MOVIS

Mobile **O**ptoelectronic **V**isual-**I**nterpretative **S**ystem for the Blind and Visually Impaired

Short Introduction

1 Principle

MOVIS in its final form will be a portable device capable of

- capturing images of the visual environment,
- analysing and interpreting these images,
- presenting the resulting analysis in form of
  - auditory output (sounds, words),
  - tactile output (relief display),
  - visual output (contrast enhanced, magnified),
  - transmittable output (telecommunication purposes).

MOVIS is designed as an aid to the orientation and mobility of the blind and visually impaired. Figure 1 shows a possible implementation. It consists of a spectacle-like device (the glasses) and a microcomputer (the pouch), connected by a cable.

Numerous additional applications can be conceived, ranging from navigation to automation.

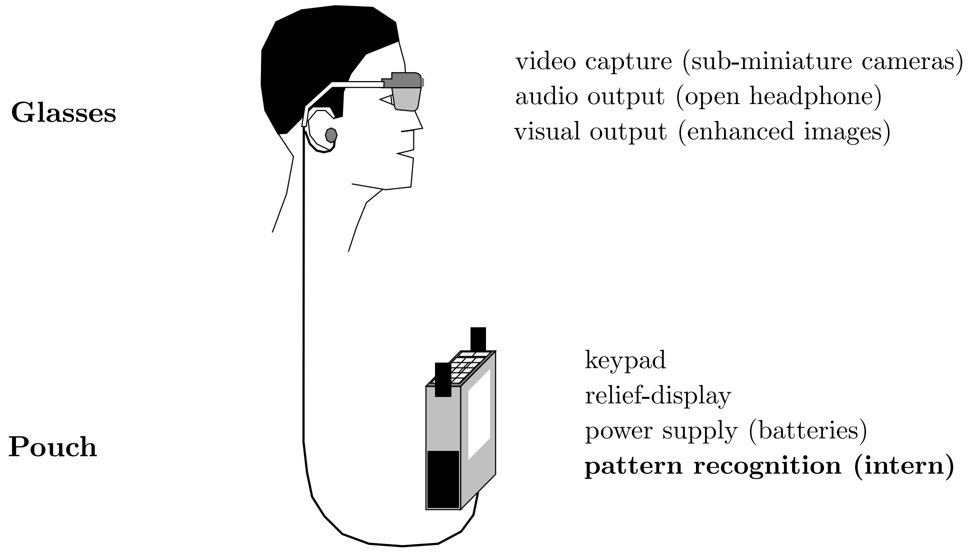


Figure 1: MOVIS in its final form.

2 Functionality

The functionality planned for MOVIS is based on an analysis of the orientational and mobility requirements of the blind and visually impaired, which was carried out during the original project MOVIS, funded by the German Ministry for Research (BMBF) as part of the project "electronic eye" (*Elektronisches Auge*). This project forms the foundation upon which a European project could be based, compare Section 5. A repeated review and refinement of the functional requirements will be part of future work.

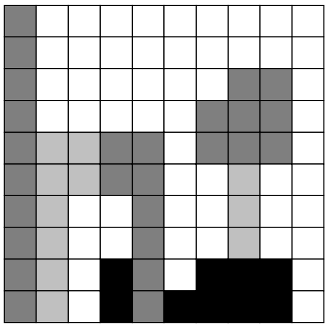
The functionality of MOVIS for perceiving and interpreting the visual environment has been designed under the assumptions that

- the cane will remain the most important tool in the near range,
- all other senses, in particular the auditory, tactile, and, if applicable, remaining visual senses should remain fully usable.

In the following sections the functionality will be described in more detail.

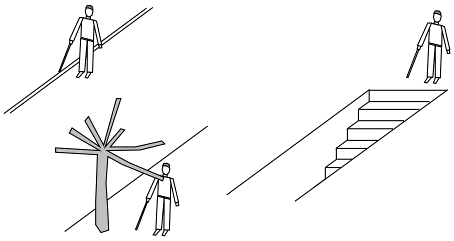
2.1 Detection and Presentation of Structural Information

Structural information — e. g. areas of different distances from the user, longer edges, or areas of different colours — will automatically be detected and presented to the user in an adequate form. This could be a relief-display, open headphone, or optionally an enhanced video-display. Further evaluation of the information presented is left to the user. Additionally, warning and guidance instructions can be given under certain circumstances.



Street-scene on the relief-display

Pedestrian and car in front of a house, archway, and tree (left to right), dark ⇒ near, light ⇒ far.



