Landmarks Detection to Assist the Navigation of Visually Impaired People

Paulo Costa[1], Hugo Fernandes[2], Verónica Vasconcelos[3], Paulo Coelho[1], João Barroso[2,4], and Leontios Hadjileontiadis[5]

[1] Polytechnic Institute of Leiria, Portugal
paulo.costa@ipleiria.pt, paulo.coelho@ipleiria.pt
[2] University of Trás-os-Montes e Alto Douro, Portugal
[3] Polytechnic Institute of Coimbra, Portugal
[4] GECAD – Knowledge Engineering and Decision Support Research Center, Porto, Portugal
[5] Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece
{Hugof,jbarroso}@utad.pt, veronica@isec.pt, leontios@auth.gr

Abstract. Assistive technology enables people to achieve independence in the accomplishment of their daily tasks and enhance their quality of life. Visual information is the basis for most navigational tasks, so visually impaired individuals are at disadvantage due to the lack of information or given insufficient information about their surrounding environment. With the recent advances in inclusive technology it is possible to extend the support given to people with visual disabilities during their mobility. In this context we propose a new algorithm to recognize landmarks suitably placed on sidewalks. The proposed algorithm uses a combination of Peano-Hilbert Space Filling Curves for dimension reduction of image data and Ensemble Empirical Mode Decomposition (EEMD) to pre-process the image, resulting on a fast and efficient recognition method and revealing a promising solution.

1 Introduction

Assistive technology enables people with disabilities to accomplish daily tasks and assists them in communication, education, work and recreational activities. Principally though, it can help them to achieve greater independence and enhance their quality of life. From the various assistive technologies available nowadays, a special focus was put on those that help blind or visually impaired people with their mobility. The World Health Organization estimates that there are 37 million blind people worldwide [1]. Blind or visually impaired people have a considerable disadvantage, as they need information for bypassing obstacles and have relatively little information about landmarks, heading and self-velocity. The main issue on using assistive technologies is to provide additional information, useful to blind people during their mobility process, i.e., walking.

Human mobility can be distinguished between Orientation and Navigation. Orientation can be thought of as knowledge of the basic spatial relationships between

J.A. Jacko (Ed.): Human-Computer Interaction, Part III, HCII 2011, LNCS 6763, pp. 293–300, 2011.

objects within the environment. Information about position, direction, desired location, route, route planning, etc., are all bound up with the concept of orientation. Navigation, in contrast, suggests an ability to move within the local environment. This navigation implies the knowledge of immediate objects and obstacles, ground morphology (holes, stairs, flooring etc.), and moving or stationary dangers.

The aim of the present work is to propose a new algorithm for the computer vision (CV) module that will be integrated in a prototype currently being developed by the SmartVision project team. The new image processing algorithm is intended to extract useful information from outdoor scenes in the University of Trás-os-Montes and Alto Douro (UTAD) campus and put the blind user correctly positioned on the sidewalk along a predefined route. In order to reduce the image complexity features for extraction/detection, some landmarks were placed along the sidewalks. The proposed algorithm uses a combination of Peano-Hilbert space filling curves [2] and Ensemble Empirical mode decomposition (EEMD) [3] for image processing and a basic correlation algorithm for template matching.

The paper is organized as follows. Section 2 presents a classification of navigation systems and related work; some projects that represent the state of the art are presented. Section 3 presents the proposed algorithm and the related techniques used. Section 4 presents and discusses the results. Finally, Section 5 concludes the paper.

2 Background and Related Work

An Electronic Travel Assistant (ETA) has to supply the visually impaired with the necessary routing information to overcome obstacles in the near environment with minimum error. Navigation systems to assist visually impaired people can be classified in three groups, based on their usage. The indoor systems are to be used in structured environments with less complex scenes, typically inside buildings or in isolated controlled campuses. The outdoor navigation systems are intended to be used in exterior open space, typically on the street. The indoor/outdoor systems can be used in both indoor and outdoor spaces, switching functionalities based on environment operation.

