

ETISEO



Internal Technical Note

Video Annotation Rules

4 – Approved

IN_ETI_1_009



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

Document Title : Video Annotation Rules

Document version : 4

Document status : Approved

Date : 23 May 2006

Availability :

Authors : Silogic

Abstract This document presents rules applied to video annotation creating ground truth and reference data.

Keyword List Meta-data, ground truth, video annotation, reference data.



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

DOCUMENT CHANGE LOG

Document Issue.	Date	Reasons for change
Version 1 Draft 1 Version 1 Draft 2	25 July 2005 9 Nov 2005	Initial version Add database annotation and composite objects annotation rules
Version 2 -0 Version 3	6 January 2006	Apply changes after seminar feedback. Add new rules during first video corpus annotation

APPLICABLE AND REFERENCE DOCUMENTS (A/R)

A/R	Reference	Title
Ref 1	IN_ETI_1_011	Data structure and output format
Ref 2	IN_ETI_1_010	Video format



Table of contents

1. INTRODUCTION	6
2. PREREQUISITE	7
3. VIDEO CLIP ANNOTATION AND DATABASE CREATION	8
4. CONTEXTUAL INFORMATION ANNOTATION	9
5. GROUND TRUTH CREATION - RULES FOR OBJECT	10
5.1. Mask.....	10
5.2. Contextual object.....	10
5.3. Static physical object of interest.....	10
5.4. Occlusion.....	10
5.4.1. General rules	10
5.4.2. Static occlusion	11
5.4.3. Dynamic occlusion	12
5.4.4. Split object due to occlusion	13
5.5. Artefacts	13
5.6. Sudden illumination changes	13
5.7. Group	14
5.8. Crowd – Multiple vehicles	15
5.9. Crowded video clip	15
5.10. Held object	16
5.11. Door of vehicle	16
5.12. Person getting in/off a vehicle	17
5.13. Multi-view annotation	17
5.14. Reappeared objects	18
6. GROUND TRUTH CREATION - RULES FOR EVENT	19
6.1. Events ontology.....	19
6.1.1. States involving one physical object of interest.....	19
6.1.2. States involving one physical object of interest and one contextual object.....	20
6.1.3. States involving several physical objects of interest.....	20
6.1.4. Events involving one physical object of interest	20
6.1.5. Events involving one physical object of interest and one contextual object.....	21
6.1.6. Events involving several physical objects of interest.....	21
6.2. Annotation of events with quantitative constraints	21
6.3. Annotation of other events	21
6.4. Specific events annotation rules	22
7. REFERENCE DATA GENERATION	24
7.1. Filtering on object size	24
7.2. Object speed computation	24
7.3. Group generation	24
7.4. Multi-view and 3D reconstruction annotation	25
7.5. Creation of events	25
7.5.1. With quantitative constraints annotation.....	25
7.5.2. That can be automatically computed	25



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

8. CONCLUSION25



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

1. INTRODUCTION

ETISEO (Evaluation du Traitement et de l'Interprétation de Séquences Vidéo, www.etiseo.net) is a research project, sponsored by the French government, in order to evaluate vision techniques for video surveillance applications.

ETISEO project aims to conceive automatic evaluation tools for vision algorithms to allow a fair and quantitative comparison between algorithm results and reference data. A large and meaningful database of video sequences and their annotations will be created.

This document presents the rules employed to create the ground truth and reference data files for video annotation. Viper-GT is used to create ground truth files and the format of both files is xml compatible with Viper-GT. Indeed, this software is flexible to data structures and it is able to read different video format (Jpeg, avi...).

This document will be diffused to the participant of ETISEO evaluation to be aware of the applied rules during the manual video annotation.



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

2. PREREQUISITE

Please report to Ref 1 document for complete information about data terminology and data structure before continuing your reading.

The data structure defines objects and events attributes that are annotated in the ground truth file by hand or filled automatically during the reference data creation task. This is developed in the rest of the document.



