A TUDOMÁNYOS KUTATÁSOK KIBONTAKOZTATÁSA
A PÁZMÁNY PÉTER KATOLIKUS EGYETEMEN
TÁMOP-4.2.1.B-11/2/KMR-2011-0002

NEW SZÉCHENYI PLAN

Pázmány Péter Catholic University

Hierarchical Visual Scene Understanding

Balázs Varga
Faculty of Information and Technology
varba@digitus.itk.ppke.hu

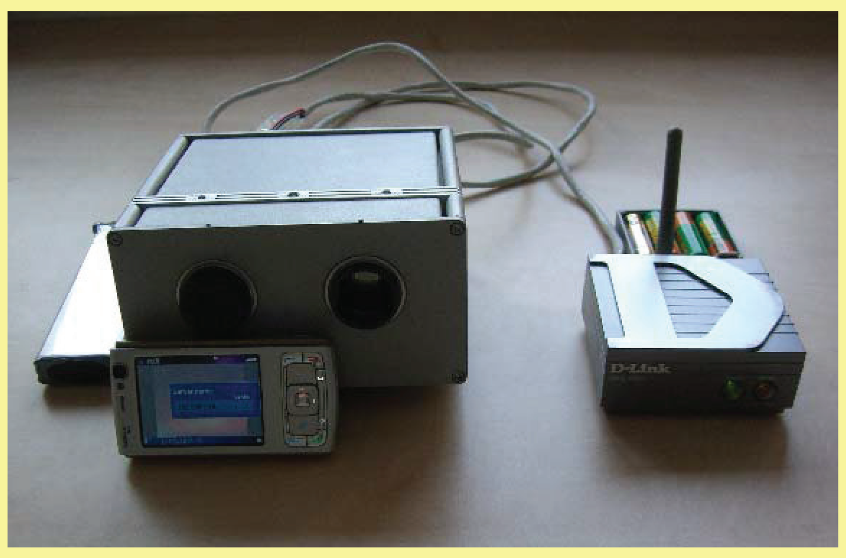
Attila Stubendek
Faculty of Information and Technology
stuat@digitus.itk.ppke.hu

Kristóf Karacs
Faculty of Information and Technology
karacs@itk.ppke.hu

The Bionic Eyeglass

The aim of our research is, to help the visually impaired people in their everyday life by constructing a mobile device, the Bionic Eyeglass. This tool can assist them at numerous tasks, by analyzing and understanding the visual environment in a way analogous to the human vision system.

This portable, low energy system is capable of processing the visual input. As the first milestone of the task, we constructed a cell phone-based navigation system. Here, the phone is wirelessly connected to a visual computing box, containing a cellular microprocessor. The device utilizes everyday tasks, such as color and pattern recognition, identification of the number of the bus, the direction of an escalator or the presence of a crosswalk; then provides voice feedback to the user.



The current system configuration

Scene understanding

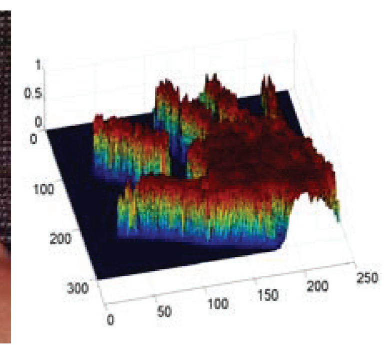
Human perception, the tracking of events, situations, scenes and the detection and recognition of substantial features are highly complex and active processes, which use various information obtained from the sensory organs. Simulating these functions and procedures, moreover using only the visual input, is a difficult task, which requires huge computational power that can be guaranteed by special many-core architectures and parallel algorithms inspired by the human neural system.

The human brain has the ability to describe various information and, on the other hand, to imagine and recognize objects without seeing them before. The sight, when orienting and searching, executes a semantic query. The control of the eye-movement is driven by the interaction between the bottom-up attention and the top-down, query-based attention mechanism. This skill, namely the merging of several different information sources, and understanding the essence is one of our main research aims.

Bottom-up approach

The bottom-up technique is based on modularity: namely that practically any entity can be reconstructed from the over segmented image. Here, the prime principle is that the visual scene seen by the camera consists of low level elements, which – by proper grouping – form the objects of the scene, such as buildings or vehicles. Element pairing is an easy task for us humans; because our visual perception system uses many tricks – such as adaptive local contrast equalization –, but plenty of difficulties arise, when we try to solve the element grouping task with an algorithmic framework. The identification of primitives is performed on the pixel level, where we have to take into account a vast amount of scene conditions, including brightness, shades, texture patterns, semantic information and so on. In our system, identification of scene objects is done by using the mean shift method; a nonparametric segmentation algorithm, which observes its feature space as a probability density function. The technique considers not only color value, but also information, resulting that the algorithm can cluster spatial differences.

Image sample



Cluster intensity map

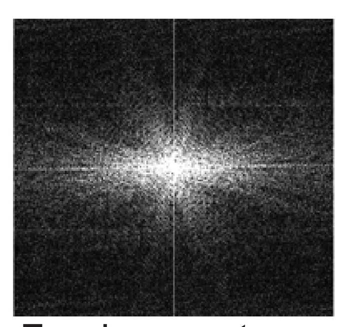
Segmented image

Top-down approach

The semantic description of an object containing information about its the shape, color, texture, and about other features of key-points can be useful in reassembling the over segmented image into a semantically segmented one. The task of the top-down methodology is to find connections between statistical, low-level and topological, higher-level descriptors. To arrive to high-level description of shapes, rotation and transformation invariant methods such as normalized statistical moments and generic Fourier descriptors can be used.

Original image

Image in the polar space



Fourier spectrum of the polar image

Práter Street 50/a., 1083 Budapest,Hungary | +36-1 886-4700 | www.itk.ppke.hu

PÁZMÁNY PÉTER CATHOLIC UNIVERSITY
1088 BUDAPEST, SZENTKIRÁLYI U. 28.
WWW.PPKE.HU
WWW.UJSZECHENYITERV.GOV.HU

National Development Agency
www.ujszechenyiterv.gov.hu
06 40 638 638

HUNGARY'S RENEWAL

The project is supported by the European Union and co-financed by the European Social Fund.