

Opening Up IIHS Video with SpokenMedia

Brandon Muramatsu

Andrew McKinney

Peter Wilkins

May 2010

Citation: Muramatsu, B., McKinney, A., Wilkins, P. (2010). Opening Up IIHS Video with SpokenMedia.
Presented at OpenCourseWare Consortium Global 2010: Hanoi, Vietnam, May 7, 2010.

Citation: Muramatsu, B., McKinney, A., Wilkins, P. (2010). Opening Up IIHS Video with SpokenMedia. Presented at OpenCourseWare Consortium Global 2010: Hanoi, Vietnam, May 7, 2010.

Unless otherwise specified, this work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License

Case Study of Using SpokenMedia for IIHS

- **Demonstrate transcripts and translations of IIHS videos**
- **Describe the process and our experiences**
 - Transcribe -> Edit -> Translate -> Present

The Indian Institute for Human Settlements (IIHS) will... “create India’s first independent National Innovation University focused on the challenges and opportunities of its urbanisation.”

*– Indian Institute for Human Settlements:
Curriculum Framework Version 3.0
January 2010*

*"The IIHS Website is our
commitment to a different
way of looking at things."*

*– Aromar Revi
5 January 2010*

*"The Institution will fail or scale
based on language."*

– Aromar Revi
5 January 2010

What did we do?



Four step process (simple)

First we used the SpokenMedia to do automatic transcription.


Next we did hand edit and translation steps.

Then we created a player for the presentation of the video and transcripts...

iihs INDIAN INSTITUTE FOR HUMAN SET

Challenges of Planning, Critical Challenges facing Cities of Global South, IIHS, and Challenges within the IIHS Curriculum Framework

Search transcript



00:15 07:36

Prof. Bish Sanyal

I think one central challenge of planning is legitimacy of government as a key institution that should be involved in bringing the change. The legitimacy of the government is under attack in many different parts of the world so that it has to be reestablished as a major planning institution. I think the second challenge for planning at least in the US or the west is that the market is

Audio English (US) Transcript English (US)

oeiT

7

Demo

How did we do it?



First we used the SpokenMedia to do automatic transcription.

How do we do it? Lecture Transcription

James Glass
glass@mit.edu



SPOKEN LANGUAGE SYSTEMS
MIT Computer Science and Artificial Intelligence Laboratory

- Spoken Lecture: research project
- Speech recognition & automated transcription of lectures
- Why lectures?
 - Conversational, spontaneous, starts/stops
 - Different from broadcast news, other types of speech recognition
 - Specialized vocabularies



9

Unless otherwise specified this work is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 3.0 United States License (creativecommons.org/licenses/by-nc-sa/3.0/us/)

Lecture Transcription

- Jim Glass and his group have years of research experience for spoken languages
- Lectures are a different type of spoken language
 - Much of the speech recognition research has focused on real time transcription of news broadcasts, or interactive voice response systems (telephone)
 - Broadcast news has something like 300 unique words in an hour long broadcast
 - Broadcast news is well structured, prepared copy (in studio via teleprompters), clear transitions between speakers, etc.
 - Lectures are conversational and spontaneous
 - Can use highly specialized vocabularies, engineering, physical sciences, mathematics

Spoken Lecture Project

James Glass
glass@mit.edu



- Processor, browser, workflow
- Prototyped with lecture & seminar video
 - MIT OCW (~300 hours, lectures)
 - MIT World (~80 hours, seminar speakers)

Supported with iCampus MIT/Microsoft Alliance funding

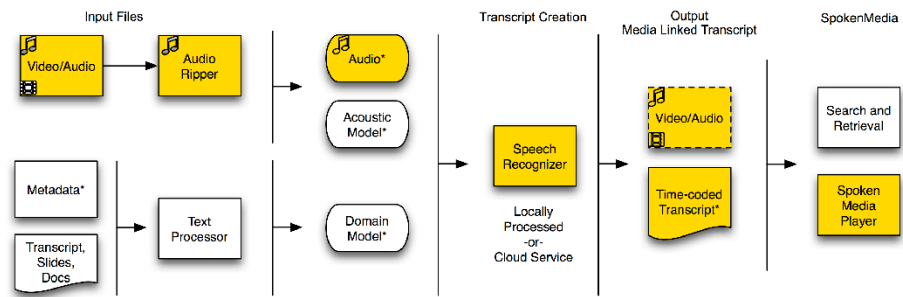


10

Spoken Lecture Project

- Supported by iCampus
- Includes the browser (which was just demo'd) the processor (back end lecture transcription) and a hand workflow to do the processing
- Approximately 400 hours of video indexed

SpokenMedia Process



We used a portion of the SpokenMedia process for the demo

We only used part of the process due to time constraints.

