



University of Amsterdam at VideoCLEF

Vid2RSS 2008 Pilot Task

Jiyin He

Xu Zhang

Wouter Weerkamp

Martha Larson



Submitted runs

- Can archival metadata achieve good performance?
- Can speech transcripts be exploited?
- Does using the embedded language help?

Run	Dutch transcripts	English transcripts	metadata
uams08m			X
uams08asrd	X		
uams08masrd	X		X
uams08asrde	X	X	
uams08masrde	X	X	X



Experimental Setup

- SVM with linear kernel
- Training data drawn from Wikipedia
 - Using the class label as query worked well
 - Mapping the class label to a Wikipedia category was less successful
- χ^2 feature selection



Results

- Metadata and metadata combined with Dutch asr were most successful
- “Music” was best class (precision 0.57 recall 0.36)

measure	uams08m	uams08asrd	uams08masrd	uams08asrde	uams08masrde
micro average precision	0.13	0.08	0.11	0.07	0.07
micro average recall	0.28	0.16	0.23	0.14	0.14
micro f-score	0.18	0.10	0.15	0.09	0.09
macro average precision	0.11	0.44	0.44	0.09	0.32
macro average recall	0.38	0.38	0.46	0.35	0.35
macro f-score	0.17	0.41	0.45	0.14	0.33



Conclusions and Outlook

- Wikipedia was fine as source of training data, but not optimal
- Transcripts for the embedded language could not be exploited
- We conjecture that the main difficulties are:
 - Mismatch between training and test data
 - High level of speech recognition error (due to lack of language detection)