

**Minutes from TRECVID 2008 Event Detection Planning Telecon**  
February 13, 2008

Jon Fiscus, NIST  
John Garofolo, NIST  
Paul Over, NIST  
Travis Rose, NIST  
Francois Bremond, INRIA  
Mert Dikmen, UIUC  
Jingen Liu, UCF  
Vasant Manohar, Matt, USF  
Michael Nechyba, Pitt Patt  
Noel O'Connor, Phil Kelly, DCU  
Dave Schipfitz, intu:Vision  
Heather Simpson, LDC  
Todd Stephenson, NRL  
Stephanie Strassel, LDC  
Murtaza Taj, OMUL  
Lexing Xie, IBM

-----  
2/13 Agenda

- Role call @ 11:00 EST
- Old Business
- Review of old action items:
  - All: Suggest tools for annotation besides VIPER, Anvil
  - All: Post suggested side information to the mailing list
  - NIST: Propose NO-SCORE regions (completed)
- NO-SCORE Mask: discussion and straw poll
  - Review the proposed no-score regions.
  - Discuss rules for implementation: (Issue 1 below)
  - Question: Should systems be allowed to use NO-SCORE data?
- New business
  - File formats: We propose to use the VIPER format for both reference annotations and system output because it will allow for future enhancements. An alternative would be to define a simpler format specifically for this task.
  - Issues to discuss from the Straw Man Proposal
    - Last paragraph of "Video Source" section states: "evaluation systems must submit output for both the devset and evalset".
- Next Telecon:
  - Feb. 20, 2008

-----  
**Review old business.**

- annotation tools

**Action item (all):**

If sites know of similar tools that may be appropriate, they can suggest these by posting to the list.

**- contextual information**

Noel, DCU:

need information about distances in 2D map vs. locations on ground plane  
-- need scale information for the map

Jon, NIST:

Does this also mean camera locations, and locations of elevator doors (for example)?

Noel, DCU:

-- yes, essentially need feature points with scale information

Other proposals:

Murtaza, OMUL:

one-way vs. two-way doors can be annotated

Paul, NIST:

Elevators: 2-way

Departure / arrival should be a one-way door

Factors of the environment can be encoded -- should we encode the intentional use of the door?

Outcome:

The direction of a door (one-way vs. two-way) should encode its intended use.

**- no-score regions**

Straw poll taken:

**Question:** Are the camera masks proposed by NIST sufficient?

**Outcome:** Majority in favor of the camera masks.

**Group discussion of the no score regions.**

Michael, Pitt Patt:

Maybe on Camera 2 the cafe region needs to be masked, but it may not matter

Noel, DCU:

Camera 3, the brightly lit area (information kiosk / currency exchange) may need to be masked

Francois, INRIA:

The specified no-score areas are ok, but we should divide the scoreable area into two areas:

-- one level is for all participants

-- second area is for more advanced researchers

Jon, NIST:

Rather than dividing the area up, it would be possible to do conditional scoring to tease out the performance difference. We want participants to process all of the data, in the scoreable area, so that we can accurately benchmark the performance.

Review of NO-SCORE handling:

- black out no-score regions for annotators
- systems: provided with a polygon mask but systems can, at their option, use the full video frame.
- Annotators make judgments on whether or not the event occurs based on the visual evidence regardless of frame or no-score boundaries. This means if the event is "bag transfer" and a person's hand is seen grabbing the bag, it is a scoreable event.

Paul, NIST:

The desire to leave this up to the annotators is to avoid the slippery slope of spatial annotation. We don't want to have extensive spatial annotation guidelines that make the annotation exceedingly difficult.

Straw poll taken:

**Question:** Is the no-score handling as stated above acceptable.  
**Outcome:** Majority in favor of the proposed no-score handling.

**New business.**

**-file formats**  
VIPER

Straw poll taken:

**Question:** Is the ViPER format suitable for reference annotation and system output?  
**Outcome:** Majority in favor of using the ViPER format

**Group discussion of VIPER tool/format.**

Stephanie, LDC:  
looking at VIPER-GT, should be possible to use

Jon:  
File format should be good for interchange

Heather:  
VIPER: can be slow -- MPEG-4 may be a possibility (in future)?

Stephanie:  
Although need to cut the video into segments, do not want to cut so small that it's counter-productive  
-- want to have right level of granularity

Comments about VIPER:  
Michael: should be okay  
-- some redundancy  
-- it's best if the annotation is consistent in where the information appears  
e.g., which frame ranges are annotated, etc.

Paul:  
NIST can suggest a profile  
-- what attributes, where the information should appear

Noel:  
A validator (for the ViPER format) would be valuable

Jon: agree, validators are useful.

**Action item (NIST/LDC):**

NIST / LDC will define a data dictionary to specify how the events are to be represented in the file format.  
NIST to provide a validator.

Straw poll taken:

**Question:** As stated in the "straw man" proposal, is it acceptable to process the entire corpus (i.e., devset + evalset)?  
**Outcome:** Majority in favor of processing the entire corpus.

**Action item (NIST):**

NIST to post address to ship drives.