Some commercial research and development (R&D) projects that currently describe the state of the art in outdoor navigation systems for assisting visually impaired people are presented as follows:

1. Navigation systems without local obstacle information: the systems BrailleNote GPS [4], StreetTalk [5], Trekker [6], NOPPA [7], Navigator [8] and Drishti [9] are GPS based systems to assist the navigation of visually impaired people. Their primary components are a PDA or Laptop especially designed/adapted for people with visual disabilities, a Bluetooth GPS receiver and specially developed software for configuration, orientation and route mapping. The output give for user interaction can be a Braille display or a speech synthesizer.
2. Navigation systems with local obstacle information provide better knowledge of the local scenario, increasing the information quality provided to the blind user to overcome local obstacles.

Several techniques are used to detect and measure object distances, like multiple ultrasonic sensors (sonar) [10] and Laser Range Scanner (LRS) [11], and computer vision (CV) techniques like Principal Component Analysis (PCA) used in ASMONC [10], the Fuzzy Like Reasoning segmentation technique used in the Tyflos system [11], the Expectation-Maximization (EM) algorithm used by Zelek [12], the stereo images for measuring distance to objects, used by Meers [13] and Hadjileontiadis [14], the Neural Network technique used in NAVI [15] and, later on in the same project, the authors also tested Fuzzy Learning Vector Quantification (FLVQ) to classify objects in the scene.

3 Peano-Hilbert and Ensemble Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) [16] is a method for breaking down the signal without leaving the time domain; it filters out functions which form a complete and nearly orthogonal basis for the signal being analyzed. The adoption of adaptive basis functions introduced by Huang et al. [16] provided the means for creating intrinsic *a posteriori* base functions with meaningful instantaneous frequency in the form of Hilbert spectrum expansion [16]. These functions, known as Intrinsic Mode Functions (IMFs), are sufficient to describe the signal, even though they are not necessarily orthogonal [16]. IMFs, computed via an iterative 'sifting process' (SP), are functions with zero local mean [16], having symmetric upper and lower envelopes. The SP depends both on an interpolation method and on a stopping criterion that ends the procedure. Some updates of the 1D-EMD have been proposed which address the mode mixing effect that sometimes occurs in the EMD domain. In this vein, 1D-Ensemble EMD (1D-EEMD) has been proposed [3], where the objective is to obtain a mean ensemble of IMFs with mixed mode cancelation due to input signal noise addition.

3.1 1D-Empirical Mode Decomposition (1D-EMD)

1D-EMD considers a signal $x(t)$ at the scale of its local oscillations [16]. Locally, under the EMD concept, the signal $x(t)$ is assumed as the sum of fast oscillations superimposed to slow oscillations. On each decomposition step of the EMD, the upper and lower envelopes are initially unknown; thus, an interactive sifting process is applied for their approximation to obtain the IMFs and the residue. The 1D-EMD scheme is realized according to the following steps [16]:

1. Identify the successive extrema of $x(t)$ based on the sign alterations across the derivative of $x(t)$;
2. Extract the upper and lower envelopes by interpolation; that is, the local maxima (minima) are connected by a cubic spline interpolation to produce the upper (lower) envelope. These envelopes should cover all the data between them;
3. Compute the average of upper and lower envelopes, $m_1(t)$;
4. Calculate the first component $h_1(t)=x(t)-m_1(t)$;
5. Ideally, $h_1(t)$ should be an IMF. In reality, however, overshoots and undershoots are common, which also generate new extrema or exaggerate the existing ones

[16]. To correct this, the sifting process has to be repeated as many times as is required to reduce the extracted signal as an IMF. To this end, $h_1(t)$ is treated as a new set of data, and steps 1-4 are repeated up to k times (e.g., $k = 7$) until $h_{1k}(t)$ becomes a true IMF. Then set $c_1(t) = h_{1k}(t)$. Overall, $c_1(t)$ should contain the finest scale or the shortest period component of the signal;

6. Obtain the residue $r_1(t) = x(t) - c_1(t)$;
7. Treat $r_1(t)$ as a new set of data and repeat steps 1-6 up to N times until the residue $r_N(t)$ becomes a constant, a monotonic function, or a function with only one cycle from which no more IMFs can be extracted. Note that even for data with zero mean, $r_N(t)$ still can differ from zero;
8. Finally,

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t), \tag{1}$$

where $c_i(t)$ is the -th IMF and $r_N(t)$ the final residue.

