|  |  |
| --- | --- |
|  | Moving Picture, Audio and Data Coding by Artificial Intelligence  www.mpai.community |

|  |  |
| --- | --- |
| **N649** | 2022/04/20 |
| **Source** | Requirements (CAV) |
| **Title** | The World Representation (WR) in MPAI-CAV |
| **Target** | MPAI Community |

# Introduction

The World Representation (WR) is the basic data format enabling a CAV to achieve its mission. It is defined as “A digital representation of the Environment produced by an Environment Sensing Technology (EST) of a CAV or an Offline Map or an integration thereof”. A WR may be a rather complex data structure that includes several elementary components called Data Types each having a Data Format.

The CAV’s WR, called Basic World Representation (BWR) results from the *integration* of the different WRs generated by different ESTs of a CAV. A CAV may produce its Full World Representation (FWR) by *integrating* the BWRs received from other CAVs in range and produce.

# The CAV’s Environment Sensing Technologies

The ESTs considered by MPAI-CAV are:

1. Global navigation satellite system or GNSS (~1 & 1.5 GHz Radio).
2. Geographical position and orientation, and their time derivatives up to 2nd order.
3. Data in the visible range, possibly supplemented by depth information (400 to 700 THz).
4. Lidar (~200 THz infrared).
5. Radar (~25 & 75 GHz).
6. Ultrasound (> 20 kHz)
7. Sound in the audible range (16 Hz to 16 kHz).
8. Other environmental data (temperature, humidity, ...)

# Requirements for WR integration

This document targets the definition of a standard WR, with the understanding that it should be possible to losslessly convert an EST-specific WR to the *standard* WR. This does not mean that each EST shall use the standard format but that the EST uses a compatible format. WR compatibility is defined as “The ability of a WR to be converted to another WR without loss of information”.

WR compatibility has the following features:

1. WR1 may include Data Types not included in WR#2 and vice versa.
2. The Data Types in WR#1 need not have the same values as those in WR#2, i.e., their values may:
   1. Not have the same Accuracy.
   2. Conflict even though their Data Types may be the same or compatible.

WR integration shall satisfy the following requirements:

1. A CAV may use a set of ESTs each providing a WR each having its own format. However, the format of the WRs shall be the same, or compatible.
2. Different ESTs may provide the same or different parameters with different levels of Accuracy.
3. Two CAVs in range may use two different EST sets. However, their BWRs shall have the same or compatible formats.
4. The BWRs of two CAVs with different positions and orientations will be different and have different levels of detail or have conflicting values, but the two BWRs shall have the same or compatible formats.
5. The formats of the digital maps shall allow for integration in a WR without loss of information.

# WR integration

Figure 1 represents the integration of different environment data provided by different ESTs of a CAV into a BWR.

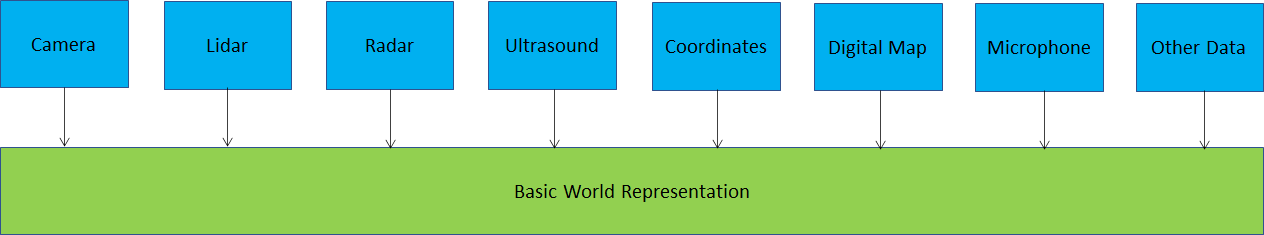


Figure 1 - Sensed WRs from different ESTs and the Basic World Representation

For the purpose of the following analysis, it is assumed that the data provided by an EST are only used to provide a WR representation where static and dynamic objects are represented according to an EST-specific technology. A simple oject representation technology is a bounding box of the appropriate size.

