

Toward Spatial Information Awareness Map (SIAM) for the Visually Impaired

Worapol Pongpech, Mohamed Bennamoun and John Williams
*School of Electrical and Electronic System Engineering
 Research Concentration in Computer Vision and Automation
 Queensland University of Technology, Brisbane, QLD 4001*

Abstract

With just one quick glance at the scene, a sighted person can obtain enough spatial information to navigate through the scene. A wall on the left and the right, a ball is situated just right in the front. This is the spatial information that helps us navigate through any unfamiliar environment. As sighted person, we are taking this spatial information for granted. This much-needed information is lost on the visually impaired and the blind. Preliminary literature investigations have shown how spatial information of the facing scene can greatly benefit individuals with impairments. This paper investigated the possibility of extracting spatial information awareness of the facing scene given its disparity image. After the whole disparity image is analysed using gradient operator, it is broken into three equal sections, foreground, mid-ground and background. Each section is analysed independently according to the needed spatial information. The relative size of the object occupied in the three sections contains needed information of the facing scene spatial information and also information regarding the type of the facing scene, close range, mid range or outer range. Experiments, which extracted spatial information awareness, were conducted on both a synthetic disparity image and real disparity images, yielding promising results.

funded. It is a research area that can benefit the blind and visually impaired. Electronic travelling aids can be categorised into 1) Obstacle Detection ETA and 2) Orientation ETA. Most of the current obstacle detection ETAs are based on sonar sensor and geared toward detecting obstacles for the user. From literature search, a few current researches on obstacle detection ETAs are based on vision sensor, and even fewer researches on orientation ETA.

It has been shown in many researches [2-5] that visually impaired individual way finding can be greatly benefited through the use of tactile map, which provides spatial information. This research motivated by getting relevant spatial information to the user. Obstacle detection information and orientation information is very useful for the visually impaired, and we are proposing an ETA that will provide the user with both obstacles detection and orientation information of the facing scene. This is achieved by providing the user with the information regarding the relationship of the open space and the standing objects in the given scene. System proposed is discussed in Section 3

The aim is to extract relevant spatial information from vision sensor and produce result that can be transferred to a binary dynamic tactile map. The spatial information extracted from a vision system will have just enough information for the user to be able to form a cognitive map of the facing environment. The provision of such information allows the user to create a path in the facing environment through which they can navigate. This spatial information intends to be provided through a binary dynamic tactile map. The user can trace his/her finger on the map and be able to obtain the approximately spatial relationship within the facing environment. Once the user has the spatial information of the facing environment, it will be much easier for them to plan their route. From the extracted information, the user will also be able to judge if the facing scene is close range scene, mid range scene or outer range scene.

1 Introduction

Today, there is an estimated 180 million people worldwide who are visually impaired. Of these, between 40 and 45 million persons are blind and, by definition, cannot walk about unaided.[1] Research in the area of electronic travelling aids (ETA) for the visually impaired is a rather small field and not well

The novelty in this research is that it is a bottom-up approach so that the outcome will be most suitable for the user. In order to construct useful electronic travelling aids (ETA) for the visually impaired, some of the questions regarding the user must be addressed. These questions are the base criteria of the electronic travelling aids for the visually impaired.

Another novelty lies in the analysis of the disparity image in three equal sections. Literature review shown no researches on the visually impaired navigation has taken this approach to extract spatial

information of the facing scene from the disparity image. After the whole disparity image is analysed, it is broken into three equal sections, foreground, mid-ground and background. Each section can be analysed independently according to the needed spatial information. The relative size of the object occupied in the three sections contains needed information of the facing scene spatial information. It is also contained information regarding the type of the facing scene, close range, mid range or outer range.

1.1 Criteria

These criteria are based on psychology researches, which are discussed in Section 2.1

- a. The map should also be in an egocentric view of the environment.
- b. The map should provide simple spatial information.
- c. The map needs to be working in a real time.
- d. The map should also provide obstacles detection and orientation.
- e. The map should be use in conjunction with the walking cane.
- f. The map should not use auditory interface to provide spatial information to the user.

2 Background and Electronic Travelling Aids Overview

In this section, the basic background relevant to the design of visually impaired navigation system is discussed. It would not be proper to design ETA (Electronic Travelling Aids) for the visually impaired without some knowledge of the visually impaired behaviour during their travelling. A good and effective design of an ETA for the visually impaired should be based on their behaviour during their travelling. The criteria, which mentioned in Section 1, are obtained from the psychology researches of the visually impaired behaviour during their travelling.

