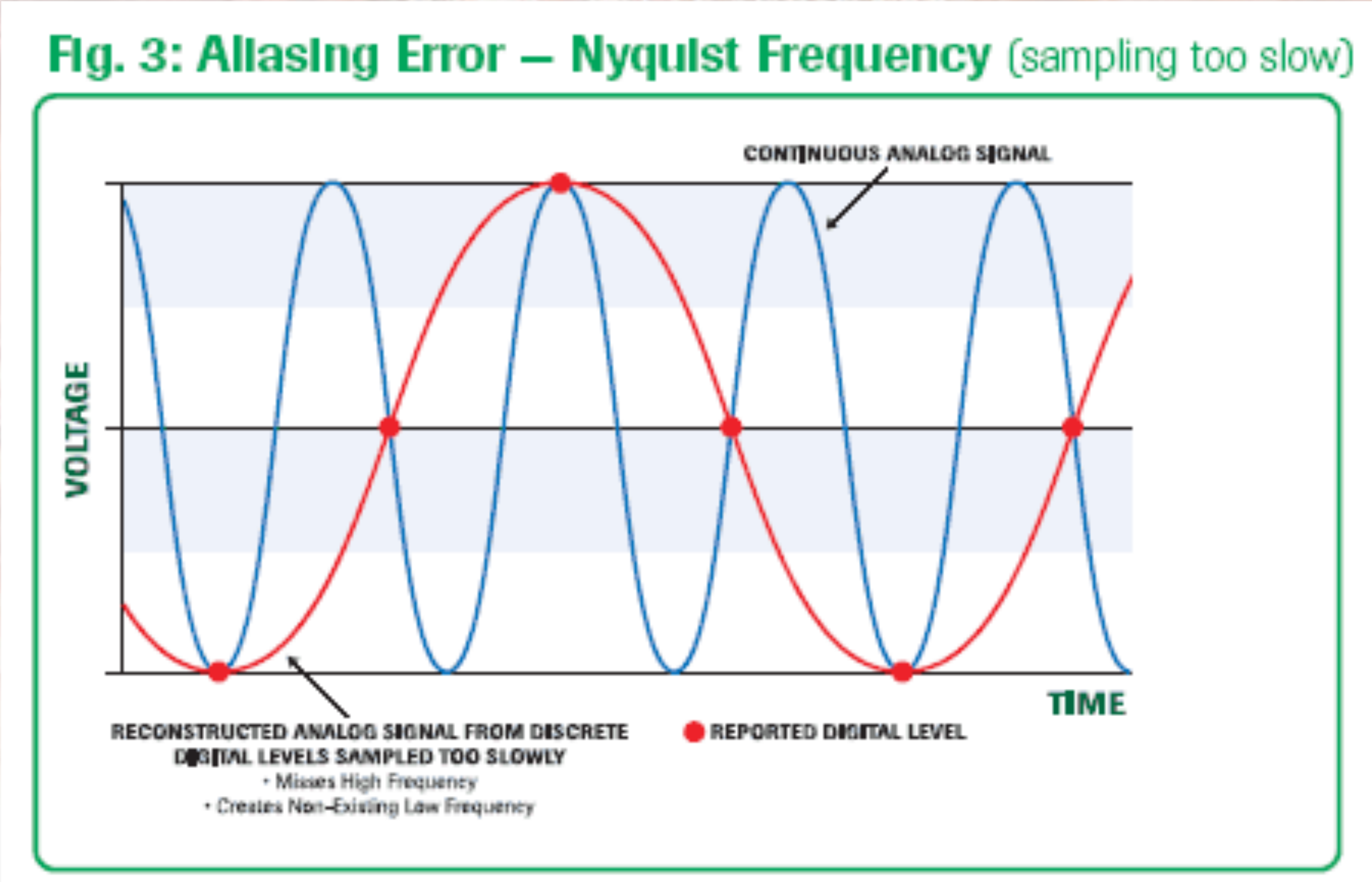


Claude E.
Shannon

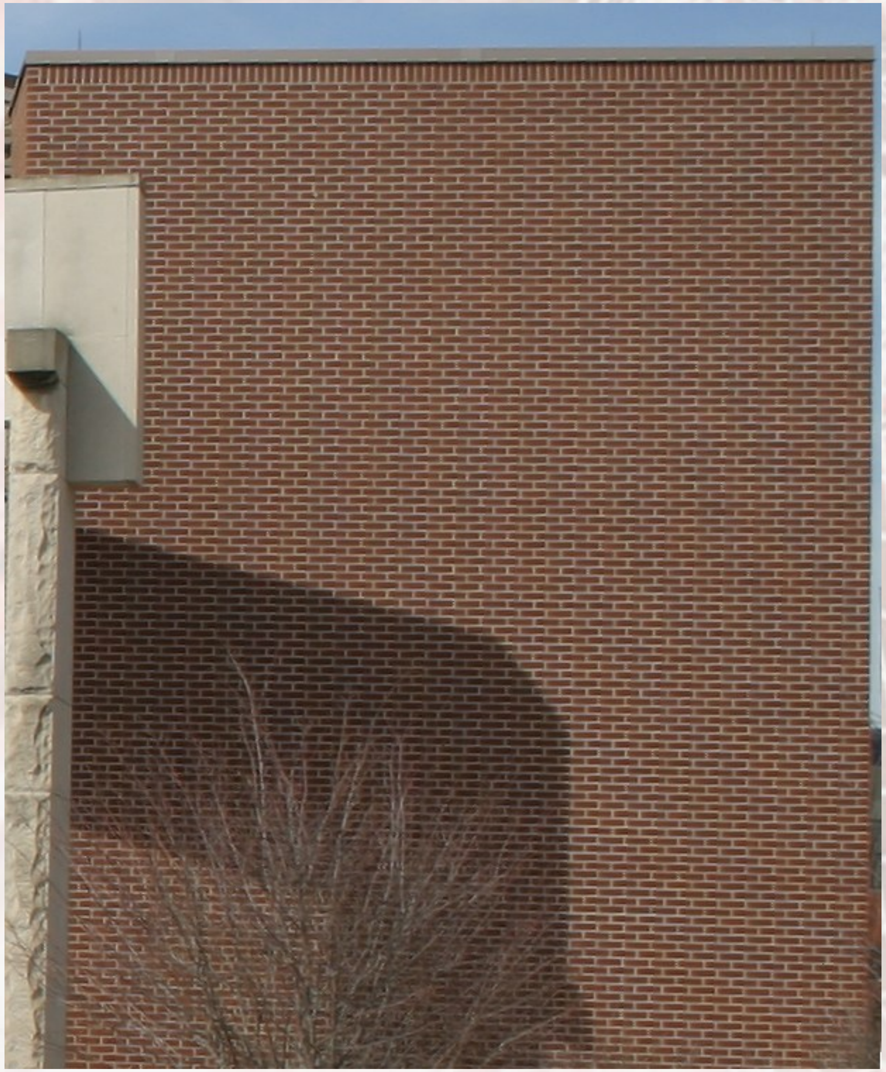
Unambiguity of time-discrete representation

- According to this sampling theorem ... a necessary and sufficient condition for fulfilling this unambiguity is that the input signal be band limited to the highest frequency of $1/(2T)=f_s/2$ [Hz]
  - the signal may have an arbitrary spectral composition, but **no spectral component may be above half the sampling frequency**.
- Thus, with **at least two samples** per sine signal period, it will be possible to reconstruct the initial signal **perfectly** for each time moment by interpolation procedure!

Example of insufficient signal sampling frequency



An example of insufficient image sampling frequency



- Original 615 x 750
- Three times decimated image 205 x 250

- Moiré pattern

Signal reconstruction

- Signal reconstruction is the reverse procedure by which the original continuous signal $x(t)$ is formed from samples $x[n]$ by the interpolation procedure

- The time discrete sequence $x[n]$ is conveniently represented by a signal in the continuous domain $x_s(t)$ consisting of an infinite sum of shifted Dirac pulses multiplied by the sample amplitudes $x[n]$.

$$
\begin{aligned}
x_s(t) = x(t) \cdot T \sum_{n=-\infty}^{\infty} \delta(t - nT) \\
= T \sum_{n=-\infty}^{\infty} x(t) \cdot \delta(t - nT) \\
= T \sum_{n=-\infty}^{\infty} x(nT) \cdot \delta(t - nT) \\
= T \sum_{n=-\infty}^{\infty} x[n] \cdot \delta(t - nT)
\end{aligned}
$$

Signal reconstruction

- The signal $x_s(t)$ has spectrum $X_s(f)$ which in the frequency range $-f_s/2$ to $+f_s/2$ **is exactly equal to the spectrum of the initial signal** $X(f)$.
- Unfortunately, the problem is that the spectrum $X_s(f)$ is an **infinite periodic function** in the variable f with period f_s, where $f_s=1/T$ is the sampling frequency.
- Removal of unwanted spectral replicas can be performed by filtering the sampled signal $x_s(t)$ with a filter $h(t)$ whose frequency response $H(f)$ has the shape of an ideal rectangle (brickwall-filter).