3. VIDEO CLIP ANNOTATION AND DATABASE CREATION

The database describes briefly the contents of the video clip. To each database component is assigned a difficulty level according to illumination conditions, presence of shadows, occlusions... and permit providing a panel of different video clip difficulty level. You will find the following attributes:

- General information:
 - Sequence id: incremental number
 - Video data set: name of the data set in which it is provided; VS1 or VS2
 - The name of the sequence following the naming convention described in Ref 2
 - The difficulty level of the video clip: computed according to following criteria (occlusion, illumination conditions...)
- Scene information:
 - Indoor or outdoor scene
 - Sequence type: road, apron, metro... (defined in Ref 2)
- Acquisition information:
 - Date and hour
 - Duration of the video clip
 - Number of frames of the video clip
 - Frame rate of the video clip
 - Camera number provided
 - Camera type: color, Infrared, grayscale
 - Camera field of view: top view or side view
 - Camera network: overlapping between camera or not
 - Camera calibration: available or not and with or without distortion information
- Light condition:
 - Light source: natural or artificial
 - Day time: day, sunrise, sunset, night
 - Weather condition: sunny, cloudy, foggy, rainy, snowy
 - Illumination condition: normal, bright or dark
 - Illumination variation: none, low or sudden illumination variation in the scene
 - Illumination variation type: darker, lighter or both
- Context:
 - Background update: empty scene available or not
 - Object number: few or crowded
 - Occlusion: presence of static or dynamic occlusion
 - Reflection: presence of reflection on the ground or on other support
 - Shadow: presence of weak or strong shadow
 - Noise: presence of noise in the video due to compression artifacts or wind in the trees...
- Content:
 - Physical objects nature: physical object of interest, contextual object
 - Physical object type: categories names that are present in the scene and annotated



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

- Physical object subtype: sub-categories names that are present in the scene and annotated
- Events: List of events that will be annotated and evaluated for this video clip.

The database is representative of various complexity levels. The choice is made to use different types of scene to evaluate algorithms in different conditions (illumination, presence of shadow or reflection...). A video clip is representative of one or several difficulties and the ground truth **does not contain** every artefacts annotation. For example, the shadows parts are for instant not annotated. This choice is made to reduce the complexity of the ground truth creation task. Never mind, this is not limitative because we will be able to distinguish degraded video clip and/or objects and adapt thresholds in consequences in the evaluation process.

The following two parts present the annotations rules use to create ground truth files.

4. CONTEXTUAL INFORMATION ANNOTATION

The contextual information annotation concerns the 2D or 3D definition of zone or equipment of interest. This annotation is provided to participant with the video data set. It is a xml file, formatted to Viper-GT and described in Ref 1.

Each zone or equipment of interest is defined by its type, its name and a polygon representing its position and size in the image.

5. GROUND TRUTH CREATION - RULES FOR OBJECT

The first part concerns the rules applied during the annotation of the physical objects.

5.1. MASK

It could be envisaged that for some sequences it will exist binary masks to define "ignored" image area that will not be annotated. These masks are communicated to participants and outputs will be filtered before the evaluation.

The axis orientation for the mask are LeftToRight for x axis and TopToBottom for the y axis.

In sequences ETI-VS2-RD-6, ETI-VS2-RD-7 and ETI-VS2-RD-8, the following mask is used: rectangle from starting point (0,0) (top left point) with size 719*170.

5.2. CONTEXTUAL OBJECT

The interval of presence of a contextual object corresponds of the time during which the object is in movement. The annotation is done only during this time for contextual objects.



Image 1. Annotation of contextual objects, door opening

5.3. STATIC PHYSICAL OBJECT OF INTEREST

We annotate as if we do not have any a-priori knowledge on the sequence, so any physical object of interest, even static ones, is annotated on the whole video clip. Nevertheless, to be able to distinguish a missed detection due to a static physical object of interest, the reference data will mention it, computing 3D object speeds.

5.4. OCCCLUSION

This section presents different occlusion cases that may be treated differently in the annotation process. Even if we distinguish diverse annotation situations, the bounding boxes are always limited by the image boundary.

5.4.1. GENERAL RULES

We define the following general rules:

- First apparition of an object: In all case, apparition along the image border, apparition after a static or dynamic occlusion, the bounding box increase with the object apparition including only the object visible part,
- All bounding box are limited by the image boundary.
- Partial or complete occlusions are distinguished by the object attribute "Occluded". (Cf. Ref 1).
- Only complete occlusions of a carried (or pushed...) object are annotated (Cf. 5.10)

We now consider in the following occlusion declinations that the object has already been seen entire one time.

5.4.2. STATIC OCCLUSION

We define a static occlusion as an occlusion of a physical object of interest by a contextual object. Different cases may be envisaged:

- The physical object of interest is partially occluded: in this case the full bounding box is drawn including hidden parts of the object,
- The physical object of interest is completely occluded but all parts of the contextual object, from which the physical object of interest can reappear, are visible (ex: a post): the full bounding box is drawn and may be interpolated between both object positions (before and after the complete occlusion) or considered as static if both positions are equals.



