FIR Pedestrian Detection and Classification

Functional and Performance Specification for Augmented Display

Document History:

Version	Description	Change Owner	Date
0.1	Initial draft version	Mark Gehrke/Claudia Zaragoza	09/11/2020
0.2	Added multi-class requirement; revised maximum range for detection/ID to be 70m per discussion with DAT; edited pedestrian definition per AR feature discussion.	Mark Gehrke	9/29/2020
1.0	Removed comments and editing notes; prepared for initial release to AR ECU vendors	Mark Gehrke	10/14/2020

1 Pedestrian Detection with Far InfraRed (FIR) video

1.1 Document Description

This document details the pedestrian detection functional engineering requirements and specification of the software service which will detect, classify and track pedestrians in FIR camera video. The output generated by this detection service will be utilized by an Augmented Reality feature to highlight the presence of pedestrians as shown in the notional example below.

Figure 1: Example FIR image scene showing notional AR Display of detected pedestrians.

1.2 ISO vehicle coordinate definition

This specification follows the ISO standard vehicle coordinates as shown in Figure 2. The origin is defined as the center of the rear axle.

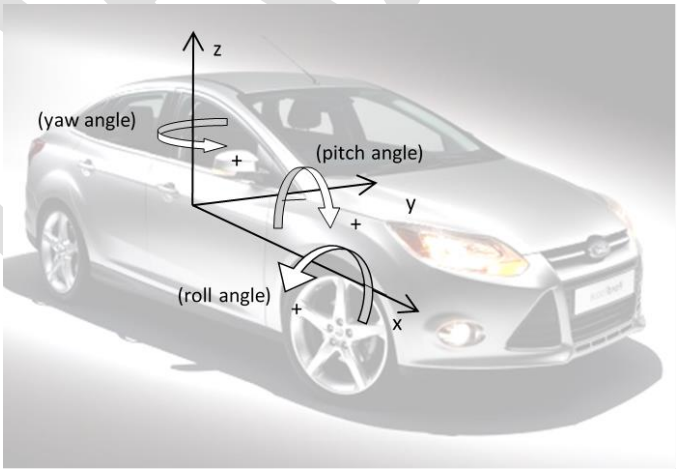

Figure 2: ISO vehicle coordinate definition

2 System Definitions

2.1 FIR_AR_PD_Def1 Host Assumptions

1. Host vehicle is assumed to be moving in the forward direction at a speed between 12kph and 100 kph. [TBR – feature input needed].

2.2 FIR_AR_PD_Def2 Ground reference plane

1. Horizontal plane, either real or imaginary, that passes through all tire contact points of a vehicle while the vehicle is in its normal ride attitude (ISO 14513:2016)

2.3 FIR_AR_PD_Def3 Definition of Pedestrian

2. Human beings located anywhere on the ground plane shall be considered pedestrian targets for the pedestrian detection algorithm to detect and classify.
3. Pedestrians include adults and children standing or prone.
4. A prone target includes pedestrians which are in any non-erect position including laying or sitting.
5. Human beings on non-powered bicycles located anywhere on the ground plane shall be considered pedestrians targets for the pedestrian detection algorithm to detect and classify.
6. Human beings using other unpowered wheeled transportation – e.g. skateboards, kick scooters, strollers – or individuals using mobility assistive equipment – walkers, canes, crutches, manual and power wheelchairs and power operated vehicles ("powered scooters") – located anywhere on the ground plane shall be considered pedestrians for the pedestrian detection algorithm to detect and classify.
7. The operators of motorcycles and mopeds located anywhere on the ground plane shall not be considered pedestrians [TBR – Feature input needed]
8. Pedestrians can be moving, with speed >0 – 14kph by walking or running, with speed > 0 - 35 kph while using non-powered wheeled transportation or stationary (0kph).
9. Pedestrians will be classified as moving or stationary based on target velocity. [TBR – feature input needed].
10. The minimum height for the detection of pedestrians is defined to be the height of a 10th percentile, standing female:
    a. A 10th percentile, adult, female is 155cm per U.S. CDC growth charts https://www.cdc.gov/growthcharts/data/set1clinical/cj41c022.pdf
11. Partially occluded pedestrians where at least two appendages visible to the camera are to be considered pedestrian targets. (see FIR_AR_PD_Req00)

