

NeoVision 2 Annotation Guidelines

Version 5.0

4/16/2012 - Public

Prepared by

VideoMining Corporation, SET Corporation

Mr. Matthew Parent

mparent@videomining.com

Phone: (814)867-8977, Ext 117

Dr. Rajeev Sharma

rsharma@videomining.com

Phone: (814)867-8977, Ext 106

Dr. Yang Ran

yan@setcorp.com

Phone: (240)965-9967

Dr. Qinfen Zheng

qzheng@setcorp.com

Phone: (240)965-9966

Table of Contents

1. Introduction	3
2. Data Types.....	3
3. Data Output Format	3
4. Object Types.....	4
5. Object Type Examples	5
6. Drawing the Bounding Box.....	11
7. Determining “Don’t Care” Frames/Regions/Objects.....	12
8. Minimum Size of Targets.....	14
9. Occlusion Handling	14
10. Annotation Quality Assurance.....	14

1. Introduction

This document explains the standards and practices followed by VideoMining Corporation and SET Corporation during video annotation for the DARPA sponsored NeoVision2 research project.

2. Data Types

Data Type	Description
BVALUE	Boolean Value. These variables hold TRUE / FALSE. The convention used is TRUE = 1 and FALSE = 0.
DVALUE	Integer Value.
LVALUE	Represented by a string but the number of unique strings is limited to a small list.
FVALUE	Floating Value.

3. Data Output Format

Data will be output in a series of CSV (Comma Separated Value) files. The csv files will be stored in a top level directory indicating the data source e.g., Tower. Within this directory, there will be a csv file named after the input image sequence ID, e.g. 001.csv. Within each csv file there will be a single row for each detected object following the format specified in Table 2. At the beginning of each csv file, the first line shows the format info as a reference.

Table 1. Sample csv file as annotation output (\\Training\\Heli\\045.csv)

Frame,BoundingBox_X1,BoundingBox_Y1,BoundingBox_X2,BoundingBox_Y2,BoundingBox_X3,BoundingBox_Y3,BoundingBox_X4,BoundingBox_Y4,ObjectType,Occlusion,Ambiguous,Confidence,SiteInfo,Version
0,701,489,870,489,870,560,701,560,Car,FALSE,FALSE,1.0,,1.1
4,452,492,621,492,621,563,452,563,Car,FALSE,FALSE,1.0,,1.1
4,861,768,1051,768,1051,839,861,839,Car,TRUE,FALSE,1.0,,1.1

Each line in Table 1 corresponds to a line in the output file. Due to the similarity in the tasks, the Helicopter and Tower scenarios will have the same output format.

Table 2. Format definition for annotation output

Attribute	Data Type	Description
Frame	DVALUE	The frame that this object is present in.
BoundingBox_Xi,i=1,2,3,4	DVALUE	4 X coordinates of the corners of the bounding box
BoundingBox_Yi,i=1,2,3,4	DVALUE	4 Y coordinates of the corners of the bounding box
ObjectType	LVALUE	May be Car, Truck, Person, Cyclist, Bus, Tractor-Trailer, Plane, Helicopter, Boat, Container, or Ignore-Region as defined in section 4.

Confidence	FVALUE	Indicates the confidence of classification. A float number from 0 to 1.0. 1.0 shows almost sure.
Occlusion	BVALUE	Indicates whether or not the object is occluded at all during this frame.
Ambiguous	BVALUE	Indicates if the ObjectType is ambiguous for this particular object/frame. When it is TRUE, the ObjectType field is ignored.
SiteInfo	LVALUE	Indicates the site location, domain info. Can be empty.
Version	FVALUE	Indicates the version number. Can be empty.