3.2 1D-Ensemble Empirical Mode Decomposition (1D-EEMD)

One of the major drawbacks of the original 1D-EMD is the appearance of mode mixing, which is defined as a single IMF consisting of signals of widely disparate scales, or a signal of similar scale residing in different IMF components. The effect of adding white noise scales uniformly through the whole time-scale or time-frequency space, will provide a reference distribution to facilitate the decomposition method. The added white noise may also help to extract the true signals in the data, a truly Noise-Assisted Data Analysis [3]. The 1D-EEMD is implemented as follows:

1. Add white noise series w(t) to the data x(t), X(t) = x(t) + w(t);
2. Decompose the X(t) data with white noise into IMFs, $X(t) = \sum_{j=1}^{N} c_j(t) + r_N(t)$;
3. Repeat step 1 and step 2 several times with different noise series $w_i(t)$, $X_i(t) = x(t) + w_i(t)$, and obtain corresponding IMFs, $X_i(t) = \sum_{j=1}^{N} c_{ij}(t) + r_{iN}(t)$;
4. Finally, the ensemble means of corresponding IMFs of the decomposition are

$$c_j(t) = \frac{1}{N} \sum_{i=1}^{N} c_{ij}(t), \tag{2}$$

where N is the ensemble members.

3.3 Peano Hilbert-Ensemble Empirical Mode Decomposition (PH-EEMD)

A space-filing curve (SFC) is a continuous scan that passes through every pixel of the image only once. In order to transform an image (2D data) on a signal (1D), the space filing curve must preserve the neighborhood properties of the pixel [17]. These curves were first studied by Peano [18] and later by Hilbert [19] and corresponding algorithms are described in [2]. The Peano-Hilbert curve has three main interesting properties: (i) the curve is continuous; (ii) a scanning curve is continuous almost everywhere; and (iii) some parts of the curve are similar with whole curve suggesting a fractal structure. The Peano-Hilbert curve is the most popular recursive SFC and is used in many applications. Fig.1 represents three Peano-Hilbert space filing square curves of area one.

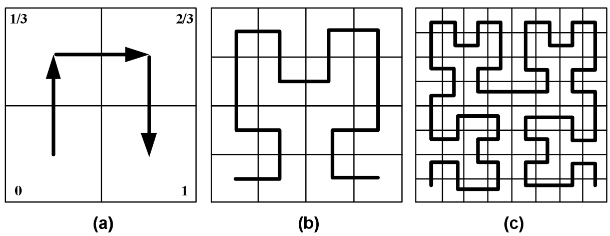

Fig. 1. Peano-Hilbert Curve: a) basic curve; b) 2 interactions; c) 3 interactions

The advantages of space filing curves were combined with the 1D-EEMD algorithm proposed by Huang for image processing. Some extensions to 2D-EMD were proposed to deal with the 2D nature of the data, but these fully 2D-EMD approaches are very time consuming processes. An algorithm for 2D application was proposed with relatively low trend in processing time. This algorithm is based on three phases:

1. Decompose the image using the Peano-Hilbert curve and get the equivalent 1D signal. For the Peano-Hilbert algorithm, a recursive function is used to get the -th order area one curve.

This procedure converts 2D data into 1D signal maintaining the local pixel spatial relations between neighbors.

2. Apply the 1D Ensemble Empirical Mode Decomposition (EEMD) to the linear signal in order to compute the 1D Intrinsic Mode Functions that carry multi-scale space-frequency information. For the EEMD white noise with amplitude of 0.1, standard deviation of the original data is added and the process is repeated 8 times.

Boundaries problems are associated to most of the data processing algorithms due to finite data samples. In the EMD algorithm this is particularly true in the interpolation procedure on the sifting process; to solve this, an even data extension method is used [16].

3. To get the 2D data decomposition the inverse procedure must be taken to reconstruct the image from the data, using the Peano-Hilbert pixel spatial relations to process the 1D IMFs back to 2D IMFs, according Fig.1.

3.4 The Proposed Algorithm for Landmarks Detection

In order to provide useful information to blind people the vision system must be able to detect relevant features in the scene and help the blind user to stay in safe paths. The first approach was intended to reduce the image complexity for the processing algorithms and to enhance the detection. Landmarks were made on sidewalks representing safe paths along the route. From several geometric marks, circles were adopted because in this application they are scale and rotational invariant. Several captured images with different circle radius were tested and to minimize the size and maximize de detection rate a 15 cm circle radius was chosen.