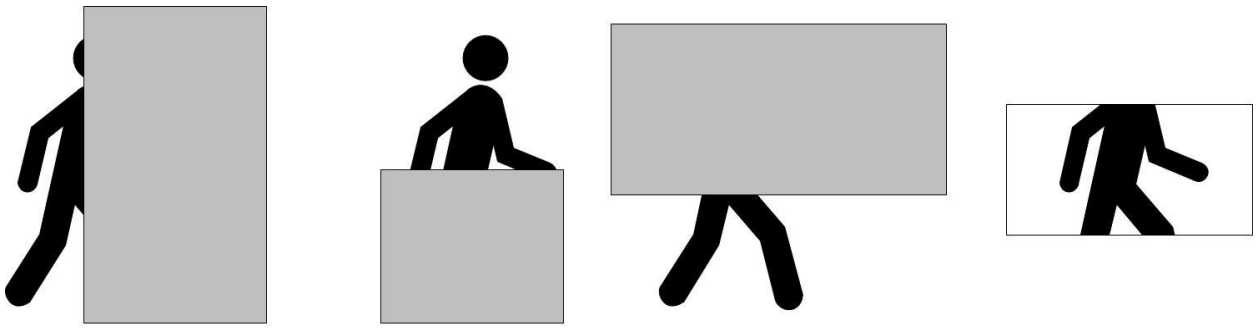

Figure 3: Examples of occluded pedestrians.

12. When any part of the pedestrian is within the maximum effective range of the FIR camera it will be considered a target if the above conditions are also met. For further definition of camera effective range (field of view) see FIR_AR_PD_Def4 **.**

2.4 FIR_AR_PD_Def4 Definition of Detection Zones

1. The Detection Zone will consist of a minimum 60-degree field of view about the centerline of the host vehicle whose vertex is located on the host centerline at an x-position value matching the entrance pupil of the camera. [TBR – feature input needed]
2. The Detection Zone will consist of four sub-zones as described in the table and figure below.

Table 1: Sub-zone range description

Sub-zone	Radial Range (meters)
Near	2* < range ≤ 25
Mid	25 < range ≤ 50
Far	50 < range ≤ 70
Outer	70 < range

*Only applies for newly detected objects. Objects should be tracked to om from the vehicle.

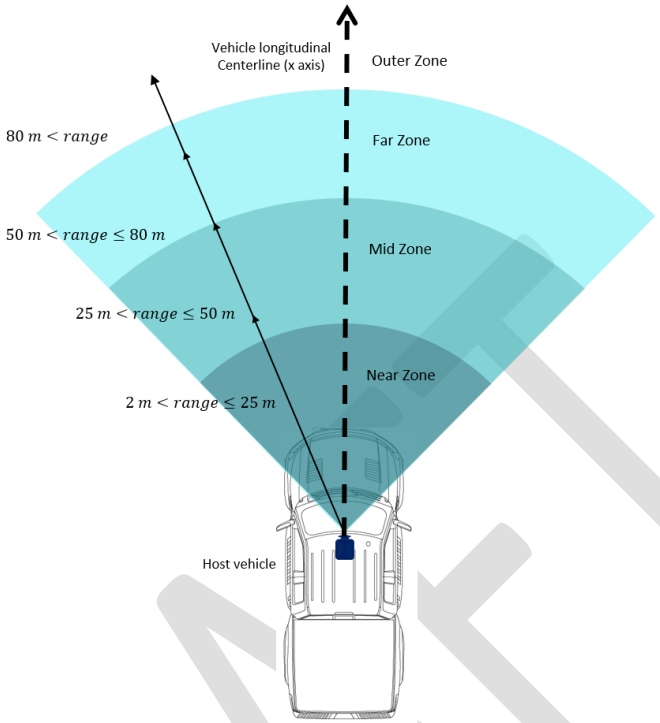

Figure 4: Representation of the Effective Camera Zones.

