

Guidelines for the TRECVID 2004 Evaluation

(last updated: Thursday, 17-Feb-2005 19:01:38 UTC)

0. Table of Contents:

- [Goal](#)
 - [Tasks](#)
 - [Shot boundary detection](#)
 - [Story segmentation](#)
 - [Feature extraction](#)
 - [Search](#)
 - [Video data](#)
 - [Information needs and topics](#)
 - [Submissions and evaluations](#)
 - [Shot boundary detection](#)
 - [Story segmentation](#)
 - [Feature extraction](#)
 - [Search](#)
 - [Milestones](#)
 - [Guideline issues and resolutions](#)
 - [Contacts](#)
-

1. Goal:

The main goal of the TREC Video Retrieval Evaluation (TRECVID) is to promote progress in content-based retrieval from digital video via open, metrics-based evaluation.

2. Tasks:

TRECVID is a laboratory-style evaluation that attempts to model real world situations or significant component tasks involved in such situations.

There are four main tasks with tests associated and participants must complete at least one of these in order to attend the workshop.

- shot boundary determination
- story segmentation
- high-level feature extraction
- search

2.1 Shot boundary detection:

The task is as follows: identify the shot boundaries with their location and type (cut or gradual) in the given video clip(s)

2.2 Story segmentation:

The task is as follows: given the story boundary test collection, identify the story boundaries with their location (time). The story classification task from 2003 with only two types left too little room to do better than a baseline

that always guesses "News" so it will not be continued.

A story can be composed of multiple shots, e.g. an anchorperson introduces a reporter and the story is finished back in the studio-setting. On the other hand, a single shot can contain story boundaries, e.g. an anchorperson switching to the next news topic.

The task is based on manual story boundary annotations made by LDC for the TDT-2 project. Therefore, LDC's definition of a story will be used in the task: A **news** story is defined as a segment of a news broadcast with a coherent news focus which contains at least two independent, declarative clauses. Other coherent segments are labeled as **miscellaneous**. These non-news stories cover a mixture of footage: commercials, lead-ins and reporter chit-chat. Guidelines that were used for annotating the TDT-2 dataset are available at <http://www ldc.upenn.edu/Projects/TDT2/Guide/manual.front.html>. Other useful documents are the [guidelines document](#) for the annotation of the TDT4 corpus and a [similar document on TDT3](#), which discuss the annotation guidelines for the different corpora. Section 2 in the TDT4 document is of particular interest for the story segmentation task.

Note: adjacent non-news stories are merged together and annotated as one single story classified as "miscellaneous".

Differences from the TDT-2 story segmentation task:

1. TRECVID 2004 uses a subset of TDT2 dataset: only video sources.
2. Video stream is available to enhance story segmentation.
3. The task is modeled as a retrospective action, so it is allowed to use global data.

There are several required runs with different inputs:

1. Required: Video + Audio (no ASR, no transcripts, etc.)
2. Required: Video + Audio + LIMSI ASR (no transcripts, etc.)
3. Required: LIMSI ASR (just the ASR)

Additional optional runs can may use other ASR and/or the closed-captions-based transcripts provided with the test data (*.tkn), etc.

With TRECVID 2004's story segmentation task, we hope to show how video information can enhance story segmentation algorithms.

2.3 Feature extraction:

Various high-level semantic features, concepts such as "Indoor/Outdoor", "People", "Speech" etc., occur frequently in video databases. The proposed task will contribute to work on a benchmark for evaluating the effectiveness of detection methods for semantic concepts

The task is as follows: given the feature test collection, the common shot boundary reference for the feature extraction test collection, and the list of feature definitions(see below), participants will return for each feature the list of at most 2000 shots from the test collection, ranked according to the highest possibility of detecting the presence of the feature. Each feature is assumed to be binary, i.e., it is either present or absent in the given reference shot.

Participants are encouraged to make their feature detection submission available to other participants for use in the search task. Donors should provide the donated detection over the search test collection in the [feature exchange format](#) by the date indicated in the schedule below..