Image 2 Example of a static occlusion where all escapes of the contextual object are visible (frame 360, 372 and 385)

- The physical object of interest is completely occluded but not all parts of the contextual object, from which the physical object of interest can reappear, are visible (ex: a wall): the object is no more annotated. The object is annotated while it is partially occluded (refer to first case).

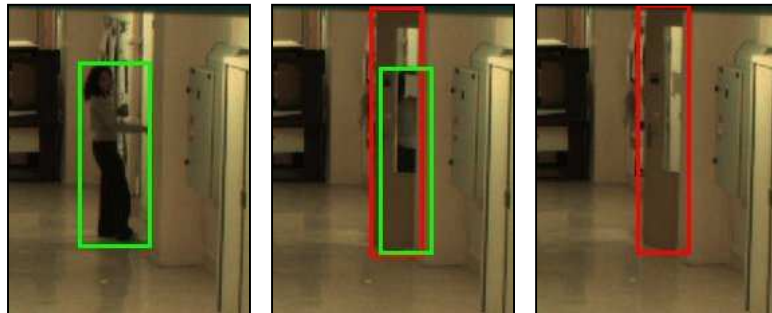


Image 3 Example of static occlusion where not all escapes of the contextual object are visible (frame 1015, 1035 and 1040)

5.4.3. DYNAMIC OCCLUSION

Any physical object of interest partially or completely occluded by another physical object of interest is considered as a dynamic occlusion. The physical object of interest is annotated as if the whole object were visible during the full dynamic occlusion.



Image 4. Example of a dynamic occlusion of a person by a vehicle



Image 5. Example of dynamic occlusion annotation: (a) the box is not occluded (b) the box is partially occluded (c) the box is completely occluded.

5.4.4. SPLIT OBJECT DUE TO OCCLUSION

If the object is split in different part due to occlusion, only one bounding box (included all sub-parts) is created.



Image 6. Example of split object annotation

5.5. ARTEFACTS

Only the part constituting the object forms the bounding box, all artefact parts are ignored (shadow, reflection...)



Image 7. Example of annotation with shadow

The qualification of artefact evaluation will be done on the whole sequence and not in the annotation task. A sequence is representative of one artefact difficulty, not each objects and frames. It could be envisaged in a second time that some artefacts are annotated in one sequence.

5.6. SUDDEN ILLUMINATION CHANGES

Some video clip contains light on/off events or quick illumination variation. Objects are still annotated during this time period.

- The physical object of interest is static before the light switch off: its bounding box is kept as the same place until the illumination becomes normal again or sufficient for human eyes and annotation,

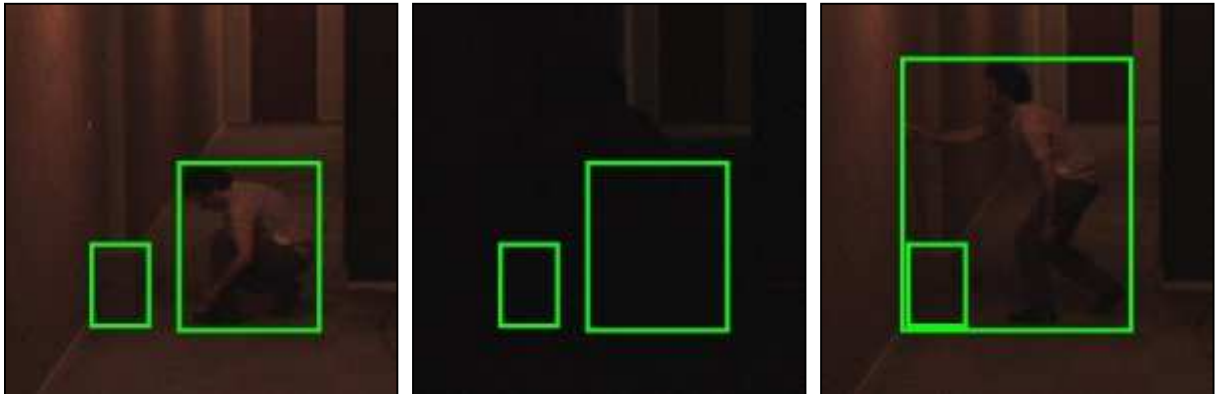


Image 8. Example of sudden illumination change annotation for static object: (a) frame 1685 (b) frame 1714 (c) frame 1715.

