A Hazard Detection and Tracking System for People with Peripheral Vision Loss using Smart Glasses and Augmented Reality

Ola Younis[1], Waleed Al-Nuaimy[2], Mohammad H. Alomari[3]
The School of Electrical Engineering
Electronics and Computer Science
University of Liverpool, United Kingdom

Fiona Rowe[4]
Institute of Psychology, Health and Society
Department of Health Services Research
University of Liverpool, United Kingdom

Abstract—Peripheral vision loss is the lack of ability to recognise objects and shapes in the outer area of the visual field. This condition can affect people's daily activities and reduces their quality of life. In this work, a smart technology that implements computer vision algorithms in real-time to detect and track moving hazards around people with peripheral vision loss is presented. Using smart glasses, the system processes real-time captured video and produces warning notifications based on predefined hazard danger levels. Unlike other obstacle avoidance systems, this system can track moving objects in real-time and classify them based on their motion features (such as speed, direction, and size) to display early warning notification. A moving camera motion compensation method was used to overcome artificial motions caused by camera movement before an object detection phase. The detected moving objects were tracked to extract motion features which were used to check if the moving object is a hazard or not. A detection system for camera motion states was implemented and tested on real street videos as the first step before an object detection phase. This system shows promising results in motion detection, motion tracking, and camera motion detection phases. Initial tests have been carried out on Epson's smart glasses to evaluate the real-time performance for this system. The proposed system will be implemented as an assistive technology that can be used in daily life.

Keywords—*Peripheral vision loss; vision impairment; computer vision; assistive technology; motion compensation; optical flow; smart glasses*

I. INTRODUCTION

Age-related macular degeneration (AMD), cataract and glaucoma are the leading causes of blindness worldwide [1]. Central vision loss is caused by AMD and cataract while glaucoma affects mainly the peripheral vision [1]. Vision problems can involve visual acuity, visual field, and colour impairments [2]. Visual acuity problems due to central causes such as refractive errors and cataract can be corrected. Visual field loss caused by brain injury or other diseases such as glaucoma is typically irreversible and non-corrected by traditional solutions as eyeglasses and lenses[3].

The human field of vision consists of different areas which are used to see varying degrees of details and accuracy about the surrounding environment. Central vision is where objects are clearly and sharply seen and used to perform most of the daily activities. This vision comprises around 13 degrees.
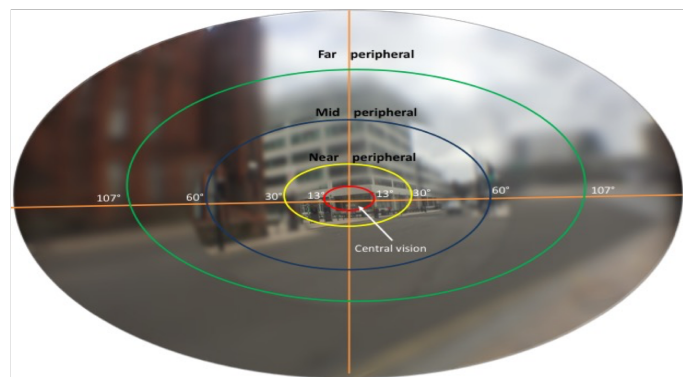

Fig. 1. Human field of view (FOV) for both eyes showing different levels of peripheral vision

The second type is the peripheral vision used to detect larger contrasts, colours and motion and extends up to 60 degrees nasally, 107 degrees temporally, 70 degrees down and 80 degrees up for each eye [3]. The human visual field of view for both eyes showing different types of peripheral vision is shown in Fig. 1. It is important to mention that human beings don't see in full resolution. Instead, we see fine details using the central vision only, whereas in the peripheral vision we see only significant contrasts, colours and recognise motion.

Peripheral vision loss is the absence of outer vision (inward, outward, upward or downward) to varying degrees while the central vision is preserved. Tunnel vision is considered to be the extreme case of peripheral vision loss, where the only part that the person can see is a small (less than 10 degrees) circle in the middle of the central vision as shown in Fig. 2. Core routines such as driving, crossing the road, reading, social activities and other daily actions may become very hard if not impossible for some people [4], [5].

Visual field tests (Perimetry) are examinations that measure visual functions for both eyes to clearly define the blind and seeing areas for each person [6]. Eye specialists interpret perimetry results manually to have an idea about a person's medical condition.

Since many people with peripheral vision loss retain some seeing areas in their visual field, a system that helps them to maximise the residual vision in daily life would be useful.

Fig. 2. Normal Vs. Tunnel vision example. The top picture shows a healthy vision, and the bottom image shows how a tunnel vision person could see the same scene.

This solution should differentiate between a person's blind and healthy areas using personal perimetry results. Furthermore, it will generate notifications if there are any potential hazards (moving or stationary) in a person's blind area.

Developing smart technology to help in healthcare systems is becoming increasingly important. Different types of wearable assistive technologies have been implemented to help people who have vision problems including devices to be worn on several body parts such as the head, chest, fingers, feet, and ears.

Information captured by head-mounted sensors such as cameras can provide a trusted input resource for processing units to define the potential hazards or threats in a person's surroundings. The considerable growth in data processing functionalities in terms of speed, power and data storage can allow people to wear assistive technology in daily life to help cope with their disabilities and defects.

In the case of vision problems, video cameras can be used to capture the surrounding environment information and send this to a processing unit where it generates feedback that enhances the awareness of surroundings. Many smart technologies have been designed to help with navigation, motion detection, quality enhancement and other visual improvements [7].