4. Object Types

- 1) **Car** The car class should be used for smaller vehicles primarily designed for transporting passengers or light cargo, including automobiles, vans, SUVs, and pickup trucks. These vehicles will typically have 4 to 6 wheels and seat 2 to 8 people.
- 2) **Truck** The Truck class should be used for vehicles primarily designed for moving materials, but smaller in size than a tractor-trailer. These vehicles will typically seat 2 to 4 people. They will typically have a smaller driver/passenger cab that has two wheels, which is permanently attached to a larger cargo container that has four to six wheels. Examples include moving trucks and tanker trucks.
- 3) **Tractor-trailer** A tractor trailer can be identified by its split body with a small cab in front attached to a large container for the main body. These will typically have 8 sets of tires under the container and 2 tires under the cab. a tractor –trailer is missing the trailer (cab only) then we will not annotate it (or make it ambiguous)
- 4) **Bus** A bus can be recognized by its large rectangular shape that may be split into multiple sections. It will typically carry more than 8 passengers.
- 5) **Container** The container class is intended to describe the shipping unit known as a freight transport container or intermodal container. Containers are typically rectangular boxes 8 feet wide, 8 to 56 feet long, and 8 to 9' 6" high. The long axis of the container will be parallel to the ground, ruling out vertically-oriented boxes such as phone booths. Containers may be stacked on top of one another. This class explicitly excludes cylindrical tanks used for intermodal shipping.
- 6) **Boat** Any water-capable vehicle that can fit 4 or more people should be marked as a boat.
- 7) **Plane** A plane will have 2 or more wings attached to a rigid body and may be located on the ground or in the air.
- 8) **Helicopter** A helicopter can be identified by the small body with rotors located directly above the body. A tail extending from the body may or may not be visible.

9) **Person** People will be annotated as a person when they are visible and not currently riding in or on a vehicle.

10) **Cyclist** A person sitting on a vehicle with 1-3 wheels should be marked as a cyclist. This includes bicycles and motorcycles. Parked cycles are not part of this class.