Audio separation, speech processing, time-coded transcript, and then presentation through a SpokenMedia player.

How did we do it?



Next we did hand edit and translation steps.

Edit & Translate: Accuracy

Automatic Transcription	Hand Transcription	Time Adjusted	Translated Hindi
I	I	I	मेरे खयाल से
think	think	think	
once	one	one	नयोजन की एक मुख्य चुनौती है
and	central		
so	challenge	central	
the	of		
challenger	planning	challenge of	
planning	is	planning	
nice	legitimacy	is	
legitimacy	of	legitimacy of	
of	government	government	सरकार की एक ऐसी मुख्य संस्थान के रूप में वैधता
government	as	as	

13

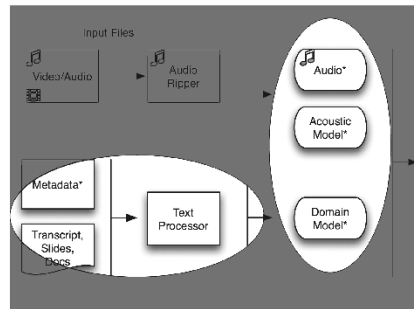
For this demo, we did computer-based automatic transcription, sent a file to IIHS for “editing” that consisted of performing a hand transcription (due to the format we sent, and the low accuracy of the automatic transcription in this case), a time alignment (though I actually feel that it’s “off” or “slow”) and then a hand translation by IIHS.

Automatic Speech Recognition Accuracy

■ Accuracy

- Domain Model and Speaker Model
- Internal validity measure
- Seed with transcript

Ongoing research by Jim Glass and his team @ MIT



Recognizer Accuracy

- Base accuracy is approximately 50% (generic domain and speaker models)
- Increase accuracy with speaker model up to 80-85%, and specific domain model
 - This approach is good for courses with multiple lectures by the same speaker
 - Domain models get more useful as more relevant text documents are indexed (keyword/noun phrase extraction)
- Initial results indicate that doing one 99% accurate (by hand/manual) transcript can help immensely for additional lectures by the same speaker
 - Better use of limited resources
- Search accuracy is closer to 90%, searches tend to be for unique words which the processor is better at recognizing

How did we do it?



Then we created a player for the presentation of the video and transcripts...

Challenges of Planning, Critical Challenges facing Cities of Global South, IIHS, and Challenges within the IIHS Curriculum Framework



Prof. Bish Sanyal

Search transcript

I think one central challenge of planning is legitimacy of government as a key institution that should be involved in bringing the change. The legitimacy of the government is under attack in many different parts of the world so that it has to be reestablished as a major planning institution. I think the second challenge for planning at least in the US or the west is that the market is

Audio English (US) Transcript English (US)



SpokenMedia today...

■ Features

SpokenMedia Player could
be used for MIT OCW Videos

- Video linked transcripts
 - Automated Lecture Transcript creation
 - Simple transcript editor (April 2010)
- SpokenMedia Player
 - “Bouncing Ball” (underline text) follow along
 - Search within a video
 - Multiple transcript language support

■ Challenges

- Accuracy (partial toolset)

What we have today

- It's not perfect, but a pretty good start
- Prototype has a number of useful features that demonstrate search interfaces and interaction interfaces

Where are we heading?

- Improved accuracy
- Search across multiple video transcripts
- New players with bookmarking, annotation, “paper-based video”
- Automate and improve processing
 - > Starting a lecture transcription service

Where are we heading?

- Transition from research project to service
- Explore new interactions—what we’re calling Rich Media Notebooks

Check it out for yourself

■ IIHS Demo:

<http://spokenmedia.mit.edu/demo/iihs/>

■ SpokenMedia Website:

<http://spokenmedia.mit.edu/>

■ Upload Videos for Automated Lecture Transcription

<http://sm.mit.edu/upload>

SpokenMedia AVAILABLE FOR
TESTING

If you want to try out
SpokenMedia, you can upload
video and we'll automatically
create a lecture transcript for you!

Upload Video

©2012 MIT. All rights reserved. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. For more information, see <http://creativecommons.org/licenses/by-nc-sa/4.0/>.

19

Thank You!

Brandon Muramatsu, mura@mit.edu

Andrew McKinney, mckinney@mit.edu

Peter Wilkins, pwilkins@mit.edu

20

Citation: Muramatsu, B., Kumar, M.S., Merriman, J., Wilkins, P. (2010). Project Greenfield: A New Way of Thinking about OpenCourseWare and Open Educational Resources for MIT. Presented at OpenCourseWare Consortium Global 2010: Hanoi, Vietnam, May 7, 2010.

Unless otherwise specified, this work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License