Computer vision algorithms and techniques have been developed that can recognise, track and classify different types of objects in real-time. A wide diversity of daily applications use these technologies such as video surveillance, augmented reality, video compression and robotic design and implementation. Due to the fast growth in smart mobile development, computer vision algorithms are now available on small, cheap

and high technology devices.

It is essential to mention the difference between virtual and augmented reality. Virtual Reality (VR) is the technology of creating virtual worlds that the user can interact with [8]. VR systems generally require a helmet or goggles. Famous examples are the Oculus Rift by Oculus [9] and HTC Re Vive by HTC [10].

Augmented Reality (AR) is the technology of superimposing computer-generated information, images or animations over a real-world images or video [11], [12]. Current AR implementations are mostly based on mobile applications. Some interesting examples of AR systems are Sony's Smart Eyeglass and the Microsoft Hololens. For more details about these examples, the reader is advised to refer to Al-Ataby et al. [13]

Both VR and AR technologies are similar in the goal of enhancing the user's cognitive knowledge but follow a completely different approach. AR systems tend to keep the user in the real world while letting them interact with virtual objects whereas a VR user is immersed in a completely virtual world. The significant difference between augmented reality systems and other systems that provide superimposition is the user's ability to interact with the computer-generated information [12].

In this work, the main aim is to develop a computer vision system to help people with peripheral vision loss. Using smart glasses and computer vision algorithms, we designed a system that recognises any moving object and classifies it to determine its danger level. Notifications appear in a person's residual field of vision in which the output is projected to. The main aim is to generate meaningful warning messages that are reliable and in the best visual position to warn the person about any possible obstacle/hazard.

This paper is structured as follows: In Section 1, we report a review of the related literature. A description of the proposed system is presented in Section 2. Exploratory evaluation experiments are presented and analysed in Section 3. Finally, research findings and conclusions and recommendations for future work are provided in Section 4.

II. LITERATURE REVIEW

Since 2001, a group at Harvard Medical School developed a device that produced an augmented reality vision for people with severe peripheral vision loss (tunnel vision) [14]. The device comprises a wide-angle camera and one display unit that projects a processed image (cartoon style) from the camera on the regular (healthy) vision. The device was tested on healthy and vision impaired people and results showed improvements of self-navigation and object finding for both cases. The authors also noted that some problems were reported by patients regarding gaze speed reduction.

In 2010 and based on the simultaneous localisation and mapping (SLAM) algorithms, a stereo vision based navigational assistive device that helps visually impaired people to scan the surrounding scene was developed in the University of Southern California, Los Angeles. Data captured by the stereo camera was processed to create tactile cues that alerted the user via microvibration motors to help avoid possible obstacles

and provide a safe route to reach the destination. This work was tested on people with vision loss and the results showed that the presented device could lead vision impaired people to avoid obstacles in their path with the minimal cognitive load. However, this device is very basic in terms of detection angle [15].

A real-time head-mounted display system with a depth camera and software to detect the distance to nearby objects was developed by a group of researchers at Oxford University [16]. The display unit was made of 24x 68 colour light emitting diodes comprised of three 60 mm LED matrices and attached to the front of a pair of ski goggles. The distance between the user and objects was captured by a depth camera. The system used an algorithm that created a depth map and then converted it to an image that the user could see after increasing the brightness of the closer objects. The system could detect objects between 0.5-8 meters. The research group performed two types of experiments; one for sighted people and the second for severely sight-impaired individuals to test their ability to walk and avoid obstacles while wearing these glasses. The authors reported that all the participants could receive response to objects in their visual field [16].

Between these project periods, many previous studies were conducted to apply computer vision concepts and techniques to help people who suffer from vision problems [17] [18] [19]. These solutions were designed to help patients to find a safe path and avoid obstacles using different types of algorithms and adequate hardware. The main objective for most computer vision systems is to highlight different types of objects around the person and prevent collisions or falls. Alarms are generated using different types of sensors like sound and vibration. Because most of these solutions have been for totally blind people, only a few of them use visual alarms.

In 1979, Netravali et al. [20] presented a recursive algorithm that minimised the prediction error of the moving object displacement estimation process for a television scene. Later in 1990, Brandt et al. [21] modelled the camera ego-motion for motion estimation and compensation. The proposed approach tracked moving objects with a moving camera by integrating background estimation techniques, Kalman filtering, autoregressive parameter estimation, and local image matching.

Moving objects in videos captured by a moving camera were positioned and tracked using a technique that applies an active contour model (ACM) with colour segmentation methods [22]. The authors used a matching approach based on an object's area such that the target feature points are tracked over time. The proposed system was tested by several experiments while mounting the video system on a helicopter or a moving car, and promising results were reported.

Vavilin and his colleagues [23] proposed an approach that tracks local image regions over time to detect moving objects and camera motion estimation. A triangular grid of feature points was composed and optimised from the first frame in the video sequence to reflect those regions with more details. Then to extract a tracking feature vector in the next frame, a colour distribution model was generated based on the neighbourhood feature points, and the grid was used to initiate the process at the new frame. A motion field, representing the camera motion parameters, was then formed based on the motion estimation from the grids of both frames.

Camera motion estimation methods have been used for vehicle tracking with moving cameras [24]. The authors proposed a background suppression algorithm to minimise the effect of strong wind and vibrations of the high pillars that mount the camera systems.

A homography transformation-based motion compensation method has been used for a moving camera background subtraction [25]. The authors calculated the movement optical-flow based on grid key-points and achieved a fast processing speed. They worked in real time with 56 frames/second with three components of background segmentation: candidate background model, candidate age, and the background model.