Section 2.1 discusses the visually impaired behaviour during their travelling. It will set clear the relationship between our criteria and their behaviour. It also shows that the visually impaired can learn more about spatial relationship of the surrounding environment through the use of a tactile map. Electronic travelling aids can be categorised into 1) Obstacle Detection ETA and 2) Orientation ETA. Section 2.1 show that the visually impaired can really benefit from tactile map user interface.

2.1 The User: The Visually Impaired.

Before further discussion on the visually impaired behaviour during their travelling, it is necessary to clarify what visually impaired is. The following definition extracted from International Classification of Diseases is given blindness and visually impaired are defined as in the International Classification of Diseases 10th edition (ICD10) [6]:

- BLINDNESS is defined as visual acuity of less than 3/60 or corresponding visual field loss in the better eye with best possible correction (cat 3,4,5)
- LOW VISION corresponds to visual acuity of less than 6/18, but equal or better than 3/60 in the better eye with best possible correction (cat1,2)

Information regarding the visually impaired behaviour during their travelling is extremely useful in the designing of electronic travelling aids for the visually impaired. The wayfinding skills of sighted individuals and visual impaired individuals are directly related to their cognitive mapping ability. Cognitive mapping ability is the ability to make informed spatial decisions through the processing and synthesis of spatial information is through various senses such as touch, hearing, smell and vision. These sensory experiences are transformed into knowing and understanding of the spatial awareness of the surrounding environment, which is known as cognitive mapping. [7] Cognitive mapping ability relies significantly on the sense of vision. Without vision sense, cognitive mapping ability is reduced to less than half of the normal mapping capability. With the use of a proper design ETA, this ability can be improved.

The wayfinding skills of the blind and visually impaired individuals are based on different kinds of information from those sighted people.[2] Walking speed and stride length of the visually impaired is less than of a sighted people since they sometime need to make contact with the surrounding to obtain

more navigating information. They rely more on the sense of hearing and the movement is focus on getting to a point and less on moving along an edge. [8] The visually impaired also used body rotation to help map objects or location in the surrounding environment in their mental map. The mental map of the surrounding is also much more relating to the visually impaired body than of a sighted person. This is called an egocentric view of the environment. [9] Thus, the visually impaired mental map of the surrounding is also in the form of egocentric view of the environment.

Due to the inability to obtain knowledge of the up coming objects or open spaces, the visually impaired need some kind of aids to provide them with the needed spatial information. Walking cane is one of the oldest and most useful travelling aids available for the visually impaired. It has a rather limit range hence will not be able to provide spatial information of the whole facing scene in advance. Researchers have been trying to replace the walking cane with electronic travelling aids. None of the ETAs can really claim to be able to replace the walking cane just yet. Thus it might be better to design an ETA to be a supplementary aids to work in conjunction with the walking cane.

Another downfall of many of the ETAs is provide the user with complex spatial information or too much information, while simple spatial information might be more suitable and more useful to the visually impaired. As it has been shown in a study that the visually impaired use more simple information more frequently than complex information. [10] A more simple spatial information can be in the form of location of open spaces in the facing scene or location of standing objects in the facing scene, which can be represented as vertical and horizontal lines on a non vibrating dynamic tactile map. Many researchers[3, 4] have also showed that the visually impaired can improve their cognitive mapping ability with the use of a tactile map.

2.2 Electronic Travel Aids for the Visually Impaired (ETA).

Electronic Travel Aids can be categorised into Obstacle detection ETA and Orientation ETA. Currently the only available commercial ETA is based on sonar sensor however many researchers are working toward using vision sensor for the scene acquisition.

2.2.1 Obstacle Detection ETA

The goal of obstacle detection ETA is to help the users avoid collision with any obstacle in their path. There are currently two types of obstacle detection

ETA, systems based on sonar sensor and systems based on vision sensor. The latter is not available commercially at the present. However, as the cost of computer and vision sensor are decreasing rapidly and wealth of information that can be obtained using vision sensor, more researches are heading toward using vision sensor instead of the sonar sensor. Since researchers are just start looking toward using vision sensors instead sonar sensor, only few obstacle detection techniques are available at the present. Some of the promising obstacle detection ETA techniques are described below.