2.5 FIR_AR_PD_Def5 Definition of Priority Zones

1. The sensor field of view shall be divided into 6 regions of interest which occupy the upper 75% of the total FOV.
2. Each ROI designates an area of specific priority. The priority depends on
    a. how close the target is to the FIR camera
    b. whether the target is in the line of travel of the vehicle
    c. if the current direction and speed of the person will cause the pedestrian to intersect the line of travel before the vehicle passes thereby creating a collision hazard

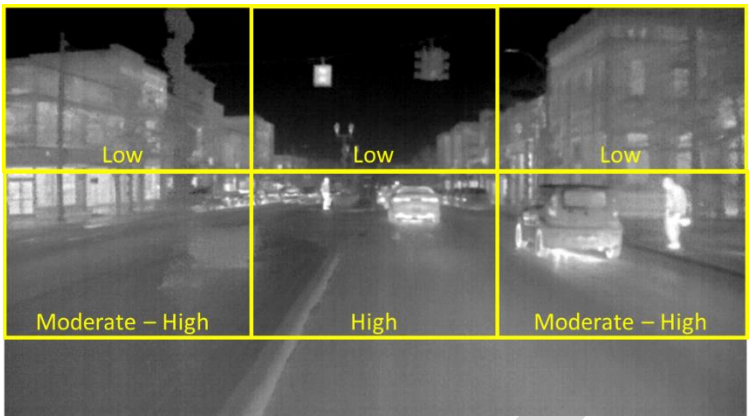

Figure 5: Representation of the Priority Zones

2.6 FIR_AR_PD_Def6 Definition of True Positives

1. For verification purposes, a true positive shall be considered when the output of the pedestrian detection algorithm meets both of the following conditions
    a. detects the object with sufficient positional accuracy such that the Intersection over Union (IOU) for the detection bounding box and truth bounding box is ≥ 0.25 for all detection zones and pedestrian types
    b. identifies an object as a pedestrian if it meets the all criteria in **FIR_AR_PD_Def3** which defines a pedestrian.

2.7 FIR_AR_PD_Def7 Definition of False Positives

1. For verification purposes, a false positive shall be considered when the output of the pedestrian detection algorithm meets either of the following conditions
    a. detects the object such that the Intersection over Union (IOU) for the detection bounding box and truth bounding box is < 0.25 for all detection zones and pedestrian types
    b. identifies an object as a pedestrian but the object does not meet all criteria in **FIR_AR_PD_Def3** which defines a pedestrian.

2.8 FIR_AR_PD_Def8 Definition of Environmental Occlusion

1. Environmental occlusion of the camera is the presence of foreign material on one or more optical elements and which degrades the image quality. Camera occlusion can result from several factors including dirt, insects, film, water and ice contamination on the window or lens In general, occlusion results in some degree of attenuation of light transmittance through the lens to the imager with the consequence being lower contrast ratios and/or lower MTF. In extreme cases the occlusion may completely block light from reaching the microbolometer imager, resulting in images with only noise content.

2.9 FIR_AR_PD_Def9 Definition of Day and Night Scenarios

1. Nominal day time environmental conditions for FIR Pedestrian Detection operation shall be defined as
    a. Minimum thermal contrast of ΔT=3°C between objects (pedestrians) and background clutter
    b. 5km visibility[1]
    c. Atmospheric conditions approximating MODTRAN mid-latitude summer[2]
2. Nominal night time environmental conditions for FIR Pedestrian Detection operation shall be defined as
    a. Minimum thermal contrast of ΔT=5°C between objects (pedestrians) and background clutter
    b. 5km visibility
    c. Atmospheric conditions approximating MODTRAN mid-latitude winter

3 Requirements

3.1 Detection of multiple classes

FIR_AR_PD_Req001	Detection and Identification of Multiple Classes
Purpose:	Allow for the detection and identification of multiple classes of objects/people. Initially only pedestrians will be considered for detection and classification, but future features may require additional classes.
Verification:	Demonstration

Requirement: The pedestrian detection service must include the capability to detect and classify multiple classes of objects, up to a maximum of 10 (TBR)

3.2 Detection of Upright Pedestrian Targets

FIR_AR_PD_Req002	Detection and Identification of Upright Pedestrian Targets
Purpose:	Detect pedestrian targets in an upright position in the FIR camera field of view.
Verification:	Test and analysis.