Description of features to be detected:

The descriptions are meant to be clear to humans, e.g., assessors/annotators creating truth data and system developers attempting to automate feature detection. They are not meant to indicate how automatic detection should be achieved.

If the feature is true for some frame (sequence) within the shot, then it is true for the shot; and vice versa. This is a simplification adopted for the benefits it affords in pooling of results and approximating the basis for calculating recall.

NOTE: In the following, "contains x" is short for "contains x to a degree sufficient for x to be recognizable as x to a human". This means among other things that unless explicitly stated, partial visibility or audibility may suffice.

Ten features were chosen from the set used in the common feature annotation for 2003 and those tested in 2003.

NOTE: Although no feature definitions were shared during the 2003 annotation effort, NIST will instruct the assessors during the manual evaluation of 2004 feature extraction submissions as follows. The fact that a segment contains video of physical objects representing the topic target, such as photos, paintings, models, or toy versions of the topic target, should NOT be grounds for judging the feature to be true for the segment. Containing video of the target within video may be grounds for doing so.

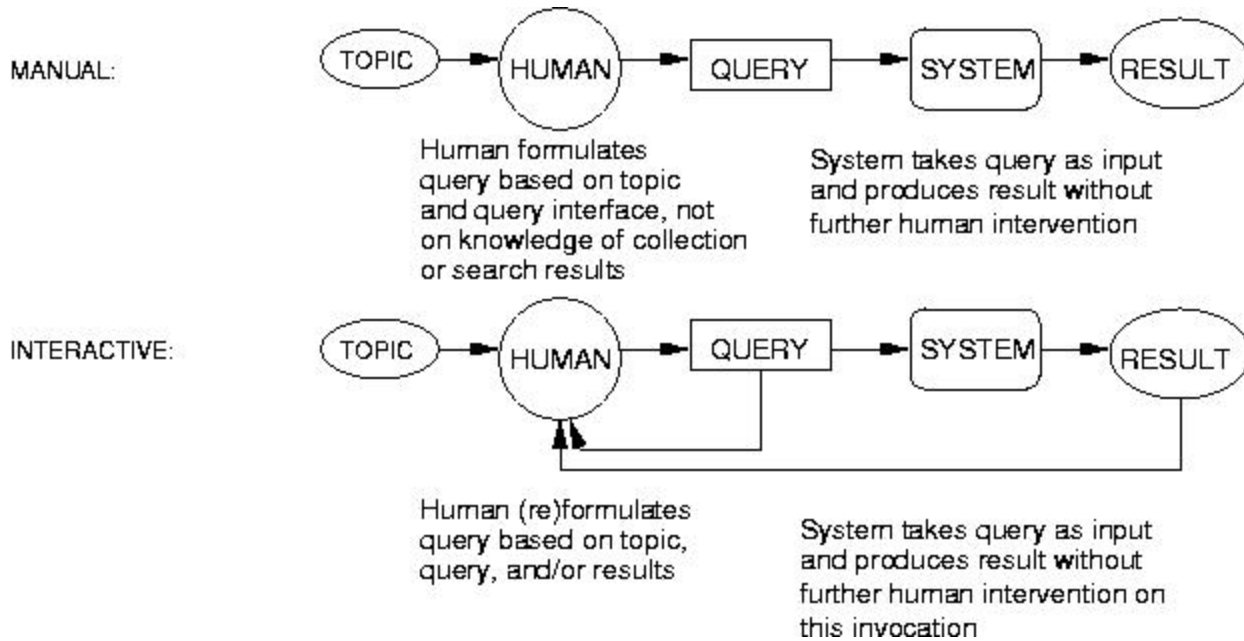
Note the in the running numbering scheme across TRECVIDs, the features for 2004 are numbered 28 - 37. Please use these numbers in submissions.

