

TREC Video Retrieval Evaluation

TRECVID 2021

Ian Soboroff¹, George Awad¹², Asad Butt¹³, Keith Curtis¹, Angela Ellis¹, Darrin Dimmick¹, Jonathan Fiscus⁴, Afzal Godil⁴

Yooyoung Lee⁴, Andrew Delgado⁴, Jesse Zhang⁴, Lukas Diduch⁴⁵, Eliot Godard⁴, Baptiste Chocot⁴, Yvette Graham⁶, Gareth Jones⁶, Georges Quénot⁷, Jeffrey Liu⁸

¹Retrieval Group / ⁴Multimodal Information Group
Information Access Division, Information Technology Laboratory, NIST

²Georgetown University

³John Hopkins University

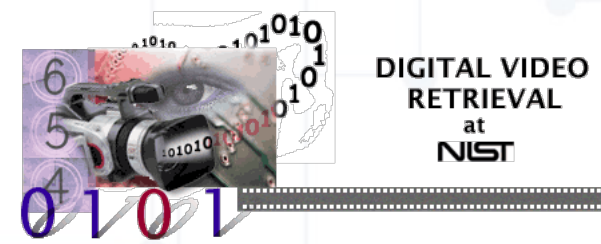
⁵Dakota Consulting, Inc

⁶Dublin City University, Ireland

⁷Laboratoire d'Informatique de Grenoble, France

⁸MIT Lincoln Laboratory

What is TRECVID ?



Workshop series
(2001 – present)

Started as a track in the TREC (Text REtrieval Conference) evaluation benchmark

Became an independent evaluation benchmark since 2003

Focus: **content-based** video analysis, retrieval, detection, etc

Provides data, tasks, and uniform, appropriate scoring procedures

Aims for realistic system tasks and test collections using:

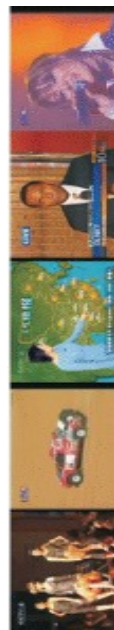
- Unfiltered data
- Focus on relatively high-level functionality (e.g. interactive search)
- Measurement against human abilities

Forum for:

- Exchange of research ideas
- The discussion of research methodology – what works, what doesn't, and why



US TV News



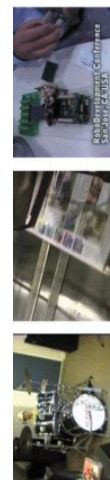
Int. TV News



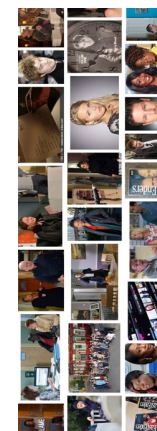
Dutch TV Prog.