Requirement:

a) The pedestrian detection and identification service must detect and classify an upright pedestrian, fully visible or occluded, as defined in **FIR_AR_PD_Def3 ,** within 120 ms (TBD) of entering the Detection Zone (**FIR_AR_PD_Def4**) with positive detection rates at each range as given in **Error! Reference source not found.**.

b) The pedestrian detection and identification service must be able to satisfy the criteria in a) above for an arbitrarily selected data set which has no common images with the training data set, has at least 250 pedestrian instances meeting **FIR_AR_PD_Def3** and includes multiple instances of at least two different modes of non-motorized wheeled transport.

c) The detection and classification will be evidenced by the output of the development data (refer **FIR_AR_PD_Req010**) and comparison with annotated data to determine true positive, false positive, and false negative detections.

[1] As defined by ICAO Annex 3 *Meteorological Service for International Air Navigation*

[2] http://modtran.spectral.com/static/modtran6/html/help_atmosphere_model.html?v=3

Table 2 Pedestrian positive detection rate on high priority zone on nominal daytime and nominal nighttime

	Positive detection rate of Pedestrians All values TBR	
Range	**Nominal Day**	**Nominal Night**
0-25m	90%	95%
25-50m	85%	90%
50-70m	60%	75%

3.3 False Positive Rate

FIR_AR_PD_Req003	**False Positive Rate**
Purpose:	Designate acceptable number of false alarms per image from the pedestrian detector
Verification:	Test and analysis

a) The rate of false positives shall not exceed 10^{-3} (TBD) per image
b) The false positive rate as given above is for the full pedestrian detection and identification service and shall include all pre and post detector/classifier false positive mitigation strategies.

3.4 Detection of Pedestrian in a non-powered wheeled transportation

FIR_AR_PD_Req04	**Detection of Pedestrian in a non-powered wheeled transportation**
Purpose:	Detect pedestrian targets using non-powered wheeled transportation within the FIR camera field of view.
Verification:	Demonstration

a) The pedestrian detection output data shall be focused on detecting and identifying human beings, regardless of his/her transportation tool, including but not limited to those given in **FIR_AR_PD_Def3**.

3.5 Multiple Detections

FIR_AR_PD_Req005	**Multiple Detections**
Purpose:	Limit detector/classifier output to single bounding box
Verification:	Demonstration

a) The pedestrian detector/classifier shall suppress multiple detections of the same object
b) Only a single bounding box shall be returned in the LROS output data structure for each detected object.
[TBR – as written this requirement is only applicable to individuals or "well-spaced" groups. The AR feature will need to define how to handle dense groups where individuals can no longer be reliably resolved]

3.6 Primary object selection

FIR_AR_PD_Req006	**Primary object selection**
Purpose:	Selection of the object with the highest threat
Verification:	Demonstration

a) The pedestrian detection service shall select the target with the highest threat in the current image frame following the priority list below (refer to **FIR_AR_PD_Def5**)
    1. Center High Priority
    2. Left or Right High Priority
    3. Left or Right Moderate Priority
    4. Any Low Priority
b) In scenarios where several pedestrians are within the same priority zone, the target with the highest threat would be the nearest to the vehicle host and/or the one moving towards the vehicle.
c) The CIPO flag shall only be set to "TRUE" for one object at each cycle.

3.7 Pedestrian Positional Accuracy

FIR_AR_PD_Req007	**Pedestrian Positional Accuracy**
Purpose:	Pedestrian position accuracy
Verification:	Test and analysis

a) The positional accuracy of the detected pedestrian shall be ≤ ± 5% true range (TBR) in the ground plane. The accuracy is determined by an average of all image plane points as shown in the Figure 6: Object Position Accuracy Measurement Points..
b) The positional accuracy shall be reported for all points.
c) The positional accuracy shall be on a calculated once the detection has occurred and averaged over 3 frames for the duration of the object track.