- 28. **Boat/ship:** segment contains video of at least one boat, canoe, kayak, or ship of any type.
- 29. **Madeleine Albright:** segment contains video of Madeleine Albright
- 30. **Bill Clinton:** segment contains video of Bill Clinton
- 31. **Train:** segment contains video of one or more trains, or railroad cars which are part of a train
- 32. **Beach:** segment contains video of a beach with the water and the shore visible
- 33. **Basket scored:** segment contains video of a basketball passing down through the hoop and into the net to score a basket - as part of a game or not
- 34. **Airplane takeoff:** segment contains video of an airplane taking off, moving away from the viewer
- 35. **People walking/running:** segment contains video of more than one person walking or running
- 36. **Physical violence:** segment contains video of violent interaction between people and/or objects
- 37. **Road:** segment contains video of part of a road, any size, paved or not

2.4 Search:

The task is as follows: given the search test collection, a multimedia statement of information need (topic), and the common shot boundary reference for the search test collection, return a ranked list of at most 1000 common reference shots from the test collection, which best satisfy the need. Please note the following restrictions for this task:

1. TRECVID 2004 will set aside the challenging problem of fully automatic topic analysis and query generation. Submissions will be restricted to those with a human in the loop, i.e., manual or interactive runs as defined below.



2. Because the choice of features and their combination for search is an open research question, no attempt will be made to restrict groups with respect to their use of features in search. However, groups making manual runs should report their queries, query features, and feature definitions.
3. Every submitted run must contain a result set for each topic.
4. One baseline run will be required of every **manual** system:
 1. A run based only on the text from the LIMSI ASR output and on the text of the topics.
5. In order to maximize comparability within and across participating groups, all manual runs within any given site must be carried out by the same person.
6. An interactive run will contain one result for each and every topic, each such result using the same system variant. Each result for a topic can come from only one searcher, but the same searcher does not need to be used for all topics in a run. [Here are some suggestions for interactive experiments.](#)
7. The searcher should have no experience of the topics beyond the general world knowledge of an educated adult.
8. The search system cannot be trained, pre-configured, or otherwise tuned to the topics.
9. The maximum total elapsed time limit for each topic (from the time the searcher sees the topic until the time the final result set for that topic is returned) in an interactive search run will be 15 minutes. For manual runs the manual effort (topic to query translation) for any given topic will be limited to 15 minutes.
10. All groups submitting search runs must include the actual elapsed time spent as defined in the videoSearchRunResult.dtd.
11. Groups carrying out interactive runs should measure user characteristics and satisfaction as well and report this with their results, but they need not submit this information to NIST. [Here is some information about the questionnaires to be used and how the team at Dublin City University will collect and distribute the data.](#)

12. In general, groups are reminded to use good experimental design principles. These include among other things, randomizing the order in which topics are searched for each run so as to balance learning effects.
-

3. Video data:

3.1 TRECVID 2004 development data

The development and test data for TRECVID 2003 plus various ancillary data created for 2003 (e.g., ASR from LIMSI) will become the development data for TRECVID 2004. TRECVID 2003 participants should already have this data. TRECVID 2004 participants who did not participate in 2003 are informed about how to get the 2003 data from LDC when they apply for TRECVID 2004. The data take up about 130 gigabytes.

The master shot boundary data was mistakenly left off of the TRECVID 2004 development data disk. That data is [available here for download](#).

The Carnegie Mellon University Informedia project has provided a large set of low-level features for the 2004 development data (i.e. for the complete TRECVID 2003 development/test data) as a common reference for TRECVID researchers. Here is a brief [README](#) and here is the link to the [features on CMU's website](#).

3.2 TRECVID 2004 test data

We will use 70 hours of video captured by the Linguistic Data Consortium during the last half of 1998 from CNN Headline News and ABC World News Tonight for test data. The video will be in MPEG-1 format. The data are estimated to take up about 80 gigabytes.

About 6 hours of test data will be randomly selected to be used solely as the shot boundary test collection. The remaining 64 hours will be used as test data for the story segmentation, feature extraction, and search tasks, i.e., the test data for those tasks will be identical to each other.

