

EBU5305

Interactive Media Design and Production

Digital Images, Video & Sound

EBU5305

Images: main points

- The points at which an **image** is sampled are known as **picture elements (pixels)**
- Colour bitmap images can be **true-colour** or **index-based**
- Images can be stored as **bitmaps** or be **vector-based**

EBU5305

Reading

- BurgChapter2.pdf p. 1-7 ; p. 46



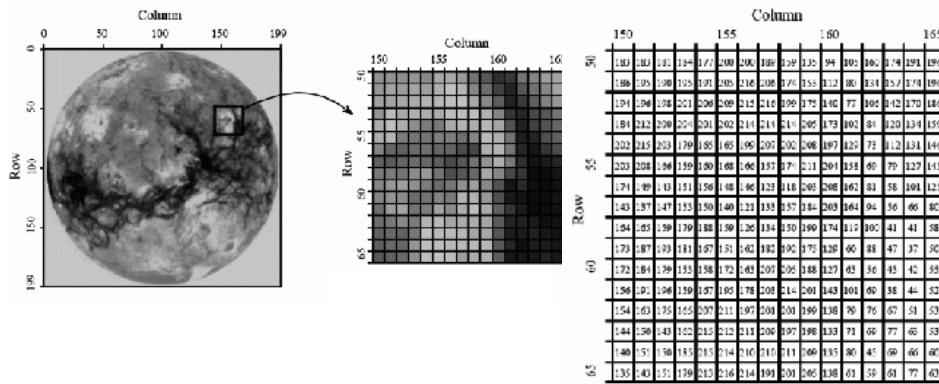
EBU5305

Image Sampling and Quantization (gray scale images)

- A digital image is represented by a matrix of numeric values, each representing a quantized intensity value. e.g. $I(r, c)$
- The intensity at each pixel is represented by an integer and is determined from the continuous image by averaging over a small neighbourhood around the pixel location.
- E.g. when 8-bit integers are used to store each pixel value, the gray scale levels range from 0 (black) to 255 (white).

EBU5305

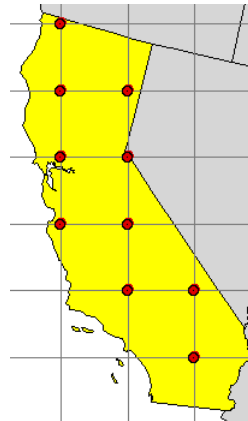
Image Sampling and Quantization (gray scale images)



EBU5305

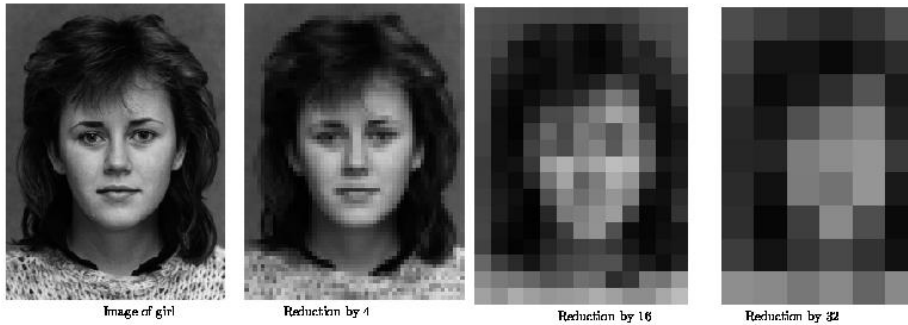
Image sampling

It is common to use a **square sampling grid** with pixels equally spaced along the two sides of the grid.



EBU5305

The effect of sampling reduction



EBU5305

Image resolution

- The number of pixels per inch is the image resolution.
- Typically, monitors have a screen resolution of 72 ppi (pixels per inch).



EBU5305

Image resolution



EBU5305

Image resolution & image size

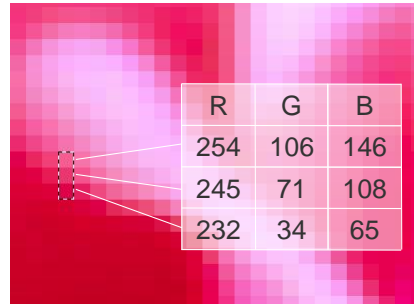


EBU5305

Image quantisation

The quantisation level is called image depth.