- The physical object of interest is in movement before the light switch off: its bounding box is interpolated until the illumination becomes normal again or sufficient for human eyes,



Image 9. Example of sudden illumination change annotation for moving object: (a) frame 2310 (b) frame 2325 (interpolated) (c) frame 2340.

In all cases, the frames containing sudden illumination variation will be annotated as such with an event and will be analysed separately during the evaluation.

5.7. GROUP

A group is formed when more than 2 objects are so close from each other that the limit between them is not designable. A group is annotated in addition to the composite object instances.

Example: The red box represents the group, which is linked with the id of the persons in blue, yellow and green.



Image 10. Example of group annotation

The different elements constituting the group are connected by their "connected-with" attribute to the physical object of interest of type group.

5.8. CROWD – MULTIPLE VEHICLES

A unique physical object of interest of type crowd or multiple vehicles is created including in its bounding box all objects.

5.9. CROWDED VIDEO CLIP

Some video clip may contain too much people and make the annotation laborious. In this case we qualify the entire video clip as crowded and only physical object of interest involved in a significant scenario are annotated. It will be communicated to participants which video clip are considered as "crowded" and which scenario will be evaluated in each of them (Cf. database). The false detections will not be evaluated in this case.



Image 11. Example of crowded scene annotation

5.10. HELD OBJECT

If a person is holding a suitcase or pushing a stroller, both physical objects of interest are created with types: person, box. The box is "connected_with" the person.



Image 12 Example of held object annotation

The carried object (ex: a bag, a person, a door) has a connection to the "holder" (resp. person, bike, vehicle).

Notes: As we mentioned in the occlusion part (Cf 5.4) the "occluded" attribute of the carried object is only annotated in case of complete occlusions.

5.11. DOOR OF VEHICLE

When the door of a vehicle is opened or closed, the vehicle bounding box is unchanged and a second object is created the time of the door movement and linked with the vehicle.



Image 13. Example of vehicle door movement

5.12. PERSON GETTING IN/OFF A VEHICLE

When the person is inside the vehicle, he is not considered as present and has no bounding box. The person is considered as present and valid when he is outside the vehicle.



Image 14. Example of annotation with person getting in a vehicle

The person in red is no more annotated when he is in the vehicle. The driver is never annotated because he does not get off the vehicle.

5.13. MULTI-VIEW ANNOTATION

It consists of annotating with same id the different entity of the same object in different views.

With same object ID in multi-view and camera calibration, we can construct 3D annotation.

