Philosys Label Editor

Why using 3D Labeling with Label Editor

June 22nd, 2015

Philosys Software GmbH
Edisonstrasse 6
85716 Unterschleissheim
Germany

Version 5.0

Why using 3D Labeling with Label Editor?

Philosys Label Editor supports 3D labeling since version 4 (LEV4). This feature allows annotation of scenes containing reference data provided by reference data sources like LiDAR with world coordinates, instead of labeling objects in 2D images with pixel coordinates.

This paper gives reasons why to use Philosys Label Editor 3D labeling for validation projects.

1 Labeling for ADAS using environmental models

Early assistant systems are based on plain 2D video image processing in order to implement their functionalities. An intermediate result of this processing often is a position and size of function-specific objects in pixels coordinates. For validation this information is compared to a corresponding label provided by an annotator and a KPI is calculated to provide information how good both match.

As ADAS are becoming more and more complex, this plain image processing and object detection is no longer sufficient. Information like position, size and movement of other objects around ego are getting more and more important. In order to track all these objects environmental models (also called “maps”) are now used to describe the current surroundings of the EGO vehicle. They contain all information available about the current environment. The information used to build up these environmental models/maps is not limited to video images, but information from other sensors like e.g. radar or ultrasound, and information from quite different sources like navigation systems, Car-2-X, etc. may be included as well. Eventually the environmental model contains a fusion of all available information, providing the basis for various different ADAS functionalities.

Since these environmental models describe the 3D scene surrounding the EGO vehicle, they are typically using a world coordinate system to describe positions of the various objects in the scene. Instead of converting this information to 2D video image coordinates for validation, it is easier and makes more sense to compare the data results directly to validation data based on the same 3D world coordinate system.

LEV4+ provides the functionality to create validation data directly with 3D coordinates and it therefore allows easy comparison between ADAS environmental model and validation data.

2 Labeling reference data

The results of early assistant systems derived from video images where compared to the annotation results derived from the same images. This way errors caused by camera or calibration errors cannot be detected. Also it isn’t possible to validate distance information derived from stereo cameras or other algorithms, because the annotator doesn’t have the ability to derive distance and other dynamic information from 2D video images.

To give better validation results, independent reference data has been introduced, mostly in the form of LiDAR. LiDAR gives a fully independent and highly accurate view on the scene seen by video based ADAS. This gives a much better validation quality and helps annotators to provide validation information that isn’t available otherwise.

LEV2+ supports feeding preprocessed reference data into the annotation process by external data interface, with sophisticated object linking for minimal user intervention.

LEV4+ supports direct display of LiDAR point clouds. The point cloud information can be used for other processing.

LEV5 supports direct display of other reference data through pluggable mixin modules, mostly compatible to EBAssist/ADTF 3DSceneDisplay.

3 Labeling on base of pre-processed reference data

With 2D annotation one can only label object visible in the current video frame. In contrast, reference data like from LiDAR, can be pre-processed by combining data of the whole scene. Objects visible only from one side may be reconstructed in 3D. So the annotator gets a full overview of the scene.

LEV4+ supports display of composite point clouds provided only once in the scene.

4 Labeling multiple camera views

Manual labeling is already time consuming for simple scenes, even if the ADAS is based on a single camera only. It gets even worse when multiple cameras are coming into the play. The video stream from each camera needs to be labeled then, and the labels of the different views need to be associated with each other. This is not only complicated and error prone for annotation, but also complicates validation considerably. With 3D annotation there exists only one object representation for an object, which can be automatically tracked through all camera views.

LEV4+ supports 3D annotation with multiple camera views.

5 Labeling stationary objects

2D annotation in the video images requires labeling of objects in every frame if there is any relative motion between the objects in the scene and EGO vehicle.

3D annotation combined with highly accurate odometric data enables one-time-only annotation of non-moving objects which is valid for the whole scene then. This considerably reduces annotation time for specific functions dealing with non-moving objects.

LEV4+ supports stationary objects

6 Labeling dynamic objects

2D annotation of moving objects is particularly complicated by the continually changing shape of objects in the video image, caused by the perspective projection. Interpolation can only help partially as it cannot take the perspective projection into account, due to missing depth information. Other impacts are pitch and roll of the EGO vehicle caused by corrugations in the road surface, leading to camera shake and thus to an undesired oscillation of the 2D object positions and shapes in the video image.

LEV4+ supports linear interpolation in 3D space to reduce the number of key frames that have to be set.

LEV4+ supports the compensation of camera shake for video images from 3D space if corresponding information is provided as input.

7 Labeling complex objects

With 2D annotation the usable shapes for object markers are very limited in complexity. With 3D annotation it is possible to use much more complex shapes to represent specific object types, as there is no perspective distortion to deal with. I.e., the mesh3D object can represent complete parts of parking areas easily. This makes labeling complex structures much easier.

LEV5 supports mesh3D objects for complex shapes.

8 Labeling by using object detection

Adjusting the shapes of objects manually with the mouse is time consuming and demanding when high accuracy is required. There are a lots of mouse movements necessary in order to finally match position and shape of the reference object. If the quality of reference data is good enough then automatic object detection enables the creation of object markers with just a few mouse clicks.

LEV5 supports object detection from point clouds and images for initial object creation. An optional programmable generic object detector allows the implementation of complex processing chains from a variety of built-in functions.