A new approach reported the use of the Color Difference Histogram (CDH) in the background subtraction algorithm [26]. This method compares colour variations between a pixel and its local neighbours, reducing the number of false detections. Then, a Gaussian membership function was used for fuzzification of the calculated difference, and a fuzzy CDH based on fuzzy c-means (FCM) clustering was implemented. The tested algorithm provided an enhanced detection performance of 0.894 Matthew's correlation coefficient (MCC) and 99.08% percentage of correct classification (PCC).

Background subtraction with an adaptive threshold value was proposed to detect moving objects on a conveyor belt[27]. The authors proposed a combined frame difference and background subtraction method with an adaptive threshold that was calculated using the Otsu method and the detection performance was improved reaching 99.6% accuracy compared with the fixed threshold methods.

A literature review for the detection of moving objects in surveillance systems considering some technical challenges such as shadows, the variation of illumination, dynamic backgrounds, and camouflage was presented [28]. An extended survey for well-known detectors and trackers of moving objects has been provided in work done by Karasulu et al. [29] covering the main ideas reported in the literature for detection and tracking in videos, background subtraction, clustering and image segmentation, and the optical flow method and its applications.

A novel navigation assistant system for blind people was implemented in work proposed by Tapu et al. [30]. The proposed system (denoted DEEP-SEE) detects both moving and stationary objects using the YOLO object recognition method [31]. Based on two convolutional networks, their system tracks the detected objects in real-time and solves the occlusion problem. The system then classifies the object based on its location, type, and distance.

III. Proposed System

This work is part of a bigger project to develop a wearable assistive technology to help people peripheral vision loss in their indoor and outdoor navigation [32], [33].

The primary goal of the proposed system is to generate a meaningful notification that is reliable and in the best visual position for the individual. Working with Epson's smart glasses (Moverio BT-200), the system processes the captured video in

real-time to generate suitable output warnings based on the object's extracted features and predefined rules. Smart glasses contain a video camera located in the right corner of the frame. The display units are integrated into the transparent lenses making the glasses capable of presenting the output without blocking the person's normal vision.

Since the users of the system are people with peripheral vision loss, their central vision is still healthy, and they can see through it. The system will superimpose their visual field with the final (most dangerous) outputs after the classification phase in order not to overwhelm them with too many alarms. Stationary obstacles located in the user's pathway are ignored because they are already evident for peripheral vision loss people. Instead, our goal is to identify and track moving objects in the user's peripheral area to generate (as early as possible) a visual notification if this object is a candidate hazard in future.

Real-time processing involves defining head motion type (static, moving or rotating) and then detect, track and classify the hazards around the person. Fig. 3 shows the main phases in this work from capturing real-time video to producing machine-learning based warnings in the person's healthy vision.
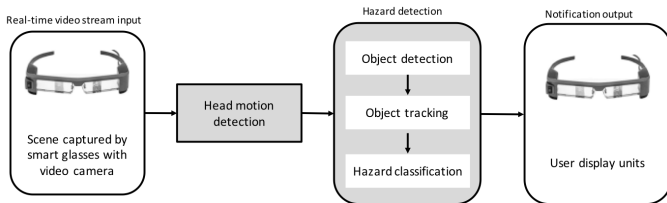

Fig. 3.   Block diagram for the proposed system

The first step is to extract frames and prepare them to be used in the head motion detection phase (HMD). This step is to define the type of head (camera) motion to (1) determine the best motion compensation technique before object detection and (2) reduce the number of false alarms due to sudden head movement. Since we have a wearable camera in this system, camera motion is often synonymous with head motion. This movement affects the whole processing phase directly from object detection to notification generation. More detail will be discussed in the following subsections.

In the object detection phase, all moving objects were detected to determine their location. Object features can't be defined directly using a single frame/image. Therefore, an object tracker is desired in this stage to build the features over time. The final phase is to decide the level of danger/risk based on the extracted features and predefined rules that will produce proper notification for each level and display them in the person's healthy visual field.

Finally, after getting the notification, its colour will vary based on the object's speed with three levels of danger:

1)   H: dangerous high level (red notification).
2)   M: dangerous medium level (orange notification).
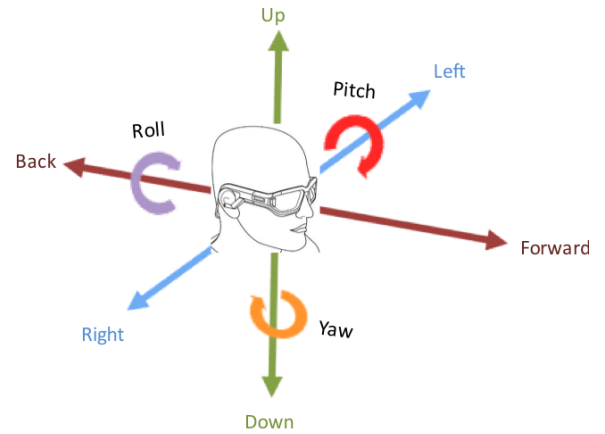3)   L: dangerous low level (green notification).

Fig. 4.   Degrees of freedom for wearable camera

A. Head Motion Detection and Optical Flow

The head motion detection phase is essential to decide if the camera is moving or not (stationary) and to detect the motion type. The output of this phase is needed to determine the best scenario for the object detection phase. In the case of a wearable camera, six degrees of freedom are expected based on head movements as shown in Fig. 4.

