
VideoCLEF 2008 Pilot

<http://ilps.science.uva.nl/Vid2RSS>

Classification of Dual Language Audio-Visual Content

Martha Larson, Eamonn Newman, Gareth Jones



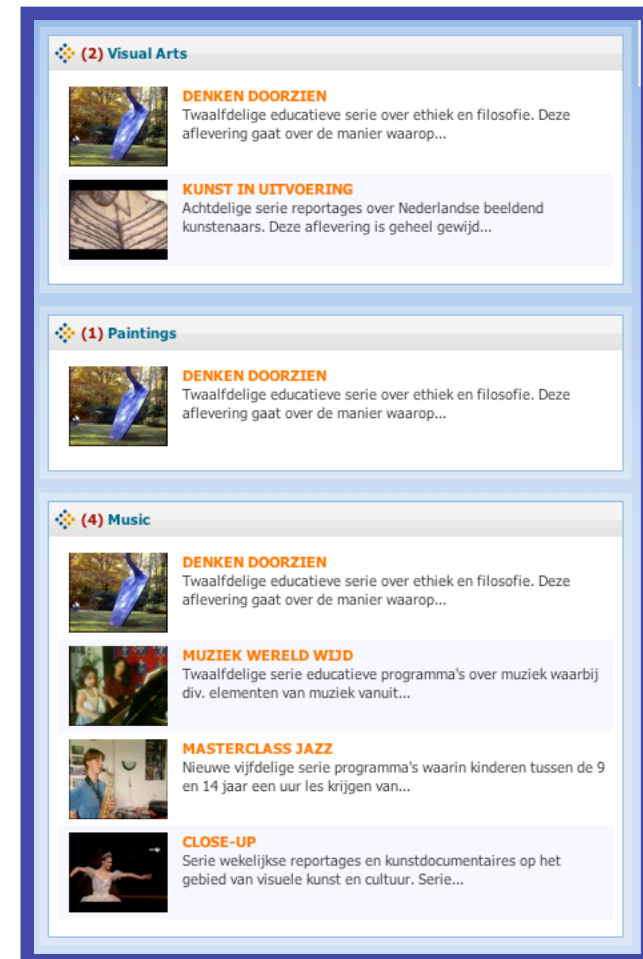
UNIVERSITEIT VAN AMSTERDAM

Goals of VideoCLEF

- Promote research on intelligent access to multimedia content in a multilingual environment
- Encourage exploitation multimodal information streams: speech transcripts, video content, metadata, ...
- Develop and evaluate multilingual video analysis tasks
- Extend the recent Cross-Language Speech Retrieval tracks into new challenges
- Be distinct from TRECVID

VideoCLEF Vid2RSS Task

- **Input:** Dual language video, including archival metadata and speech recognition transcripts in Dutch and English
- **Output:** Series of topic feeds (in RSS format) containing videos; one feed per thematic class



The screenshot displays a video feed interface with three thematic classes, each containing video thumbnails and titles:

- (2) Visual Arts**
 - DENKEN DOORZIEN**: Twaalfdelige educatieve serie over ethiek en filosofie. Deze aflevering gaat over de manier waarop...
 - KUNST IN UITVOERING**: Achtdelige serie reportages over Nederlandse beeldend kunstenaars. Deze aflevering is geheel gewijd...
- (1) Paintings**
 - DENKEN DOORZIEN**: Twaalfdelige educatieve serie over ethiek en filosofie. Deze aflevering gaat over de manier waarop...
- (4) Music**
 - DENKEN DOORZIEN**: Twaalfdelige educatieve serie over ethiek en filosofie. Deze aflevering gaat over de manier waarop...
 - MUZIEK WERELD WIJD**: Twaalfdelige serie educatieve programma's over muziek waarbij div. elementen van muziek vanuit...
 - MASTERCLASS JAZZ**: Nieuwe vijfdelige serie programma's waarin kinderen tussen de 9 en 14 jaar een uur les krijgen van...
 - CLOSE-UP**: Serie wekelijkse reportages en kunstdocumentaires op het gebied van visuele kunst en cultuur. Serie...

Data I

- 50 dual language videos (30 hours) from The Netherlands Institute for Sound and Vision (Beeld en Geluid)
- Videos are episodes of Dutch television shows, mostly documentaries
- Dutch is the main (matrix) language; English is an embedded language
- Embedded language is spoken mainly by interviewees

Data II

- Videos are accompanied by Dutch-language **archival metadata records**
- Metadata includes series title, episode title, description, date of broadcast and other production information
- **Speech recognition transcripts** in MPEG-7 format supplied by the University of Twente (both Dutch and English transcripts)
- Shot-level keyframes supplied by Dublin City University

Subtasks of Vid2RSS

Classification Task (Main Task)

- Assign videos to thematic classes using speech recognition transcripts only (required)
- Use combination of metadata and speech recognition transcripts to perform classification

Translation Task

- Translate output RSS-feeds (e.g., into English)

Keyframe Extraction Task

- Select a keyframe to provide a semantic representation of the entire video to be used to depict the video in the feed.

