Introduction to Computer Vision

Jean Ponce, Mathieu Aubry, Gül Varol, Karteek Alahari

Class notes by Antoine Groudiev

Last modified 10th October 2024

Contents

Abstract

This document is Antoine Groudiev's class notes while following the class *Introduction to Computer Vision* (Introduction à la vision artificielle) at the Computer Science Department of ENS Ulm. It is freely inspired by the class notes written by Jean Ponce.

1 Introduction to Computer Vision

2 Camera Geometry

3 Camera Calibration

3.1 Affine models: weak perspective projection

4 Image processing using filters and convolutions

An image can be interpreted either as a continuous function $f(x, y)$ or as a discrete array $F_{u,v}$.

4.1 Filters and convolution

4.1.1 Basic filters

An image can be blurred using a filter, by replacing a point by the average of its neighbors. Blurring an image gives a smoother image, making it easier to compute derivatives.

4.1.2 Convolutions

Given two integrable functions $f, g : \mathbb{R} \to \mathbb{R}$, we can define their convolution as:

$$\begin{aligned} f * g : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{-\infty}^{+\infty} f(x-t)g(t)dt \end{aligned}$$

Note that $f * g = g * f$ using a change of variable.

This is the definition of the convolution from a continuous perspective. When dealing with images, we want to apply the convolution to a discrete array.

$$R_{i,j} = (F * G)_{i,j} = \sum_{u,v} F_{i-u,j-v} G_{u,v}$$

Convolution follow basic properties:

Commutativity $f * g = g * f$

Associativity $(f * g) * h = f * (g * h)$

Linearity $(af + bg) * h = af * h + bg * h$

Shift invariance $f_\dagger * h = (f * h)\dagger$

Note that is the only operator that is both linear and shift-invariant.

The convolution can be differentiated:

$$\frac{\partial}{\partial x}(f * g) = \frac{\partial f}{\partial x} * g$$

Gaussian filters are used to blur images.

4.2 Computing derivatives

4.3 Edge detection

4.4 The Canny edge detector

4.5 Denoising, sparsity and dictionary learning

5 Edge detection

6 Radiometry and Color

7 Color perception and Two-view geometry

8 Epipolar Geometry and Binocular Stereopsis

9 Markov random fields

10 Recovering structure from motion

11 Mean-shift algorithm for segmentation

12 Multi-view object models

13 Neural Networks for Visual recognition

14 Learning methods