UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

COMPUTER VISION

Willie Brink
Applied Mathematics, Stellenbosch University

July 2017

Foreword

Computer Vision is concerned with the development of mathematical techniques and models for interpreting and understanding visual data. The development of this field was perhaps sparked by the visual perception of the human brain, and ongoing efforts aim to mimic its remarkable capabilities.

Data is usually provided in the form of a digital image, or a sequence of images, and a typical list of tasks for a computer vision system may be:

- to segment an object from a cluttered background;
- to reconstruct certain properties of the object (such as its 3D shape, surface reflectance parameters, colour distribution, etc.) by also considering and modelling the process of image acquisition;
- and to recognize or classify the object.

Modelling the visual world with all of its rich complexities can be extremely difficult. The observed colour of a pixel in the image can be ascribed to many factors such as the viewing angle, the reflectance properties of the material, the presence of unknown light sources, atmospheric conditions, camera effects such as electronic noise and lens distortion, etc., and extracting meaningful information from the image is often challenging. There is also the issue of ambiguity (an image is produced essentially by projecting a three-dimensional scene onto a two-dimensional plane, so that large and distant objects appear similar in size as small objects close to the camera). A further complication is often the sheer size of the data — a single 640×480 colour image, for example, is comprised of about a million single-byte integer values.

Amidst these challenges Computer Vision continues to be a very active research field. Cameras and other imaging devices are popular choices for sensors due to their versatility, their passiveness (a scene can be measured without interference), the large amount of data one gets from a single sample and the high sampling rate, and the fact that cameras are relatively cheap and readily available. Moreover, application areas that benefit from computer vision systems are numerous. Some are listed below.

- **medical applications:** extracting information from X-rays, CAT, MRI, microscopy images, etc.

- **security:** intent recognition for crime prevention; biometrics such as face or fingerprint recognition for access control
- **surveillance:** monitoring activities in sensitive or unmanned areas
- **traffic management:** monitoring and responding to the flow of traffic
- **production industry:** automatic inspection and quality control; measurement; optical sorting of items on a conveyor belt
- **robotics:** obstacle detection, navigation, mapping, path planning, etc.
- **military:** missile guidance systems; battlefield awareness
- **quality of life:** driver assistance; intent recognition
- **photography:** automatic face and smile detection; panoramic stitching; high dynamic range imaging
- **sports:** decision support systems such as the Hawk-Eye in cricket; real-time graphics augmentation in broadcasts
- **web applications:** large scale content-based image searching; SafeSearch; automatic image tagging
- **support for visual effects:** in The Matrix, Avatar, The Hobbit, etc.
- **gaming:** Nintendo Wii; Sony EyeToy; Xbox Kinect

This module will not attempt to cover the entire scope of Computer Vision. We will begin with a brief overview of image processing, and the concept of image features and how corresponding points in different images of the same scene can be identified. We will then introduce projective geometry as an elegant means of modelling pinhole cameras, and also what is often called "two-view geometry" to enable the reconstruction of 3D coordinates from two spatially separate views of the same scene or object (much like the brain's ability to estimate depth from data captured by the two eyes). The second part of the module will touch on more modern focus areas of Computer Vision that involve the application of machine learning. These include object segmentation, recognition and detection.