An Integrated Assistance Tool for Visual Impairment

Kristóf Karacs
Faculty of Information Technology, Pázmány Péter Catholic University
Budapest, Hungary

***Abstract*— Although vision restoration is slowly becoming a reality through the various advancements in medical research and technology, the applicability of most methods are fairly limited as they only work in a few eye conditions and cannot provide help for most visually impaired people. Intelligent assistive devices however, albeit without offering visual experience, have three distinct advantages: they can be used by anyone, they have a huge potential cost advantage, and they can be brought to market much faster. We present a portable assistive device that processes the visual image flow provided by a built-in camera, and provides interactive detection and recognition functions. Pattern detection is designed to be universal, building on structural embeddedness and morphologic saliency including the Gestalt principles. Shape recognition in turn is trained for typical patterns appearing in specific situations proposed by blind experts. Testing is carried out regularly on blind subjects using experimental prototypes.**

I. Introduction

A visual assistive device that can intelligently recognize important aspects of the environment can increase the independence and autonomy of blind and visually impaired people. Important end-user aspects include portability, robust operation and ease of use. Commercially available assistive technology typically comes in the form of target tools that aim to solve a single problem. However users are reluctant to carry around multiple devices, thus they prefer an integrated device. This gave the basis for the idea of a universal device built on a versatile platform. As smart phones not only contain cameras, but several other sensors, and most people carry one anyhow, they are a very good platform for this purpose. There are certain drawbacks as well including the touchscreens that are less than ideal for people with visual impairments, but there are appropriate solutions to address these issues.

However, in spite of the enormous growth of computational power available in mobile and embedded systems and the major advances in scene understanding in recent years, realization of mobile devices capable of real time image flow understanding still requires a careful balancing of flexibility and functionality against power consumption, as well as the adequate choice of the algorithmic approach. In algorithm design it becomes more and more important to pay attention to data parallelization, and the use of operators that can be mapped to topographic processors.

Although mobile internet connectivity is becoming a commodity and it can solve the computational complexity problem by allowing the use of remote resources, it increases the dependence on external factors and the sources of potential errors. Thus we aim at a solution that, similarly to the brain, is built on power efficient local computing and can be used at any point of the world regardless of the external infrastructure.

II. System design

We have followed hierarchical principles in our system through all levels, since human visual processing and cognition has been shown long ago to have a hierarchical structure that spans from retinal processing to cortical representation. The advantage of the hierarchical blob based structural representation is that it allows for the incorporation of background knowledge, such that extension towards higher level understanding will be possible through a cognitive model.

The system is composed of four levels. At first the input image flow is segmented into homogeneous regions. In general, this is a color based segmentation, but for some applications a binary thresholding might also suffice. In the second stage blobs of interest are extracted and organized into layers by morphologic processing to form perceptually coherent groups. Based on object properties pools of candidates are selected.

In the third step selected shapes are classified by different classifiers trained for specific tasks providing confidence values for the potential output classes. Finally, the votes on the outputs are combined using a mixture of experts network to maximize the likelihood of the final decision.

III. Experimental results

We have built a prototype system to test the algorithms using a mobile phone as the frontend that can stream the image flow to the back end processor and communicate the results to the user.

We have chosen banknote recognition as one of the most requested functions as a test scenario. The tests were conducted with blind subjects, under medical control. On a test including 108 Hungarian Forint bills and 372 recognition events, the system achieved a 99.6% percent recognition rate at the machine level, and 100% rate at the human level.

The support of the Swiss Contribution, the Bolyai János Research Scholarship of the Hungarian Academy of Sciences, and the grant no. KAP 1.2.14/008 from the Faculty of Information Technology and Bionics of the Pázmány Péter Catholic University are gratefully acknowledged.