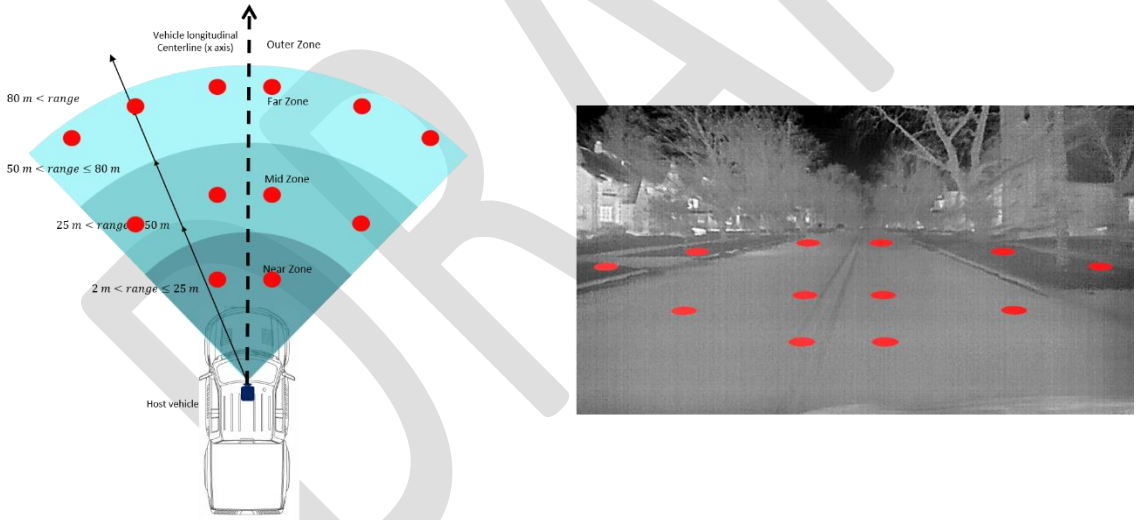

Figure 6: Object Position Accuracy Measurement Points. TBR – update with image using production FOV

3.8 Pedestrian Tracking Accuracy

FIR_AR_PD_Req008	**Pedestrian Tracking Accuracy**
Purpose:	Required parameters for tracking occlusion pedestrians
Verification:	Test and analysis

a) The pedestrian detection and identification software service shall track each object identified as a pedestrian such that the bounding box at the tracked location maintains 0.25 IOU with the manually annotated ground truth bounding box on a frame-by-frame basis.

3.9 Tracking of Occluded Pedestrians

FIR_AR_PD_Req009	Tracking of Occluded Pedestrian Targets
Purpose:	Tracking of pedestrian targets transitioning from visible to occluded
Verification:	Test and analysis

a) The output shall include a flag indicating that the target has become occluded and a timer indicating the amount of time from initial occlusion
b) Targets identified as pedestrians must be tracked into and through occlusion for a minimum of 1000ms [TBD – feature input needed].
c) If the target emerges from occlusion within the minimum dead-reckoning time, the object detection camera should maintain the unique target id assigned to the pedestrian before encountering the occlusion.
d) A target which has become occluded and is subject to dead-reckoning will be identified as such in the LROS output specific to this object detailed in requirement.
e) Tracking shall be maintained if a pedestrian is initially detected and then become obscured by the hood of the host vehicle.
f) If the pedestrian is tracked via a method other than dead-reckoning this info will be recorded in the LROS output specific to this object