The shot boundary test data (~ 5 gigabytes) will be shipped by NIST to participants on DVDs (DVD+R). Distribution of the remaining test data will be handled by LDC using loaner IDE drives, which must be returned or purchased within 3 weeks of loading it on their system unless they have gotten an exemption from LDC in advance.. The only charge to participants for test data will be the cost of shipping the drive(s) back to LDC. More information about the data will be provided on the TRECVID website starting in March as we know more.

Note: Participating groups from TRECVID 2003 who received a loaner drive and have not returned or bought the drive, are not eligible to participate in TRECVID 2004.

2004 data license agreement for active participants

In order to be eligible to receive the test data, you must have completed the following form and faxed it (Attention: Lori Buckland) to  fax number in the US.

- CNN/ABC data for TRECVID 2004
 - [form](#)

3.3 Ancillary data associated with the test data

Provided with the ABC/CNN MPEG-1 test data (*.mpg) on the loaner drive will be a number of other datasets.

Closed-captions-based transcript

A closed-captions-based transcript will be provided. The transcript will contain simple tokens (*.tkn) with no other information.

ASR output from LIMSI-CNRS:

Jean-Luc Gauvain of the [Spoken Language Processing Group at LIMSI](#) has graciously agreed to donate ASR output for the test collection. Be sure to credit them for this contribution by a non-participant.

J.L. Gauvain, L. Lamel, and G. Adda.
The LIMSI Broadcast News Transcription System.
Speech Communication, 37(1-2):89-108, 2002.
ftp://t1p.limsi.fr/public/spcH4_limsi.ps.Z

Common shot boundary reference and keyframes:

A common shot boundary reference will again kindly be provided by Georges Quénot at CLIPS-IMAG. Keyframes will also be selected for use in the search and feature extraction tasks.

The emphasis in the common shot boundary reference will be on the shots, not the transitions. The shots are contiguous. There are no gaps between them. They do not overlap. The media time format is based on the Gregorian day time (ISO 8601) norm. Fractions are defined by counting pre-specified fractions of a second. In our case, the frame rate will likely be 29.97. One fraction of a second is thus specified as "PT1001N30000F".

The video id has the format of "XXX" and shot id "shotXXX_YYY". The "XXX" is the sequence number of video onto which the video file name is mapped, this will be listed in the "collection.xml" file. The "YYY" is the sequence number of the shot. Keyframes are identified as by a suffix "_RKF" for the main keyframe (one per shot) or "_NKRF" for additional keyframes derived from subshots that were merged so that shots have a minimum duration of 2 seconds.

The common shot boundary directory will contain these file(type)s:

- **shots2004** - a directory with one file of shot information for each video file in the development/test collection
 - xxx.mp7.xml - master shot list for video with id "xxx" in collection.xml
- collection.xml - a list of the files in the collection
- README - info on the segmentation
- time.elements - info on the meaning/format of the MPEG-7 MediaTimePoint and MediaDuration elements

Low-level features for the 2004 test data - from CMU

The Carnegie Mellon University Informedia project has provided a large set of low-level features for the 2004 test data as a common reference for TRECVID researchers. Low-level feature OCR files are included. Here is the link to the [features](#) on CMU's website

3.4 Restrictions on use of development and test data

Each participating group is responsible for adhering to the letter and spirit of these rules, the intent of which is to make the TRECVID evaluation realistic, fair and maximally informative about system effectiveness as opposed to other confounding effects on performance. Submissions, which in the judgment of the coordinators and NIST do not comply, will not be accepted.

Test data

The test data shipped by LDC cannot be used for system development and system developers should have no knowledge of it until after they have submitted their results for evaluation to NIST. Depending on the size of the team and tasks undertaken, this may mean isolating certain team members from certain information or operations, freezing system development early, etc.

