

Development and Evaluation of Indexed, Captioned, Searchable Videos for STEM Coursework

[Tayfun Tuna](#)¹, Jaspal Subhlok,¹
[Lecia Barker](#)², Varun Varghese,¹
Olin Johnson¹, Shishir Shah,¹



¹ University of Houston Department of Computer Science

² University of Texas School of Information

- Usage of video to deliver coursework online: commonplace
- Evidence for video usage:
 - MIT OpenCourseware ,
 - Apple's iTunes University
 - YouTube Edu



Khan Academy

2,718 Videos

6,707,463 Views



UC Berkeley

2,444 Videos

5,087,826 Views



MIT

1,781 Videos

3,288,603 Views



Stanford University

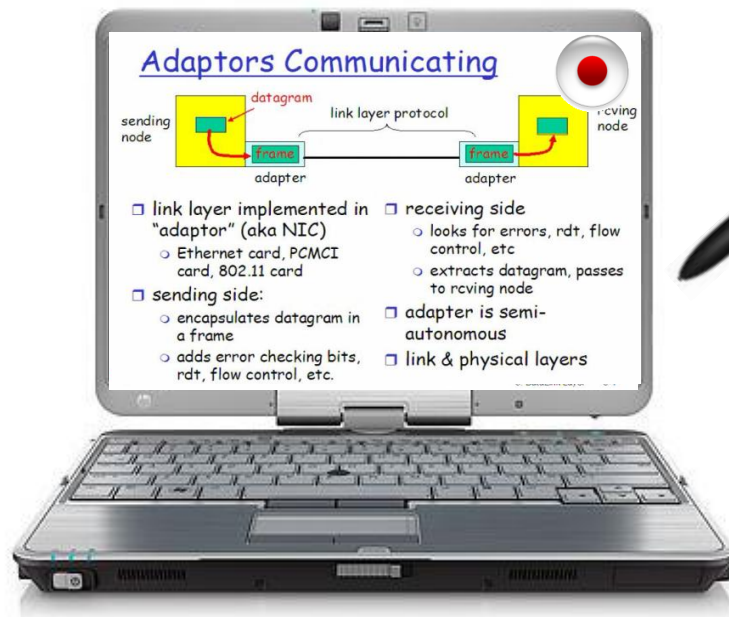
1,390 Videos

2,907,013 Views

➤ Classroom lecture videos widely used at University of Houston

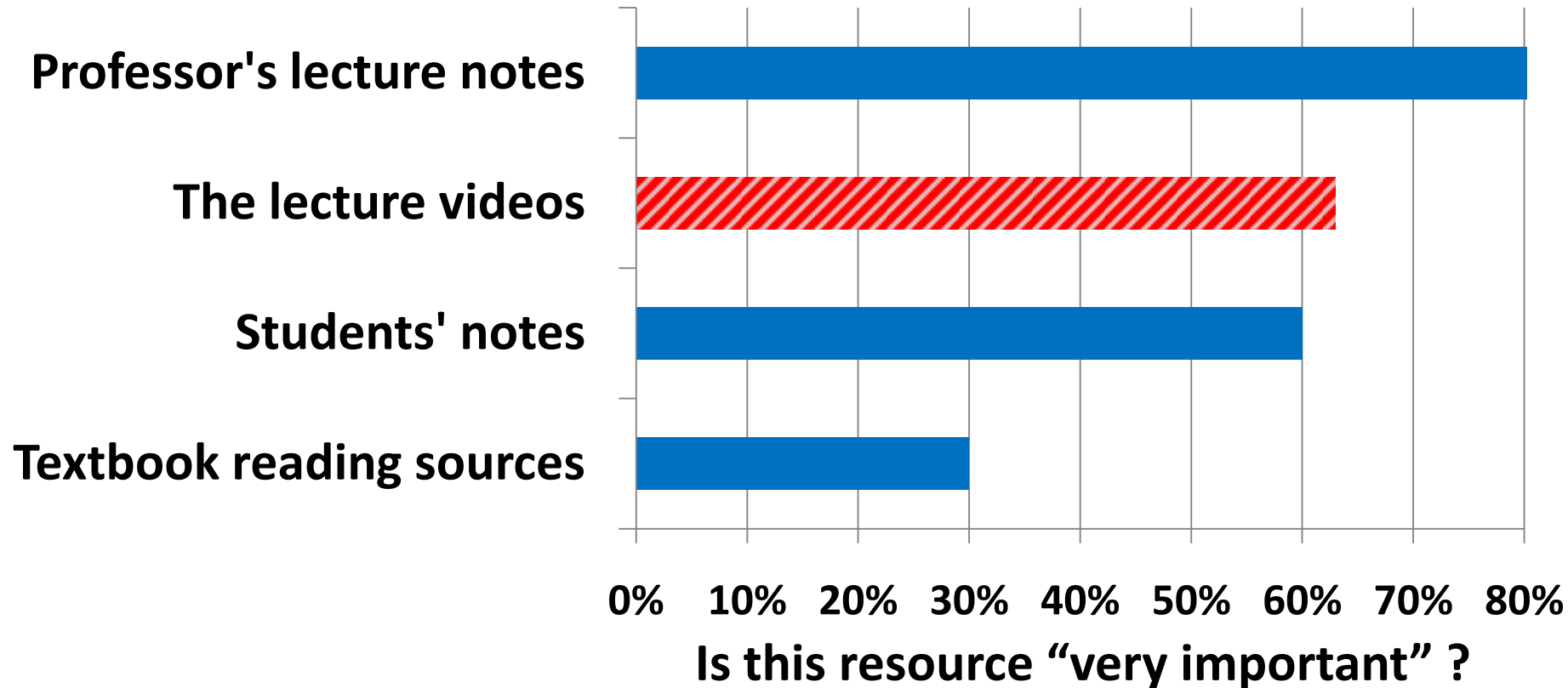


➤ Tablet PC recordings



Surveys, focus groups:

➤ Video lectures : powerful and versatile resource



➤ The biggest weakness of the video format : *Inability to quickly access the content of interest*

➤ ICS Videos

- **Indexing:** Segmented videos
- **Search:** Keyword search in video
- **Captioning:** Scrolling text for audio

Integrated in **ICS Video Player**

➤ **Project Goal:** *Quick access to video content*

ICS Video Player

linear ✕ 🔍

CHAPTERS

1 | 00:06 | Index 1

Computation of Wraparound Convolution

2 | 04:18 | Index 2

DFT Computation of Wraparound Convolution

3 | 07:59 | Index 3

Example

☞ One of the simplest types of linear convolutions is the **operation** (or averaging filter).

☞ Conceptually, each image pixel is replaced by the average within a square "window":

Input image Output image

☞ This may be expressed, **at most points** (without proving it here) as the wraparound convolution of the image with another image of a square with intensity $1/M$, where $M = \#$ pixels in the square:

Input image Image of square

14

Example
 One of the simplest types of linear convolutions is the operation (or averaging filter). Conceptually, each image pixel is replaced by the average within a square "window":

Linear Convolution by Zero Padding
 This may be expressed, at most points (without proving it here) as the wraparound convolution of the image with another image of a square with intensity 1/M, where M = # pixels in the square.