2.2.1.1 Sonar Sensor Obstacles Detection System

Sonar sensor is one of the most common forms of distance measurement used in Mobile Robotics and a variety of other applications. The principle is simple to understand, a speaker (Transducer) is used to emit a short burst of sound (Ping). The sound wave travels through the air and reflects off a target back to the Transducer (Echo). By measuring the Time of Flight between Ping and Echo detection, one can calculate the distance between the target and transducer.

2.2.1.2 Vision Sensor Obstacles Detection System.

Vision sensor is becoming more common sensor of choice for mobile robotics. As vision sensor can capture much more information in just one single frame. A vision sensor is basically a single camera used to capture the image needed to analyse.

2.2.2 Orientation ETA.

The aim of orientation navigation ETA is to provide current location and best route to the user destination. With the widely available and more reliable of global positioning system (GPS) and geographic information system (GIS), more researchers are turning to the aids of GPS and GIS to provide orientation and location information for the visually impaired. A brief discussion of the current researches on orientation travelling aids for the visually impaired is discussed below.

3 Spatial Information Awareness Map (SIAM)

The ultimate goal is to provide the user with a Spatial Information Awareness Map (SIAM). The SIAM is intended to provide users with spatial information awareness, in particular the relationship between an open space and standing object, of the facing scene. The information is provided to the user with symbolic horizontal and vertical lines, representing the open space and obstacles in the facing scene

3.1 Disparity Image and Gradient Operators

The disparity image is usually obtained from two cameras, referred to as stereo vision. The idea of using stereo vision stemmed from human visual system, which use two eyes for depth perception. Stereo vision uses two cameras to observe the environment, finds the same object in each image, and measures depth to the object by triangulation; that is, by intersecting the lines of sight from each camera to the object. Finding the same object in each image is called matching and is the fundamental computational task underlying stereo vision. The disparity (d) is defined as a difference between the column coordinates of the pixel locations of corresponding match pixels in the left and right image of a rectified stereo pair.

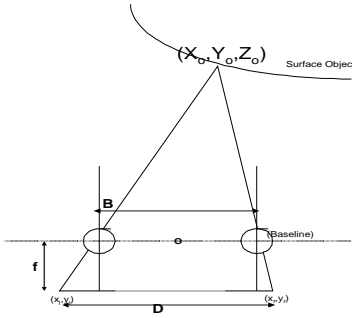


Figure 1: Stereo imaging geometry.

With a parallel stereo set up in Fig 1, the disparity (D) is

$$(x_r - x_l) = \frac{-Bf}{Z} \quad (1)$$

$$D(x, y) = x_r - x_l \quad (2)$$

B is the baseline length and f is the focal length. Z is the depth of the object, and $D(x, y)$ is the disparity value at coordinate x and y . Also x_r and x_l represent coordinate of the matching pixel from the right camera and left camera respectively.

Gradient operators are categorised under local pre-processing method. They are based on local derivatives of the image function. Gradient operators can be interpreted as high pass filter as they suppress low frequencies of the image function. Since they are based on derivatives, which proportion to changes in image function, they are often use to locate rapid change in the image.

Gradient of an image $I(x, y)$ at coordinate (x, y) is given as a vector

$$\nabla I = \begin{bmatrix} \frac{\partial I(x, y)}{\partial x} \\ \frac{\partial I(x, y)}{\partial y} \end{bmatrix} \quad (3)$$

Gradient of the Disparity Image is

$$\nabla D(x, y) = \left[\begin{array}{c} \frac{\partial D(x, y)}{\partial x} \\ \frac{\partial D(x, y)}{\partial y} \end{array} \right] \quad (4)$$

The direction of the gradient is denoted as $\theta(x, y)$

$$\theta(x, y) = \arctan \frac{\partial D(x, y) / \partial y}{\partial D(x, y) / \partial x} \quad (5)$$

The problem we aim to solve is: given a disparity image of a scene, what kind of spatial information can be extracted for the visually impaired? As mention earlier there are two types of spatial information that are really valuable to the user: Empty space in the scene and Standing Object in the scene.

Using (1) and (2), it is clear that the disparity is inversed proportion to the depth as given in (6).

$$D(x, y) = \frac{-Bf}{Z} \quad (6)$$

Using the relationship between the disparity and the depth, the characteristics of open spaces and standing objects can be clearly stated. Assuming that the camera is optical axis parallel to the scene, the open space depth (Z) is increasing away from the camera and toward the top of the image. Since the disparity is inversely proportional to the depth, open spaces in the disparity must decrease as they move away from the camera and toward the top of the disparity image. However, the disparity image is discrete, so instead of having a smooth region of decreasing disparity representing the open space, a decreasing regions of constant piece-wise disparity are provided. By nature standing object must have a constant depth region. According to (6), this is represented as a constant disparity region in the disparity image.