Object characteristics is defined as:

1. Object identifier
2. Status
   1. Static
      1. Digital map
      2. No
   2. Dynamic
   3. Unknown
3. Digital map object
4. Position dynamics
5. Orientation dynamics
6. Shape
7. Semantics

The process whereby an EST creates its WR unfolds as follows:

1. The EST uses proprietary technologies to represent objects in the sensed Environment with their position, orientation and shape.
2. If the position/orientation of the object at time t is compared with the static objects around the identified position of the scene at time t-Δt andand possibly preceding instants:
   1. Unchanged, the object characteristics are inherited.
   2. Changed
      1. compare with dynamic objects at time t-Δt
         1. if object characteristics match, inherit object characteristics
         2. if do not match, new object characteristics are assigned
      2. The position/orientation/size of the object is compared with the object(s) present in the digital map covering that position.
      3. If a match between the candidate static object and the previously is found, the candidate static object is labelled as the object in the digital map.
      4. If a match not found, the object is compared with candidate static objects at that position identified at time t-Δt.
      5. If a match is found, the object inherits the characteristics of the previously identified object.
   3. Changed:
      1. If a match with previously identified objects is found,
      2. If a match is not found, the object is labelled as a new candidate static object first identified at time t.
      3. The object is labelled as the object in the digital map. If not it is labelled as a candidates static object at time t.
3. Each EST has a WR containing static and dynamic objects that are built using the WRs developed at previous time instants.
4. Each dynamic object in the WR of each EST has position/orientation and time derivatives of the objects.
5. The CAV estimates the position/orientation of each dynamic object from the WR of each EST at time t+Δt and compares the estimate with the dynamic objects of the scene at time t+Δt.
6. A bounding box is created with a label expressing the confidence of the origin of the object (i.e., existing or new) depending on the match between the object estimated at time t+Δt and the object estimated at time t:
7. The Basic World Representation AIM of Figure 1 produces the CAV’s BWR by integrating the different box-based dynamic object representations.

The bounding boxes may be an intermediate step toward creating a semantic description of the static and dynamic objects. In later stages boxes may be expanded to more complex representations.

# Types of bounding boxes

The boxes of each EST can be 2D, 2.5D or 3D depending on the specific EST.

**2D Basic World Representation**:

1. Static environment: parametric free space representation or alternative representations.
2. Dynamic environment: object-based representation or alternative representations.

**2.5D Basic World Representation**:

1. Static components of the scene
   1. Grid-based (elevation maps or stixel world).
   2. Object-based for traffic poles and signals (e.g., Stixel world, Multi-level surface map).
2. Object-based for the dynamic parts (e.g., Stixel world, Multi-level surface map).

The **3D Basic World Representation** requirements are:

1. Static components of the scene
   1. Voxel grids, meshes (?).
   2. Object-based for traffic poles and signals (voxel grids, meshes?).
2. Dynamic components of the scene (voxel grids, meshes?)

# BWR requirements

The BWR and FWR requirements are

1. The BWR shall include all available information, to be integrated with BWRs coming from other CAVs, that enable a CAV to define a Path in the Decision Horizon Time.
2. The BWR shall include
   1. All sensed Audio and Visual information.
   2. All sensed environmental data (e.g., weather, temperature, air pressure, ice and water on the road, wind, fog etc.)
3. The BWR shall allow for easy refinement of its data based on other WRs.
4. Each Visual Object in the BWR shall be described by
   1. Its ID.
   2. Its State.
   3. Its physical characteristics, e.g., static, or dynamic.
   4. Its bounding box (as a minimum), its full shape if known, semantics, flags (e.g., warning).
   5. Its semantics (e.g., other CAVs or other objects).
   6. The estimated accuracies of object parameters.
5. All traffic signalisation, e.g., roads and lane geometry, and lane-specific traffic rules shall be described with a topology graph.
6. Each Audio Object in the BWR shall be described by
   1. Its ID.
   2. Its State.
   3. Its physical characteristics, e.g., static, or dynamic.
   4. Its semantics (e.g., other CAVs or other objects).
   5. The estimated accuracies of object parameters.
7. The BWR should allow for easy validation of a trajectory.
8. The BWR shall have a scalable representation, i.e., allowing for:
   1. Transmission of part of the BWR based on required Level of Detail.
   2. Increase of the Environment complexity.
   3. Fast access to critical data.
   4. Fast access to data required by different AIMs.
   5. Incremental refinement of Object and Scene.
   6. Updates for Object and Scene from one snapshot to another.
   7. Deliberative and reactive actions.
9. The BWR shall associate Audio objects with Visual Objects whenever possible.

# FWR requirements

The FWR requirements are:

1. The FRW shall be an extension of the BWR.
2. The FWR shall indicate
   1. Where the ego BWR has major discrepancies with the BWR received from other CAVs.
   2. The identity of the CAV with which the discrepancy has been detenced.

# Terminology

|  |  |
| --- | --- |
| **Term** | **Definition** |
| Accuracy | The precision with which a Data Type measures an entity. |
| Data Format | The standard digital representation of Data. |
| Data Type | An elementary component of a Data Format. |
| Environment Sensing Technology | One of the technologies used by the Environment Sensing Subsystem to sense the Environment. |
| Format compatibility | The ability of a Data Format to be losslessly converted to another Data Format. |
| World Representation | A digital representation of the Environment produced by a CAV Environment Sensing Technology or an Offline Map. |