Signal reconstruction

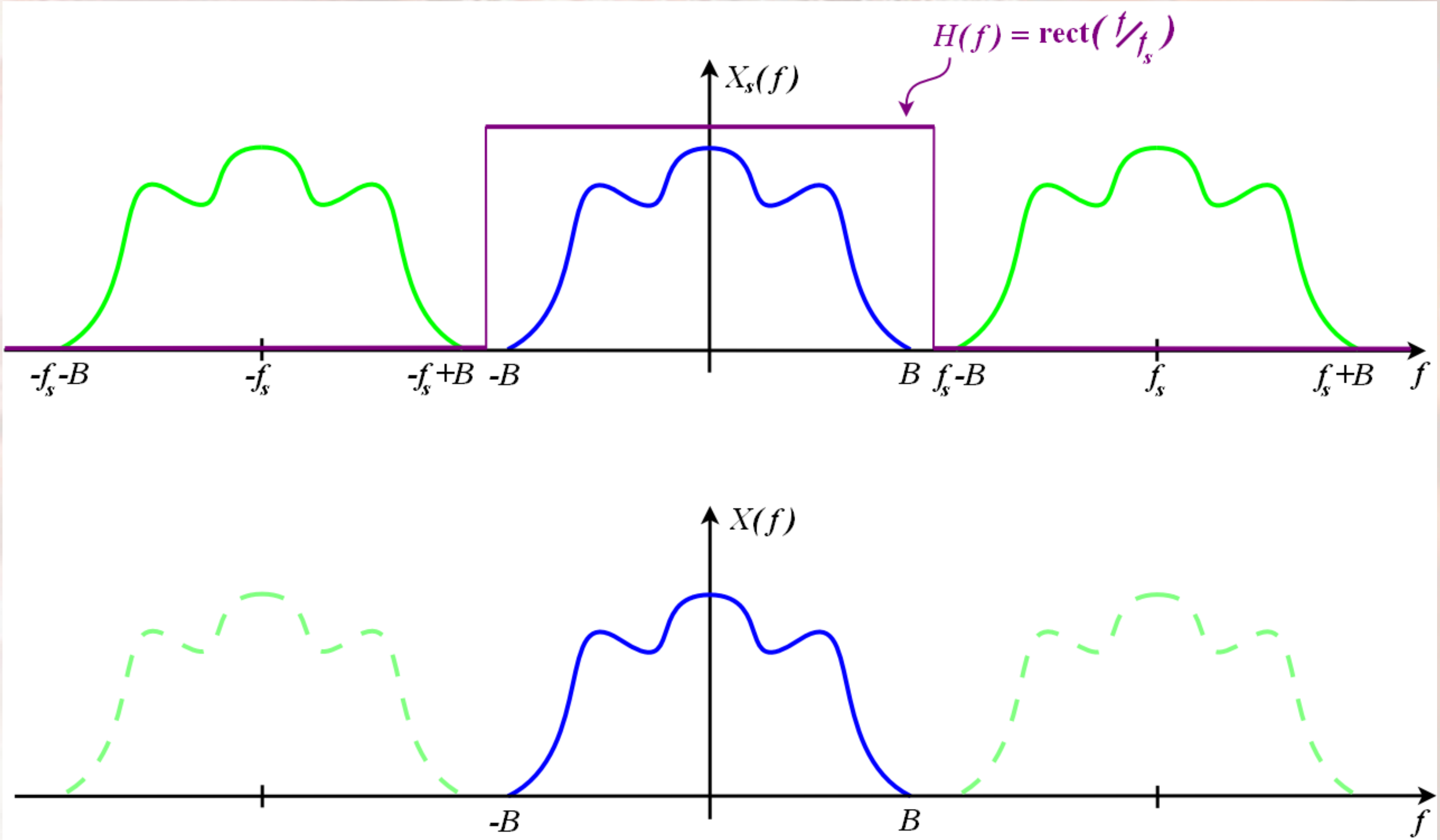
- Filter *H*(*f*) is called the ideal reconstruction filter:

$$H(f) = \mathrm{rect}\left(\frac{f}{f_s}\right) = \begin{cases} 1 & |f| < \frac{f_s}{2} \\ 0 & |f| > \frac{f_s}{2} \end{cases}$$

- Since the signal reconstructed by applying the ideal reconstruction filter *H*(*f*) at the positions of the samples *x*(*nT*) will be exactly equal to the samples of the time-discrete sequence x[n], it is also called the **interpolation filter**.

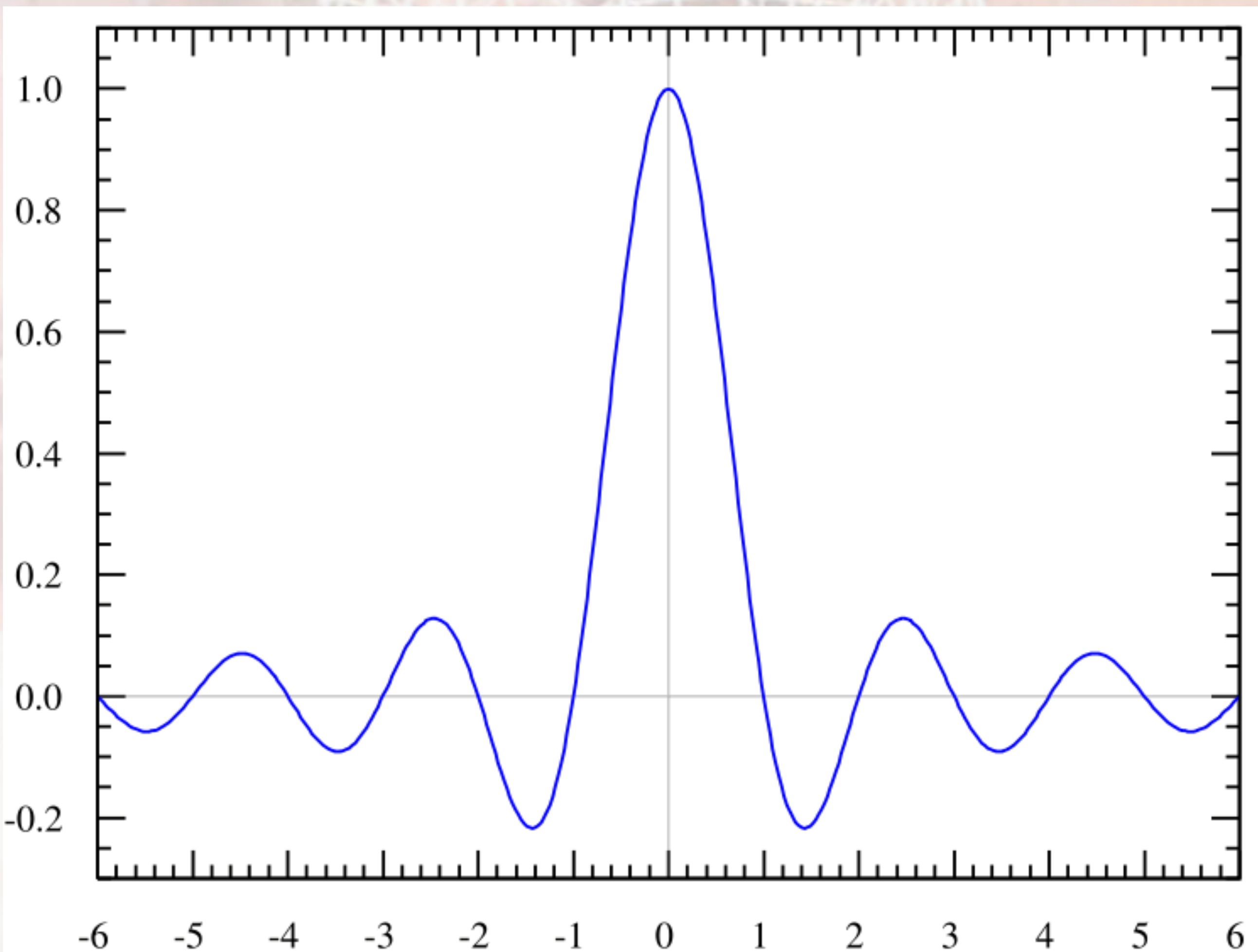
Signal reconstruction

- Operation of interpolator *H*(*f*) in the frequency domain



Impulse response of an ideal interpolator

- Sinc function for T=1 $h(t)=\sin(\pi t)/(\pi t)$



Impulse response of an ideal interpolator

$$
\begin{aligned}
h(t) &= \mathcal{F}^{-1}\{H(f)\} \\
&= \int_{-\infty}^{\infty} H(f) e^{i2\pi ft}\, df \\
&= \int_{-\infty}^{\infty} \operatorname{rect}\left(\frac{f}{f_s}\right) e^{i2\pi ft}\, df \\
&= \int_{-f_s/2}^{f_s/2} e^{i2\pi ft}\, df \\
&= \frac{1}{i2\pi t} e^{i2\pi ft}\Big|_{-f_s/2}^{f_s/2} \\
&= \frac{1}{\pi t} \frac{\left(e^{i\pi f_s t} - e^{-i\pi f_s t}\right)}{2i} \\
&= \frac{\sin(\pi f_s t)}{\pi t} \\
&= f_s \operatorname{sinc}(f_s t)
\end{aligned}
$$