Guiding along a line, warnings "Attention, head!", "Subway!"

Figure 2: Presentation of a relief-map and warning and guidance instructions.

Examples are (compare Figure 2):

- The presentation of distances using a relief-display (tactile display).
- The presentation of structural information (long edges, areas of different colours) using either the relief-display or sounds and words.
- The classification of the environment into guidelines and potential hazards, and the computation of specific advice.

2.2 Recognition of Special Objects

MOVIS will contain models for a range of interesting objects. These models can be used to recognise the objects they represent independent of perspective distortions. Possibilities range from simple objects like bollards or traffic-lights to complex ones like cars. Currently implemented or partly implemented are e. g. telephone booths, pedestrian crossings, and a small number of street-signs (e. g. bus-stop). These were implemented as part of the original MOVIS.

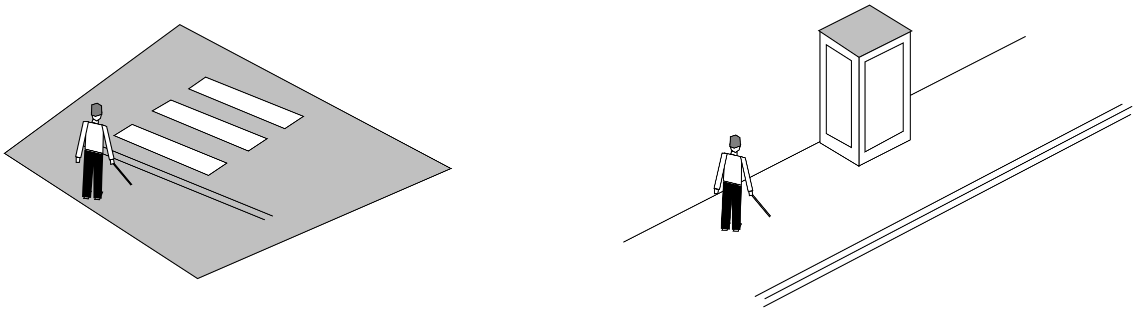


Figure 3: Recognition of special objects (a pedestrian crossings and a phone booth).

2.3 Recognition of Trained Landmarks

With this functionality the user can train MOVIS to recognise landmarks, i.e. views which may help the user to navigate. Typically, training is achieved by taking one or more snapshots of a view which will then be processed by MOVIS to be recognisable as a landmark. Examples could be the entrance of the post-office or a particular shop. MOVIS will recognise the landmark once the user is at a spot sufficiently close to the original position used for training.

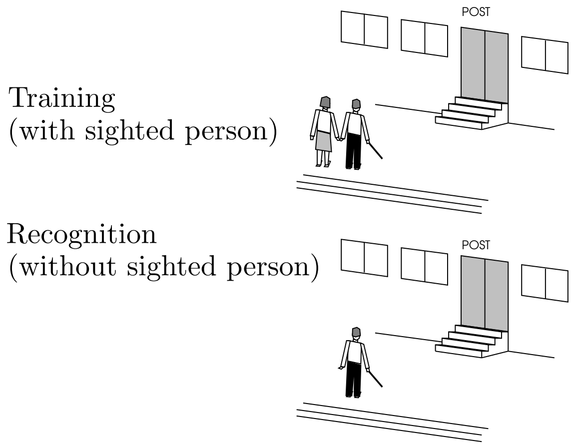


Figure 4: Recognition of trained views.

2.4 Localisation and Recognition of Icons and Text

MOVIS will allow the localisation and recognition of standardised icons and simple texts. The primary goal is to support the recognition of signs used for the orientation in public areas. Examples are given in Figure 5.

- Localisation and recognition of selected icons, e. g. DIN 66 079, ISO 7001.
- Localisation and recognition of simple texts, e. g. house-numbers, street-names, bus-lines.

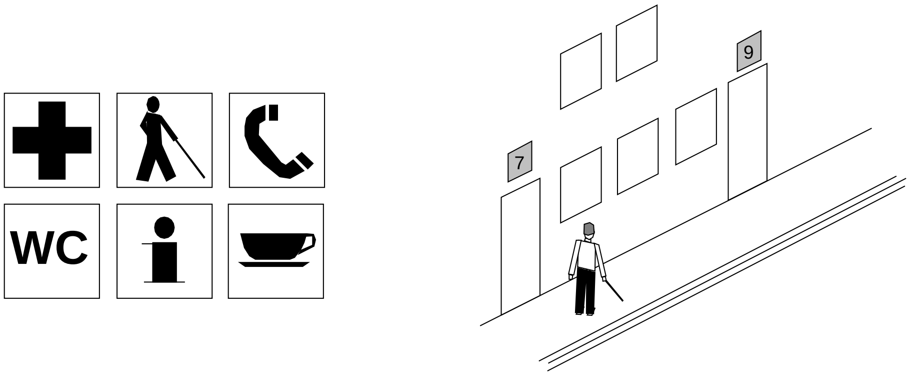


Figure 5: Recognition of icons and house-numbers.