Security Cameras



Web Video



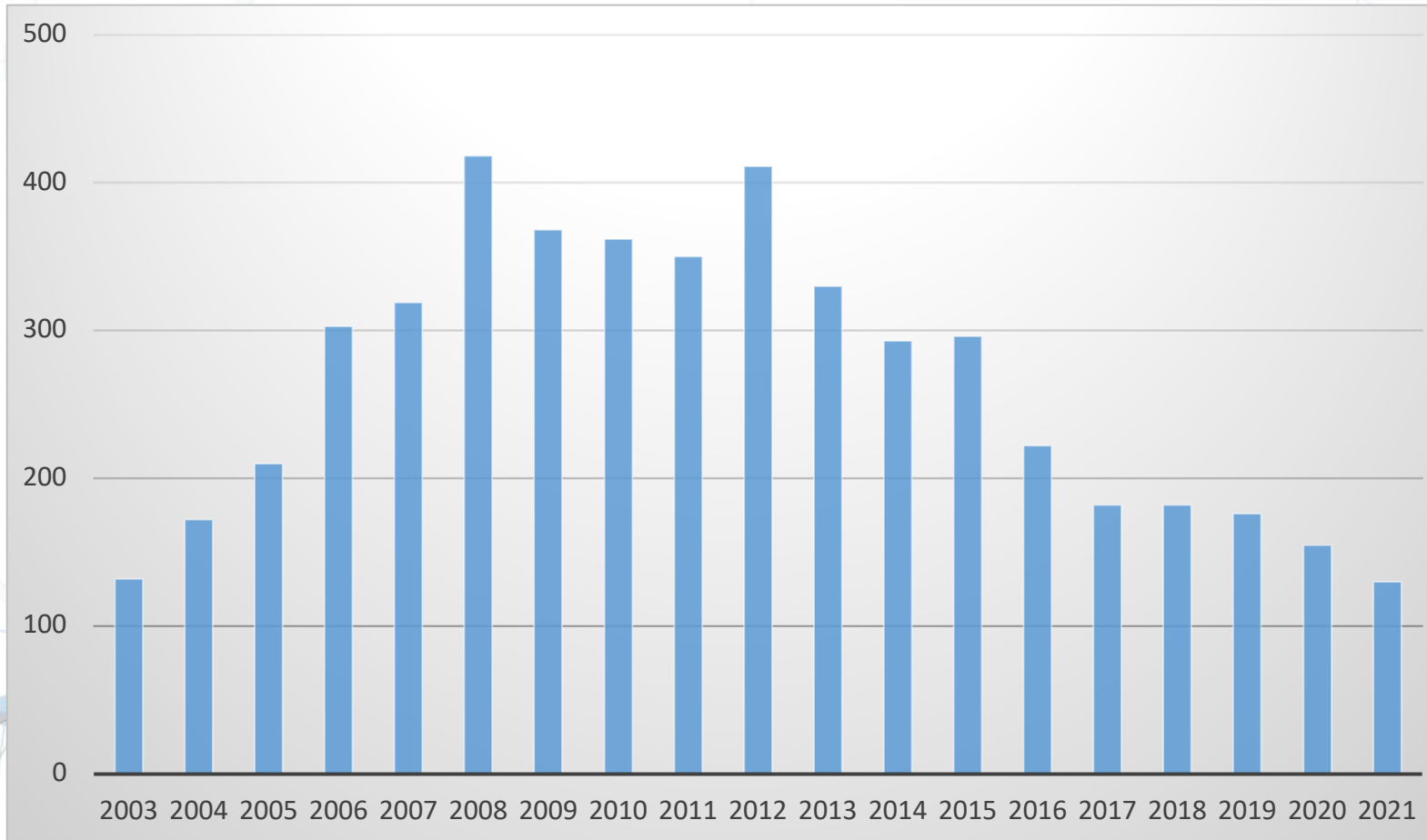
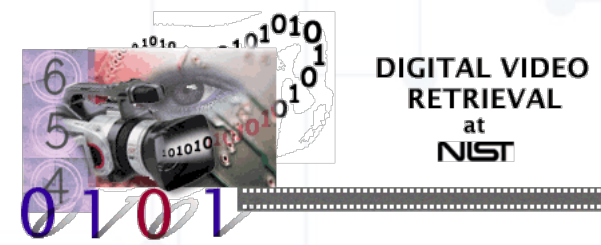
Soap Opera