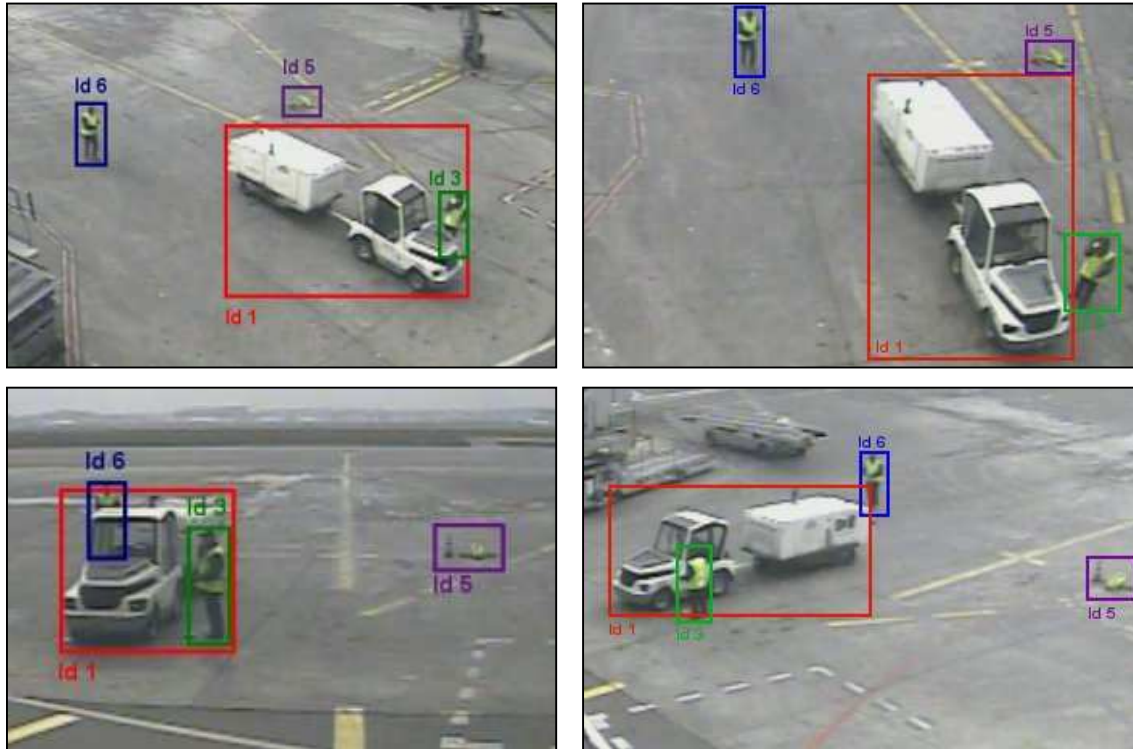


Image 15. Example of multi-view annotation

5.14. REAPPEARED OBJECTS

We will distinguish both cases:

- Multi-camera mode: an object reappearing in a camera view keeps the same id as its first apparition in the same camera view. This id also corresponds of the same object id in the other views because this is the way the multi-camera is annotated (Cf. 5.13).

Notes: When the object reappeared it keeps also its previous size even if we do not see it entire.

- Mono-camera mode: an object reappearing in a camera view is annotated with a new id. Its "link" attribute for its first frame of reappearance is equal to its previous id. Its "father" is its previous id in the scene. The object has as many different id as it is reappearing several times in the camera view and each different id is "linked" with the previous one. Example: A person who has consecutively the id 3-4-6-7 => the father of 4 is 3, the father of 6 is 4 and the father of 7 is 6.

Notes: In this case, when the object is reappeared, it is considered as a new one and its bounding box size increases with its apparition (Cf. section 5.4.1). Its previous size is not kept.



6. GROUND TRUTH CREATION - RULES FOR EVENT

The list of events present in each video sequence and expected to be recognised is diffused to participants with the video data set. Nevertheless, we define below the ontology of events that may be present in the database. The "understanding" annotation consists of defining the start and end point of these events and gives eventually the list of physical objects and contextual objects involved in, according to the "tracking" annotation (Id of physical objects of interest) and the "context" annotation (Id of contextual objects; Cf. part 4).

6.1. EVENTS ONTOLOGY

An event is characterized by the physical objects involved in. They are six types of physical objects which can be involved in an event. Each of these physical objects are referenced by a key letter in the following enumeration:

- One person (**p**).
- Group of persons (**g**) containing 2 to 7 persons.
- Crowd (**c**) which is above 7 persons.
- One portable object (**o**).
- One vehicle (**v**).
- Multiple vehicles (**m**).

An event can also involve contextual objects which are described either in the annotation or in the description of the empty scene. These contextual objects are also referenced by a key letter:

- One zone (**z**).
- One piece of equipment (**e**).

In the following, we are enumerating all the events which can be of interest in ETISEO. Depending on the properties of the physical objects which are highlighted in the event, a key letter is added to the event. There are three possible properties:

- **k**: related to kinematics properties.
- **s**: related to size(height, width) properties.
- **h**: related to shape properties.

We have ordered the events of interest depending on whether they are:

- states or events.
- involving one or several physical objects.

6.1.1. STATES INVOLVING ONE PHYSICAL OBJECT OF INTEREST

stopped(p or g or v or m)	k
walking(p)	k
running(p)	k
standing(p)	s/h
sitting(p)	s/h
lying(p)	s/h



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

crouching(p) s/h
moving(p or v) k

waiting(p or g)
fighting(g)
overcrowding(c)
unbalanced_floor_occupation(c)
abandoned_baggage(o)

6.1.2. STATES INVOLVING ONE PHYSICAL OBJECT OF INTEREST AND ONE CONTEXTUAL OBJECT

inside_zone(p or g or v or m, z) k
stays_inside_zone(p or g or v or m, z) k
stays_at(p, e) k
close_to(p or g or v, e or z or o) k
far_from(p or g or v, e or z or o) k
going_toward(p or v, e or z or o) k
going_backward(p or v, e or z or o) k
blocking_an_exit(g, z) k
access_to_forbidden_area(p or g or v, z)

empty_area(z)