FAQs

Why Dual Language Video?

With appropriate access techniques, information seekers can find spoken content in their own language which is embedded in an archive with an unfamiliar matrix language.

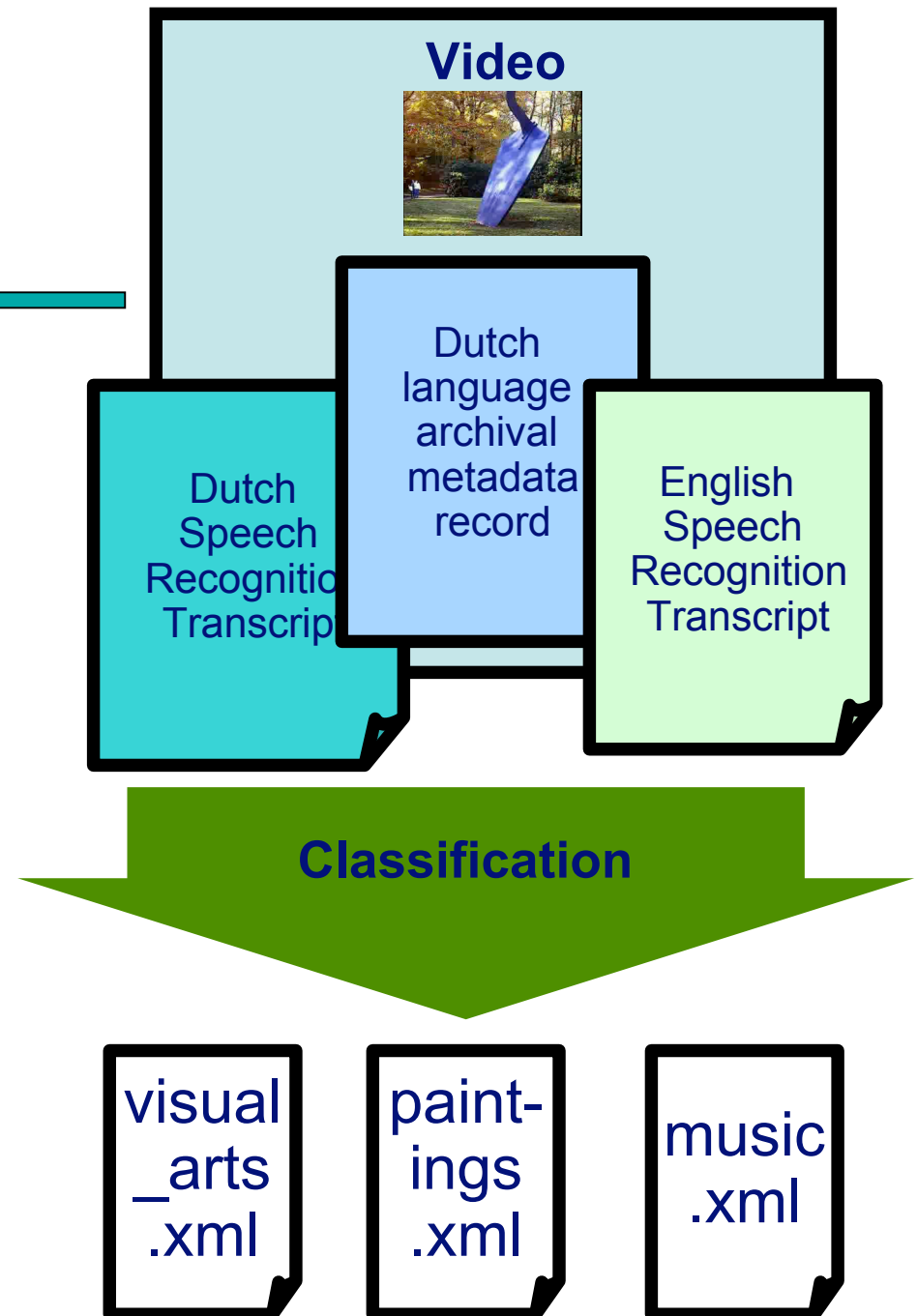
Why RSS?

Task results in RSS-format can be directly visualized in a feed reader. They can be immediately assessed by end-users, e.g., archive staff. RSS-format is trivial to generate.

Classification

Motivation

- Thematic subject labels encode high-level semantics
- The subject labels have known utility for search
- Ground truth available
- **Labels:** Archeology, Architecture, Chemistry, Dance, Film, History, Music, Paintings, Scientific research and Visual Arts



Classification Challenges

- Variation of vocabulary
- Interviewees do not necessarily use topic-specific vocabulary
- Speech recognition errors
- Participants must collect their own training data
- Not a challenge: Feed generation

(3) Art

The Glass Photos
Photo negatives found in the Amsterdam City Archive by 19th Century painter G.H. Breitner

Tormented Genius
This episode explores the topic of tormented genius in particular the link between creativity and...