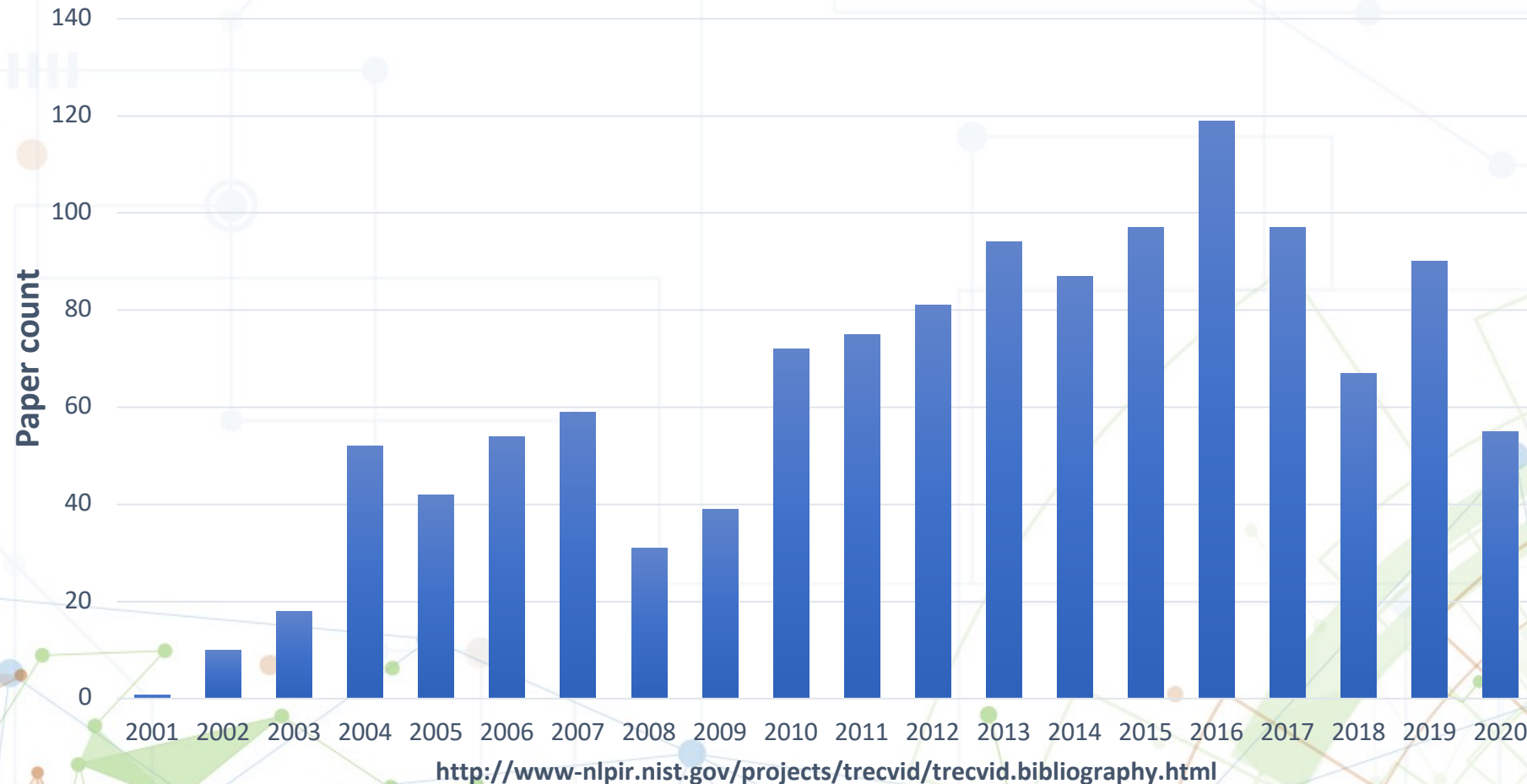
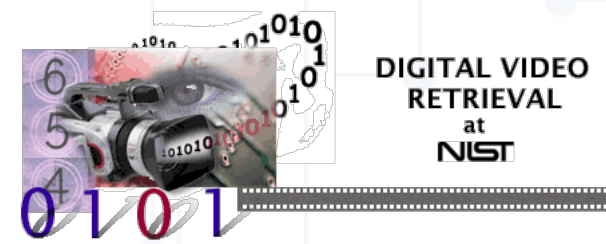
Civil Air Patrol

Datasets (2003 – 2021)

Annual Workshop Authors



Bibliography



TRECVID 2021 Tasks & Data



Ad-hoc Video
Search (AVS)

Instance
Search (INS)

Video to Text
(VTT)

Video
Summarization
(VSUM)

Activities in
Extended
Videos (ACTEV)

Disaster Scene
Description &
Indexing (DSDI)

Vimeo CC
(V3C1)

1000 h testing data,
~1 M shots, title,
keywords,
description

~2000 h dev data

20 *new* textual
queries

20 progress queries
[fixed 2019 – 2021]

BBC
Eastenders

*464 h testing data
471,527 shots,
transcripts*

*1 episode dev video +
external web resources*

*20 *new* image/video
Person + Action
queries*

*20 progress queries
[fixed 2019-2021]*

Vimeo CC
(V3C2)

1977 Short Vimeo
testing video clips

10,862 dev videos and
captions from Twitter
Vine and Flickr

Two tasks:
- Fill-in-the-blanks
- Generate automatic
video description

BBC
Eastenders

Summarize important
character events
during 8 – 12 weeks of
episodes.