Participants may use donated feature extraction output from the test collection but incorporation of such features should be automatic so that system development is not affected by knowledge of the extracted features. Anyone doing searches must be isolated from knowledge of that output.

Participants cannot use the knowledge that the test collection comes from news video recorded during the first half of 1998 in the development of their systems. This would be unrealistic.

Development data

The development data is intended for the participants' use in developing their systems. It is up to the participants how the development data is used, e.g., divided into training and validation data, etc.

Other data sets created by LDC for earlier evaluations and derived from the same original videos as the test data cannot be used in developing systems for TRECVID 2004.

If participants use the output of an ASR system, they must submit at least one run using that provided on the loaner drive from LDC. They are free to use the output of other ASR systems in additional runs.

If participants use a closed-captions-based transcript, they must use only that provided on the loaner drive from LDC.

Participants may use other development resources not excluded in these guidelines. Such resources should be reported at the workshop. Note that use of other resources will change the submission's status with respect to system development type, which is described next.

In 2003 a group of participants creating and sharing annotation of the development data for TRECVID 2003. See the [Video Collaborative Annotation Forum webpage](#) for details. The set of collaborative annotations created for TRECVID 2003 is part of the development data for 2004. If you use the collaborative annotations, please include the following citation in any publications, presentations, etc.:

C.-Y. Lin, B. L. Tseng and J. R. Smith, "Video Collaborative Annotation Forum: Establishing Ground-Truth Labels on Large Multimedia Datasets,"
NIST TREC-2003 Video Retrieval Evaluation Conference, Gaithersburg, MD, November 2003.
<http://www-nlpir.nist.gov/projects/tvpubs/papers/ibm.final.paper.pdf>

In order to help isolate system development as a factor in system performance each feature extraction task submission, search task submission, or donation of extracted features must declare its type:

- A - system trained only on common development collection and the common annotation of it. In 2004, a Type A system may also use annotation based on truth data created at NIST for the 2003 topics run against the 2003 test data.
- B - system trained only on common development collection but not on (just) common annotation of it
- C - system is not of type A or B

3.5 Data license agreements for active participants

- CNN/ABC data for TRECVID 2004
 - Coming soon

4. Information needs and topics:

4.1 Example types of informations needs

I'm interested in video material / information about:

- a specific person
- one or more instances of a category of people
- a specific thing
- one or more instances of a category of things
- a specific event/activity
- one or more instances of a category of events/activities
- a specific location
- one or more instances of a category of locations
- combinations of the above

As an experiment, NIST may create a topic of the form "I'm looking for video that tells me the name of the person/place/thing/event in the image/video example"

Topics may target commercials as well as news content.

4.2 Topics:

The topics, formatted multimedia statements of information need, will be developed by NIST who will control their distribution. The topics will express the need for video concerning people, things, events, locations, etc. and combinations of the former. Candidate topics (text only) will be created at NIST by examining a large subset of the test collection videos without reference to the audio, looking for candidate topic targets. The goal will be to create about equal numbers of topics looking for video of person, things, events, locations. As part of this process NIST will examine a log of almost 13,000 actual queries logged and provided by the BBC for the test time period. Accepted topics will be enhanced with non-textual examples from the Web if possible and from the development data if need be. The goal is to create 25 topics.

* Note: The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology

- Topics describe the information need. They are input to systems and guide to humans assessing relevance of system output
- Topics are multimedia objects - subject to the nature of the need and the questioner's choice of expression
- As realistic in intent and expression as possible
- Template for topic:
 - Title
 - Brief textual description of the information need (this text may contain references to the examples)
 - Examples* of what is wanted:
 - reference to video clip
 - Optional brief textual clarification of the example's relation to the need
 - reference to image
 - Optional brief textual clarification of the example's relation to the need
 - reference to audio
 - Optional brief textual clarification of the example's relation to the need

* If possible, the examples will come from outside the test data. They could be taken from various stable public domain sources. If the example comes from the test collection, the text description will be such that using a quotation from the test collection is plausible, e.g., "I want to find all the OTHER shots dealing with X." A search for a single shot cannot be described with an example from the target shot.