The head can move in a forward/backward, left/right and up/down translation. In terms of rotation, pitch motion represents the rotation around the x-axis, yaw rotation is a movement around the y-axis, and finally, a roll is a rotation around the z-axis. In this work, we will cover all translation motion types (left/right, up/down and forward/backwards). Pitch rotation is considered to be similar to the up/down type, while yaw rotation is deemed to be the same as the left/right motion. The mentioned motion types can be summarised as follows:

1)   Stationary camera (S): static background, moving objects.
2)   Translation/Rotation Right (TRR), Moving Translation/Rotation Left (TRL): background change in horizontal direction.
3)   Translation/Rotation Up (TRU), Moving Translation/Rotation Down (TRD): background change in vertical direction.
4)   Moving Forward (MF) or Moving Backward (MB): fast changes in the background and foreground.

In the case of a stationary camera, moving objects can be detected using traditional foreground segmentation methods. In the case of moving camera, motion compensation step is needed before background subtraction to distinguish between real and artificial movement. Finally, the forward/backwards moving camera case requires advanced motion estimation and compensation algorithms before the object detection phase which will be covered in our future work.

Optical flow methods are used to calculate motion vectors (velocity and direction) for some predefined key-points. The algorithm determines the head case every half a second to be used in the second half for the detection and tracking

processes. A Neural Networks classifier has been used for camera motion type classification using the calculated average velocity and direction. Each frame has been divided into nine subregions. The main aims of segmenting frames into nine subregions are to simplify the motion flow calculations, to reduce the effect of moving objects, and to provide a better representation for the camera motion using more key-points that are widely spanning all subregions. The NN model uses eighteen inputs (nine speed - direction pairs for the corresponding sub-regions) and six targets (static, left, right, up, down and forward). Several experiments were carried out to find the optimum NN configuration, and the six head motion cases were detected with 95% average accuracy.

B. Motion Detection using Stationary Camera

Object detection phase is where all critical objects are defined by their location to be tracked and classified later. This step needs the output from the previous phase (Head motion detection) to determine the best technique for moving object detection. Background subtraction method was used in the case of a stationary camera to model the static background and segment the foreground.

The Gaussian mixture-based background/foreground segmentation algorithm [34] was used to model the background and detect the moving objects. After applying the foreground mask on each input frame, moving objects were displayed as white blobs in the foreground image. Useful features (centre, size, location) were extracted after contouring the detected objects to be used in the tracking process. Fig. 5 shows the mentioned steps.
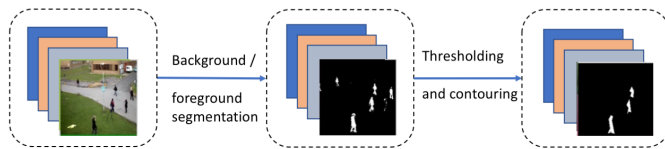

Fig. 5.   Foreground detection using Mixture of Gaussians Segmentation.

C. Motion Compensation for Moving Camera

In the moving camera rotation scenario, motion compensation step was performed before detecting the moving object. The motion caused by the camera was compensated using a homography matrix (H) that aligns the previous frame with the current one. The first step is to define key-points in the current frame (I_{t-1}) to track their corresponding location in the next frame (I_t). Shi and Tomasi corner detection algorithm [35] was used to find the most prominent points in each frame. Point quality measure is calculated at every source frame pixel using the cornerMinEigenVal. The corresponding location for the detected points was calculated using Lucas-Kanade optical flow in pyramids [36].

After defining the new location for each point in the frame (I_{t-1}), a perspective transformation between the two frames was calculated to determine the homography matrix (H). This matrix was used to compensate the camera motion by aligning the first frame to the second frame using the flowing equation:

$$\hat{I}_{t-1} = H I_{t-1} \qquad (1)$$

The result of (1) is shown in Fig. 6 (c). Black sides (right and top) represent the translation that occurred due to camera motion. The new images were almost identical and the frame subtraction method will detect moving object clearly as shown in Fig. 6 (e).

Fig. 6.   Moving object detection after motion compensation. (a) frame (I_{t-1}) (b) frame (I_t) (c) the warped frame using the homography matrix H calculated based on the optical flow from the two consecutive frames (d) the thresholding result for frame subtraction ($c - b$) (e) the final output where moving object with maximum area is detected, red arrows show the optical flow results for the detected points.

It is worth mentioning that multiple noise results were expected because of the accuracy of the homography matrix used for translation. This accuracy has a strong correlation with the number of the key points used to compute the optical flow which is a trade-off between accuracy and computation load. Additional threshold based on blob's area was applied to extract the significant objects only.

D. Motion Tracking and Classification

In this part, the goal is to track the detected objects and extract motion features. Moving objects in the first and second camera motion scenarios (stationary and rotation) were tracked. Since the system had recognised the moving objects in the previous phases, the approximate location for each object is known.

For each tracked object in each frame, the position, age (the appearance time in terms of the number of frames), current location, velocity (magnitude (V) and direction (θ), and the change of area have been defined.