3 Description

3.1 General Specifications

While there exist preliminary implementations of some of the MOVIS functionalities, a detailed specification will be worked out as part of the proposed project. Under the aspect of user-friendliness, the following objectives are of particular importance:

- low weight
- small dimensions
- pleasing look
- can be worn comfortably and securely
- robustness against mechanical strain
- robustness against varying lightening conditions
- easily usable
- reasonable battery-life.

3.2 Input

The numerous functions intended for MOVIS require the possibility of some interaction between MOVIS and the user. The foremost method of input will be a keypad on the pouch. We place large importance on the user-friendliness of all interactions with MOVIS (with an eye towards the usability of MOVIS by the elderly). Alternative methods of interaction will be explored, speech-recognition being one of the more obvious possibilities.

In normal operation the user will select only one particular functionality — or at most a small set of functions — at any given time. This guarantees a selective output which will not overburden the user's remaining senses.

3.3 Output

Means of output are an open headphone, a tactile (relief) display and optionally enhanced visual output for the visually impaired, making use of the remaining sense of vision. Using an open headphone means that any ambient noise can still be heard and reacted upon by the user. For the same reason the audible output itself will be reduced to the bare minimum, usually simply naming distance and direction to selected objects or landmarks.

4 Exemplary Use

In the following, the use of MOVIS will be illustrated using a constructed (and surely incomplete) example. It is assumed that all landmarks which are needed in this example have already been learned in previous sessions. The example demonstrates the different tasks used for orientation and mobility on a day off. The example might however just as well be applied to a working day.

1. The user activates MOVIS. The system checks whether the glasses are worn correctly, whether all components are in working order, and calculates the remaining battery-life.
2. The user employs MOVIS to check for the ambient lighting conditions (*brightness*: dusk, dawn, night?) and the condition of the ground (*colour*: wet, snow-covered?).
3. The user activates the *warning*-function and checks for obstacles in his path using the *relief-display*.
4. The user starts his walk and activates the special object class recognition functionality *pedestrian crossing* to locate a crossing and change to the other side of the street.
5. In passing an archway he uses the *relief-display* combined with the structure recognition functionality *vertical edges* to corroborate the evidence given by his cane and ear.
6. A *warning sound* points out an obstacle nearby. The special object recognition functionality *vehicle* identifies the obstacle as a delivery-van parked on the pavement, with the loading ramp sticking out into the user's path. The obstacle has to be bypassed on a new route.
7. The unexpected detour results in a loss of orientation, often a critical situation for blind persons. With the help of the *relief-display* in conjunction with the functions *vertical lines*, *horizontal lines*, *colour*, and with the help of special object recognition the user is able to gain back his orientation.
8. The previously trained *landmark post-office* is recognised once the user is in its vicinity, and direction and distance are announced by artificial speech.
9. A telephone booth is found by activating the special object class *phone booth*.
10. After leaving the post-office, the user decides to look for the house number 3 in Xyz-road. Once the street sign is found using the functionality *sign reading*, the user then searches for the correct house relying on the *relief-display* combined with the functionalities *vertical lines* and *reading*.
11. Once the user has left the house, he decides to rest on a bench in the park. The functionality *colour* allows him to identify the beginning of the lawn and to some extent appreciate the flowers by the way. He finally finds a bench with the help of the special object class *bench*. The *warning*-function prevents a collision with a branch hanging into the path.
12. From here he intends to take the bus number 123 to the train-station. The next bus-stop is easily identified using the special object class functionality *bus stop*, and the correct line using the functionality *sign reading*.
13. The entrance to the train-station is found by recognising the landmark *train-station*.
14. Using the functionality *icons* the user is able to find the toilet, a fast-food restaurant, the ticket counter, and finally the correct platform.
15. etc.

5 Stages of Development

5.1 Current State

As mentioned before, the development of the envisioned system MOVIS 2000 will be based on the original prototype MOVIS developed in an earlier project. Work in this project was mainly concerned with the recognition of user-trainable landmarks and special object classes.