The proposed CV phases are described below:

1. Decompose the captured image with Peano-Hilbert Ensemble Empirical Mode Decomposition
2. Image filtering to eliminate higher frequencies containing noise and fine details. This process is achieved in the EEMD reconstruction phase (1) by eliminating the first two IMFs according to a criterion of root mean square error (RMSE) minimization. Fig.2. represents the *RMSE* during the reconstruction phase of the image which was corrupted by Gaussian white noise ($\sigma = 0.1$). The minimum error occurs when removing the first two IMFs, where the reconstruction is $x(t) = \sum_{i=3}^{N} c_i(t) + r_N(t),$ were i starts at 3 IMF.

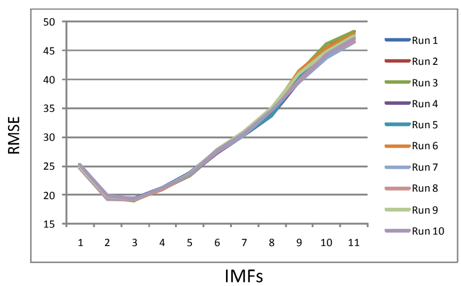

Fig. 2. Reconstruction error of a image corrupted with white noise

3. Define a region of interest (ROI) near the blind user; in our case we chose to analyze the first half of the image and a quarter image size for each side of the blind user. Data outside the ROI area is set to zero.
4. Perform data binarization of the ROI image with a global threshold using Otsu's method [20], followed by Canny edge detection.
5. Finally the ROI image is passed onto a circle detection procedure using simple correlation template matching.

4 Experimental Results and Discussion

In order to test the proposed algorithm for the detection of landmarks within the sidewalk, a set of different images were captured on UTAD campus (outdoor

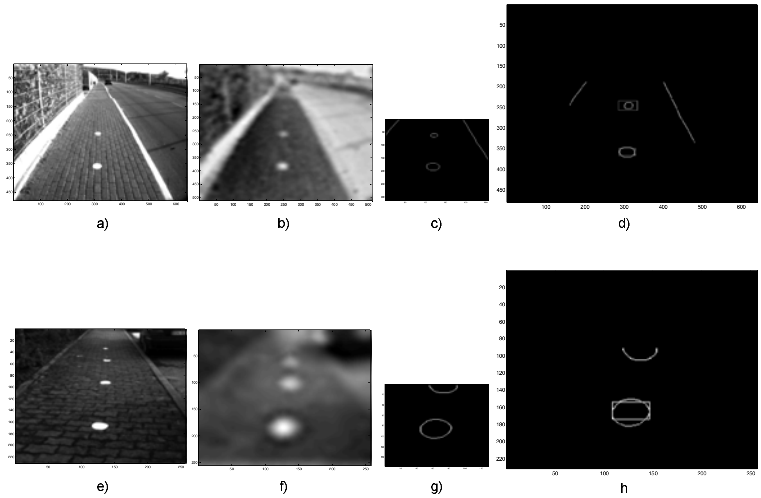

Fig. 3. Results of the proposed algorithm for landmarks detection

scenario). Two images that represent different sections of the campus are depicted in Figs. 3a) and 3e), with the latter representing a more difficult task for any image processing algorithm. Fig. 3b) and Fig. 3f) are the corresponding EEMD filtered images. As it can be seen from these figures the higher frequencies were removed and, with this procedure of prior image binarization, the appearance of small artifacts was minimized. Fig. 3c) and Fig. 3g) show the ROI near the user that are processed in circle detection. We consider that the user is centered at the bottom of image. Finally, Fig. 3d) and Fig. 3h) represent the circle detection results of the respective images; all detected circles are marked with a rectangle. In order to improve the visualization of the results, these two images are presented at a bigger scale.

From the outputs of circle detection in the image, the blind user must head in the direction of the nearest circle. This ensures that s/he will not get out of safe path. Based on the relative position of the nearest circle to the blind user it is possible to compute the trajectory correction and output it to the blind user. The interface to the user is made through five microvibrators, corresponding to five directions, i.e., left, left-diagonally, straight, right-diagonally and right.