1) Object tracking and feature extraction: For all objects detected in each frame, the new positions were compared with the old ones for both directions (x and y) to decide if the new object is a new one or an old object with different location. Consider the object $P_{i,t}$ where i is the object number or ID and t is the time or frame number. If you have the following objects in the first frame $P_{1,1}, P_{2,1}, P_{3,1}$, and $P_{4,1}$ and the objects $P_{1,2}, P_{2,2}, P_{3,2}, P_{4,2}$, and $P_{5,2}$ in the second frame, then to check the tracking possibility for object $P_{3,2}$, you compare its position over the horizontal dimension $P_{3,2}(x)$ and the vertical dimension $P_{3,2}(y)$ to that of all objects in the previous frame within assumed windows w_x and w_y, respectively. So, for any object $P_{b,t}$ in frame t to be a tracked version of the object $P_{b,t-1}$ in frame $t - 1$ you should have:

$$|P_{b,t}(x) - P_{b,t-1}(x)| < w_x \quad AND \quad |P_{b,t}(y) - P_{b,t-1}(y)| < w_y$$
$$(2)$$

Otherwise, the object will be considered as a new object and stored to be tracked in the following frames.

Since not every moving object is considered as a hazard, it is important to check the motion model of the moving object before tracking it. To test if the object is moving towards the centre of view (approaching) or away from the centre of view (receding), the average rate of change of the tracked object's area has been defined as:

$$\Delta A(P_{b,t}) = \frac{(A(P_{b,t}) - A(P_{b,t-1})) + \Delta A(P_{b,t-1})}{2} \quad (3)$$

where $\Delta A(P_{b,t})$ is the area of the object $P_{b,t}$, $A(P_{b,t-1})$ is the area of the same object in the previous frame and $\Delta A(P_{b,t-1})$ is the latest update for the object area difference compared to the last frame. When the object is detected for the first time, $\Delta A(P_{b,t})$ will be zero and then this value is sequentially updated in a cumulative manner.

Fig. 7 shows an example of a series of sequential frames selected from a public dataset [37]. The top table shows the extracted features for the tracked objects, while the bottom pictures show the tracking output.

In this example, a moving object has been seen from frame 90-94. For each tracked object, its age, location, speed, direction, and area are updated as long as it is detected from the previous phase. No tracking output generated before frame 91 because the age of the object is 1, meaning that this is the first time for the object to appear. It is important to mention that the object was moving very fast in this example. This explains the big bounding box around the detected object that refers to a significant difference between the consecutive frames.

To find the direction of movement for each object, the motion angle was calculated using the changes in the x and y axis. After this step, the direction of interest (DOI) has been defined based on the object's current location and object direction over time. Since not all moving objects have the same priority, only objects approaching the user had been considered. Fig. 8 shows the DOI in each quadrant. Red arrows represent the high priority direction, while orange arrows represent low priority direction.

2) Hazard classification rules: Our main aim of this work is to enhance the quality of life for people with peripheral vision loss. Therefore, it is necessary to classify the moving objects that were detected and tracked before displaying the notification for the user. For a moving object to be classified as a hazard, the following rules are applied:

1) The object should be in the user's visual field for sufficient time (Object's age >1).
2) The object should move at a significant speed (Object's speed > predefined threshold).
3) The object should move towards the user (Object has a DOI).
4) The object is approaching the user (Object's change of area >0).

| | | | Object change of area |
| Frame Number | Object ID | Object Age (frames) | Object current location | | + : object is approaching − : object is receding |

F#	ID	Age	Location		Speed	Direction	ΔA
90	0	2	[252	134]	846.936	-157.073	0.796773
91	0	3	[203	137]	1471.91	-177.085	1.27666
92	0	4	[142	150]	1859.64	-167.426	0.441651
93	0	5	[72	177]	2245.45	-159.269	0.508205

Frame 90 | Frame 91
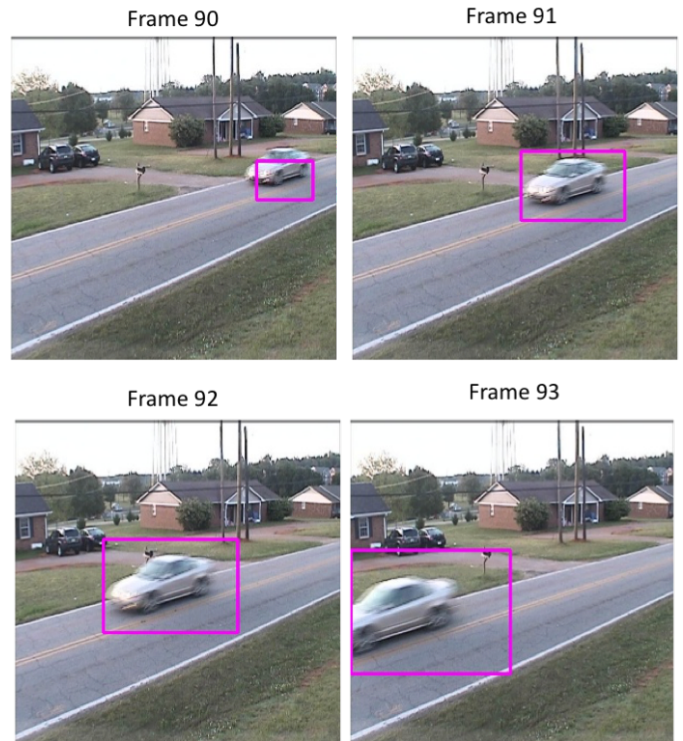
Frame 92 | Frame 93

Fig. 7. Tracking example. Top table is the tracking extracted features. bottom images show the tracking output

