**Michal PODPORA**

Opole University of Technology

# Computer vision in parallel computing

***Abstract**. Computer vision is known to be a CPU power consuming task. Real-time video capturing, filtering and converting take so much time that creating a substantial artificial intelligence algorithm seems to be unrealisable. In this article author considers using a cluster in a computer vision project, advantages and disadvantages of this solution. He explains some ideas about minimization of amount of data and other methods of speeding up the algorithm.*

**Keywords:** parallel computing, computer vision, signal processing, neural networks

## Introduction

Almost every single being on our planet uses sight as one of its most important senses. We already know how the retina works and more or less how the brain acquires and identifies the visual "data", but our knowledge is still insufficient to build a complete real model.

Many scientists and hobbyists have struggled against lack of CPU power for all the real-time video capturing, filtering, converting, processing, etc. not mentioning e.g. pattern recognition algorithms.

## Why parallel computing

Nowadays, CPU's are much more efficient, but a single PC still seems to be not enough for such a complex job. That's why the author has chosen a different solution (surprisingly not a popular one in computer vision) – to use a cluster instead of a "single" PC.

## Computation time vs. communication time

Writing an application for a cluster is always a great challenge. Moreover, computer vision is based on real-time video stream, whereas one of the main problems of parallel computing is minimization of communication time between nodes of a cluster. This objective can be achieved in many ways. First of all, the computer with the capturing device should be a part of the cluster (as input node) – otherwise, the communication between this computer and the cluster would be a bottleneck of entire process. Another idea is to reduce all time-consuming communication by eliminating video streams and still frames and reduction of amount of data sent between nodes.

## Communication – no frames?

Using video stream as input signal does not mean sending and receiving video data between all nodes of a cluster. In fact, video data can be processed on the input node "in real-time", while all other nodes can work asynchronically [2] basing on reduced and selected data from the input node. It means that reactions of application can possibly happen a moment after specific stimulus – like in real world – it is possible to catch a sight of something, and a moment later to "get the visual" (e.g. when travelling by bus).

The solution of this problem is brought by the nature – eye is only a complex receptor: photoreceptor nerve cells in retina convert light to electrical signals sent to the brain via optic nerve. This process should not be compared with a video camera. The border between a video camera and a PC should rather be understood as the border between pupil and retina than between optic nerve and brain. The captured signal should not "fall" directly into the neural network. It should be converted to a form that would be comprehensible for a neural network.

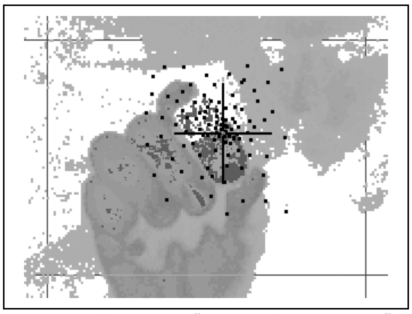

Fig.1. An object being "observed" by the "yellow spot"

Not all details in the field of vision are watched with the same attention at the same time. Retina has a small depression called yellow spot with a rod-free region (fovea centralis) with more closely packed cones providing the sharpest and most detailed information [1]. This sentence is the key to reducing the amount of data. Let's define a "virtual yellow spot" – a point of the most detailed information. It is visualized on (1) as the big black cross. Neural network could modify spots' coordinates to "watch" details of the scene. Inputs of the first layer should get value (of e.g. colour or movement) in a specified point in the neighbourhood of the yellow spot – on (1) shown as black points.

## Cluster virtual topology and communication model

Designing a parallel algorithm requires a due consideration to the broad lines of the clusters' topology. Defining virtual connections between nodes can appreciably speed up (or slow down if done improperly) whole application. Topology proposed by the author for this particular use is modeled on the biological specificity of "understanding the vision". Besides the input node (which captures and converts video stream to a form acceptable for inputs of a 1st layer of a neural network) there are other nodes in the cluster – getting data from the input node in a particular purpose: to process color or shape or movement – just like nerve centers in a brain. These nodes can not only influence coordinates of the "yellow spot" but also generate some output information about the input. Use of an asynchronous algorithm [2] for the communication also speeds up overall performance.

REFERENCES

[1] Hecht E., Optics, *Addison Wesley*, 4th edition (2002)

[2] Karbowski A., Niewiadomska - Szynkiewicz E., Obliczenia równoległe i rozproszone, *Oficyna Wydawnicza Politechniki Warszawskiej*, (2001)

***Author**: Michal Podpora, Opole University of Technology, Poland 46-022 Kepa ul.Luboszycka19a, ravyr@klub.chip.pl*