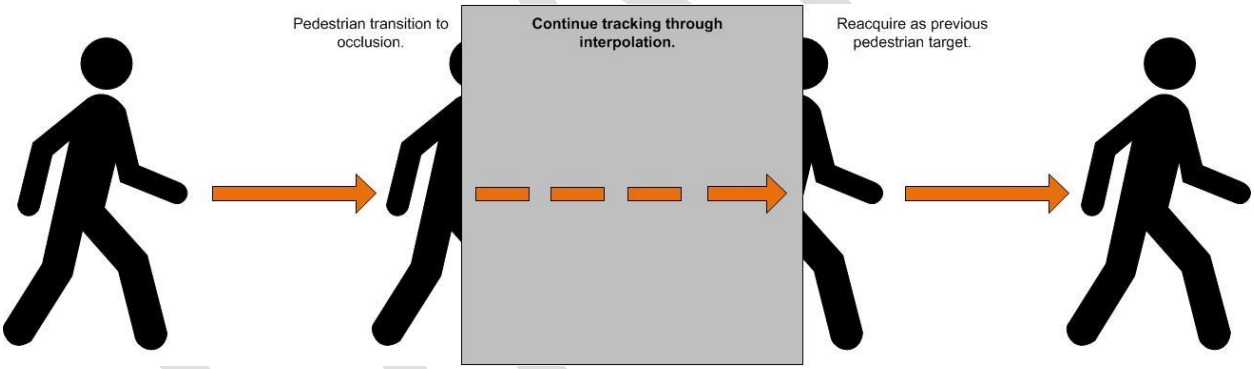

Figure 7: Tracking of Pedestrian Target Through Occlusion.

3.10 Frame update rate

FIR_AR_PD_Req010	Frame update rate
Purpose:	Verify system speed
Verification:	

a) The detection and identification service shall update the LROS data structure with new state vectors at a minimum of 30 Hz. The updated state vectors may be the output of the detector/classifier, tracker, or both.
b) The tracker shall operate with latency ≤33ms, including updating LROS structure, issue of CAN messages and/or embedded image metadata

c) The detection/classification service shall operate with latency ≤66ms, including updating LROS structure, issue of CAN messages and/or embedded image.
d) The nominal and worst-case frame processing cycle time shall be measured and documented in the release notes.

FIR_AR_PD_Req010	**Object Output Parameter Criteria**
Reference:	FNV2 DAT2 LROS CS ver 011 Release (Stage I) - 20190506
Purpose:	Object Output Parameters and Format
Verification:	

a) The output interface will be an LROS compliant format. The pedestrian detection/identification service shall make all detection information available in this format.
b) The output objects may be included in a fusion map that may or may not be owned by the FIR_AR_PD algorithm provider.
c) The output data shall contain but not limited to the following LROS entries.

Signal Name	Description
< object target id>	Unique target id
<CIPO>	Flag that indicates when the target is priority
< object classification>	Target classification (pedestrian)
< object x position>	x coordinate at the closest point on object projected on ground plane
< object y position>	y coordinate at the closest point on object projected on ground plane
< object width>	Target width
< object height>	Target height

d) Additional metadata not included in the LROS format may be published via CAN network.

4 Appendix A: Requirements for support functions needed by pedestrian detection/classification and other AR services and running in parallel with those services.

4.1 Data Preprocessing

FIR_AR_Support_Req001	Data Preprocessing
References:	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Apply image processing to camera data to enhance detection/classification performance or visual appearance of the data

a) The AR ECU shall apply basic image processing to camera data to improve performance of computer vision algorithms or the visual appeal of the images.
b) The operations will be prescribed by the algorithm/feature owners and may include but are not limited to
   i. MTF correction
   ii. Noise reduction
   iii. Contrast enhancement
   iv. Spatial filtering
   v. 16bit to 8bit data type conversion

4.2 The Vehicle shall issue a Camera occlusion warning to the customer when the AR ECU detects a Occlusion Detection Warning

FIR_AR_Support_Req002	Occlusion Detection Warning
References:	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Detect conditions when driver action (e.g. cleaning windshield) could be taken to restore detection performance. While this condition is true, the customer will be warned of limited visibility and the object detection functionality shall be restricted.

a) The Vehicle shall issue a Camera occlusion warning to the customer when the AR ECU detects a camera obstruction which significantly degrades target detection performance. Conditions include:
   - Full Solid Occlusion –A full solid occlusion condition is detected
   - Partial Solid Occlusion - Detection of dark spots in the image that do not change intensity for a long duration while vehicle speed exceeds a minimum speed threshold.
   - Blurred Image - Detection of low image contrast or blurriness in the image for a long duration while vehicle speed exceeds a minimum speed threshold and the wipers are off and a low sun condition is not detected and near range objects/lines/signs are not detected

b) A strategy shall be implemented to minimize the occurrence of false occlusion detections. This strategy shall be agreed upon by the supplier(s) and Ford.
c) A strategy shall be implemented to ensure that the occlusion detection fault does not cycle between TRUE and FALSE. This strategy shall be agreed upon by the supplier(s) and Ford.