IV. EXPERIMENTAL RESULTS AND EVALUATION

A. Motion Compensation and Object Detection Evaluation

Since the purpose of this system is to detect moving objects for people with vision impairment using smart glasses, the performance of the proposed system should be tested on a moving camera video. To test the effectiveness of the motion compensation method, we applied it on a video [38] containing scenes from a continually moving camera that rotates horizontally and vertically on the side of a street. Different types of moving objects appeared in this video such as cars, pedestrians, bikes and others. A total of 3650 frames (30 frames/second) were used to evaluate moving object detection with and without motion compensation. Detection after post-processing (performing some morphological transformations to filter out small noises) was considered to optimise the detection process. Moving object detection with rotating camera using the motion compensation method has provided good results. Around 48% of the detected objects have been filtered out without affecting the detection accuracy.
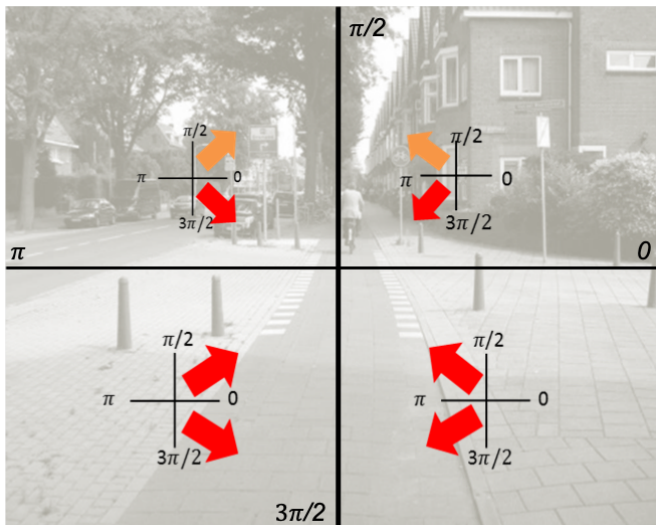
Fig. 8.   Direction of interest example

TABLE I.   PERFORMANCE COMPARISON FOR MOVING OBJECT
DETECTION ALGORITHMS

Algorithm name	Recall	Specificity	FPR	FNR	F score
St-Charles et al. [39]	0.698	0.991	0.009	0.302	0.462
Maddalena et al. [40]	0.856	0.680	0.320	0.144	0.037
Allebosch et al. [41]	0.918	0.922	0.078	0.082	0.584
Sajid et al. [42]	0.577	0.995	0.006	0.423	0.512
Chen et al. [43]	0.797	0.979	0.021	0.203	0.386
Charle et al. [44]	0.831	0.963	0.037	0.169	0.348
Gregorio et al. [45]	0.336	0.998	0.002	0.664	0.322
Varadarajan et al. [46]	0.641	0.928	0.072	0.359	0.247
Kurnianggoro et al. [47]	0.713	0.983	0.017	0.287	0.329
Our work	**0.928**	**0.978**	**0.022**	**0.072**	**0.629**

Public available dataset from changeDetection1 [37] was used to evaluate object detection after motion compensation. For this purpose, the sequence (continuousPan) was used under the category PTZ. This sequence was chosen because it contains scenes from a continuously moving camera. The camera is panning horizontally at slow speed. Moving objects (such as cars and trucks) were seen moving fast. The sequence contains 1700 frame (480 x 704) and a detection rate of 93% has been achieved. Performance comparison also provided in Table I. The used performance metrics are Recall, Specificity, False positive rate, False negative rate, F-score, and Precision. The results show that this method is very competitive and highly sensitive. The rate of relevant detection overall detection is the best compared with other algorithms. It is important to mention that in this project, the accuracy of the detection location is not very sensitive. It is important to detect an approximate location which is as close as possible to the real moving object. This explains the high recall rate for this test comparing to other work.

B. Motion Tracking and Classification Evaluation

Initial evaluation experiments were carried out to test the motion tracking method using a moving camera. The same dataset was used in previous phases [37] to check the performance of the motion tracking and hazard classification. The video contains 1700 frames (704 x 480) taken by a rotating camera to the side of a road. The speed of moving objects was significantly high.

A total of 204 moving objects were detected. The tracking method tracked 162 objects correctly with a tracking accuracy of 79%.
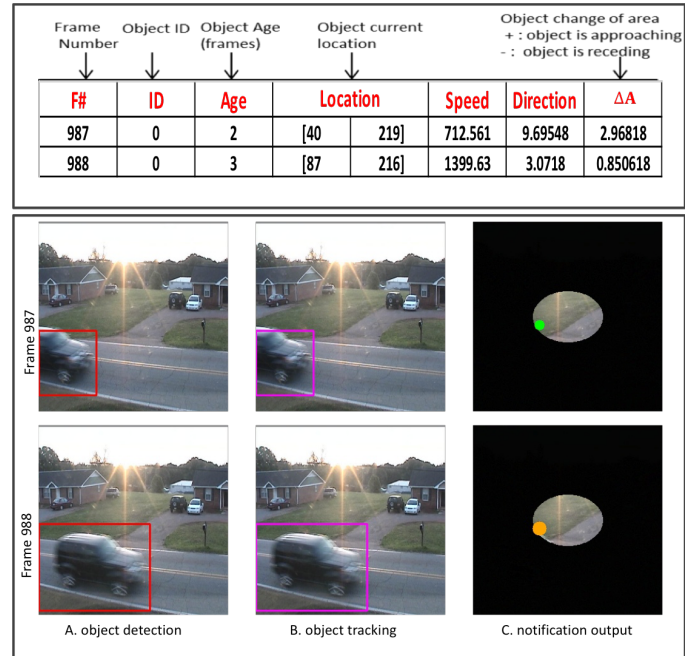

Fig. 9.   Notification output example based on a predefined hazard classification rules

In Fig. 9, an example of a notification output generated based on the predefined rules mentioned in sub-section 2.4. The left images show the motion detection output (red rectangle) with a car moving towards the centre. The purple rectangles in the middle images refer to the tracking outputs. Finally, the right images show an example of a tunnel vision visual field (when a person loses vision in the peripheral visual fields while retaining vision in the central regions only).

