**TOPIC: A VOICE NAVIGATION SYSTEM USING NEURAL NETWORK SYSTEMS**

**REVIEW OF RELATED WORKS**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **S/NO** | **AUTHOR NAMES** | **YEAR** | **TOPICS** | **PROBLEM STATEMENT** | **AIM** | **OBJECTIVES** | **METHOD USED** | **RESULT**  **ACHIEVED** | **LIMITATIONS** |
|  | 1. Saleh Shadi 2. Saleh Hadi 3. Mohammad Amin Nazari 4. Wolfram Hardt | 2019 | Outdoor Navigation for Visually Impaired based on deep learning | Recently, several outdoor navigation systems have been developed to aid blind people in navigation. Unfortunately, most of them are not easy to use due to complicated configurations and dependencies. Other solution strategies have various dependencies such as sensors and other devices that are not suitable for visually impaired people to carry around for purposes of navigation. On the other hand costs play a major role in the development of outdoor navigation devices. In conclusion, the navigation system for the visually impaired must be portable, independent, user-friendly, lightweight and cost-efficient for the blind to use as a navigation system. However the visually impaired people hands should be free for other operation such as holding their personal properties and guide cane to avoid some sudden collision. | The aim of this work is to implement a novel outdoor navigation solution for visually impaired using deep learning | 1. To propose an efficient novel solution based on the smartphone camera and deep learning algorithms to detect and recognize different objects and obstacles.  2. Where segmentation-based object recognition is implemented to detect and recognize objects based on pixels with specific segmentation color and class probabilities.    3. The segmented predication color for an object and class probabilities in a single neural network directly is produced from the full image in one evaluation. Because the detection pipeline from a single neural network, it is possible to optimize them directly as end-to-end detection performance. The images of the environment are continuously captured in front of the user, then object detection and image processing are performed to recognize the objects with the estimated distance, in order to deliver voice comments to visually impaired people about the objects or obstacles in front of them. | In this study, the proposed solution should be able to run in real-time on the mobile phone with high accuracy and with minimal size and resource usage. The mobile application uses a smartphone camera to capture and deliver sequence input images to the deep learning model, which perform object recognition with distance estimation and provide voice comments that help the visually impaired understand the objects and obstacles in the outdoor environment. Furthermore, visually impaired people will receive more information (such as type of the different obstacle objects) about their environment and the mobile application will help them not only navigate in an unfamiliar environment, as well it will provide more information about the various obstacles with the estimated distance. Normally, visually impaired people have different behaviors when they walk in some new areas. Therefore, our application provides two different scenario modes. The first mode is called stable mode, in which the walking human is in a completely unknown environment at first. The walking speeds of the human are usually reduced, and the human needs a navigation system with high accuracy of obstacle detection to provide safe navigation with especial voice comments to describe this environment. Visually impaired people will therefore be able to become familiar with the environment very quickly. In the second scenario, it is outdoors in crowded environments where many obstacles and objects are encountered such as people, cars, fences, and buildings, in this case, the probability of users colliding with these obstacles is high. Thus, the navigation system requires high detection speed of the obstacles in the real system in order to inform the user about the different detected object. Therefore, this mode is referred to as Fast Mode according to the previous specifications of a navigation system. Our model is trained on the pre-defined dataset that meets the blind person's needs in order to identify the set of objects needed for blind navigation with very high accuracy and recognition speed for individual walking situations. | As the result showed, visually impaired people will be more alerted to their surrounding environment and the proposed system will assist them not only in navigating in an unfamiliar environment but also in getting more information about the various obstacles they may face them with the estimated distance. | Currently, this system has been developed to navigate the visually impaired people with voice commentary and guidance, and the accuracy of the captured object with distance estimation was good and limited by a specific list of objects necessary for this system. For the future, the system should be expanded to include a larger number of objects with a larger dataset for the recognition of outdoor and indoor objects as well. The system is able to inform the visually impaired persons about different type of object. Therefore, impaired people understand what the objects are around and able to find the objects that they need in indoor as well as outdoor. The calculated distance should be improved, and the error minimized. Therefore, the depth information should be estimated in the meantime with objects recognition using a single monocular camera and a light neural architecture to predict pixel-wise depth map. |