4.3 Camera Occlusion Detection

FIR_AR_Support_Req003	Camera Occlusion Detection
References:	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Detection of camera occlusion

a) For purposes of the object detection system, occlusion shall be defined as the point at which motion in the scene can no longer be determined due to occlusion of the field of view.
b) Motion shall be defined as vehicle speed is ≥ 1 kph for 1 second
c) High Contrast transitions, ≥ 2:1 ratio of adjacent pixels, shall be used to define features in the scene

4.4 Camera Occlusion Cleaning

FIR_AR_Support_Req004	Camera Occlusion Cleaning
References:	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Clearing the occlusion detection flag via removal of occluding media

a) The occlusion detection signal, FIR_Blocked_Stat, shall be set to FALSE when the motion in the scene can again be detected.

4.5 Camera Occlusion Cleaning

FIR_AR_Support_Req005	Camera Occlusion Detection Execution Rate
References:	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Define occlusion detection execution rate

a) Occlusion detection shall be run at a rate no less than 3Hz

4.6 Camera Occlusion Detection Operation

FIR_AR_Support_Req006	Camera Occlusion Detection Operation
References:	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Define occlusion detection operating parameters

a) Occlusion detection shall operate in all gear positions

4.7 Camera Occlusion Reporting

FIR_AR_Support_Req007	Camera Occlusion Detection Reporting
References:	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Adjust performance requirements for occluded environmental conditions

a) When experiencing occluded environmental conditions as defined in **FIR_AR_Support_Req001** the object detection/classification service must identify and report via the LROS output specific to this object that an occluded condition exists.

4.8 Detection of FIR Camera Misalignment

FIR_AR_Support_Req008	Detection of FIR Camera Misalignment
References	CGEA 1.3 CADS3.5 Object Sensing Functional Specification ES-LB5T-19H423-BAC_Near_Range_Camera_OD_and_Classification FNV2 DAT2 LROS CS ver 011 Release (Stage I)
Purpose:	Adjust performance requirements for occluded environmental conditions

a) The AR ECU shall have the ability to detect and compensate for camera misalignment. If the misalignment is larger than what can be compensated for the customer shall be warned of a sensor malfunction.

b) Misalignment conditions which shall be detected include the following:
   - In-Service Camera Misalignment within FOV: Camera is mechanically within acceptable FOV. This condition shall be detected and compensated for. No DTC shall be stored.
   - In-Service Camera Misalignment outside FOV: Camera is mechanically outside of the acceptable FOV. This condition shall be detected and shall result in a stored DTC.
   - The camera is in an unaligned state (i.e. has not performed factory or service alignment). This condition shall result in a stored DTC.
   - Camera service or factory alignment is in-progress. This condition shall result in a stored DTC.

c) If any of the above misalignment conditions are detected, an appropriate camera warning message shall be displayed. Specifically:
   - If autocorrection of a Camera Misalignment causes functionality to be temporarily disabled (e.g. active SPC disables object and lane detection), then the customer shall be warned that the camera is temporarily unavailable (The LROS ES signal, LROSStatus.VisionAlignmentIncomplete, shall be set to TRUE).
   - If a Camera Misalignment outside FOV fault is detected, then the customer shall be warned that camera service is required (The LROS ES signal, LROSStatus.VisionAlignmentOutOfRange, shall be set to TRUE)
   - If the camera is in an unaligned state, then the customer shall be warned that camera service is required (The LROS ES signal, LROSStatus.VisionAlignmentNotStarted, shall be set to TRUE)
   - If the service alignment is in-progress, then the customer shall be warned that camera is temporarily not available (The LROS ES signal, LROSStatus.VisionAlignmentIncomplete, shall be set to TRUE).

d) Once detected, this fault condition shall remain set continuously across ignition cycle(s) until the AR ECU determines that the camera is no longer misaligned.