Using a 300 x 300 frame size, the first output appeared as a green circle in the bottom left part of the oval shows that there is a hazard to the bottom left location of the person's visual field. In the following frame (988), the age and speed of the danger increased. Thus, the size and colour of the output were updated to reflect these changes. The top tables show the extracted features for the tracked object.

V.   CONCLUSION

In this work, a novel, wearable hazard warning system to help people with peripheral vision loss who are unable to see using their peripheral vision is presented. The proposed system implements real-time computer vision techniques to detect, track and classify moving objects in the peripheral area with different scenarios for different camera motion states. Head motion detection was used to decide if the camera is stationary or moving (forward, rotation up, down, right, or left). The output from this step was used to select the suitable motion compensation method for the moving objects detection phase.

Moving hazard detection with rotating camera using the motion compensation method has provided good results. Motion compensation is a necessary step for the moving camera

scenario to distinguish between real and artificial motion caused by the camera movement. This difference was used to track real moving objects and reduce false detection due to camera motion. Moving object detection rate of 93% has been achieved.

The detected moving objects were tracked and their motion features were extracted. Tracking accuracy of 79% was obtained. The extracted features are object age, location, speed, direction, and area change rate. To minimise the number of notifications displayed in the user's visual field, the extracted features were used to classify the objects based on predefined rules and then, notifications are displayed based on the classification result. The work is tested on smart glasses (Epson Moverio BT-200). The initial experiments showed relatively slow performance, but we are in the process of testing our system on the latest smart glasses available in the market.

In this work, we choose to use smart glasses because we believe that including the video capturing unit, the processing unit and the display unit in one wearable platform will help the user to navigate easily. Furthermore, because people with peripheral vision loss retain healthy vision in their central visual field, it is essential to keep the existing visual case and add to it the needed information. This work will be developed further in our future work to provide a wide range of warnings and notifications for visually impaired people using more extracted features and machine-learning classification methods.

REFERENCES

[1] J. M. J. Roodhooft, "Leading causes of blindness worldwide," *Bull Soc Belge Ophtalmol*, vol. 283, pp. 19–25, 2002.

[2] R. R. A. Bourne, J. B. Jonas, S. R. Flaxman, J. Keeffe, J. Leasher, K. Naidoo, M. B. Parodi, K. Pesudovs, H. Price, R. A. White, T. Y. Wong, S. Resnikoff, and H. R. Taylor, "Prevalence and causes of vision loss in high-income countries and in eastern and central europe: 1990–2010," *British Journal of Ophthalmology*, vol. 98, no. 5, pp. 629–638, 2014.

[3] H. Strasburger, I. Rentschler, and M. Juttner, "Peripheral vision and pattern recognition: A review," *Journal of Vision*, vol. 11, no. 5, pp. 13–13, 2011.

[4] M. Hersh and M. A. Johnson, Eds., *Assistive Technology for Visually Impaired and Blind Peoplel*, 1st ed. Springer-Verlag London, 2008.

[5] M. Ervasti, M. Isomursu, and I. I. Leibar, "Touch-and audio-based medication management service concept for vision impaired older people," in *RFID-Technologies and Applications (RFID-TA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 244–251.

[6] B. Nayak and S. Dharwadkar, "Interpretation of autoperimetry," *Journal of Clinical Ophthalmology and Research*, vol. 2, no. 1, pp. 31–59, 2014.

[7] B. Woodrow and C. Thomas, "Fundamentals of wearable computers and augmented reality," *Lawrence Erlbaum Associates, Inc*, pp. 27–31, 2000.

[8] R. A. Earnshaw, *Virtual reality systems*. Academic press, 2014.

[9] V. Oculus *et al.*, "Oculus rift," *Available from WWW:¡ http://www. oculusvr. com/rift*, 2015.

[10] L. Prasuethsut, "Htc vive: Everything you need to know about the steamvr headset," *Retrieved January*, vol. 3, p. 2017, 2016.

[11] W. Barfield, *Fundamentals of wearable computers and augmented reality*. CRC Press, 2015.

[12] S. K. Ong and A. Y. C. Nee, *Virtual and augmented reality applications in manufacturing*. Springer Science & Business Media, 2013.

[13] A. Al-Ataby, O. Younis, W. Al-Nuaimy, M. Al-Taee, Z. Sharaf, and B. Al-Bander, "Visual augmentation glasses for people with impaired vision," in *Developments in eSystems Engineering (DeSE), 2016 9th International Conference on*. IEEE, 2016, pp. 24–28.

[14] F. Vargas-Martín and E. Peli, "Augmented view for tunnel vision: Device testing by patients in real environments," in *SID Symposium Digest of Technical Papers*, vol. 32, no. 1. Wiley Online Library, 2001, pp. 602–605.

[15] V. Pradeep, G. Medioni, and J. Weiland, "Robot vision for the visually impaired," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 2010, pp. 15–22.

[16] S. L. Hicks, I. Wilson, L. Muhammed, J. Worsfold, S. M. Downes, and C. Kennard, "A depth-based head-mounted visual display to aid navigation in partially sighted individuals," *PLoS ONE*, vol. 8, no. 7, pp. 1–8, 2013.