3.2 Regional Processing

Regional processing is introduced to ensure that the system can operate as close to real time as possible. This idea is illustrated in Fig 3. Where the disparity image is broken into 3 equal, the top region, middle region and bottom region. Information contained in each of the regions can be analysed according to their immediate needs. Information contained in each of the region can be analysed according to theirs immediate need. Information regarding flat surface is much more useful for the users in the bottom region than the top region. Size of the object in each region and how much it is occupied other regions contain information regarding the type of scene: close range, mid range and outer range. According to perceptive rule of image, the top region is always the farthest region.

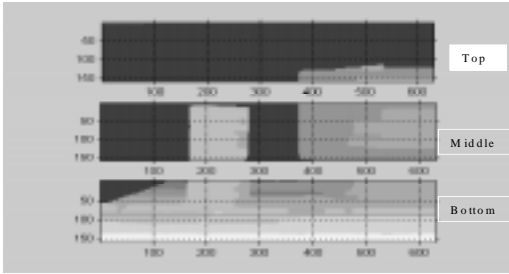


Figure 2: Regional disparity image

By product of the regional processing are also very valuable spatial informations for the visually impaired. After apply gradient operator on the given disparity image, two types of spatial information can be extracted 1) Empty space region and Standing object region.

3.2.1 Empty Space Region

Referring to (1) and (2), the empty spaces in the disparity image can be identified from regions of decreasing piece-wise disparity in the y direction. Due to the physical nature of an open space region, its piece-wise disparity is always decreasing in the y direction unless it is a ceiling, in a which case the opposite is true. Gradient operators, which are able to detect and locate rapid change, are used to extract empty space regions from the disparity image. Two cases are used to detecting open space, derivative in x direction and in y direction.

1) Derivative in y direction detecting changes in disparity

$$\Delta_y D(x, y) = D(x, y) - D(x, (y-1)) \neq 0 \quad (7)$$

Changes in disparity

2) Derivative in x direction

$$\Delta_x D(x, y) = D(x, y) - D((x-1), y) = 0 \quad (8)$$

No changes in disparity

For open space, there should be no changes along x direction. For open space, detecting along y direction results in multiple horizontal lines shown in Fig. 3

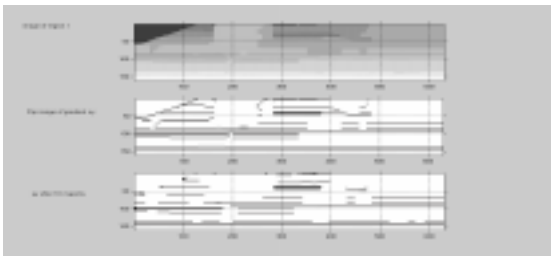


Figure 3: Extracted result of the bottom region in y direction.

Mathematical morphology technique is also used to

help clean up isolated pixels.

3.2.2 Standing Object

Standing object in the scene can be identified by constant disparity in the x and y direction. It is not necessary to be able to extract the shape of the standing object since only its spatial relationship in the scene is needed. Gradient operators are used to extract vertical lines represent standing object in the scene. As mention in Section 3.1, the disparity image is discrete in nature so its derivatives in x direction must be approximated by differences and is given by

$$\Delta_x D(x, y) = D(x, y) - D((x-1), y) \quad (9)$$

One cases is used to detect standing object, derivative in the x direction

1) Derivative in x direction detecting changes in disparity

$$\Delta_x D(x, y) = D(x, y) - D((x-1), y) \neq 0 \quad (10)$$

Changes in disparity

For standing object, detecting along the x direction results in multiple vertical lines shown in Fig. 4

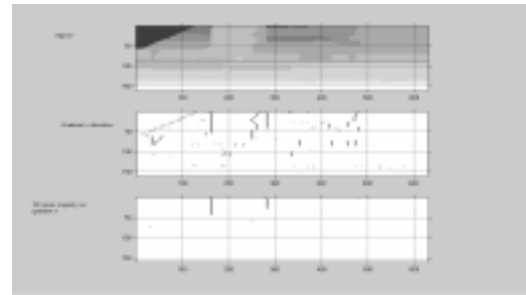


Figure 4: Extracted result of the bottom region in x direction.