- Interpolation procedure

$$
\begin{aligned}
x(t) &= h(t) * x_s(t)\,, \\
&= h(t) * \sum_{n=-\infty}^{\infty} T \cdot x[n] \cdot \delta(t - nT) \\
&= \sum_{n=-\infty}^{\infty} x[n] \cdot T \cdot [h(t) * \delta(t - nT)] \\
&= \sum_{n=-\infty}^{\infty} x[n] \cdot T \cdot h(t - nT) \\
&= \sum_{n=-\infty}^{\infty} x[n] \cdot (T f_s) \operatorname{sinc}(f_s(t - nT)) \\
&= \sum_{n=-\infty}^{\infty} x[n] \cdot \operatorname{sinc}\left(\frac{t - nT}{T}\right)
\end{aligned}
$$

Signal reconstruction

- In real MM systems, the reconstruction process is performed within a **D/A converter**.
- Due to the noncausality of the sinc function, the actual interpolators can **only approximate** the ideal interpolation filter $H(f)$.
- A significant step forward in the accuracy of the approximation was achieved by applying **oversampling procedures** for implementation of A/D and D/A converters, and especially by using structures based on the **Delta-Sigma modulator**.

What have we learned?

- Signal sampling
- Sampling frequency
- The condition for unambiguous sampling
- Nyquist-Shannon sampling theorem
- Insufficient sampling frequency error
- Signal reconstruction
- Interpolation / reconstruction filter
- Ideal reconstruction - sinc filter
- D/A converters

Signal digitization - quantization

prof.dr.sc. Davor Petrinović

Signal quantization

- The process by which analog signal samples of infinite precision are converted into discrete value numbers (integers).
- In MM systems, quantization is typically performed on two levels:
  - immediately **after sampling** as part of A/D input signal conversion and
  - **within lossy signal coding procedures**, where the digital signal is further transformed for the purpose of data compression.

Signal quantization

- For now, we will deal exclusively with this first type of quantization within the A/D converter.
- The result of this quantization is a digital media signal of nominal (complete) accuracy, which is called **raw signal**.
- This signal is further transformed by means of coding procedures into **encoded signal** from which, depending on the encoder type, it is possible to reconstruct the initial signal of nominal accuracy either approximately or perfectly without any loss.

Signal quantization

- The most common type of quantizer used in A/D converter of MM system is a **linear or uniform quantizer**:
  - this means that the spacing Δ of all quantization levels is equal i.e.
  - there is an approximately linear relationship between the amplitude of the input signal and the output codes.
- The uniform quantization operation can be mathematically written as follows :

$$x_q = \Delta \cdot \mathrm{round}(x / \Delta)$$

Signal quantization

- A common name for this type of analog signal representation is **linear PCM**, where **PCM** stands for **P**ulse **C**ode **M**odulation.
- The name is inherited from the past, when as an alternative to the PCM other signal modulation techniques were also considered :
  - pulse-amplitude modulation, PAM
  - pulse-width modulation, PWM
  - pulse-position modulation, PPM
  - sigma-delta modulation, ΣΔ
  - adaptive-delta modulation, ADM.

Signal quantization

- However, ... physically achievable uniform quantizer is characterized by two basic parameters :
  - **spacing of quantization levels** , Δ
  - but also, **the total number of quantization levels**, *N.*
- The range from the lowest to the highest input voltage is equal to the product $N\Delta$
  - and is called **the input dynamics** of the A/D converter.

Quantizer overexcitation (saturation)

- The input signal by amplitude and offset must be adjusted to the input dynamics of the converter:
  - The A/D converter cannot represent input voltage values outside this range.
  - The phenomenon of getting out of dynamics is called **saturation or overexcitation**.
  - It occurs either below the lowest input voltage corresponding to code 0 or above the highest input voltage corresponding to code N-1.

Signal quantization

- To facilitate implementation, the total number of quantization levels of the A/D converter is usually the integer power of 2 , ... $N=2^b$
  - The output value can therefore be represented as a *b*-bit binary number.
- The output width *b* is also called the A/D converter **resolution** (in bits).
- In case the input dynamics is A (in Volts), then the distance of two adjacent quantization levels is equal to $\Delta=A\cdot 2^{-b}$. This quantization step defines the so-called quantizer **granulation error**.

Signal quantization

- Converters adapted to **bipolar signals** are often used in digital signal processing systems
- For them, the input dynamics is symmetric around zero .... +/- $A/2$ [V]
- The output code of bipolar converters is usually represented in a b-bit **two’s complement format**
  - input voltage 0V gives code 0 ,
  - input voltage –A / 2 gives the most negative code $-2^{(b-1)}$,
  - while input voltage $(2^{(b-1)}-1)/(2^{b})*A$ gives code $2^{(b-1)}-1$.
- The 8-bit bipolar converter has codes from -128 to 127

Signal quantization

- Resolution directly limits the accuracy of A/D converters
  - because all input voltages whose absolute deviation from the i-th quantization level is less than half of the steps $\Delta/2=(A\cdot 2^{-b})/2$ will be represented with the same output code (i-1).
- In that sense, the quantization operation is extremely nonlinear and has no inverse ... once quantized, the signal can no longer be corrected (dequantized)!

Quantization error

- For the purpose of objective estimation of quantization error, it is convenient to determine the ratio of signal variance σ_x^2 and quantization error variance σ_e^2.
- Samples of the signal $x[n]$ can be viewed as the realization of a random process x.
- Also, quantization error sequence $e[n]=x[n]-x_q[n]$ can be viewed as the realization of the random process e.

Quantization error

- variances are then found as :
- ... $\sigma_x^2 = E((x - E(x))^2)$
  $\sigma_e^2 = E((e - E(e))^2)$
- The operator $E(\circ)$ is the statistical expectation operator.
- The quantization error e is always contained within one quantization interval $\pm \Delta/2$.
- Determining statistical expectation assumes knowledge of probability density functions for processes x and e denoted by $f_x(x)$ and $f_e(e)$, i.e., ***pdf*** – ***p****robability* ***d****ensity* ***f****unction.*

Quantization error

- $f_x(x)$ depends on the input signal, while the error probability density $f_e(e)$ depends on the parameters of the quantizer, but also on the input signal from which it originated.
- It is often assumed that the processes x and e are statistically independent, which significantly facilitates the analysis.
- Also, it is often assumed that process e has a uniform probability density in the interval $\pm\Delta/2$, i.e. that all errors in that interval are equally probable, which is a reasonable assumption for a large N
- ... $f_e(e) = 1/\Delta$

Quantization error

- With a uniform error distribution, the expectation of quantization error is zero due to symmetry, i.e., the quantizer is not biased :

$$E(e) = \int_e e \cdot f_e(e) \cdot de$$

$$= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} e \cdot de = 0$$

- For simplicity, it can be assumed that the input signal also has a zero mean value *E*(*x*)=0

Quantization error

- the expressions for variances are therefore simplified : ... $\sigma_x^2 = E(x^2)$, $\sigma_e^2 = E(e^2)$

  and correspond to signal and error energies.
- The variance of the quantization error with uniform $f_e(e)$ is therefore :

$$\sigma_e^2 = E(e^2) = \int_e e^2 \cdot f_e(e) \cdot de = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} e^2 \cdot de$$

$$= \frac{1}{\Delta} \left. \frac{e^3}{3} \right|_{-\Delta/2}^{\Delta/2} = \frac{\Delta^2}{12}$$

Quantization error

- As a first example, let us select the input signal $x[n]$ in the form of a cosine signal of the maximum amplitude A/2 :

$$x[n] = \frac{A}{2}\cos(\omega_0 n)$$

- The argument $\phi = \omega_0 n$ can be treated as a random variable with uniform probability density $f_\phi(\phi) = 1/(2\pi)$ on the interval $[0,2\pi]$, and x as a random process in the variable ϕ: $x = A\cos(\phi)/2$.
- Let us now determine the process variance of x ...