5. Submissions and Evaluation:

Please note: Only submissions which are valid when checked against the supplied DTDs will be accepted. **You** must check your submission. Various checkers exist, e.g., [the one at Brown University](#), [Xerces-J](#), etc.

The results of the evaluation will be made available to attendees at the TRECVID 2004 workshop and will be published in the final proceedings and/or on the TRECVID website within six months after the workshop. All submissions will likewise be available to interested researchers via the TRECVID website within six months of the workshop.

5.1 Shot boundary detection

- Participating groups may submit up to 10 runs. **Please note** that an additional attribute has been added to the shotBoundaryResult.dtd.
Optionally, groups may submit an additional 10 runs which, unlike the original required runs, ALL use as additional system input the [ASR provided by LIMSI for the shot boundary test collection](#). All runs will be evaluated.
- Here is a [DTD for shot boundary results on one video file](#), one for [results on multiple files](#), and a [small example](#) of what a site would send to NIST for evaluation. Please check your submission to see that it is well-formed
- Please send your submissions (up to 20 runs) in an email to over@nist.gov. Indicate somewhere (e.g., in the subject line) which group you are attached to so that we match you up with the active participant's database.
- Automatic comparison to human-annotated reference. Software for this comparison is available under [Tools used by TRECVID](#) from the main TRECVID page.
- Measures:
 - All transitions: for each file, precision and recall for detection; for each run, the mean precision and recall per reference transition across all files
 - Gradual transitions only: "frame-recall" and "frame precision" will be calculated for each detected gradual reference transition. Averages per detected gradual reference transition will be calculated for each file and for each submitted run. [Details are available.](#)
- **Information on processing complexity** - all groups who submitted shot boundary determination runs have been asked to provide some standard information on the processing complexity of each run to NIST by 17 October. [The details are documented here.](#)

5.2 Story segmentation

Submissions

- Participating groups may submit up to 10 runs. All runs will be evaluated.
- The task is defined on the search dataset, which is partitioned into a development and test collection (cf. [Section 3](#)).
- The reference data is defined such that there are no gaps between stories and stories do not overlap.
- The evaluation of the story segmentation task will be defined on the video as present in the test data. The notion of "clipping points" from 2003 will be incorporated in the evaluation software so submissions need not be concerned with it. Please submit the start of the first story in the video and the end of the last as boundaries. Use the timing from the MPEG-1 test data.
- Here is a [DTD for a story segmentation submission from a group](#) and a partial example of [a segmentation submission](#). Please check your submission to see that it is well-formed. Stories within a run result must be in chronological sequence with the earliest at the beginning of the file. Submissions should include boundaries for all the videos in the test set.

- Please send your submissions (up to 10 runs) in an email to Joaquim.Arlandis@nist.gov. Indicate somewhere (e.g., in the subject line) which group you are attached to so that we match you up with the active participant's database.

Evaluation

- Since story boundaries are rather abrupt changes of focus, story boundary evaluation is modeled on the evaluation of shot boundaries (the cuts, not the gradual boundaries). A story boundary is expressed as a time offset with respect to the start of the video file in seconds, accurate to nearest hundredth of a second. Each reference boundary is expanded with a fuzziness factor of five seconds in each direction, resulting in an evaluation interval of 10 seconds.
- A reference boundary is detected when one or more computed story boundaries lies within its evaluation interval.
- If a computed boundary does not fall in the evaluation interval of a reference boundary, it is considered a false alarm.
- Story boundary recall= number of reference boundaries detected/ total number of reference boundaries
- Story boundary precision= (total number of submitted boundaries minus the total amount of false alarms)/ total number of submitted boundaries

Comparability with TDT-2 Results

Results of the TRECVID 2003/4 story segmentation task cannot be directly compared to TDT-2 results because the evaluation datasets differ and different evaluation measures are used. TRECVID 2003/4 participants have shown a preference for a precision/recall oriented evaluation, whereas TDT used (and is still using) normalized detection cost. Finally, TDT was modeled as an on-line task, whereas TRECVID examines story segmentation in an archival setting, permitting the use of global information. However, the TRECVID 2003/4 story segmentation task provides an interesting testbed for cross-resource experiments. In principle, a TDT system could be used to produce an ASR+CC or ASR+CC+Audio run.