6.1.3. STATES INVOLVING SEVERAL PHYSICAL OBJECTS OF INTEREST

close_to(p1, p2 or g or v) k
far_from(p1, p2 or g or v) k
going_toward(p1 or g1, p2 or g2 or v) k
walking_with(p1, p2) k
following(p1, p2 or g) k
queuing(p1, p2 or g) k
tailgating(p1, p2 or g) k
meeting(p1, p2 or g) k

holding(p, o: suitcase) h
pushing(p, o: stroller) h
moving(p, o) h

6.1.4. EVENTS INVOLVING ONE PHYSICAL OBJECT OF INTEREST

sits_down(p) k
falls_down(p) k
stands_up(p) k
jumps(p) k

rapid_increase_of_crowding_level(c)



6.1.5. EVENTS INVOLVING ONE PHYSICAL OBJECT OF INTEREST AND ONE CONTEXTUAL OBJECT

enters_zone(p or g or v, z)	k
exits_zone(p or g or v, z)	k
changes_zone(p or g or v, z1, z2)	k
opens(p, e: Door)	s,k
closes(p, e: Door)	s,k
moves_close_to(p or g or v, e or z or o)	k
goes_away_from(p or g or v, e or z or o)	k
backward_escalator(p, e: Escalator)	

6.1.6. EVENTS INVOLVING SEVERAL PHYSICAL OBJECTS OF INTEREST

meet(p1 or g1, p2 or g2 or v)	k
go_away(p1 or g1, p2 or g2 or v)	k
exchange_object(p1, p2, o)	h
picks_up(p, o)	h
puts_down(p, o)	h
exits(p, v)	k
enters(p, v)	k

6.2. ANNOTATION OF EVENTS WITH QUANTITATIVE CONSTRAINTS

Some events, as “moves_close_to” or “running”, correspond to geometric or temporal characteristic of the object in the environment, where quantitative constraints are needed.

Examples:

1. Events = Person 1 is moving toward an equipment. She becomes “close to” the equipment when distance between the middle of the base of the bounding box projected in 3D an equipment 3D polygon is under 2 meters.
2. State = Person 3 is running. Constraint: speed of person 3 is over 2 m.s^{-1} .

In this configuration, events are not annotated manually but corresponding starting and ending time are automatically computed during the creation of the reference data. We define constraints for each kind of events that are possible to express with quantitative constraints. The advantage of this method is that annotation is less subjective.

6.3. ANNOTATION OF OTHER EVENTS

Some events can't be clearly identified based on geometric or temporal simple relation. The annotation of this kind of events, as a person getting in a car, is more subjective and is leave under the choice of the annotator. The evaluation step will take into account this limitation (subjectivity), with less restrictive constraint on time intervals. Nevertheless, we have defined conventions to be coherent from one sequence to another.



6.4. SPECIFIC EVENTS ANNOTATION RULES

Please find below rules defined for specific events annotation.

First of all we define some convention to determine an event validity intervals according to an information X:

1. If information occurs at frame T, the expression "+ N frames" signifies that event validity interval is set to [T, T+N]
2. If information occurs at frame T, the expression "- N frames" signifies that event validity interval is set to [T-N, T]
3. If information occurs at frame T, the expression "+/- N frames" signifies that event validity interval is set to [T-N, T+N].
4. The event takes place from information T until information T', its validity interval is set to [T, T'].

Notes: Each event validity interval is limited by the presence intervals of all objects which are involved in. The event validity interval is shifted (case 3) or cut (case 1, 2 and 4) if one of the involved physical objects is not present during the whole period. For example in case 3, if one involved object is not present on interval [T-N; T] the event validity interval is shifted on interval [T; T+2N].

Following the previous time conventions, please find below, some rules which were used to annotate some specific events.

- **picks_up:** 1st frame the carried object is "connected_with" the holder +/- 10 frames (or if the carried object is not visible on interval [T-10], the event is valid when the 1st frame carried object exists and is "connected_with" + 20 frames)
- **puts_down:** last frame "connected_with" +/- 10 frames
- **exchange_object:** the second person holds the object + 30 frames
- **gets_in:** person gets_in vehicle. True from the first movement of the vehicle door until the last frame the person is visible. If there is no door, the event is valid when the last frame the person is visible -10 frames
- **gets_out:** person gets_out vehicle. True from the first frame the person is visible until the last frame the door is in movement. If there is no door, the event is valid when the first frame the person is visible +10 frames
- **gets_on:** person gets_on bike. 1st frame "biking" ("connected_with") - 20 frames
- **gets_off:** person gets_off bike. 1st frame after "biking" ("connected_with") + 20 frames
- **door_control:** 1st frame the person is in contact with the door control button +/- 10 frames
- **inside_zone:** True when vehicle centre of gravity is inside the zone OR when a person has at least one foot in the zone.