Quantization error

- variance follows as:

$$\sigma_x^2 = E(x^2) = \int_{\phi} x(\phi)^2 f_{\phi}(\phi) d\phi = \frac{A^2}{4 \cdot 2\pi} \int_0^{2\pi} \cos(\phi)^2 d\phi =$$

$$= \frac{A^2}{8\pi} \int_0^{2\pi} \left(\frac{1}{2} + \frac{\cos(2\phi)}{2} \right) d\phi = \frac{A^2}{8}$$

- If we now express the input dynamics A as $A=2^b\Delta$ ...

$$\sigma_x^2 = \frac{2^{2b}\Delta^2}{8}$$

Quantization error

- ... we can finally determine the required quotient of variances, which is most often expressed in logarithmic measure expressed in decibels :

$$SQNR(b) = 10\log_{10}\frac{\sigma_x^2}{\sigma_e^2} = 10\log_{10}\frac{\dfrac{2^{2b}\Delta^2}{8}}{\dfrac{\Delta^2}{12}} =$$

$$= 20b\log_{10}2 + 10\log_{10}\frac{3}{2}$$

$$= 6.02\cdot b + 1.76\ \ [dB]$$

Quantization error

- The magnitude of SQNR(b) represents the logarithmic relationship between signal level and quantization error level, and the abbreviation SQNR comes from the English name: ***S**ignal-to-**Q**uantization-**N**oise-**R**atio.*
- The expression shows that the quantization error is **reduced by 6dB for each additional bit** of the quantizer resolution!
- The offset of the affine function of the quantization error model depends on the properties of the signal and its *pdf* function $f_x(x)$, and for the cosine signal it is equal to 1.76dB.

Quantization error

- Suppose that the input signal of arbitrary waveform has a known variance σ_x^2 , i.e., it has an effective (rms) value of σ_x
- If we want to prevent saturation of the output of the A/D converter, then the peak value of this signal must not be higher than A/2.
- The quotient of the peak and rms values of the signal is known as the **crest factor** δ :

$$\delta = \frac{A/2}{\sigma_x}$$

Quantization error

- The input dynamics that avoid saturation is therefore : $A = 2\delta\sigma_x$, so the quantization step must be:

$$\Delta = 2\delta\sigma_x \cdot 2^{-b}$$

- Now, we can directly determine the SQNR ratio:

$$SQNR(b) = 10\log_{10}\frac{\sigma_x^2}{\sigma_e^2} = 10\log_{10}\frac{\sigma_x^2}{\frac{4\sigma_x^2\delta^2}{12}2^{-2b}} =$$

$$= 20b\log_{10}2 + 10\log_{10}\frac{3}{\delta^2}$$

Quantization error

- The resulting expression is identical to the example for the cosine signal, except for the model offset which is now dependent on the crest factor δ.
- For the cosine signal, the factor δ is sqrt(2), so it is obvious that we got the expected result.
- The factor δ is limited from below with a minimum value of δ =1 for the rectangular (or DC) signal, for which the highest SQNR is achieved for the selected resolution *b*.
- The biggest problems for the quantizer are the signals with large "peaks" and low energy (large δ).

An example of a digitization process

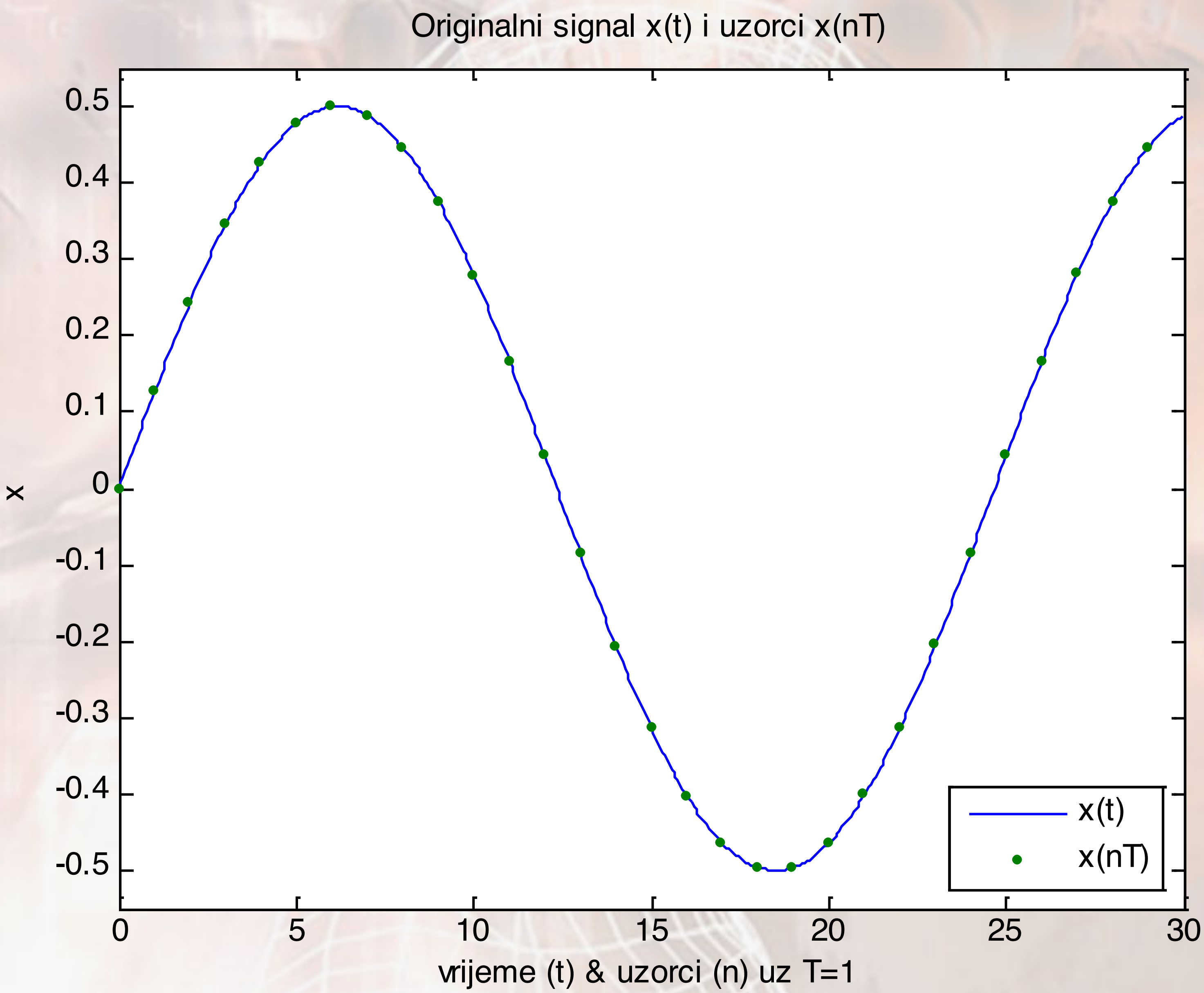
- Cosine-shaped input signal $x[n]=A/2\cos(\omega_0 n)$
  - converter input dynamics is A=1,
  - randomly selected frequency ω_0 in the interval 0 to 1,
  - signal duration of ½ seconds,
  - sampling frequency of fs=8000 Hz.
- Let us simulate:
  - signal sampling,
  - b-bit signal quantization, (b=4, 8 and 1-16)
  - approximate signal reconstruction, and finally
  - determine the actual and expected *SQNR*.