image queries of 5
surprise characters

Two tasks:
- Given predefined
questions
- unknown questions

VIRAT

*10 h multi-camera
video surveillance
dataset*

*Detection and
Localization of
activities*

35 activities evaluated

LADI

20,000+ dev images
from airborne low
altitude disaster
events

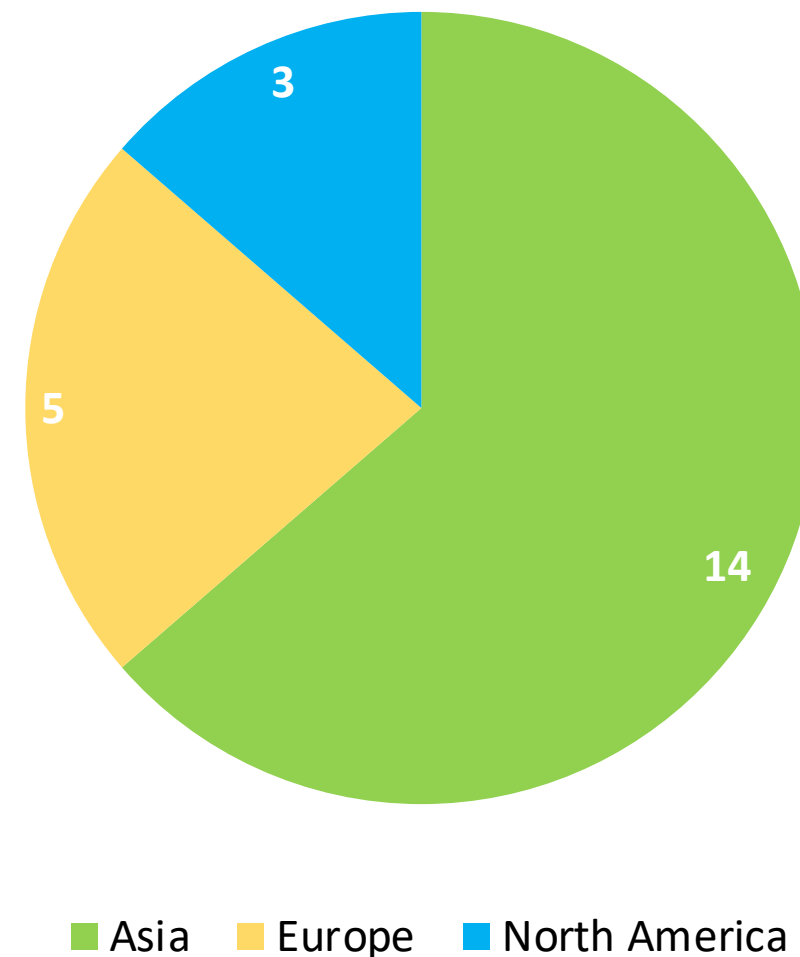
~6 h testing data from
natural disaster event

32 known features
identified by search
& rescue personnel

TRECVID 2021 Finishers (22 / 39)

Groups Finished	Task Code	Task Name
8	AVS	Ad-hoc Video Search
6	ACTEV	Activities in Extended Video
5	VTT	Video to Text Description
3	VSUM	Video Summarization
3	INS	Instance Search
3	DSDI	Disaster Scene Description & Indexing

Unique finishing teams



TRECVID 2020 Finishers (22 / 39)



Team	Organization
VCL_CERTH	Centre of Research and Technology Hellas, Information Technologies Institute
Adapt_team	ADAPT Research Centre
BUPT_MCPRL	Beijing University of Posts and Telecommunications
VIREO	Singapore Management University and City University of Hong Kong
INF	Carnegie Mellon University
DMT_CUC_01	Communiation University of China
FIU_UM	Florida International University ; University of Miami
kindai_ogu_osaka	Kindai University; Osaka Gakuin University; Osaka University
EURECOM	EURECOM
GodSpeed	Kuaishou Tech
ITI_CERTH	Information Technologies Institute, Centre for Research and Technology Hellas
kslab	Nagaoka University of Technology
NII UIT	National Institute of Informatics, Japan; University of Information Technology, VNU-HCMC, Vietnam
PKU_WICT	Peking University
RUC_AIM3	Renmin University of China

TRECVID 2020 Finishers (22 / 39)



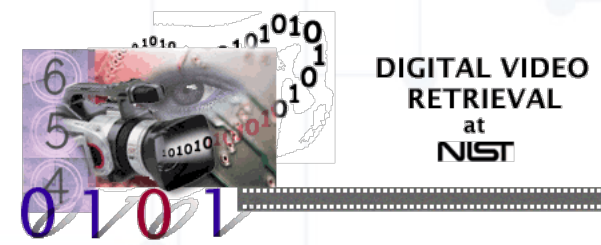
Team	Organization
RUCMM	Renmin University of China
UEC	The University of Electro-Communications, Tokyo
TokyoTech	Tokyo Institute of Technology
MMCUniAugsburg	University of Augsburg
WasedaMeiseiSoftbank	Waseda University; Meisei University ; SoftBank Corporation
WHU_NERCMS	Wuhan University
UCF	University of Central Florida