Deteriorating Art
Research about the cause of deterioration of art and the role of art historians, restoration and...

(4) Music

DENKEN DOORZIEN
Twaalfdelige educatieve serie over ethiek en filosofie. Deze aflevering gaat over de manier waarop...

MUZIEK WERELD WID
Twaalfdelige serie educatieve programma's over muziek waarbij drie elementen van muziek vanuit...

MASTERCLASS JAZZ
Nieuwe vijfdelige serie programma's waarin kinderen tussen de 9 en 14 jaar een uur les krijgen van...

CLOSE-UP
Serie wekelijkse reportages en kunstdocumentaires op het gebied van visuele kunst en cultuur. Serie...

(2) Visual Arts

DENKEN DOORZIEN
Twaalfdelige educatieve serie over ethiek en filosofie. Deze aflevering gaat over de manier waarop...

KUNST IN UITVOERING
Achtdelige serie reportages over Nederlandse beeldend kunstenaars. Deze aflevering is geheel gewijd...

(1) Paintings

DENKEN DOORZIEN
Twaalfdelige educatieve serie over ethiek en filosofie. Deze aflevering gaat over de manier waarop...

RunID	micro-averaged f-score	macro-averaged f-score	feature language	test doc rep	site
CUT-C1R1 [▲]	0.15	0.27	en/nl	asr	CUT
CUT-C1R2	0.11	0.14	en/nl	asr	CUT
CUT-C2R1	0.13	0.26	en/nl	asr/md	CUT
CUT-C2R2	0.13	0.17	en/nl	asr/md	CUT
dcu_run1 [▲]	0.41	0.54	nl	asr	DCU
dcu_run2 [▲]	0.25	0.47	en	asr	DCU
dcu_run3 [▲]	0.28	0.58	nl	asr	DCU
dcu_run4	0.28	0.59	en	asr	DCU
dcu_run5	0.29	0.43	nl	md	DCU
MIRACLE-CNL	0.46	0.49	nl	asr	MIRACLE
MIRACLE-CNLEN	0.39	0.27	nl/en	asr	MIRACLE
MIRACLE-CNLMeta [▲]	0.47	0.47	nl	asr/md	MIRACLE
uams08m	0.18	0.17	nl	md	UAms
uams08asrd	0.10	0.41	nl	asr	UAms
uams08masrd	0.15	0.45	nl	asr/md	UAms
uams08asrde	0.09	0.14	nl/en	asr	UAms
uams08masrde	0.09	0.33	nl/en	asr/md	UAms
SINAI-Class-I	0.51	0.49	nl	asr	SINAI
SINAI-Class-II	0.53	0.51	en	asr	SINAI
SINAI-Class-I-Trans	0.10	0.40	nl	md	SINAI

Classification: What worked

- Archival metadata and/or Dutch speech recognition transcripts
- Wikipedia, but also general Web, as source of training data
- k-NN/1-NN classifier achieved good precision
- Simplistic retrieval approach: using class labels as queries and video as documents

Classification: Lessons learned

- Task is not trivial
- Archival metadata and speech recognition transcripts both good feature sources
- Features from speech transcripts of the embedded language (here, English) not helpful
- Performance on certain classes (e.g., Music) was quite acceptable
- Need an evaluation metric that captures human intuitions of performance

Translation

- Carried out by Chemitz Univeristy of Technology
- **What worked:** Google's AJAX language API
- **Assessment:**
 - Translation evaluated with 3 human assessors rating adequacy and fluency of the translations
 - 2.8 for adequacy (on scale of 1-5)
 - 3.5 for fluency (on scale of 1-5)
- **Lesson Learned:** Translation of sufficient quantity to make Dutch-language episode descriptions accessible to non-Dutch speaking English speakers

Keyframe Extraction I

- Carried out by MIRACLE
- Keyframe was selected from set of keyframes provided (one per shot)
- **What worked:** MIRACLE chose the keyframe whose speech transcript was most representative for the episode

Keyframe Extraction II

- **Assessment**
 - 5 human assessors chose keyframe better representative of video episode
 - Choice was between manually selected baseline and automatically selected keyframe
 - In 44% of the cases, automatic keyframe was chosen by human
- **Lesson Learned:** Automatic keyframe selection competitive with manual keyframe selection

Outlook

- Vid2RSS **Classification** scale-up:
more data, more classes
- New tasks under consideration
 - **Favorite filtering:** *Topic independent selection of videos the user's prefer*
 - **Non-Dutch quote retrieval:** *Mining the collection for statements useful to non-Dutch speakers*
 - **Personalized keyframe selection:** *Choosing representative keyframes most useful to a particular information seeker*
 - **Finding related resources:** *Identifying information from non-Dutch sources to support understanding of the video*
- Please join us at Friday's breakout session!