Quantization simulation in Matlab

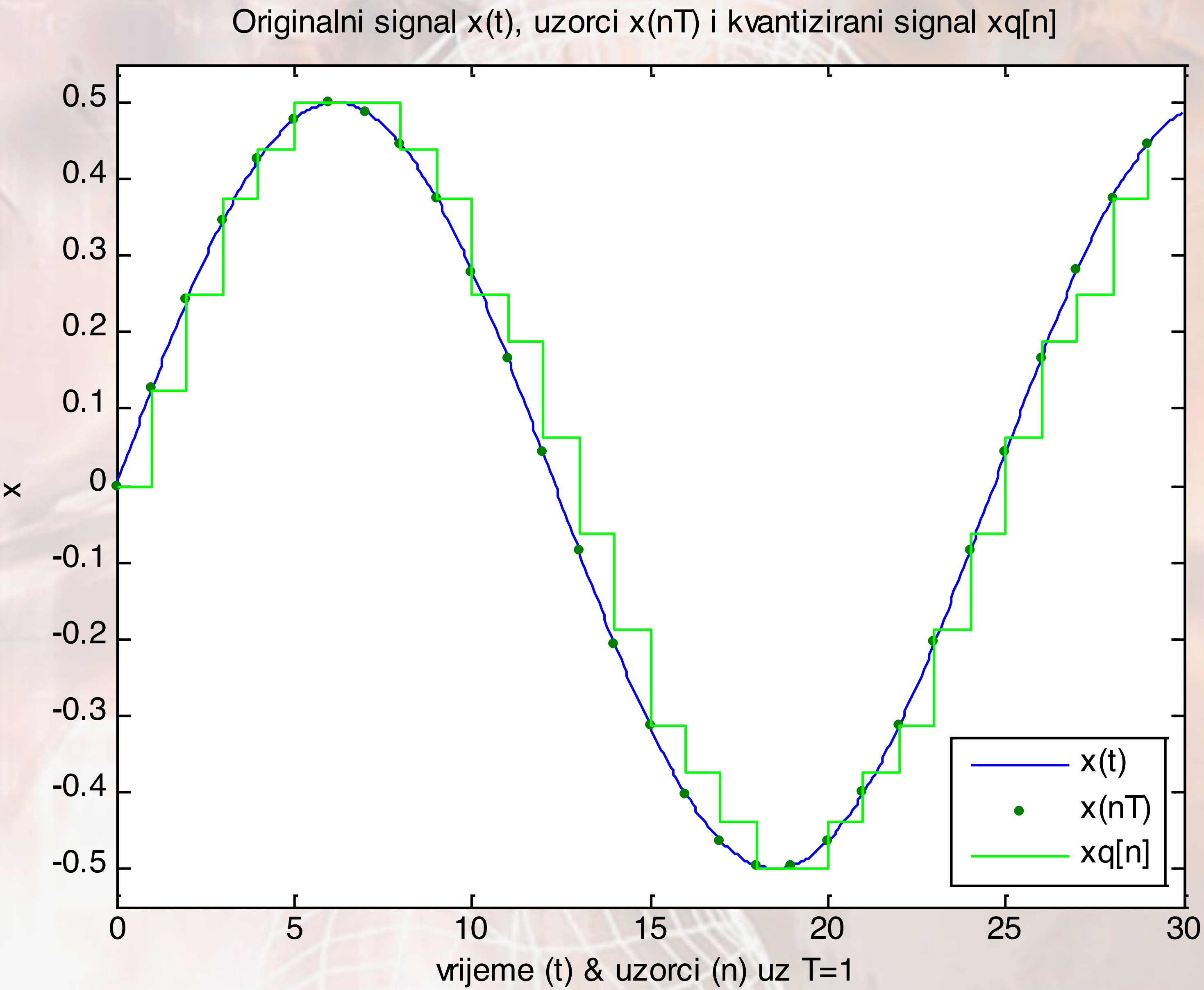
```
fs=8000;        % frekvencija otipkavanja
N=fs/2;         % broj uzoraka signala od 1/2 sec
n=[0:N-1];      % vremenska os
A=1;            % dinamika pretvornika

while (1)       % ponavljaj za razne frekvencije
   w=rand;      % slucajna frekvencija (0..1 [rad])
   y=A/2*cos(w*n);      % vrem. diskretni signal
   bv=[1:16];           % za cijeli niz rezolucija
   for b=bv,            % za sve rezolucije
      D=A*2^-b;         % razmak kvant. nivoa
      yq=D*round(y/D);% simuliraj kvantizaciju
      er=y-yq;          % pogreska kvantizatora
      E_er=mean(er.^2);   % ocekivanje kvadrata pogreske
      E_y=mean(y.^2);     % ocekivanje kvadrata signala
      SNR(b)=10*log10(E_y/E_er); % Odnos signal sum [dB]
   end;         % od petlje po rezolucijama
   % Ocekivani odnos signal sum za zadanu rezoluciju
   SNR_oc=20*bv*log10(2) - 10*log10(2/3);
end;
```