The feasibility of this approach was demonstrated with a sample implementation, in which the the *mobile components* — glasses and pouch, comparable to the envisioned final stage — were connected to a *stationary computer* using a 30 m cable.

The main reason for this was the high amount of raw computing power required for the computer vision and pattern recognition tasks of MOVIS. In 1997 this computing power was not easily available in mobile form. The bottle-neck of computer power is also one of the reasons for the intermediate stage proposed in the next section.

5.2 Wireless Connection

Since field tests are urgently needed but the necessary computational power is currently not available in portable form, especially for some of the more advanced functionalities, we suggest an interim solution as displayed in Figure 6.

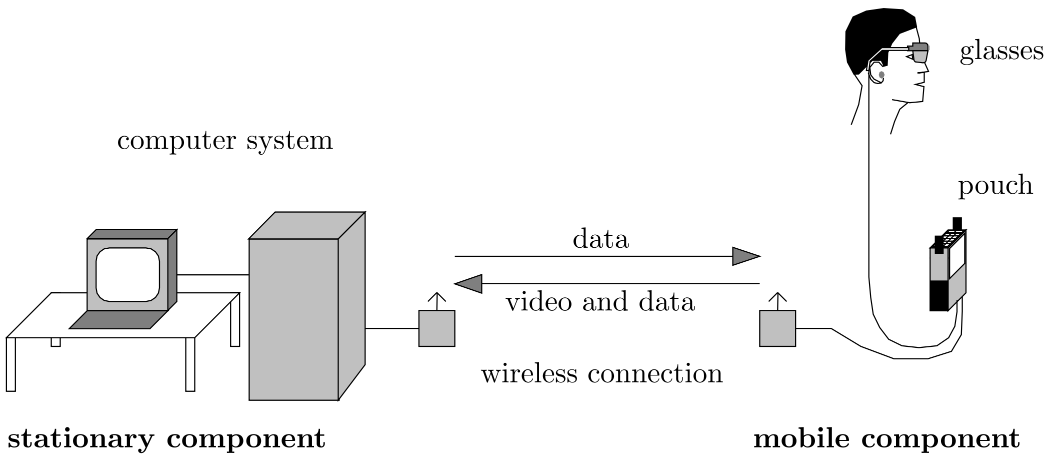


Figure 6: Wireless connection between MOVIS and a base-station.

The images captured in the mobile component are transmitted to a stationary unit via wireless communication, where the actual analysis of the images takes place. The mobile component itself will also contain a small computer, enabling it to provide partial autonomy and in particular deal will all aspects of user-interaction (sound and speech output, control of the tactile and optionally visual display). More and more functionality will be moved into the mobile component as more powerful processors become available. For the end-user, the wireless solution will operate very similarly to the proposed final system.

5.3 The Final Product

The rapid progress in the development of ever faster and more powerful computers will soon allow the complete integration into a portable product as mentioned in Section 1. Until then, products would have to rely on an intermediate solution as mentioned in Section 5.2.

One possible extension would be the use of one stationary unit by several mobile ones. This assumes that the mobile component can mostly act autonomously, and interaction with the stationary component is only sporadically needed for more complex tasks. Sharing the fixed costs for a stationary unit among multiple end-users could serve as the basis for an economically feasible product.

Products could also possible be based on separate solutions for a particular problem. One example could be a mobile component incorporating only the functionality mentioned in Section 2.1 (structural information, including guidance and warning).

6 Current and Former Partners

The partners involved in the original basis-project MOVIS (compare Table 1) have proven the viability of the approach with the implementation of a prototype, details of which can be found in their final report[1].

Nr.	Partner	Main Task
	Industrial Partners	
P1	JURCA Optoelektronik GmbH 63110 Rodgau Dr. Arnulf Reuter	• Coordination • Hardware-development of the pouch and glasses
P2	CORRSYS GmbH 35578 Wetzlar Dr. Norbert Lauinger	Hardware optical correlation systems
P3	empirica GmbH 61231 Bad Nauheim Dr. Rainald von Gizycki	• user surveys • user tests
	Universities	
P4	Institute of Pattern Recognition and Computer Vision University of Freiburg 79110 Freiburg i. Br. Prof. Dr. Ing. Hans Burkhardt	algorithms and software
P5	Artificial Intelligence Lab University of Hamburg 22527 Hamburg Prof. Dr. Bernd Neumann	algorithms and software

Table 1: Current and Former Partners and their main area of contribution.

[1] Available in German from the partners P1 or P4.