5. Object Type Examples

Plane



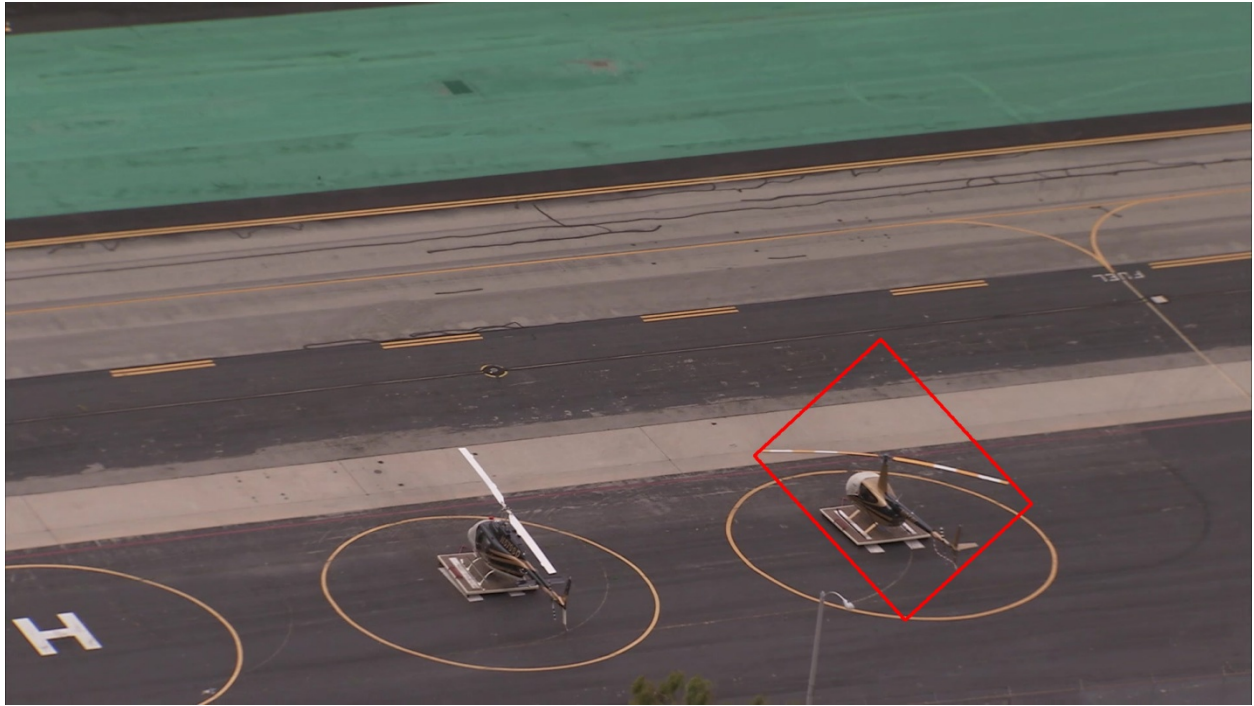
Car



Ambiguous



Helicopter



Cyclist



Boat



Tractor-Trailer



Human



Bus



Truck



Container



6. Drawing the Bounding Box

The bounding box should be oriented so that it conforms as well as possible with the boundary of the object. The simplest way to do this is to minimize the amount of non-object pixels contained within the bounding box. Some examples are given below. For objects on the ground one edge of the bounding box will likely orient with the ground plane. For objects that are partially out of frame draw the bounding box so that none of the vertexes are out of frame. Due to the oriented nature of the bounding box, this may result in a visible portion of the object being placed outside of the bounding box.



Figure 1: The image on the left shows an incorrect bounding box drawn while the image on the right shows the bounding box drawn correctly. The same frame has been used for both images.



Figure 2: A more extreme example highlighting the difference between a good and bad bounding box. The bad bounding box is on the left side of the image.

7. Determining “Don’t Care” Frames/Regions and Ambiguous Objects

Sometimes we want to exclude certain areas from annotation because they contain too many targets, making annotation too difficult. Examples include a full parking lot or an expressway with a traffic jam. To address this problem, Don’t Care Regions (DCRs) are used to identify rectangle areas in frames which can’t be spatially annotated and which are to be eliminated entirely from the mapping and scoring process. Detected objects which fall inside a DCR or whose area is contained primarily within a DCR will be eliminated prior to the mapping/scoring process and will thus not generate false alarm errors.



Figure 3: Example of “Don’t Care Region” (DCR).

“Ambiguous Objects” are identifiable objects like golf carts or baby strollers that do not fit into a NeoVision2 object class. They are referred as objects that cannot be classified into the current categories. The value of corresponding label is marked as blank in the annotation results.



Figure 4: Example of “Ambiguous Object”.

8. Minimum Size of Targets

Objects smaller than 30 linear pixels will not be annotated.

9. Annotation Density

Due to the limitation of resource and time, annotation will not be performed at every frame in each domain. For some of the sequences, one out of every N frames will be annotated. Frames that are not annotated will not be evaluated.

10. Occlusion Handling

If an object is occluded it should still be annotated so long as the object is still recognizable. If more than 33% of the object is not visible then the object will be marked as occluded.

11. Annotation Quality Assurance

Quality is assured throughout annotation through several distinct methods that enable errors and questions to be caught and handled quickly. The annotators, also called image data operators (IDO), are supervised by an image data supervisor (IDS). The annotation QA process is described below.

- 1) The IDS frequently patrols the annotation area to make sure none of the IDOs are having any problems.
- 2) After initially completing an image sequence the IDO will review their work before marking the sequence as completed.
- 3) Once a sequence has been marked as complete the IDS will perform random spot checks on the data to ensure that the data is accurate. If an IDO is consistently making mistakes the IDS will re-train them and keep an especially close eye on their work until performance is improved.
- 4) If an IDO is unsure of the best approach to take in a particular situation they will do their best to perform the most correct annotation and leave a comment explaining their uncertainty. The IDS reviews the comments every evening and discusses them in a group setting with all IDOs the next morning.
- 5) 10% of sequences will be annotated by two IDOs independently. The results for these overlapping sequences will be compared and any differences will be resolved by the IDS. Upon discovering differences among annotators the IDS will provide any necessary clarifications to the IDOs in a group setting. For objects without significant differences the IDS will choose one to use for the final output.
- 6) When the IDS is unsure of the best approach to any ambiguity in the project the issue will be raised to the project manager. At this point, the project manager may decide to raise the issue with the rest of the NeoVision Test Team.

Project Oversight:

Dr. Rajeev Sharma

Project management, Planning, Formatting and Tools Analysis:

Mr. Matthew Parent