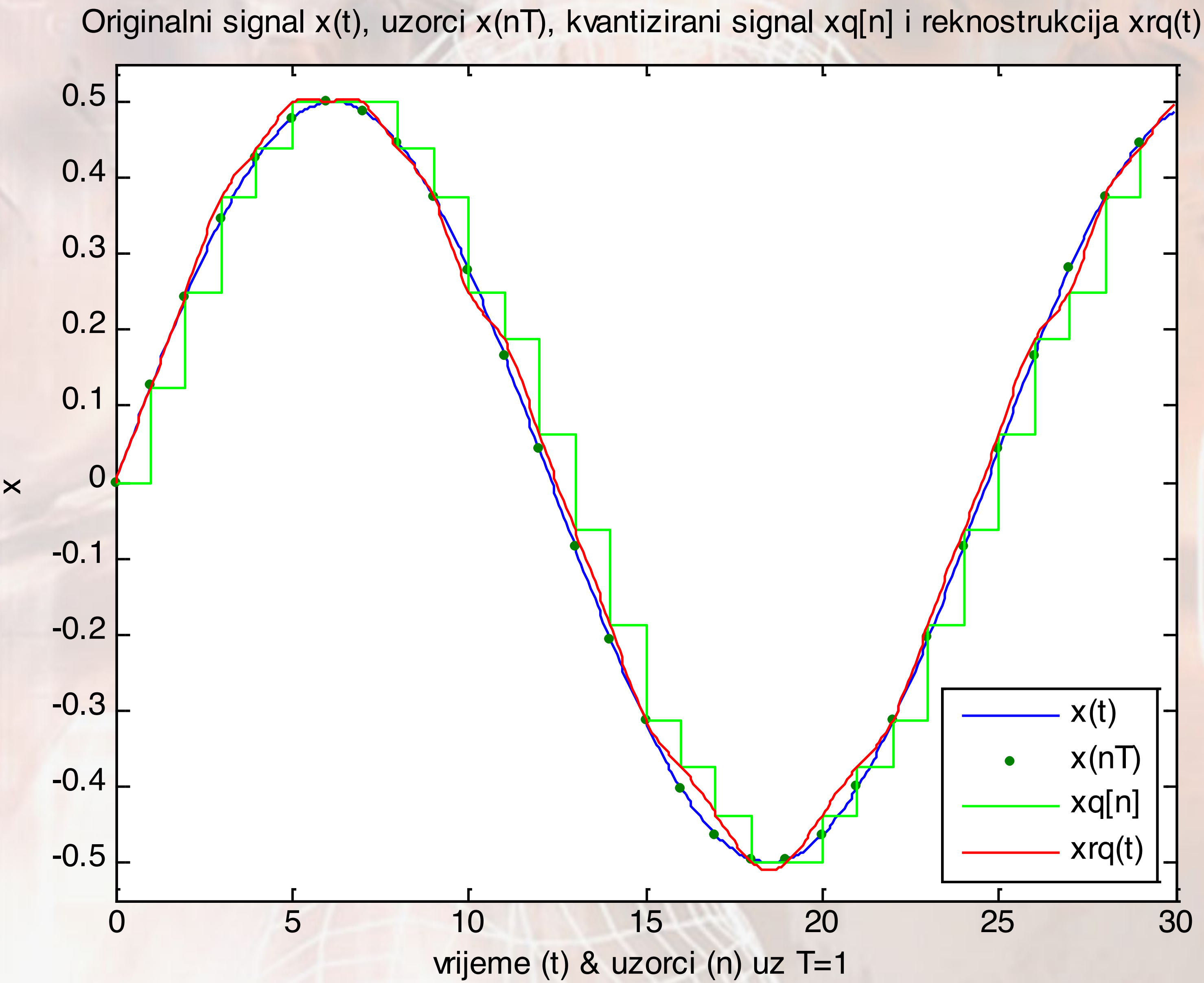
Example 1 - sampling



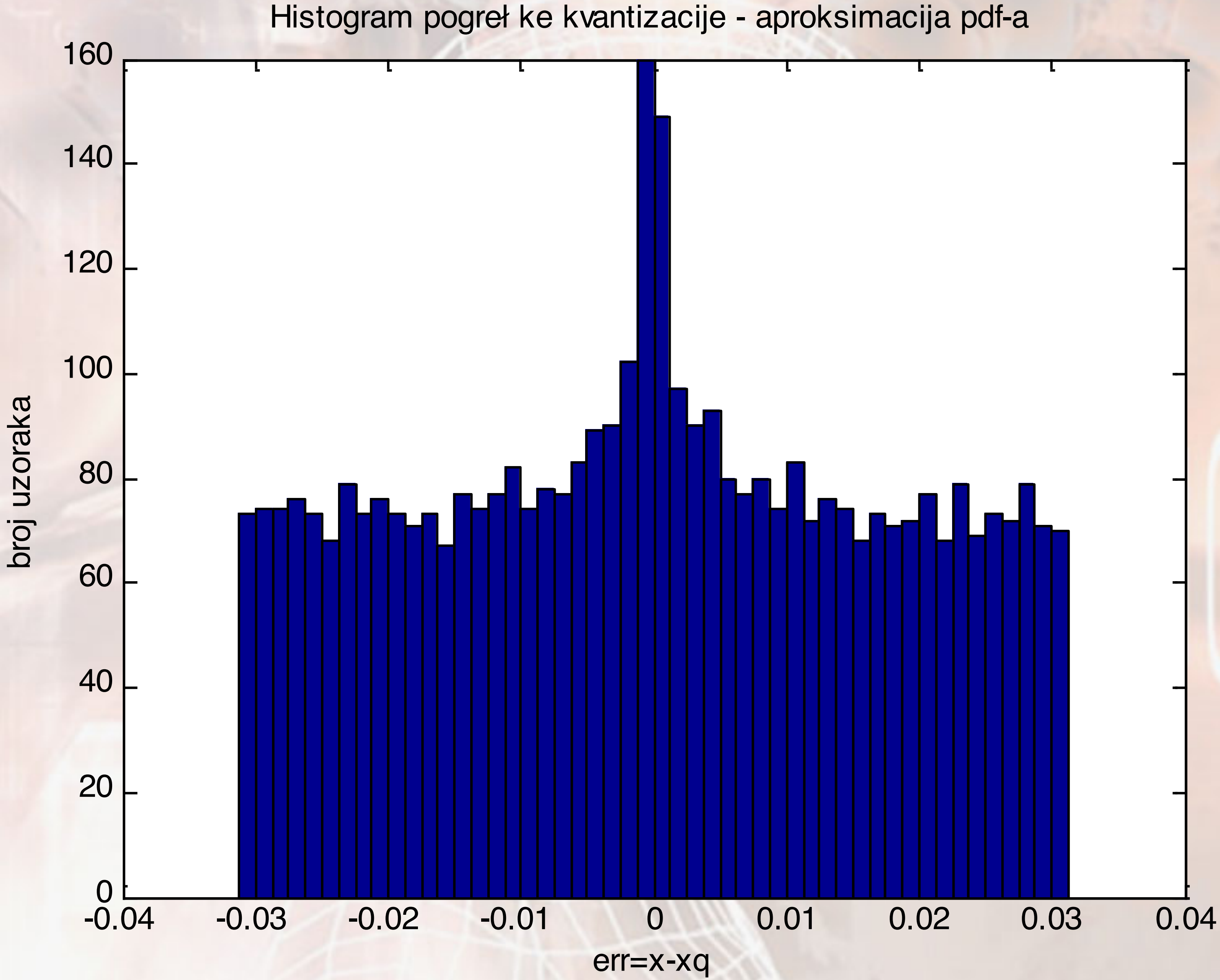
Example 1 - quantization, *b*=4



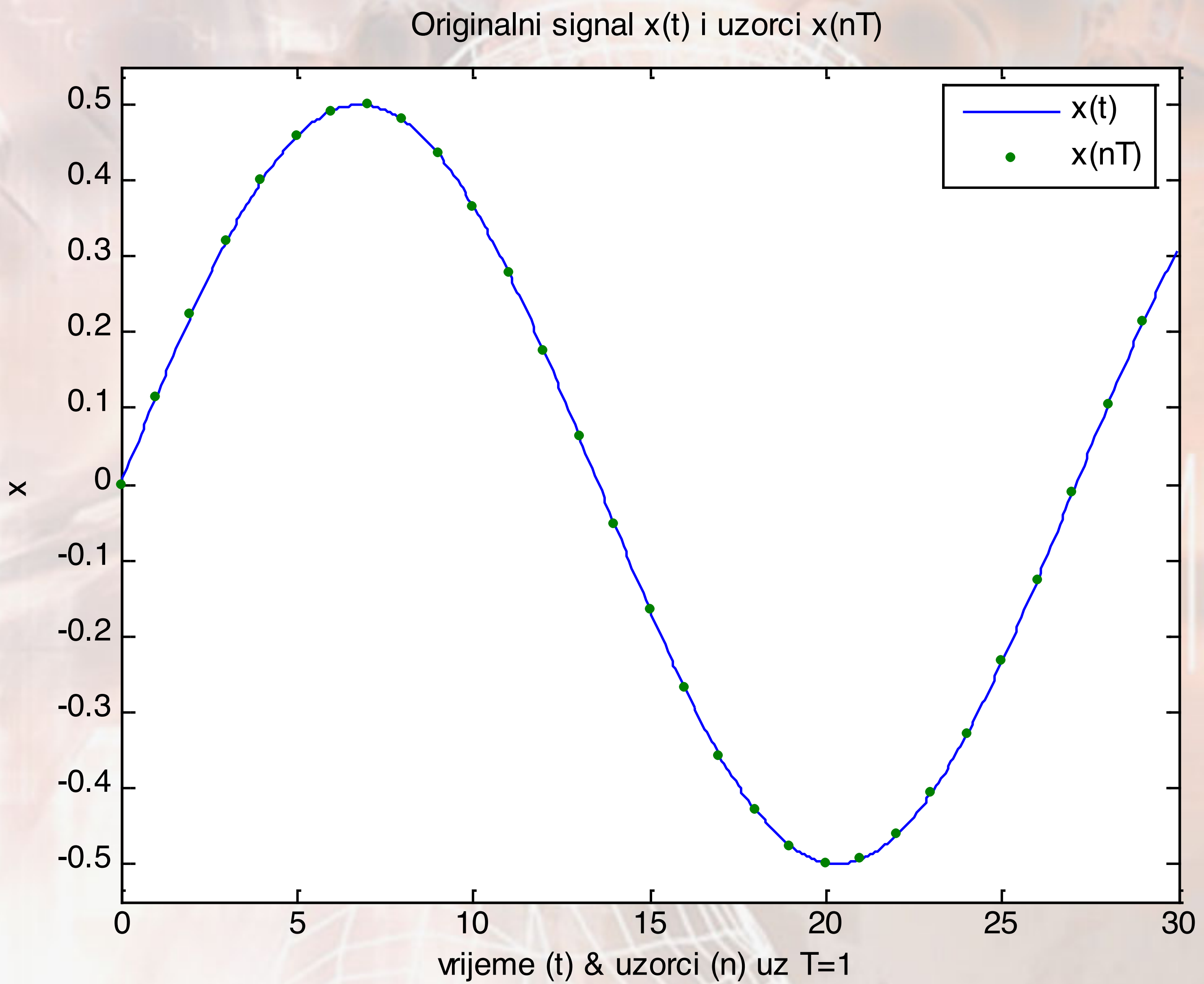
Example 1 - reconstruction, *b*=4



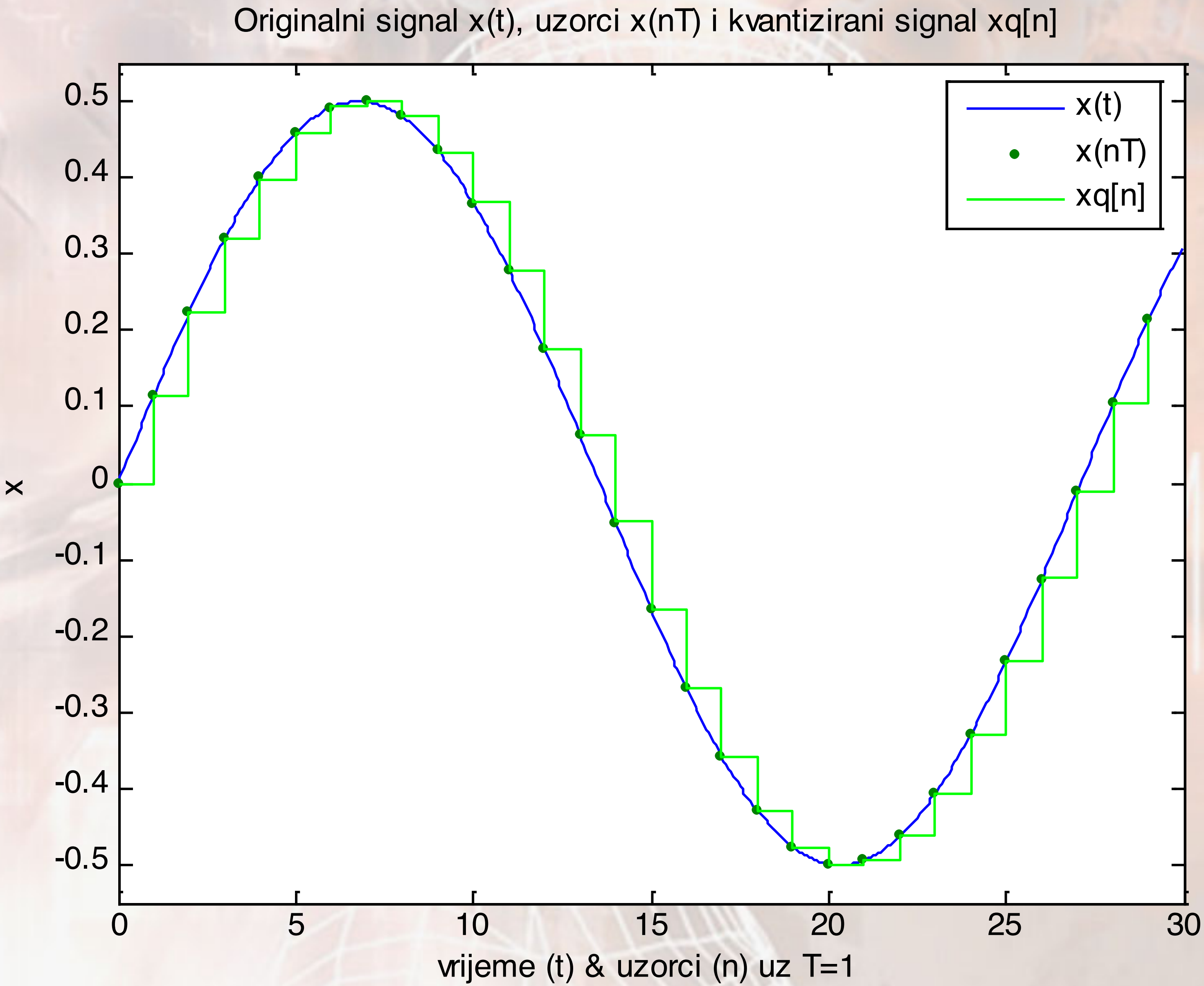
Example 1 – error histogram, *b*=4



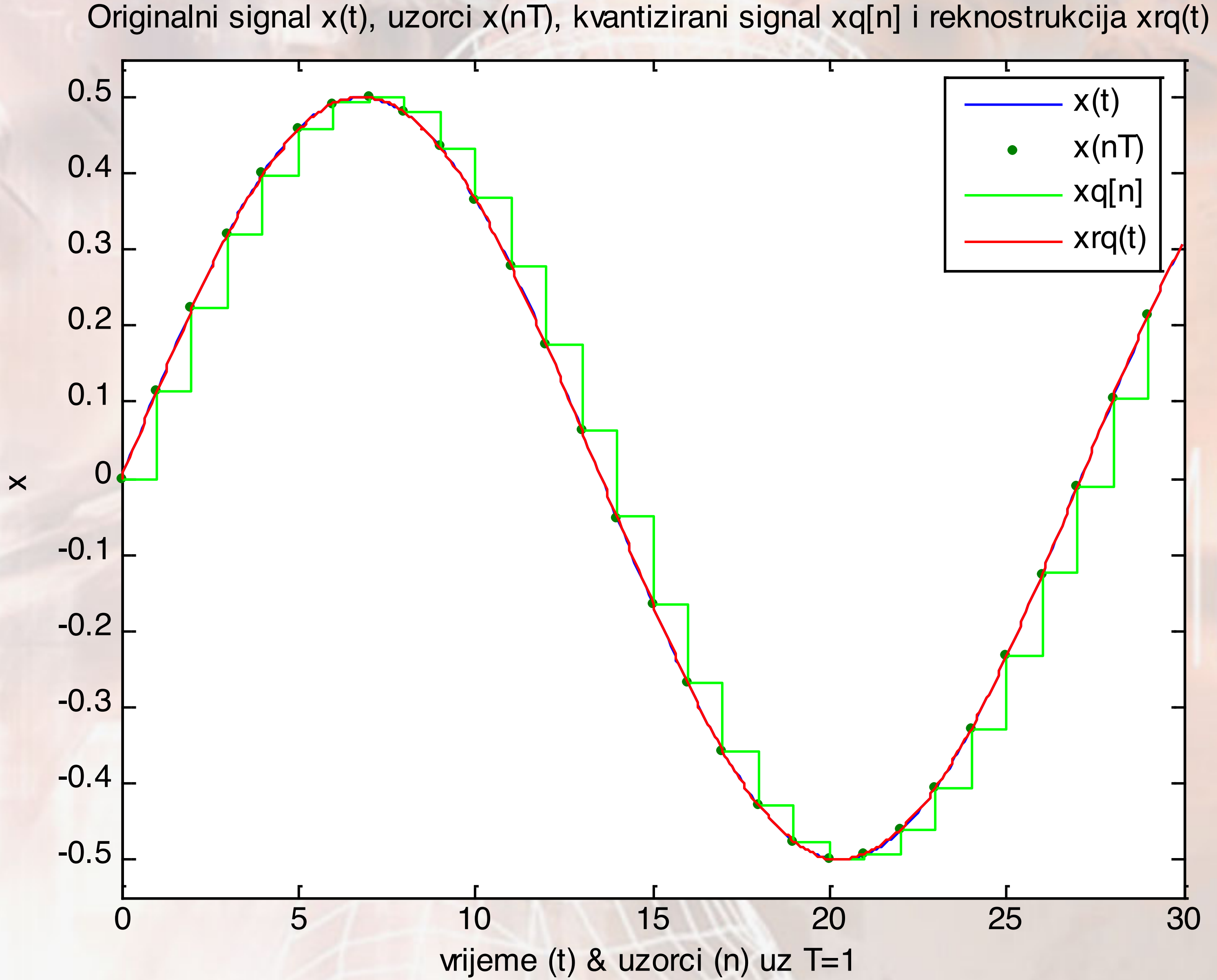
Example 2 - sampling



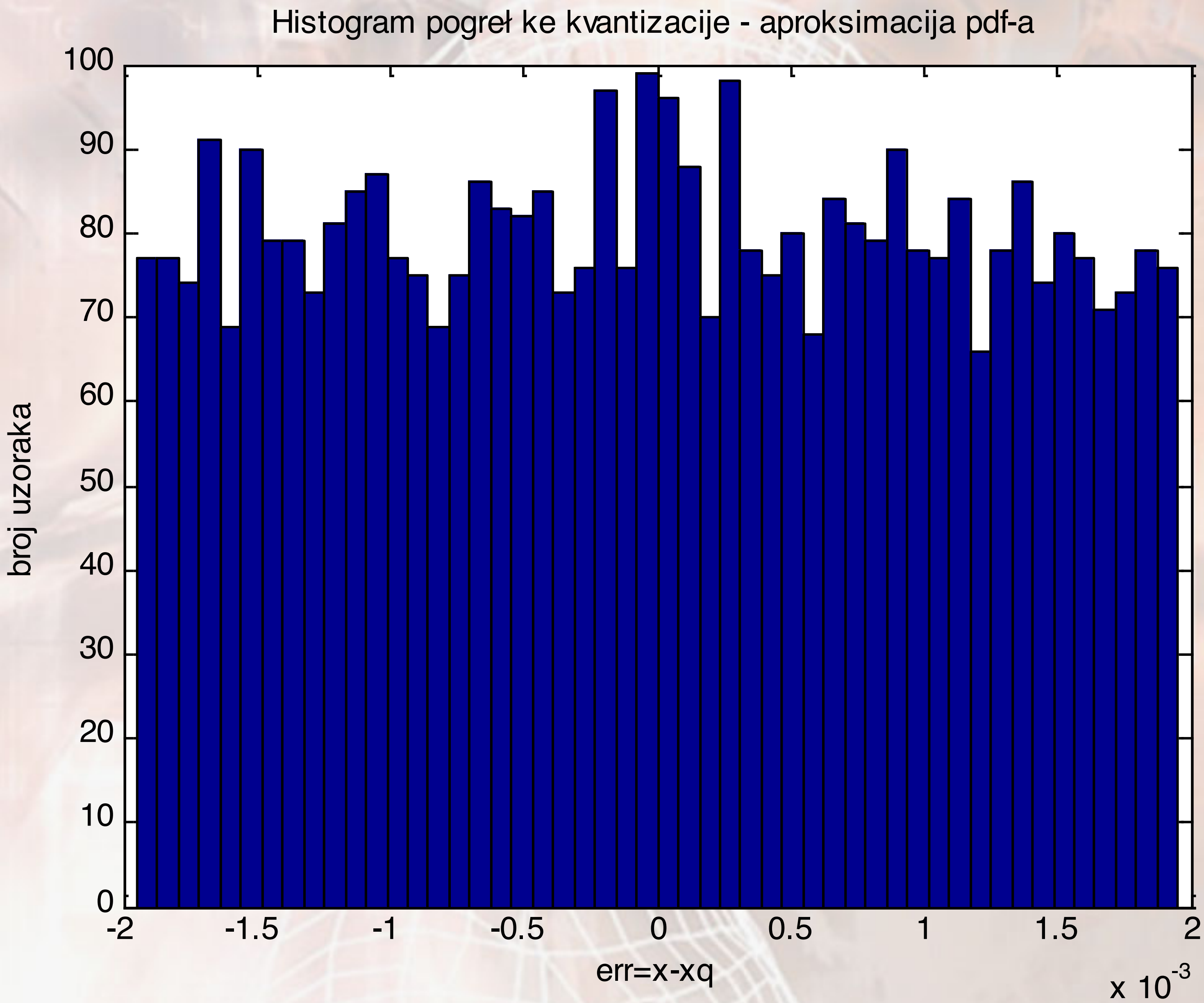
Example 2 - quantization, *b*=8



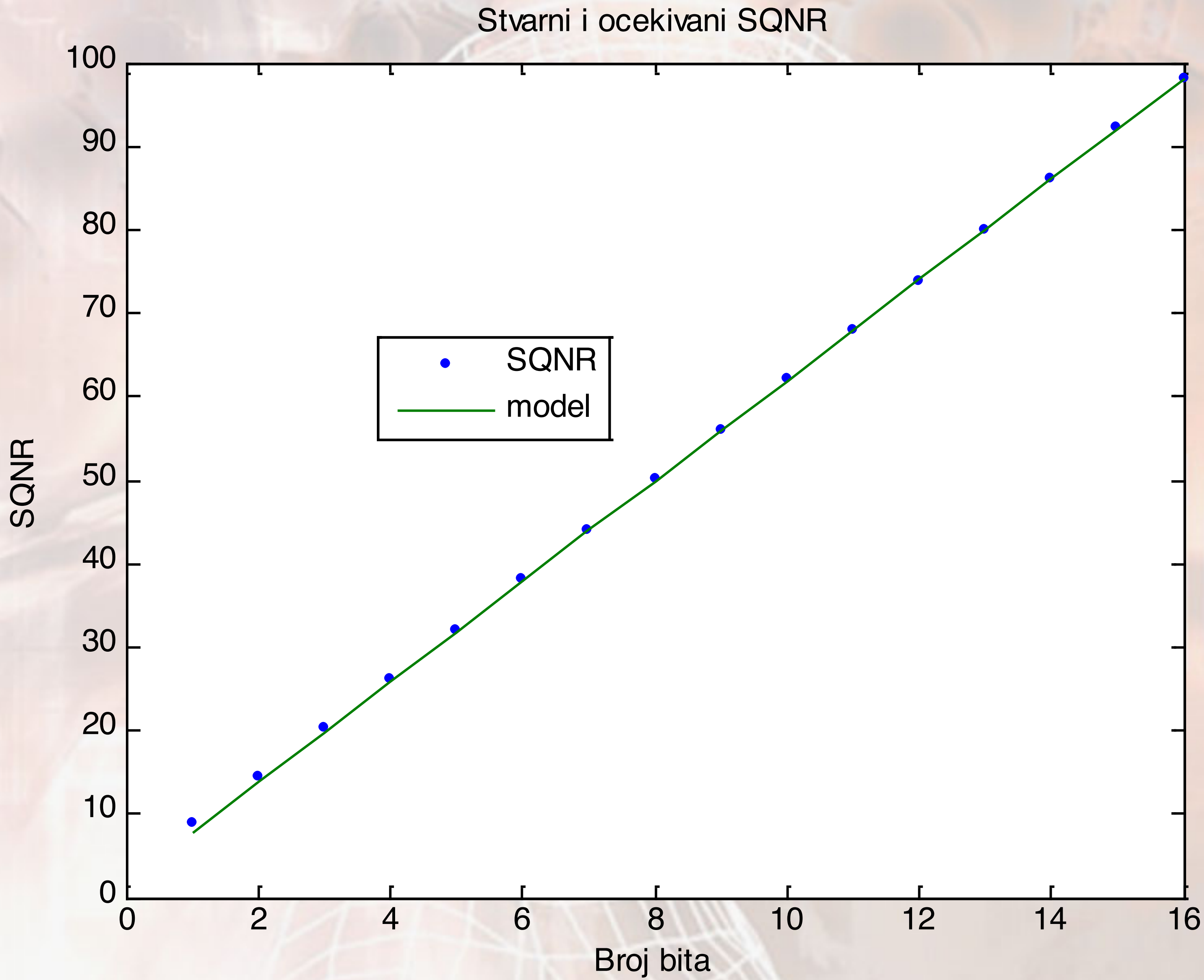
Example 2 - reconstruction, *b*=8



Example 2 – error histogram, *b*=8



Example 3 - quantization, b = 1 to 16



Discussion

- The model predicts SQNR extremely well for a larger number of bits *b* (for higher resolutions).
- Also, the histogram of the quantization error is closer to the assumed uniform distribution for a larger number *b*.
- It is shown that for higher resolutions, the behavior of quantizers can be very accurately predicted by an appropriate analytical model.
- This approximation is the basis for the so called **High-rate quantization theory**.

What have we learned?

- What is quantization
- Uniform quantizer, linear PCM
- Quantization step, resolution and dynamics
- Bipolar quantizers, output codes
- A/D conversion accuracy
- Quantization error
- Statistical analysis of quantization error
- Signal to quantization noise ratio
- Quantization error model
- Examples and simulations of quantization

Data rate of media signals – raw format

prof.dr.sc. Davor Petrinović

Media data rate

- Consider the data rate of media signals of nominal accuracy, or **raw signal data rate** for certain signal types.
- Data rate *R* is measured as the number of bits (or bytes) per second of a media signal or per media object.
- For one-dimensional signals, the data rate is equal to the product of the A/D converter resolution and the sampling frequency :
- ..... $R = b \cdot f_s$ [bit/s].

Media data rate

- In the case of two-dimensional signals, in addition to the amplitude resolution of an individual sample, we also define the **spatial resolution**, i.e., the dimension of the *M*x*N* matrix of the basic elements of the image called **pixels**.
- For monochrome photographs, we measure the data rate as the total number of bits per image :
- ... $R=b \cdot M \cdot N$ [bit/image],
- while for a monochromatic video signal, we also multiply this measure by the number of frames per second f_r (i.e., the ***frame rate***):
- ... $R=b \cdot M \cdot N \cdot f_r$ [bit/s].

Multichannel signals

- There are also **multi-channel** media signals.
  - The stereo audio signal has two channels, while the surround 5.1 audio signal has a total of 6 parallel channels.