5.3 Feature extraction

Submissions

- Participating groups may submit up to 10 runs. All runs will be evaluated.
- For each feature in a run, participants will return at most 2000.
- Here is a [DTD for feature extraction results of one run](#), one for [results from multiple runs](#), and a [small example](#) of what a site would send to NIST for evaluation. Please check your submission to see that it is well-formed
- Please send your submission in an email to Cedric.Coulon@nist.gov. Indicate somewhere (e.g., in the subject line) which group you are attached to so that we match you up with the active participant's database. Send all of your runs as one file or send each run as a file but please do not break up your submission any more than that. A run will contain results for all features you worked on.

Evaluation

- The unit of testing and performance assessment will be the video shot as defined by the track's common shot boundary reference. The submitted ranked shot lists for the detection of each feature will be judged manually as follows. We will take all shots down to some fixed depth (in ranked order) from the submissions for a given feature - using some fixed number of runs from each group in priority sequence up to the median of the number of runs submitted by any group. We will then merge the resulting lists and create a list of unique shots. These will be judged manually down to some depth to be determined by NIST based on available assessor time and number of correct shots found. NIST will maximize the number of

shots judged within practical limits. We will then evaluate each submission to its full depth based on the results of assessing the merged subsets. This process will be repeated for each feature.

- If the feature is perceivable by the assessor for some frame (sequence) however short or long then, then we'll assess it as true; otherwise false. We'll rely on the complex thresholds built into the human perceptual systems. Search and feature extraction applications are likely - ultimately - to face the complex judgment of a human with whatever variability is inherent in that.
- Runs will be compared using precision and recall. Precision-recall curves will be used as well as a measure which combines precision and recall: **(mean) average precision**(see below under Search for details).

5.4 Search

Submissions

- Participating groups may submit up to 10 prioritized runs. All runs will be evaluated.

Each interactive run will contain one result for each and every topic using the system variant for that run. Each result for a topic can come from only one searcher, but the same searcher does not need to be used for all topics in a run. If a site has more than one searcher's result for a given topic and system variant, it will be up to the site to determine which searcher's result is included in the submitted result. NIST will try to make provision for the evaluation of supplemental results, i.e., ones NOT chosen for the submission described above. Details on this will be available by the time the topics are released.

- For each topic in a run, participants will return the list of at most 1000 shots. Here is a [DTD for search results of one run](#), one for [results from multiple runs](#), and a [small example](#) of what a site would send to NIST for evaluation. Please check your submission to see that it is well-formed
- Please send your submission in an email to over at nist.gov. Indicate somewhere (e.g., in the subject line) which group you are attached to so that we match you up with the active participant's database. Send all of your runs as one file or send each run as a file but please do not break up your submission any more than that. Remember, a run will contain results for all of the topics.

Evaluation

- The unit of testing and performance assessment will be the video shot as defined by the track's common shot boundary reference. The submitted ranked lists of shots found relevant to a given topic will be judged manually as follows. We will take all shots down to some fixed depth (in ranked order) from the submissions for a given topic - using some fixed number of runs from each group in priority sequence up to the median of the number of runs submitted by any group. We will then merge the resulting lists and create a list of unique shots. These will be judged manually to some depth to be determined by NIST based on available assessor time and number of correct shots found. NIST will maximize the number of shots judged within practical limits. We will then evaluate each submission to its full depth based on the results of assessing the merged subsets. This process will be repeated for each topic.
- Per-search measures:
 - average precision (definition below)
 - elapsed time (for all runs)
- Per-run measure:
 - mean average precision (MAP):

