

# **System for spatial orientation support of blind people, using the conversion of 2D images into a spatial sound**

M.Sc Robert Świta

Koszalin University of Technology Faculty of Electronics and Computer Science

## **PhD DISSERTATION SUMMARY**

PhD thesis is dedicated to the development of support system for spatial orientation of blind people. The work results in transformation method of images recorded by the camera into the information cues about the details of the object, concerning its shape and location. Developed algorithms are also implemented in a computer program. There exist two different mechanisms for identifying the scene - the first, major one, is a transformation based on the characteristics of the shape and position of the central subject of the scene into a spatial sound and second - auxiliary - is based on the contextual recognition of objects from selected scenes by trained ANN preceded by PCA modules.

Dissertation is divided into eight chapters, arranged in a manner consistent with the data processing order.

After the introduction, presenting an overview of existing systems of hearing images and the concept of the proposed solution, in the second chapter, the basics of color theory and color space used for comparing pixels during image segmentation is presented.

The third chapter discusses ways to interpolate images and shows the method for creating color palettes and implementation of the class *Pixelmap* that enables convenient use of the images and palettes of pseudocolors. The basic methods of automatic white balance were presented and proposed combination of their advantages in a new solution.

Chapter four begins with mathematical background for principal components analysis PCA and SVD matrix factorization. PCA method enables reduction and filtering of data by determining features that are mutually linearly independent and is widely used in the processing systems. Alongside well-known and widely used algorithms, suggestions and author's modifications of these algorithms were presented. An essential part of chapter four, however, is a description of region segmentation methods and hybrid segmentation, which is a combination of region growing and watershed segmentation methods. The result of these algorithms was compared with other types of segmentation (not necessarily working in real time) using a UC Berkeley project BSDS500.

The PCA analysis was used during automatic object recognition, after segmentation. Implementation uses two neural network modules preceded by a PCA blocks, and is described in chapter five.

Chapter six presents the developed method of sound generation based on the image of central segment. This segment is divided into columns. The height of the columns determines height of the tones of the musical MIDI instrument. Columns are processed sequentially,

synthesizing a unique melody. This way the size of the object can be presented. Existence of sound sources at the columns centers location is simulated using HRTF filtration. The shape and location of the object can be determined by an analysis or recognition of melody and virtual sound source position.

The application is also equipped with a learning module. Description of the module is a topic for chapter seven. Embedded editor allows creation of arbitrarily complex shapes thanks to the grouping and ungrouping of simple shapes. Shapes and their groups can be transformed by 2D affine transformations and converted into spatial sound.

Chapter eight is a collection of gathered conclusions resulting from the study, implementation and experiments followed by summary.

Appendix A contains a list of available instruments in the General MIDI standard.