5 Conclusion and Future Work

In the presented work a Computer Vision module for the SmartVision project was proposed. For an efficient assistance to a blind user's navigation the CV module must detect accurately specific features in the environment. Due to very different scenarios that can be found in outdoor navigation the sidewalks were marked with landmarks to improve the CV feature detection efficiency. For landmark detection the Peano-Hilbert Ensemble Empirical Mode Decomposition Template Matching method was implemented and the system has proved to be able to detect the defined landmarks and provide valid and simple instructions to the blind user.

Further work is needed to enhance the method accuracy and future improvements will still continue to use PH-EEMD image analysis. Range image information (disparity map) will be integrated into the CV module to provide obstacle detection features.

References

1. WHO: Magnitude and causes of Visual Impairment. Fact sheet no. 282 (November 2004)
2. Allaart, P.C., Kawamura, K.: Dimensions of the coordinate functions of space-filling curves. Journal of Mathematical Analysis and Applications 335, 1161–1176 (2007)
3. Wu, Z., Huang, N.E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. Centre for Ocean-Land-Atmosphere Studies, Calverton, Md, USA (2005)
4. GPS, `http://www.senderogroup.com` (last viewed on Janury 2011)
5. StreetTalk, `http://www.freedomscientific.com` (last viewed on January 2011)
6. Trekker, `http://www.humanware.ca` (last viewed on Janury, (2011)
7. Ari, V., Sami, K.: NOPPA - Navigation and Guidance System for the Blind. VTT Industrial Systems, ITS World 2004 (2004)
8. Kowalik, R., Kwasniewski, S.: Navigator – A Talking GPS Receiver for the Blind. In: Miesenberger, K., Klaus, J., Zagler, W.L., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 446–449. Springer, Heidelberg (2004)
9. Ran, L., Helal, S., Moore, S.: Drishti: An Integrated Indoor/Outdoor Blind Navigation System and Service. In: Proceedings of IEEE International Conference on Pervasive Computing and Communication 2004, Orlando, pp. 23–31 (2004)
10. Molton, N., Se, S., Brady, J.M., Lee, D., Probert, P.: A stereo vision-based aid for the visually impaired. Image and Vision Computing 16, 251–263 (1998), 0262-8856/98/$19.00 q 1998 Elsevier Science B.V. All rights reserved
11. Bourbakis, N.G., Kavraki, D.: Tyflos - A 2D Vibration Array for Sensing Dynamic Changes and 3D space for Blinds' Navigation. In: Proceedings of the 5th IEEE Symposium on Bioinformatics and Bioengineering (BIBE 2005). IEEE, Los Alamitos (2005), 0-7695-2476-1/05 $20.00 © 2005
12. Zelek, J.: Seeing by touch (haptics) for wayfinding. International Congress Series, vol. 1282, pp. 1108– 1112 (2005), 0531-5131/ D 2005, Elsevier B.V. All rights reserved
13. Meers, S., Koren, W.: A Vision System for Providing 3D Perception of the Environment via Transcutaneous Electro-Neural Stimulation. In: Proceedings of the Eighth International Conference on Information Visualisation (IV 2004), pp. 1093–9547. IEEE, Los Alamitos (2005), 1093-9547/04 $ 20.00
14. Hadjileontiadis, L.: - SmartEyes: an enhanced orientation and navigation system for blind or visually impaired people. Computer Society International Design Competition, Final Report (April 2003)
15. Nagarajan, R., Sainarayanan, G., Yacoob, S., Porle, R.: An Improved Object Identification For NAVI. Artificial Intelligence Research Group, School of Engineering and Information Technology, University Malaysia Sabah (2004)
16. Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Yen, N., Tung, C., Liu, H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of The Royal Society Of London series A-Mathematical Physical and Engineering Sciences 454 (1971), 903–995 (1998)
17. Dafner, R., Cohen-Or, D., Matias, Y.: Context-based Space Filling Curves. In: Gross, M., Hopgood, F.R.A. (eds.) EUROGRAPHICS 2000, vol. 19(3) (2000)
18. Peano, G.: Sur une courbe qui remptit toute une aire plane. Math. Ann 36, 157–160 (1890)
19. Hilbert, D.: Aber die stetige Ubbildung einer Linie aufdie Flachenstuck. Math. Ann. 38 (1891)
20. Sezgin, M., Sankur, B.: Survey over Image Thresholding Techniques and Quantitative Performance Evaluation. Journal of Electronic imaging 13(1), 146–165 (2004)