  - Color images and color video signals also use multiple channels to describe color. As a rule, three channels are used for each pixel, and the most common representation is in the **RGB** (**R**ed-**G**reen-**B**lue) color space.
- For multi-channel signals, the data stream must be additionally multiplied by the number of used channels *C*.

Amplitude and time resolution required

- The required amplitude resolution is determined by the signal properties, but more importantly:
  - **perceptual (in)significance of quantization error**.
- Analogously, the sampling frequency is determined by the **highest perceptually significant frequency** contained in the media signal.
- Due to competition, manufacturers are trying to increase both the amplitude resolution and the sampling frequency even far above the perceptually based thresholds!

Amplitude resolution required

- When discussing the required amplitude resolution, other sources of error should be included in the analysis:
  - noise in the input signal itself,
  - noise of the MM signal sensor,
  - noise introduced in the process of pre-processing the analog signal,
  - nonlinearity of all input circuits and A/D converters,
  - cross-talk from other signal sources.
- For a high-resolution converter to make sense, the combined effect of all these influences must be less than half the quantization step!

Amplitude resolution required

- To illustrate, ... modern A/D converters used in sound cards for recording audio signals on a PC support an amplitude resolution of *b*=24 bits per sample.
- For a standard input level voltage of 1V, the quantization step of the 24-bit resolution is only 0.1µV, which is even a few orders of magnitude below the typical amount of the other errors listed earlier!
- The user can be very happy if the most significant 16 bits of this 24-bit word really carry useful information about the input audio signal ☺

Time resolution required

- For the same example, modern A/D and D/A audio converters use a sampling frequency of 192k samples per second.
- Such D/A converters at their output can generate a high-quality analog signal with a bandwidth of at least 0.8*f_s/2, which is about 77kHz.
- Note that even young people with exceptional hearing have difficulty hearing signals whose frequency is higher than 16kHz.
- Obviously, there is at least a fivefold redundancy in time resolution here too!

Time resolution required

- The signal from the D/A converter still has a long way to go to the listener's ear :
  - sound card output circuits,
  - cables leading the signal to the audio amplifier,