Mathematical morphology technique is also used to help clean up isolated pixels.

4 Results



Figure 5: Test Disparity Image

We now present results from applying the previous described algorithm to another real disparity

image. It is shown on Fig. 6 that the scene is a mid-range scene. Information extracted from the bottom region present in Fig 7 shown multiple horizontal lines suggests the users the presence of open space in the scene. Fig 7 shows the extracted multiple horizontal lines from the bottom region of the disparity image in Fig. 5, which suggest the open space region.

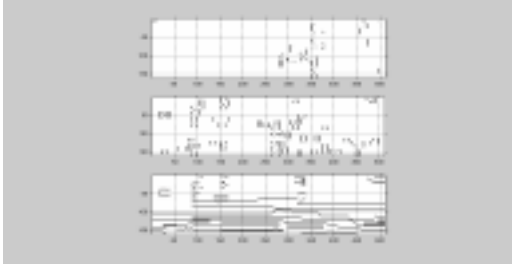


Figure 6: Mid-range Scene

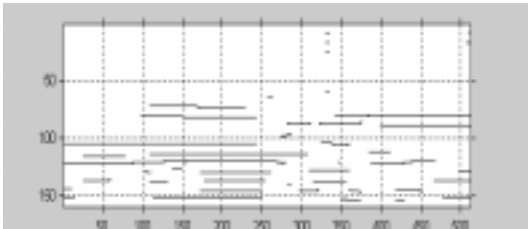


Figure 7: Bottom Region Extraction Result

5 Conclusions and Future Works

This paper presented a novel hybrid spatial information for the visually impaired navigation. In order for the system to operate as close as real time as possible, a computational efficiency gradient operators was used to extract open spaces and standing objects from a given disparity scene. The idea was tested on a real scene disparity image and the results are very promising. Granted that the gradient operators are not without fault, the novelty of the paper is lied in the hybrid spatial information provided for the users and regional processing of the given disparity map.

More work is also being done on integrating the whole system: acquisition, processing and user interface. Dynamic tactile map and a real time stereo matching, which are crucial to the system will be investigated further. The processing part of the system, which extracted spatial information awareness, will also be under more investigated.

Reference

- [1] WHO Statistical Information System, "Data on Visual Impairment," vol. 2001: World Health Organization, 2001.
- [2] S. Ungar, A. Espinosa, M. Blades, E. Ochaita,

and C. Spencer, "Blind and visually impaired people using tactile map.," *Proceeding of the Association of America Geographers Annual Meeting.*, 1996.

- [3] R. Passini and Proulx G, "Wayfinding without vision: an experiment with congenitally, totally blind people," *Environment and Behavior*, vol. 20, pp. 227-252, 1988.
- [4] R. Edwards, S. Ungar, and M. Blades, "Routes descriptions by visually impaired and sighted children from memory and from maps," *Journal of Visual Impairment and Blindness.*, 1996.
- [5] S. U. M. Angeles Espinosa, Esperanza Ochaita, Mark Blades and Christopher Spencer, "Comparing Methods for Introducing Blind and Visually Impaired People to Unfamiliar Urban Environments," *Journal of Environmental Psychology*, pp. 227-287, 1998.
- [6] International Classification of Diseases, "ICD10 Definition,".: Available from :<http://www.who.int/pbd/pbl/img/icd10.gif>, 1999.
- [7] R. D. Jacobson, "Cognitive Mapping without sight: Four Preliminary Studies of Spatial Learning," *Journal of Environmental Psychology*, vol. 18, pp. 289-305, 1998.
- [8] G. Jansson, "Travelling Without Vision : On The Possibilities of Cognitive and Perceptual Compensation," in *Communication and Handicap: Aspects of Psychology Compensation and Technical Aids*, E. Hjelmquist and L. G. Nilsson, Eds. Holland: Elsevier Science Publishers, 1984, pp. 103-114.
- [9] M. Brambring, "Electronic Spatial Sensing for the Blind: Contribution from Perception, Rehabilitation and Computer Vision," presented at Proceeding of the NATO Advanced Research Workshop on Visual Spatial Prosthesis for the Blind, Lake Arrowhead, USA, 1984.
- [10] B. L. Bentzen and P. A. Mitchell, "Audible Signage as a Wayfinding Aid: Verbal landmark verses Talking Signs," *Journal of Visual Impairment and Blindness*, vol. 89, pp. 494-505, 1995.