| 2. | Huck vale | 2014 | A speech recognition system using neural network systems |  | The aim of this work is to implement a speech recognition system using neural network systems |  |  | Speech recognition used a new Viterbi net architecture which  recognized the input patterns and provided an accuracy of recognition rate more than 99% on a large speech database | System is used  for isolated  word recognizer |
| 3. | Fausset | 2014 | A voice navigation system using neural network systems |  | The aim of this work is to implement a voice navigation system using neural network systems |  |  | The model named “Listen, Attend and Spell”(LAS), literally “listens” to the acoustic signal, pays “attention” to different parts of the signal and “spells” out the transcript one character at a time | The recognition sentence is small |
| 4. | Jiang Ming | 2017 | A voice navigation system using neural network systems |  | The aim of this work is to implement a voice navigation system using neural network systems |  |  | Author implemented a pre-trained deep neural network using the hidden markov model (DNN-HMM) hybrid architecture which is used to train the DNN to produce the better recognition results of large vocabulary speech  database | The algorithm concluded a lower accuracy of 82% |
| 5. | Bor-Shing Lin 1  Cheng-Che Lee 1  Pei-Ying Chiang 2 | 2017 | Simple Smartphone-Based Guiding System for Visually Impaired People | Visually impaired people usually have problems walking and avoiding obstacles in their daily lives. Traditionally, such people use guide canes to detect obstacles in front of them. Thus, visually impaired people cannot exactly know what types of obstacles are in front of them and must only depend on guide canes and experiences to walk safely and in the desired path. Furthermore, when in unfamiliar environments, visually impaired people often require assistance in the form of volunteers to guide them through the surrounding environment. Visually impaired people cannot entirely depend on a guide cane to become familiar with their surroundings or react quickly to unforeseen circumstances. When encountering obstacles, visually impaired people must only rely on their experiences to react because they may not know what the obstacles are. However, some obstacles on the road cannot be predicted, such as a parked bicycle or a resting dog. For visually impaired people, reacting quickly to avoid such obstacles by relying solely on experiences is difficult. Therefore, such obstacles could engender danger for an unaccompanied visually impaired person. | The aim of this work is to implement a simple smartphone-based guiding system for solving the navigation problems for visually impaired people. | 1. To design a simple smartphone-based guiding system for solving the navigation problems for visually impaired people and achieving obstacle avoidance.  2. To enable visually impaired people to travel smoothly from a beginning point to a destination with greater awareness of their surroundings.  3. A computer image recognition system and smartphone application was integrated to form a simple assisted guiding system. Two operating modes, online mode and offline mode, can be chosen depending on network availability. When the system begins to operate, the smartphone captures the scene in front of the user and sends the captured images to the backend server to be processed. The backend server uses the faster region convolutional neural network algorithm or the you only look once algorithm to recognize multiple obstacles in every image, and it subsequently sends the results back to the smartphone. The results of obstacle recognition in this study reached 60%, which is sufficient for assisting visually impaired people in realizing the types and locations of obstacles around them. | The proposed navigation system employs a smartphone to continually capture images of the environment in front of a user and perform image processing and object identification to inform the user of the image results.  A university scene was used as an example to illustrate our proposed system. When a visually impaired user is walking on a campus, two scenarios may occur. In the first scenario, the visually impaired user comes to the campus for the first time. Hence, the user is not familiar with the university’s environment and thus walks at a reduced speed. In this scenario, the user requires high recognition accuracy.  Accordingly, the system mechanism called “stable mode” is provided for achieving high recognition accuracy. Under the stable mode, the user can become familiar with the environment quickly by obtaining precise notifications with higher recognition rates than the other mode.  In the second scenario, the visually impaired user is walking through a crowded area or an area with many obstacles such as bicycles parked randomly in front of the university building or dogs resting in the user’s path. The user has a high probability of colliding with some of the obstacles. In this scenario, the system requires a high recognition speed to alert the user about the environment in front of him or her.  