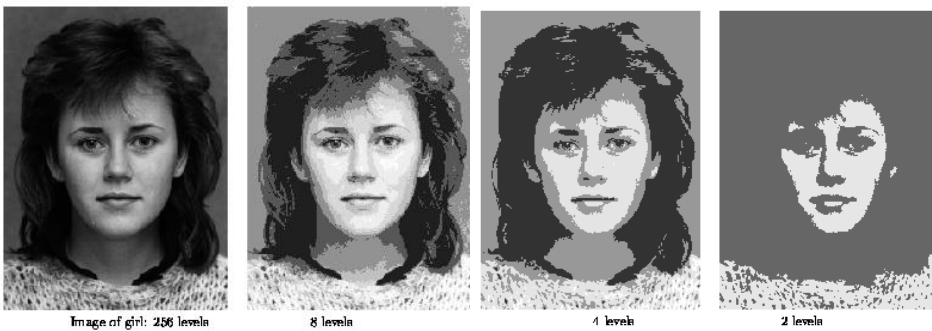
1	1	1	1	1
1	2	2	2	1
1	2	5	2	1
1	2	2	2	1
1	1	1	1	1



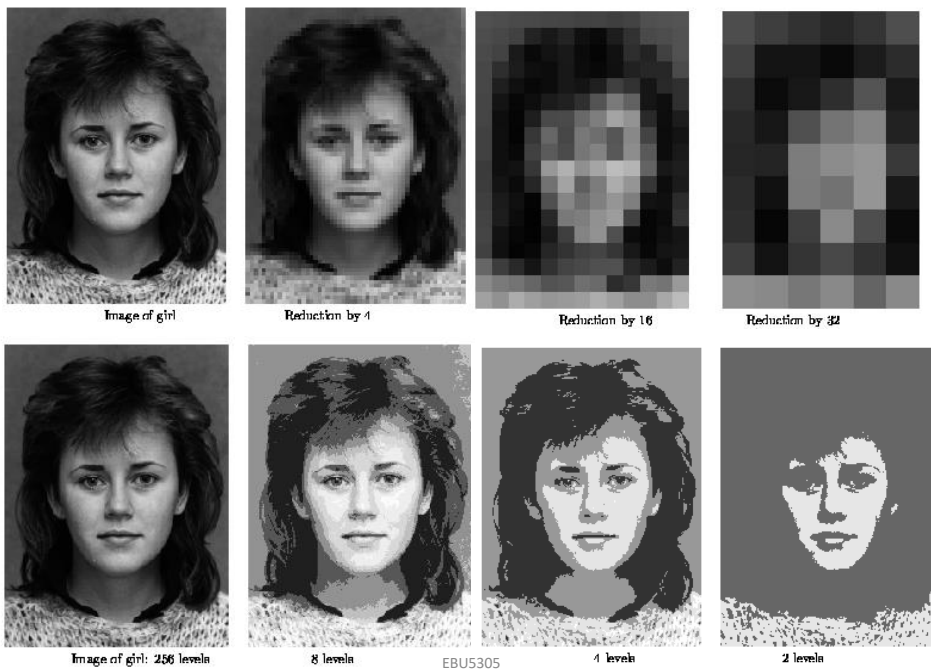
EBU5305

12

The effect of quantisation reduction



EBU5305



Summary so far ...

- The points at which an image is sampled are known as picture elements (pixels)
- The number of pixels per inch is the image resolution
- The image resolution varies with its size
- The quantisation rate is the image depth

EBU5305

Images: main points

- The points at which an image is sampled are known as picture elements (pixels)
- Bitmap images can be **true-colour** or **index-based**
- Images can be stored as bitmaps or be vector-based