Non-interpolated **average precision**, corresponds to the area under an ideal (non-interpolated) recall/precision curve. To compute this average, a precision average for each topic is first calculated. This is done by computing the precision after every retrieved relevant shot and then averaging these precisions over the total number of retrieved relevant/correct shots in the collection for that topic/feature or the maximum allowed result set (whichever is smaller). Average precision favors highly ranked relevant

documents. It allows comparison of different size result sets. Submitting the maximum number of items per result set can never lower the average precision for that submission. The topic averages are combined (averaged) across all topics in the appropriate set to create the non-interpolated **mean average precision (MAP)** for that set. (See the [TREC-10 Proceedings appendix on common evaluation measures](#) for more information.)

6. Milestones:

The following was the schedule for 2004.

12. Jan

NIST sent out Call for participation in TRECVID 2004

16. Feb

Applications for participation in TRECVID 2004 due at NIST

1 Mar

Final versions of TRECVID 2003 papers due at NIST

Training data (i.e., 2003 development and/or test data) available from LDC

30. Apr

Guidelines complete

12. Jul

Test data, common shot reference, and key frames available on IDE harddrives

16. Jul

Shot boundary test collection DVDs shipped by NIST

13. Aug

Search topics available from TRECVID website.

16. Aug

Shot boundary detection submissions due at NIST for evaluation.

23. Aug

Feature extraction task submissions due at NIST for evaluation.

Feature extraction donations due at NIST

25. Aug

Feature extraction donations available for active participants

7. Sep

Story segmentation submissions due at NIST for evaluation

27 Aug

Results of shot boundary evaluations returned to participants

30. Aug - 15. Oct

Search and feature assessment at NIST

20. Sep

Results of feature extraction evaluations returned to participants

22. Sep

Search task submissions due at NIST for evaluation

23. Sep

Results of story segmentation evaluations returned to participants

17. Oct

Information on [processing complexity](#) of shot boundary runs due at NIST

18. Oct

Results of search evaluations returned to participants

24. Oct

[Speaker proposals](#) due at NIST

31. Oct

[Notebook papers](#) due at NIST

8. Nov

Workshop registration closes

10. Nov

Copyright forms due back at NIST (see Notebook papers for instructions)

15,16 Nov

TRECVID Workshop at NIST in Gaithersburg, MD

19. Nov

[Workshop papers publicly available \(slides added as they arrive\).](#)

[Some feedback from the workshop discussions](#)

1. Mar 2005

Final versions of TRECVID 2004 papers due at NIST

7. [Guideline issues and resolutions:](#)

8 Contacts:

- Coordinators:
 - [Alan Smeaton](#), and [Wessel Kraaij](#),
- NIST contact:
 - [Paul Over](#),
- Email lists:
 - Information and discussion for active workshop participants
 - trecvid2004@nist.gov
 - [archive open to active participants only](#)
 - NIST will subscribe the contact listed in your application to participate when we have received it. Additional members of active participant teams will be subscribed by NIST if they send email to [Lori Buckland](#), indicating they want to be subscribed, the email address to use, their name, and providing the TRECVID 2004 active participant's password. Groups may combine the information for multiple team members in one email.

Once subscribed, you can post to this list by sending you thoughts as email to trecvid2004@nist.gov, where they will be sent out to everyone subscribed to the list, i.e., the other active participants.

- General (annual) announcements about TRECVID (no discussion)
 - trecvid@nist.gov
 - [open archive](#)
 - If you would like to subscribe, logon using the logon to which you would like trecvid email to be reflected. Send email to lori.buckland@nist.gov and ask her to subscribe you to trecvid. This list is used to notify interested parties about the call for participation and broader issues. Postings will be infrequent.

Last updated: Thursday, 17-Feb-2005 19:01:38 UTC

Date created: Monday, 19-Nov-01

For further information contact [Paul Over](#), [NIST contact](#),