Accordingly, the system mechanism called “fast mode” is provided. This mode provides the user with a high quantity of obstacle alerts within a short period. The processing time of the fast mode is shorter than that of the stable mode. According to the preceding scenarios, the proposed system provides stable and fast modes for corresponding circumstances. Through one of these two modes, users can obtain information with varying degrees of accuracy and recognition speed for individual walking situations. | To meet various user needs to a sufficient extent, according to the answers obtained from the interviews with visually impaired people, this system is designed to recognize seven common obstacle types. In this study, the system was trained using an Internet model that had been trained by an image dataset called Image Net [32]. We intend to train the system again using images in Pascal [33]; however, Pascal has the following three drawbacks. First, piers are not present among the 20 image categories in Pascal. Second, although images in Pascal were collected from various locations worldwide, none have been captured from experimental environments such as the ones used in this study. Third, the images in Pascal were captured using high-resolution cameras, whereas the images recognized by the system in this study are captured using smartphones. The difference in scale between images captured using a high-resolution camera and those captured using a smartphone would influence the recognition results. To overcome these drawbacks, images captured by ourselves were used in this study to train the classifier, thereby enabling the training model to be closer to a real situation. Therefore, 1710 images were taken in and around the campus, and each obstacle in these images was labeled with its location and type. This study adopted mean average precision (mAP) to evaluate the effectiveness of depth learning. Additional 800 images captured in and around campus were used for testing. We conducted the training process 10,000, 20,000, 30,000, and 40,000 times to evaluate the differences in the mAP values. Figure 15 shows the results of using only Pascal images for training. *Sensors* **2017**, *17*, 1371 15 of 22 *4.2. Recognition Result and Recognition Rate* To meet various user needs to a sufficient extent, according to the answers obtained from the interviews with visually impaired people, this system is designed to recognize seven common obstacle types. In this study, the system was trained using an Internet model that had been trained by an image dataset called ImageNet [32]. We intend to train the system again using images in Pascal [33]; however, Pascal has the following three drawbacks. First, piers are not present among the 20 image categories in Pascal. Second, although images in Pascal were collected from various locations worldwide, none have been captured from experimental environments such as the ones used in this study. Third, the images in Pascal were captured using high-resolution cameras, whereas the images recognized by the system in this study are captured using smartphones. The difference in scale between images captured using a high-resolution camera and those captured using a smartphone would influence the recognition results. To overcome these drawbacks, images captured by ourselves were used in this study to train the classifier, thereby enabling the training model to be closer to a real situation. Therefore, 1710 images were taken in and around the campus, and each obstacle in these images was labeled with its location and type. This study adopted mean average precision (mAP) to evaluate the effectiveness of depth learning. Additional 800 images captured in and around campus were used for testing. We conducted the training process 10,000, 20,000, 30,000, and 40,000 times to evaluate the differences in the (mAP) values. Figure 15 shows the results of using only Pascal images for training. | In the future, to provide information on more types of obstacles and more accurate recognition, a broader range of obstacle images and a high-end server equipped with a more powerful graphics processing unit could be used to increase the number of recognition categories and the recognition rate. On the other hand, the feature recognition has extensibility. More recognition services of other object types can be added when the computing ability of smartphone improves. These improvements could provide more information without needing wireless connections. |
| 6. | Kai-Wei Chiang | 2004 | INS/GPS Integration Using Neural Networks for Land Vehicular Navigation Applications | Most of the positioning technologies for modern land vehicular navigation systems have been available for 25 years. Virtually all of the systems augment two or more of these technologies. Typical candidates for an integrated navigation system are the Global Position System (GPS) and Inertial Navigation Systems (INS). The Kalman filter has been widely adopted as an optimal estimation tool for the INS/GPS integration, however, several limitations of such multi-sensor integration methodology have been reported; such as the impact of INS short term errors, model dependency, prior knowledge dependency, sensor dependency, and linearization dependency. To reduce the impact of short term INS sensor errors, the bandwidth of true motion dynamics would be identify by spectrum analysis and the first generation denoising algorithm that used the Discrete Wavelet Transform (DWT) would be apply to identify the limitations of the existing denoising algorithm. Consequently, this research proposed the cascade denoising algorithm to overcome the limitations of existing denoising algorithms. | The aim of this research is to implement INS/GPS Integration using neural networks for land vehicular navigation application | 1. To develop a novel wavelet denoising algorithm, cascade denoising algorithm, and develop a conceptual intelligent navigator that consists of ANNs based INS/GPS integration architectures for next generation land vehicular navigation systems.  2. Ultimately, the conceptual intelligent navigator is expected to overcome or, at least, reduce the limitations of the conventional Kalman filter based INS/GPS integration algorithms.  