⏮ ⏪ ⏩ ⏭
🔊 🔇
⏮ ⏪ ⏩ ⏭

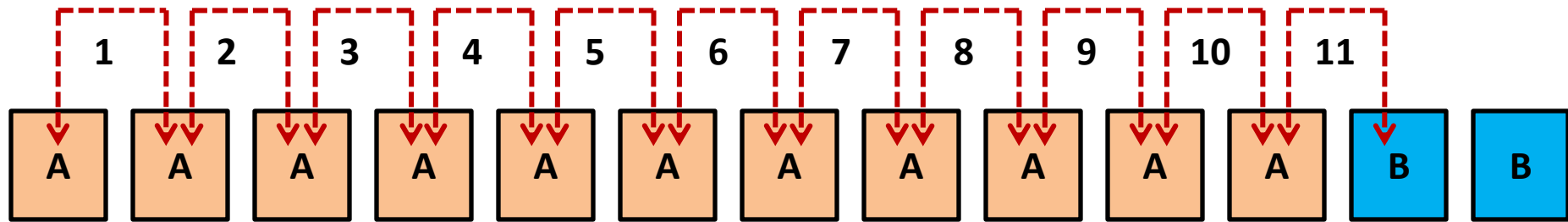
INDEXING

➤ **Indexing** : dividing video into segments;
User can start playback from any index frame

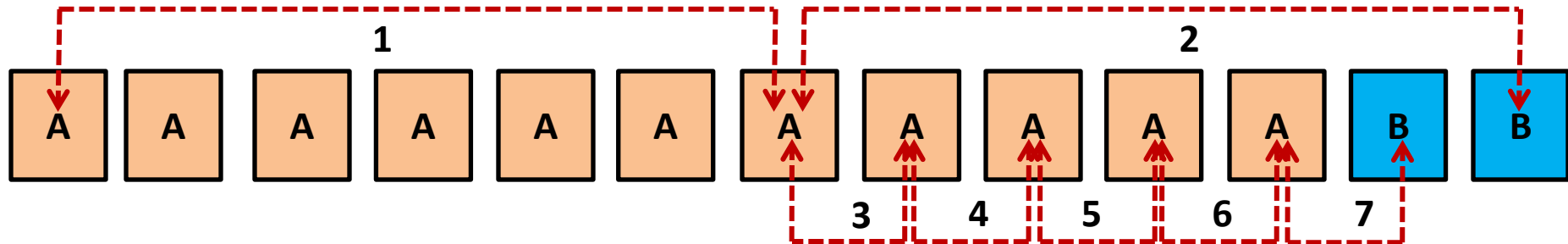
➤ Indexing Tasks :

- identify all ***transition points (TP)***
- select subset of TPs as **index points**

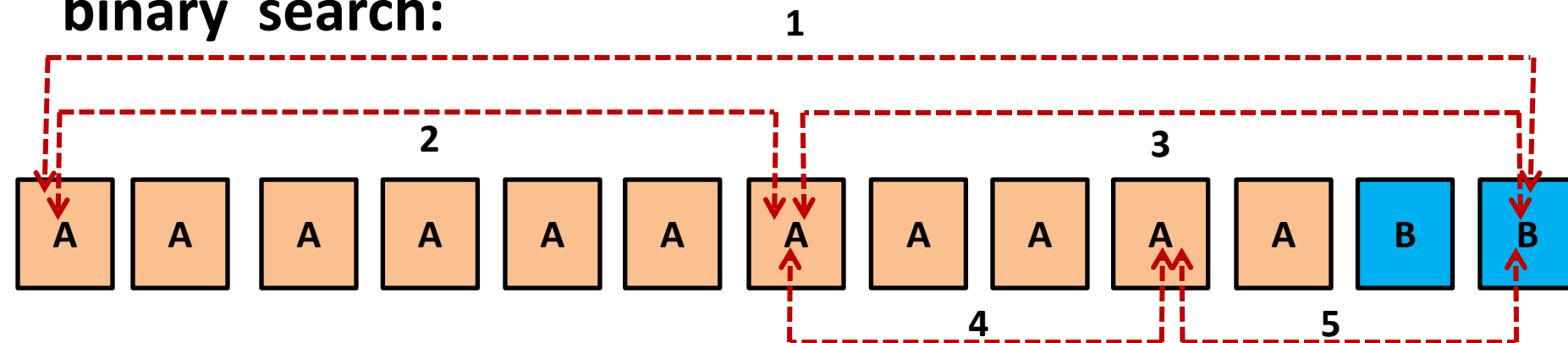
sequential:



sequential (with jumping):