EBU5305

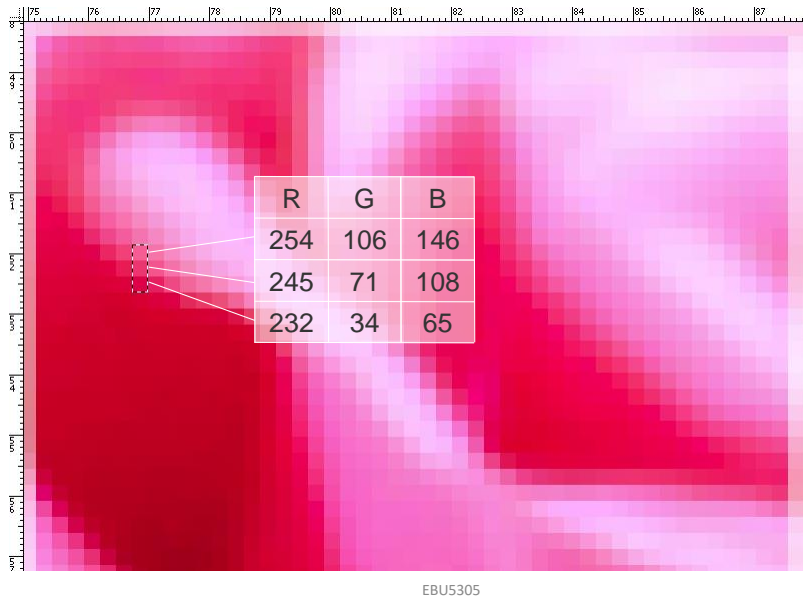
Stored Image Formats

- **RGB formats**, also known as true-colour, use 8 bits of data for each Red, Green, and Blue value. Together, this forms a 24-bit pixel palette which has 16.7 million colours.
- **Indexed formats** are mapped to a smaller colour palette (CLUT) : 256-colours or less (normally). The indexed image's palette contains all of the colours that are available for the image.

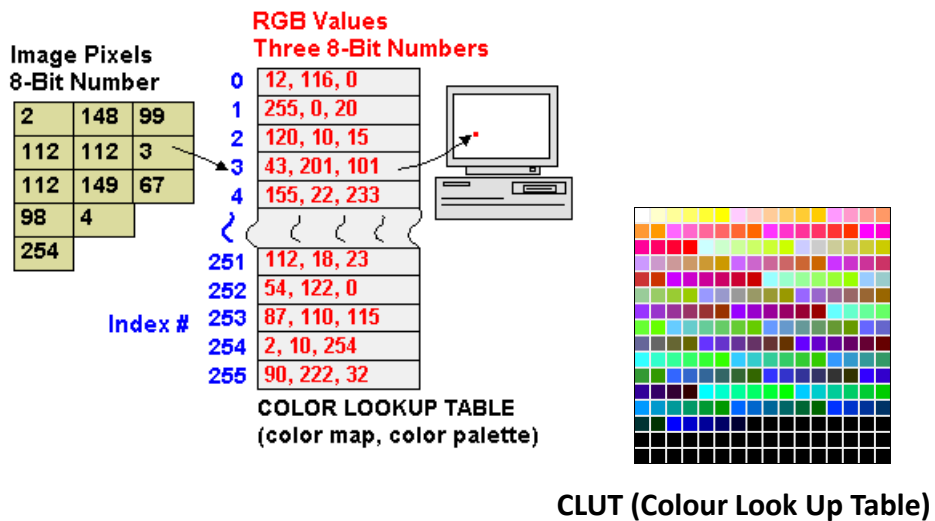


EBU5305

RGB format (true colour)

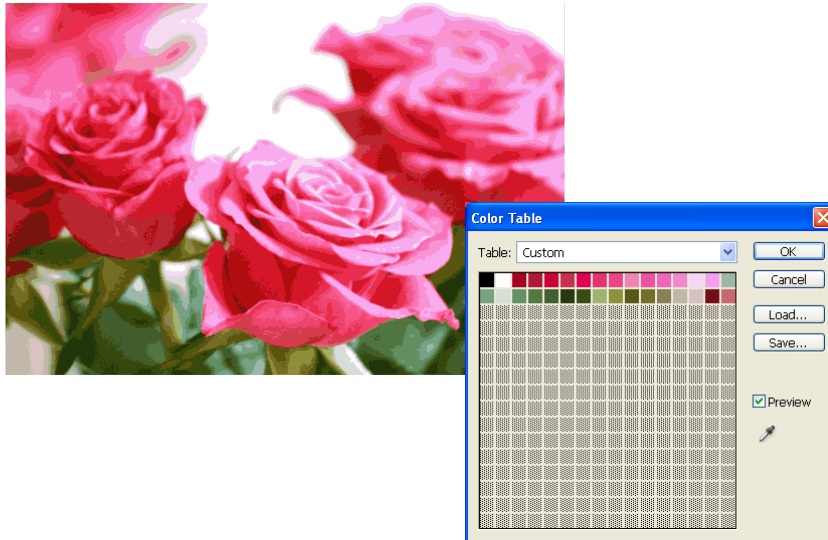


Indexed format



EBU5305

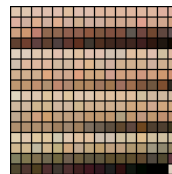
32 Colours (Indexed format)



EBU5305

Indexed formats

8bits/256 colours



2bits/4 colours



EBU5305

Question 1



A 200 by 300 pixel image is stored in different formats. Calculate its size in Kbytes in the following cases:

- True colour image
- Greyscale (8 bits image)
- 4 bits indexed format

EBU5305

Question 2



A Colour Look-Up Table with 256 entries is used.

- Calculate the reduction in the size of the bitmap when a CLUT is used instead of true colour.
- Comment on how use of the CLUT will affect the appearance of the image on the screen.
- If the first entry in the CLUT has value 0,0,0, what colour would be stored there?

EBU5305

Summary so far ...

- Bitmap images can be true-colour or index-based
- A true colour image pixel is represented with 24 bits (8 bits for each colour channel)
- 16.7 million colours can be represented with 24 bits
- A grayscale image has an image depth of 8
- An indexed image uses a Colour Look Up Table (CLUT)
- In an indexed image, each pixel is represented by an index, which refers to a colour in the CLUT
- The number of bits needed for each pixel depends on the size of the CLUT

EBU5305

Images: main points

- The points at which an image is sampled are known as picture elements (pixels)
- Bitmap images can be true-colour or index-based
- Images can be stored as **bitmaps** or be **vector-based**

EBU5305

Bitmapped Vs. Vector-based

- All images display on a computer screen as a grid of pixels of various colours. The image files that contain these images store that image data in one of two fundamentally different ways:
 - Bitmapped image files store image data as a **map of individual pixels** (e.g. GIF and JPEG).
 - Vector-based image files store image data as a set of **mathematical formulas** that instruct the computer how to draw the image. **Flash is a vector-based format.**
- Some file formats, such as PNG, contain both bitmapped and vector-based image data.

EBU5305

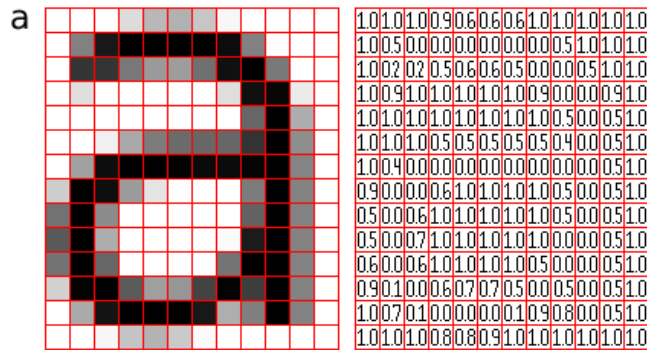
Vector-based image

```
draw circle
  center      0.5, 0.5
  radius      0.4
  fill-color   yellow
  stroke-color black
  stroke-width 0.05
draw circle
  center      0.35, 0.4
  radius      0.05
  fill-color   black
draw circle
  center      0.65, 0.4
  radius      0.05
  fill-color   black
draw line
  start       0.3, 0.6
  end         0.7, 0.6
  stroke-color black
  stroke-width 0.1
```



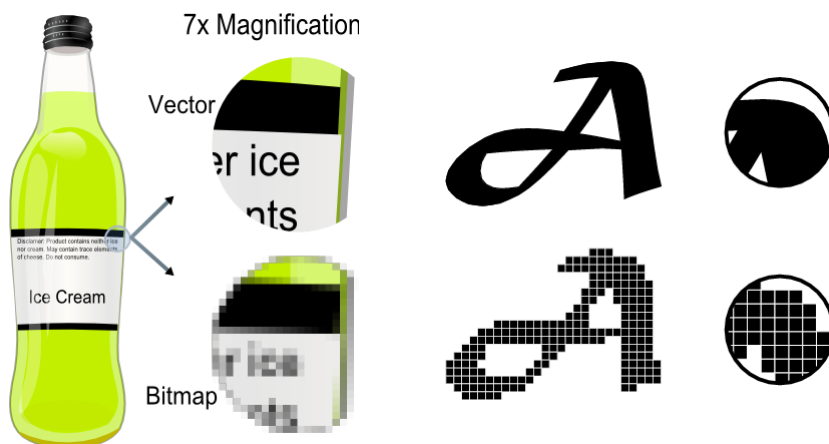
EBU5305

Bitmapped image (raster)



EBU5305

Bitmapped Vs. Vector-based



EBU5305

Summary so far (images)

- The points at which an **image** is sampled are known as **picture elements (pixels)**
- Bitmap images can be **true-colour** or **index-based**
- Images can be stored as **bitmaps** or be **vector-based**

EBU5305

Video: main point

- A video is a **sequence of images**

Video - definitions

- **Video** is the technology of electronically capturing, recording, processing, storing, transmitting, and reconstructing a [sequence of still images](#) representing scenes in motion.
- **Frame rate**: the number of still pictures per unit of time of video.
- **Analog video**: video recording method that stores continuous waves of red, green and blue intensities.
- **Digital video**: video recording system that works by using a digital rather than an analog video signal.
- Video formats, storage formats, interlacing, television ...

32

Refresh rate and frame rate

- The **refresh rate** is the number of times in a second that the display hardware draws the data (i.e. repeated drawing of identical frames).
- The **frame rate** measures how often a video source can feed an entire frame of new data to a display.
- Typical frame rates: 25fps (interlaced), 30fps (progressive).

33

Question



A 30fps digital video uses 352 by 255 pixels video frames with a pixel depth of 8.

- i) Calculate the size of 1 second of data.
- ii) What compression ratio would be needed to transmit 1 second of data in real-time over a 64 Kbps communication channel?

34

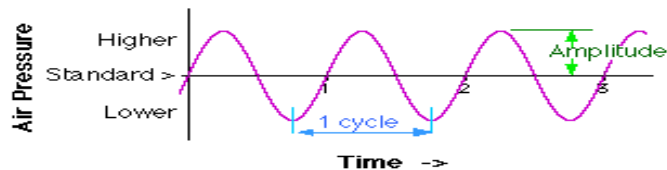
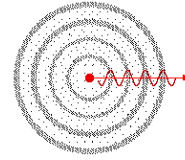
Sound: main points

- A sound is characterised by its frequency (**pitch**) and amplitude (**loudness**)
- **CD standard quality** is 44,100 Hz (sampling) and 16 bits (quantisation)
- **Speech signals** contain 3 types of sound, some of them are used for speech recognition
- **MIDI format** for music stores information such as instrument specification, beginning and end of a note, basic frequency, etc.

35

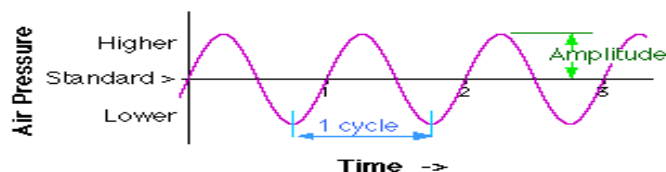
Sound

- Sound is a physical phenomenon produced by the **vibration of matter**, such as a violin string, or a block of wood.
- As the matter vibrates, **pressure variations** are created in the air surrounding it.
- This alteration of high and low pressure is propagated through the air in a **wave-like motion**.