TRECVID 2020 Non-Finishers (17 / 39)



Team	Organization
PicSOM	Aalto University
ARETE	Arete_Associates
zju_om_center	Bingjing Institute of Zhejiang University; Zhejiang University; Carnegie Mellon University
DMT_CUC02	Communiation University of China
95	CUC
IMFD_IMPREESE	Univesity of Chile Millennium Institute Foundational Research on Data
drylwlsn_visual	drylwlsn_visual
fhg.iais.nm.map	Fraunhofer IAIS (NetMedia)
DVA	IIST DECU ISRO
chandra	Individual
LINKS	Links Foundation
RMIT_GORSE	RMIT
SejongRCV	Sejong University
Mitju	Tianjin University
TMGO	Tongji University, Nanjing University of Information Science and Technology
Ok	University of Science and Technology of China and Institute of Automation, Chinese Academy of Sciences
YildizTeam	Yildiz Technical University

Support



The running of TRECVID 2021 has been funded directly by:

- National Institute of Standards and Technology (NIST)
- Intelligence Advanced Research Projects Activity (IARPA)
- Defense Advanced Research Projects Agency (DARPA)

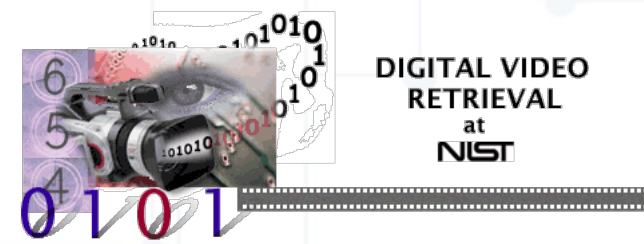
TRECVID is only possible because of the additional efforts of many individuals and groups around the world.



NIST Assessors

TRECVID 2021

Resources & Contributions



- Luca Rossetto and Heiko Schuldt (University of Basel) who Provided the V3C dataset, master shot reference and metadata
- Georges Quénot at LIG who provided the master shot reference for the IACC.3 video
- The LIMSI Spoken Language Processing Group and Vocapia Research provided ASR for the IACC.2-3 videos and HAVIC data
- Koichi Shinoda of the TokyoTechCanon team agreed to host a copy of the IACC.2 data
- Noel O'Connor and Kevin McGuinness at Dublin City University along with Robin Aly at the University of Twente worked with NIST and Andy O'Dwyer plus Rob Cooper at the BBC to make the BBC EastEnders video available for use in TRECVID
- Yvette Graham (DCU, Ir) for supporting the video-to-text Direct Assessment annotations
- Andrew Weinert and Jeffrey Liu (MIT Lincoln Laboratory) for supporting the LADI dataset, and Jarlath O'Neil-Dunne for contributing drone footage for DSDI task.

Agenda at a glance



1

Day 1 (Tuesday)

Adhoc Video Search

Video-to-Text

2

Day 2 (Wednesday)

Keynote Talk

Activities in Extended Videos

3

Day 3 (Thursday)

Disaster Scene Description & Indexing

Keynote Talk

Video Summarization

4

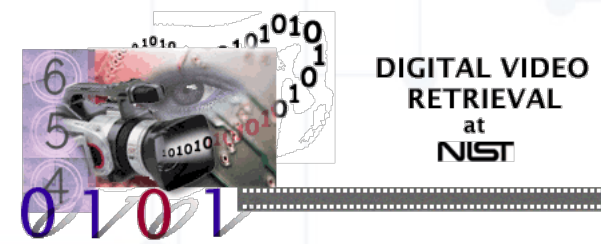
Day 4 (Friday)

Instance Search

Deep Video Understanding

Workshop Conclusion

Miscellany



- OpenMFC Sessions running everyday after TRECVID
- Slack
 - 2021 TrecVid Track Channels
 - Agenda/Schedule Channel
 - BlueJeans & gather town (breaks & after workshop hours) Meeting links
 - Random
- Online workshop proceedings:
 - Teams papers, task information, datasets, ground truth, etc) can be found at:
 - <https://www-nlpir.nist.gov/projects/tvpubs/tv.pubs.21.org.html>