[17] R. Manduchi, J. Coughlan, and V. Ivanchenko, "Search strategies of visually impaired persons using a camera phone wayfinding system," *Computers Helping People with Special Needs*, pp. 1135–1140, 2008.

[18] Y. Tian, X. Yang, C. Yi, and A. Arditi, "Toward a computer vision-based wayfinding aid for blind persons to access unfamiliar indoor environments," *Machine Vision and Applications*, vol. 24, no. 3, pp. 521–535, 2013.

[19] V. Ivanchenko, J. Coughlan, and H. Shen, "Crosswatch: a camera phone system for orienting visually impaired pedestrians at traffic intersections," in *International Conference on Computers for Handicapped Persons*. Springer, 2008, pp. 1122–1128.

[20] A. N. Netravali and J. D. Robbins, "Motion-compensated television coding: Part i," *The Bell System Technical Journal*, vol. 58, no. 3, pp. 631–670, March 1979.

[21] A. v. Brandt, *Object Tracking and Background Estimation with a Moving Camera*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, pp. 186–191.

[22] C.-F. Chen and M.-H. Chen, *Target Tracking and Positioning on Video Sequence from a Moving Video Camera*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, vol. 4319, pp. 523–533.

[23] A. Vavilin, L.-M. Ha, and K.-H. Jo, *Camera Motion Estimation and Moving Object Detection Based on Local Feature Tracking*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, vol. 7345, pp. 544–552.

[24] P. Mazurek and K. Okarma, *Background Suppression for Video Vehicle Tracking Systems with Moving Cameras Using Camera Motion Estimation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, vol. 329, pp. 372–379.

[25] L. Kurnianggoro, Wahyono, Y. Yu, D. C. Hernandez, and K.-H. Jo, "Online background-subtraction with motion compensation for freely moving camera," in *Intelligent Computing Theories and Application: 12th International Conference, ICIC 2016, Lanzhou, China, August 2-5, 2016, Proceedings, Part II*, D.-S. Huang and K.-H. Jo, Eds. Cham: Springer International Publishing, 2016, pp. 569–578.

[26] D. K. Panda and S. Meher, "Detection of moving objects using fuzzy color difference histogram based background subtraction," *IEEE Signal Processing Letters*, vol. 23, no. 1, pp. 45–49, Jan 2016.

[27] D. Tripathy and K. G. R. Reddy, "Adaptive threshold background subtraction for detecting moving object on conveyor belt," *International Journal of Indestructible Mathematics and Computing*, vol. 1, no. 1, pp. 41–46, 2017.

[28] P. A. Pojage and A. A. Gurjar, "Review on automatic fast moving object detection in video of surveillance system," *International Journal of Scientific Research in Science and Technology(IJSRST)*, vol. 3, no. 3, pp. 545–549, 2017.

[29] B. Karasulu and S. Korukoglu, *Moving Object Detection and Tracking in Videos*. New York, NY: Springer New York, 2013, pp. 7–30.

[30] R. Tapu, B. Mocanu, and T. Zaharia, "Deep-see: Joint object detection, tracking and recognition with application to visually impaired navigational assistance," *Sensors*, vol. 17, no. 11, p. 2473, 2017.

[31] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[32] O. Younis, W. Al-Nuaimy, M. A. Al-Taee, and A. Al-Ataby, "Augmented and virtual reality approaches to help with peripheral vision loss," in *2017 14th International Multi-Conference on Systems, Signals & Devices (SSD)*. IEEE, 2017, pp. 303–307.

[33] O. Younis, W. Al-Nuaimy, F. Rowe, and M. H. Alomari, "Real-time detection of wearable camera motion using optical flow," in *2018 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2018, pp. 1–6.

[34] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 2, Aug 2004, pp. 28–31 Vol.2.

[35] J. Shi and C. Tomasi, "Good features to track," in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Jun 1994, pp. 593–600.

[36] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, ser. IJCAI'81. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1981, pp. 674–679.

[37] N. Goyette, P. M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "Changedetection.net: A new change detection benchmark dataset," in *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2012, pp. 1–8.

[38] iLuvTech. (2016, 3) 4k street view, hongdae, korea. [Online]. Available: https://youtu.be/qA2W4hLh6Gc

[39] P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin, "A self-adjusting approach to change detection based on background word consensus," in *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on*. IEEE, 2015, pp. 990–997.

[40] L. Maddalena and A. Petrosino, "A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection," *Neural Computing and Applications*, vol. 19, no. 2, pp. pp.179–186, 2010.

[41] G. Allebosch, F. Deboeverie, P. Veelaert, and W. Philips, "Efic: edge based foreground background segmentation and interior classification for dynamic camera viewpoints," in *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 2015, pp. 130–141.

[42] H. Sajid and S.-C. S. Cheung, "Background subtraction for static & moving camera," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 4530–4534.

[43] Y. Chen, J. Wang, and H. Lu, "Learning sharable models for robust background subtraction," in *Multimedia and Expo (ICME), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1–6.

[44] P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin, "Subsense: A universal change detection method with local adaptive sensitivity," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. pp.359–373, 2015.

[45] M. De Gregorio and M. Giordano, "Change detection with weightless neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 403–407.

[46] S. Varadarajan, P. Miller, and H. Zhou, "Spatial mixture of gaussians for dynamic background modelling," in *Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on*. IEEE, 2013, pp. 63–68.

[47] L. Kurnianggoro, Y. Yu, D. C. Hernandez, K.-H. Jo *et al.*, "Online background-subtraction with motion compensation for freely moving camera," in *International Conference on Intelligent Computing*. Springer, 2016, pp. 569–578.