3. As each of these limitations contributes to certain amount of positional error accumulation during GPS outage, therefore, the proposed new algorithms including cascade denoising algorithm and the conceptual intelligent navigator are expected to reduce the impact of these limitations by reducing the positional error accumulation during GPS outage. | ANNs was chosen in this research for building the foundation for developing the conceptual intelligent navigator that has an artificial intelligent based INS/GPS data fusion and navigation algorithm. The requirement of such an algorithm is to overcome the limitations of current INS/GPS data fusion which is based solely on the Kalman filter.  ANNs have been extensively studied with the aim of achieving human-like performance, especially in the field of pattern recognition and robot control and navigation [Mandic and Chambers, 2001]. ANNs are composed of a number of nonlinear computation elements which operate in parallel and are arranged in a manner reminiscent of biological neural interconnections. In addition, ANNs are designed to mimic the human brain and duplicate its intelligence by utilizing adaptive models that can learn from the existing data and then generalize what it has learnt [Ham and Kostanic, 2001]. | The results Presented strongly indicate the potential of including the conceptual intelligent navigator as the core navigation algorithm for the next generation land vehicular navigation system that uses a low cost MEMS IMU integrated with a GPS receiver operating in SPP mode. The most important factor that affects the performance of the conceptual intelligent navigator is the accumulation of navigation knowledge. Theoretically, if enough navigation knowledge can be acquired in one or fewer field tests, the conceptual intelligence might be able to operate in full prediction mode for every new navigation mission. However, the knowledge accumulation should be conducted whenever new navigation knowledge is acquired as the true motion dynamics of the vehicle operating in real life is far more complicated. With the presence of the conceptual intelligent navigator, the traditional navigator that uses a Kalman filter should be regarded as an optimal estimator, instead of a navigator, as it doesn’t have any ability to store and generalize the navigation knowledge that it has learned. In contrast, the conceptual intelligent navigator has the ability to generate, store and generalize the navigation knowledge it has learned. | Thus, the challenge is to acquire large amounts of navigation knowledge and develop an efficient database management facility to accelerate the knowledge retrieval and accumulation process.  The limitations of traditional navigator (KF) • Its position error grows with time. Thus, it can not provide acceptable navigation solutions during long GPS outages (>2 minutes). • It requires a human expert to tune the optimal parameters of the Kalman filter (i.e., Q and R matrices). In addition, these parameters are sensor dependent. • The parameters are not adaptive once the Kalman filter starts navigation. In other words, unless its parameters are optimal, the Kalman filter might not be able to provide reasonable positioning accuracy during certain GPS outages. This issue might not be critical for a navigation grade IMU due to its superior quality; however, it is an important factor that affects the performance of low cost MEMS IMUs due to their poor stability. • A Kalman filter generates navigation solutions by receiving information from INS and GPS as it is an optimal estimator in nature. However, it can not recall the experience and the navigation knowledge which are important to a real “navigator”.  The limitations of the conceptual intelligent navigator • As the training target information is provided by GPS, the navigation states that can be provided by the conceptual intelligent navigator are limited. For example, it can not estimate sensor errors explicitly and it cannot provide roll and pitch information as a single GPS receiver can not provide such information for training. • Currently, the optimal ANN architecture (i.e., number of hidden neurons) of the conceptual intelligent navigator is decided empirically. This is the fundamental problem of a MFNN. In other words, the number of hidden neurons is fixed during data processing. For application like INS/GPS 223 integration, it might require a more flexible approach to adjust the number of the hidden neurons as the navigation knowledge accumulates continuously. • It requires more storage space to store the navigation knowledge. As the accumulation of the navigation knowledge makes the conceptual intelligent navigator different from the traditional navigator, it is the price to pay. Given the fact that the incorporation of artificial intelligence to the navigation algorithm is new to the navigation community, it needs more extensive research to accelerate wider inclusion of such an idea to commercial products. In fact, using artificial intelligence for mobile robot navigation has been studied extensively in robotic engineering related research works since the field of artificial intelligence started. Therefore, developing a new artificial intelligent INS/GPS integration architecture that can overcome some of the limitations of the traditional navigator in a land vehicle environment is a huge challenge. Unlike other applications in signal processing and control, the INS/GPS integration application requires that synaptic weights stored as part of the navigation knowledge to be adaptive, so that the conceptual intelligent navigator can be adapted to the latest sensor error characteristic and vehicle motion dynamic variations. In other words, the stored navigation knowledge is expected to be adjustable during navigation to ensure the learning process can be conducted continuously. |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |