

Computer Vision

Diego Oniarti

Anno 2024-2025

Contents

| | | |
|----------|-------------------|----------|
| 1 | 24-02-2025 | 2 |
| 2 | 28-02-2025 | 3 |

1 24-02-2025

Main topics of the course

1. acquisition
2. motion detection
3. motion analysis
4. stereo/multi -view
5. 3D point cloud
6. feature extraction / classification

Evaluation

The written exam will be 40% of the vote.

You can choose between a project and an oral exam. If you're not satisfied with the result of the written you can take an oral later.

Reading groups are also a thing.

- 24 mar: teams + project ideas
- 31 mar project titles assignment for those who haven't chosen one

For the project we'll use python, openCV, and ffmpeg.

You can deliver the project and written exam in different sessions. But the written exam expires in 1 year.

2 28-02-2025

Bayer Pattern

is a pattern used in camera sensors to optimize the distribution of colors. Since the human eye is more sensitive to green light, the pattern is composed of a checkerboard pattern where one color is green and the other is divided between red and blue.

Quantization

Usually we use 8bpp (bits per pixel) because it is byte aligned and because it's plenty enough for the human eye. At lower bpp, contouring appears.

Video

Static images lose the temporal and movement information, so we need videos. The frame rate of an image must be compliant with the thing that is being captured. With an high enough rate we can ensure a smooth transition between frames without losing information.

This is the reason video-cameras usually have a lower resolution than photo-cameras. With too high of a resolution, there is too much information that needs to be processed and it can't be done at a fast enough rate.

Humans also focus less on image quality while watching a video, as they're more captivated by the evolution of events than the single frames.

Relevant features

The relevant features in an image are color, edges, and contrast. In a video the features are the same but also their progression through time.

Image compression

Images take up a lot of space and videos take up even more. Compression standards exist to reduce the amount of data required.

Compression requires there to be an **encoder** and a **decoder**. Some examples are JPEG, MPEG, and DIVX. Both visualization and processing are executed on the uncompressed image, since humans can't visualize raw compressed data, and filters can't work on the compressed image.

Some compression algorithms are lossy while some other are lossless.

Histogram

is a simple way to describe the color distribution of a picture by approximating a probability function.

$$hist(p) = \frac{\#pixels : I(x,y) = p}{N \cdot M} \approx f(p)$$

Where N, M are the size of the picture in pixels.

Various filters can be applied to an image by manipulating the histogram with operations like stretching and thresholding.

We can equalize an histogram defining a partial sum $CHist_I(p) = \sum_{k=0}^p hist(k)$ e assegnando $hist_{eq}(p) = \frac{CHist(p) - CHist_{min}}{M \cdot N - 1} \cdot 255$.

Even equalizing we can not get to a flat histogram, but we can do our best to get to that point.

Edge extraction

blah blah