Internal Technical note Video Annotation Rules

Vers : 4 - Approved

Ref : IN_ETI_1_009

Date : 23 May 2006

- **empty_area**: complementary to event inside_zone.
- **enters_zone**: first part of the vehicle in the zone until the whole vehicle is inside zone OR first foot inside zone for a person +/- 5 frames
- **exits_zone**: first part of the vehicle outside zone until the whole vehicle is outside zone OR last foot inside zone for a person +/- 5 frames
- **changes_zone** (A to B): union of "exits_zone A" and "enters_zone B" OR if zone B is not visible (wall and door), the event is valid from the first frame the door is in movement until the last frame the person is visible OR if zone A is not visible, the event is valid from the 1st frame the person is visible + 20 frames.

Notes: The events enter or exit a zone are not annotated if the object is not in contact with a border of the zone (example: a person getting off a car doesn't "enter" in the zone below the vehicle).

- **close_to**: distance between both physical objects gravity centres is less than 1.50 meters
- **waiting**: the person is stopped for more than 250 frames until he is out of field of view OR until a new person arrives and stops close_to him.
- **goes_away_from**: end of close_to until distance is more than 5 meters AND person speed is not null.
- **meeting**: people stopped close_to from each other + 40 frames
- **splitting_up**: first frame after persons were close_to until distance between persons is more than 5 meters.
- **abandoned_baggage**: The person who puts down the baggage is out of field of view of the camera until a person is holding the baggage.
- **crossing**: the event "person crossing a road" is valid when the person movement direction is perpendicular to the road axis and the person is inside the zone defining the road boundaries.
- **overtaking**: the event vehicle A is overtaking vehicle B is valid when vehicle A and B are in movement and at least one part of vehicle A is between vehicle B front and back wheels.



7. REFERENCE DATA GENERATION

Referenced data are composed of:

- Ground truth data,
- Computed Data.

Computed data are extracted combining:

- Ground truth data concerning several objects (ex: dynamic occlusion),
- Ground truth data and contextual data (ex: calibration/ 3D position).

7.1. FILTERING ON OBJECT SIZE

The first step of the creation of the reference data from the ground truth is the filtering on the bounding box size. The validity interval of an object is reduced while the object size is too small (under a fixed threshold according to the image size) in the border frames of its interval. The ground truth should contain small bounding box to permit filtering during the reference data generation process.

Example: The interval of presence for Person 5 is [4500,4700]. If size of (Id 5, frame 4500) < Threshold, the new interval is [4501, 4700]. The process is recursive until size of (Id 5, frame N) > Threshold. We process similarly for the high limit (4700).

7.2. OBJECT SPEED COMPUTATION

(CF. 5.3 Static physical object of interest)

After a certain period of time of immobility, the object is set to static. This annotation includes the computation of bounding boxes speed. We evaluated the speed using two successive positions of the bounding box and the frame rate of the sequence. The object is immobile when its speed (thresholded) equals 0.

7.3. GROUP GENERATION

(CF. 5.7 Group)

The generation of the object group is done automatically with a spatial constraint application between objects.



7.4. MULTI-VIEW AND 3D RECONSTRUCTION ANNOTATION

(CF. 5.13 Multi-view annotation)

This task concerns the 3D objects position computation. It combines calibration information with context and multi-view ground truth.

7.5. CREATION OF EVENTS

7.5.1. WITH QUANTITATIVE CONSTRAINTS ANNOTATION

As exposed in section “6.2 Annotation of events with quantitative constraints”, some events can be automatically computed if they are based on quantitative constraints.

Example: walking, running, closed to something or someone. The reference data creation tool generates these events.

7.5.2. THAT CAN BE AUTOMATICALLY COMPUTED

Events like “going toward”, “go away”, “following” can be also automatically generated.

8. CONCLUSION

Present document constitutes basic rules that were applied to the first video data set annotation. If there are evolutions during second data set annotation, the participants will be informed. All document updates are available on the web site.