36

Characteristics of Sound Waveforms

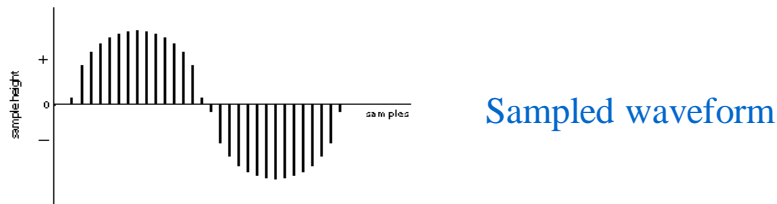


- ◆ **Frequency** determines the pitch (higher frequency = higher pitch)
 - Infra-sound: from 0 to 20 Hz
 - **Human hearing frequency range: 20 Hz – 20 kHz**
 - Ultrasound: from 20 kHz to 1 GHz
- ◆ **Amplitude** of the wave determines the volume or intensity (a property subjectively heard as loudness).

38

Computer Representation of Sound - Sampling -

- A computer measures the amplitude of the waveform at regular time intervals to produce a series of number (**sampling**). This is done by an **ADC** (*Analog-to-Digital Converter*)
- **Sampling rate**: the rate at which a waveform is sampled.
e.g. *the CD standard sampling rate of 44100 Hz means that the waveform is sampled 44100 times / second.*



39

Computer Representation of Sound - Quantization -

- **Quantization**: the resolution or quantization of a sample value depends on the number of bits used in measuring the height of the waveform (*usually 8-bit or 16-bit*)

3-bit quantization



40

Reminder: Nyquist theorem

Sample twice as often as the highest frequency you want to capture

Let f be the frequency of a sine wave. Let r be the minimum sampling rate that can be used in the digitisation process such that the resulting digitised wave is not aliased. Then:

$$r = 2f$$

r is called the **Nyquist frequency**.

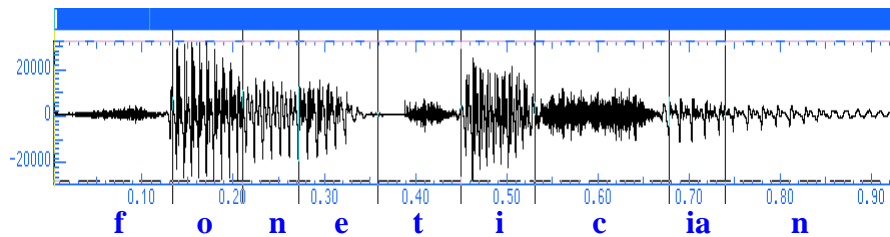
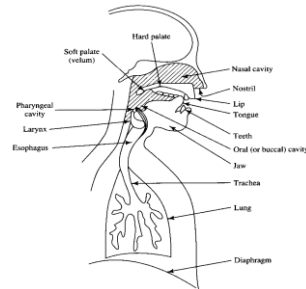
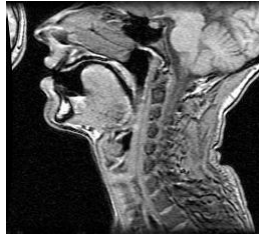
42

Sound: main points

- A sound is characterised by its frequency (pitch) and amplitude (volume)
- CD standard quality is 44,100 Hz (sampling) and 16 bits (quantisation)
- **Speech signals** contain 3 types of sound, some of them are used for speech recognition
- MIDI format for music stores information such as instrument specification, beginning and end of a note, basic frequency, etc.

44

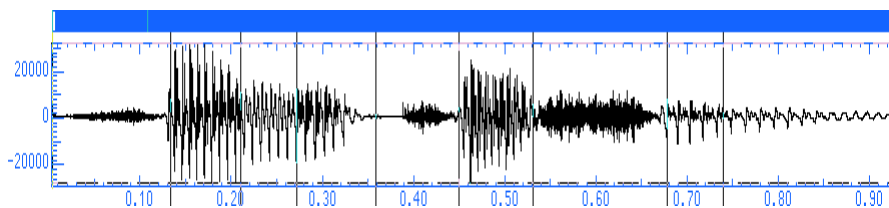
Speech



45

Types of Speech Sounds

- **Voiced sounds** : the vocal chords are vibrated, which can be felt in the throat. All vowels are voiced.
- **Fricatives** (unvoiced sounds) : a consonant, such as *f* or *s* in English, produced by the forcing of air through a constricted passage.
- **Plosives** (also unvoiced sounds) : a speech sound produced by complete closure of the oral passage and subsequent release accompanied by a burst of air, as in the sound (*d*) in *dog*.