  - audio amplifier input circuits,
  - power amplifier,
  - speaker cables,
  - speaker crossover,
  - the speaker drivers themselves together with the speaker box,
  - and finally, the acoustic space of reproduction.

Time resolution required

- Each element along this signal pathway with its linear and nonlinear properties affects the quality of the reproduced signal.
- The most significant is the influence of the frequency characteristics of individual elements, whose upper limit frequency must be at least an order of magnitude higher than the highest signal frequency of 77kHz.
- Such properties are met only by specific highly professional audio equipment.
- There is also the question of the justification and safety of such reproduction, from the point of view of ultrasonic pollution!

Media data rate

- As an example of an extremely high data stream media signal, consider the **Full-HD video signal**.
- From the point of view of the display device (Full-HD screen), the parameters of this signal are :
  - C=3 (three color channels per pixel),
  - MxN=1920 x 1080 (spatial resolution),
  - b = 8 (or 10) (amplitude resolution of each channel),
  - f_r=60 (frames per second),
- ... which gives ***R*=3Gbit/s** (b=8) or 3.7Gbit/s (b=10).

Media data rate

- **Blue-Ray Disc** (BD) was used as a medium to store Full-HD quality video signals
- The capacity of the dual layer BD is 50GBy
- If the previously described Full-HD signal were saved in raw format accuracy directly on the disc, then the maximum duration of the video on one BD disc would be **only 2 minutes and 13 seconds!**
- Currently, the fastest BD recorders achieve a write speed of 216Mbit/s in 6x mode, which is **14 times lower** than the required rate of R=3Gbit/s.

Media data rate - discussion

- From the described example it is obvious that no operations on such a signal are possible without significant compression of its content!
- Currently, the only useful operation that can be performed on such "raw" data stream is its transmission via the HDMI digital transmission standard, which supports max. speed of 3.96Gbit/s.
- A newer version of HDMI 1.3 supports an increased transfer rate of max. 10.2Gbit / s, which enables increased amplitude resolution (30, 36 or even 48 bits per pixel), ... so-called “***Deep color***”

What have we learned?

- Definition of media data rate
- Rate of one-dimensional signals
- Rate of two-dimensional signals
- Multichannel media signals
- Amplitude resolution required
- Time resolution required
- Sources of errors in MM signal chain
- Full-HD signal data rate for raw format - need for compression.