Large-scale Computer Vision

Ideas for research directions at NLeSc

Elena Ranguelova

Introduction

Computer vision (CV) is the science and technology of machines that can see. As a scientific discipline, CV is concerned with the theory for building artificial systems that obtain information from images.

Why is automating vision a very hard problem, while for humans seeing seems to be easy? The short answer is: because *seeing* is not *perceiving* (Figure 1).

Figure 1 Human vision vs computer vision

From projects and proposals at NLeSc, where the data for the object of scientific research are captured as 2D/3D images, some main applications of CV can be identified:

- **Where is my object? (Localization)**. For example, if the object of my study is a freely moving animal, while my camera is fixed somewhere in its habitat, can a CV system find automatically where an animal appears on the recorded video? Technically, the problem is how to automatically find the object of interest or reduce the data to be processed further.
- **Is my object the same? (Identification)** For example, if I am studying a specific animal, named King Kong, which shares habitat with other animals of the same species, can a CV system find me only the images where King Kong appears on? Technically, the problem is to (semi-) automatically determine if the study object is the same in multiple instances of photographing it.
- **What is my object? (Classification)** For example, if King Kong is a gorilla sharing a habitat with other gorillas and chimps, I might like to separate the data of the chimps from those of the gorillas, and even identify new animals. Hence, the problem is to (semi-) automatically classify the study object to one of possible categories.
- **Large scale.** The common challenge for these questions is how to efficiently answer them from *large scale* scientific images/videos collections.

Research directions

NLeSc should conduct (scientific domains-driven) CV research, for example within eStep. The research efforts should focus around the relevant research questions, presented above: localization, identification and classification. The CV researchers in academia are focused mostly on large commercial applications like organizing large photo collections, autonomous driving, etc. There is still not enough effort directed

towards the other domain sciences (except from the medical imaging) where NLeSC can contribute. Examples of such domains are given below:

Animal biometrics

The subject of the field is to produce computerized systems for phenotypic measurement and interpretation. The main questions are: how to profile species, individuals and animal behavior by representing phenotypic appearance. Figure 2 illustrates the main components of a biometric system.

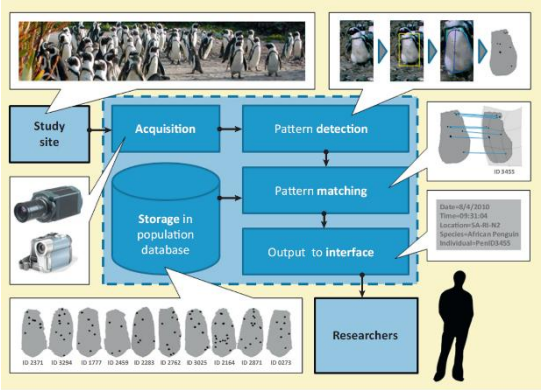

Figure 2 Main components of an animal biometric system.

Plant identification

Similar field is automatic plant identification. This is an example of the Classification task (What is my object?).

Computer forensics

Many questions in digital forensics can be answered (or assisted) by CV technologies, such as face detection and identification, same object/scene identification, image/video categorization, camera identification etc.

Social signal processing

One of the research topics is the development of video surveillance algorithms to help studying social interactions. Technologically the problems are those of gesture (frequency) analysis, gait, pose and emotion recognition, geometric configuration of people, etc.

When working in domain sciences projects we should focus on the domain-driven problems. We can foresee the need for CV research which will be applicable for the above domains:

- Improve the previously developed MSSR or develop new salient region detector. Such a detector can be used to answer the localization and identification problems.
- Explore and build expertise in (using) Convolutional Neural Networks (CNNs) within project Sherlock and the activities of the MLCV interest group.
- Explore and build expertise in large scale frameworks for CV.
- If suitable project appears explore the possibilities in analyzing data captured as Point clouds. Can the CV algorithms be applied directly on PC or new or adapted algorithms are needed?