46

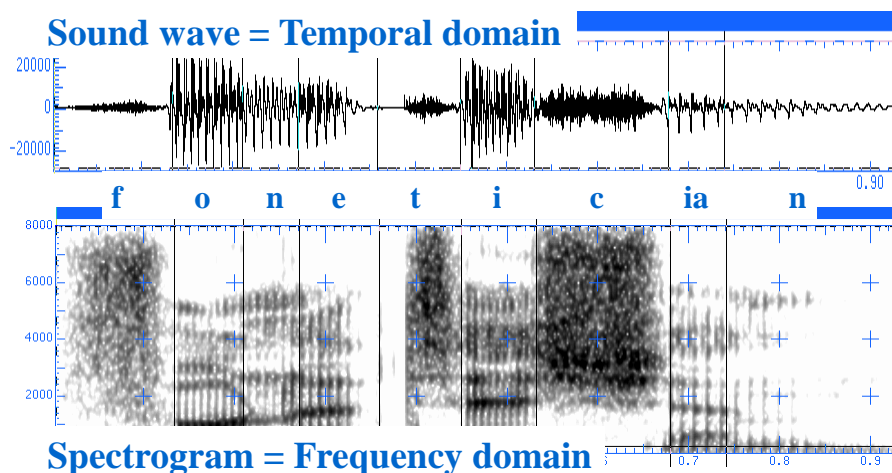
Voiced Speech Sounds

Voiced speech sounds have two properties which can be used in speech processing:

- Speech signals show during certain time intervals almost **periodic behaviours**. These signals are *quasi-stationary signals* for around 30 ms.
- The spectrum of speech signals (voiced sounds) shows characteristic maxima. These maxima, called **formants**, occur because of resonances of the vocal tract.

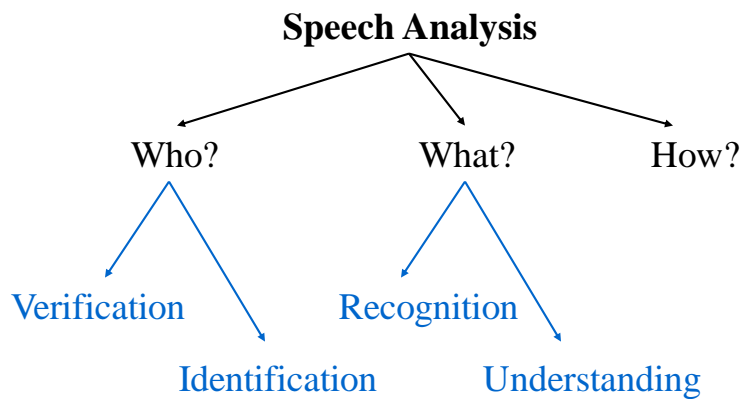
47

Temporal and Frequency Domains



48

Speech Processing Applications



49

Sound: main points

- A sound is characterised by its frequency (pitch) and amplitude (volume)
- CD standard quality is 44,100 Hz (sampling) and 16 bits (quantisation)
- Speech signals contain 3 types of sound, some of them are used for speech recognition
- **MIDI format** for music stores information such as instrument specification, beginning and end of a note, basic frequency, etc.

50

Music

MIDI (Music Instrument Digital Interface) : interface between electronic musical instruments and computers (a small piece of equipment that plugs directly into the computer's serial port and allows the transmission of music signals).

A MIDI interface has two different components:

- Hardware connects the equipment (MIDI port, MIDI cable)
- A data format encodes the information travelling through the hardware. Instead of individual samples (cf audio format), the encoding includes: **instrument specification**, beginning and end of a **note**, basic **frequency** and sound **volume**.

51

MIDI Data Format

- The data are grouped into **MIDI messages**
- Each MIDI message communicate one **musical event** between machines.
e.g. when a musician presses a piano key, the MIDI interface creates a MIDI message where the beginning of the note with its stroke intensity is encoded. This message is transmitted to another machine. In the moment the key is released, a corresponding signal (MIDI message) is transmitted gain.
- The MIDI standard identifies 128 instruments (including noise effects) with unique numbers (e.g. 40 for the violin).

52

Question



Why are MIDI encoded music signals very small?

53