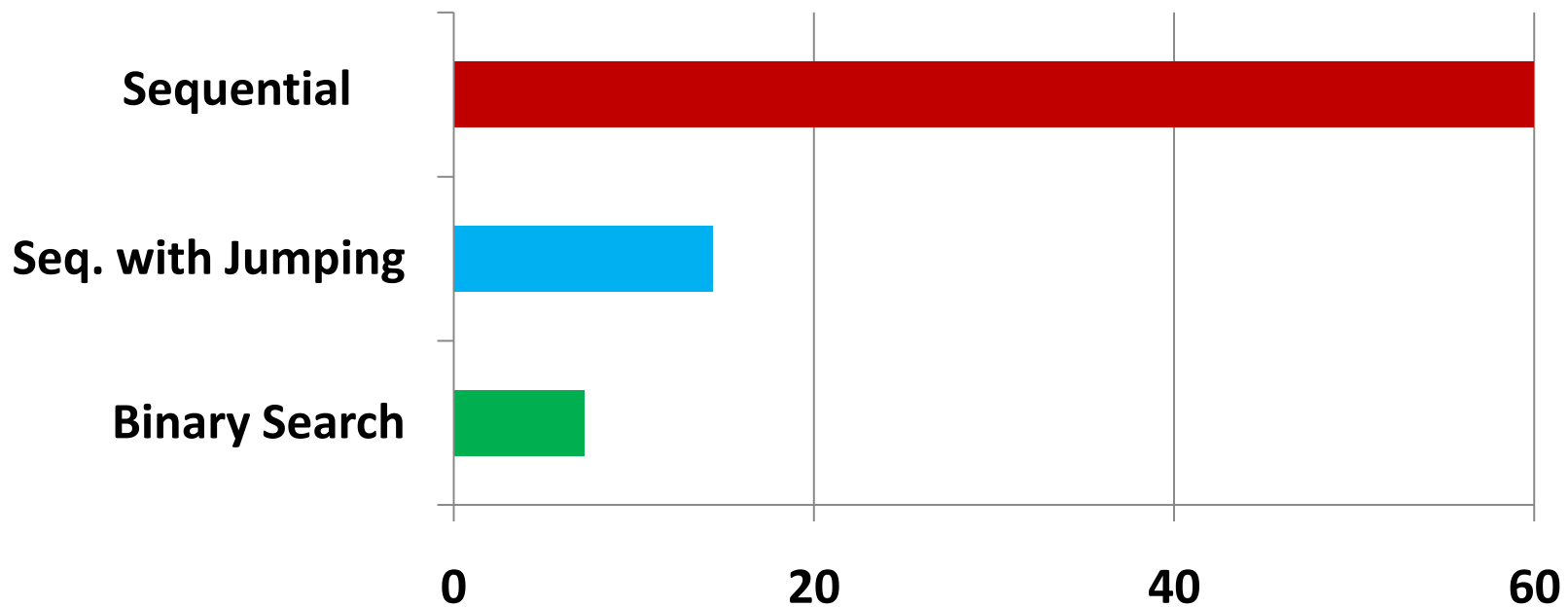


binary search:



Detection and Execution Performance

-Almost all Transition Points are detected by all alg.



Average time (mins) for indexing an hour of video

Transition points → Index points

- Large number of transition points hard to manage: 10s-100s
- Automatically select
 - fixed number of index points (15)
 - roughly uniformly spaced

KEYWORD SEARCH

ICS Video Player

linear

CHAPTERS

1 | 00:06 | Index 1

2 | 04:18 | Index 2

3 | 07:59 | Index 3

Example

One of the simplest types of **linear** convolutions is the **operation** (or averaging filter).

Conceptually, each image pixel is replaced by the average of the values within a square "window":

This may be expressed, **at most points** (without proving it here) as the wraparound convolution of the image with another image of a square with intensity $1/M$, where $M = \#$ pixels in the square:

Input image

Output image

average of values within square window

Input image

Image of square

Matches: Keyword (1)

Hide Keyword Matches

Matches: Keyword (1)

Hide Keyword Matches

14

- Keyword Search requires text detection in video frames
- Can be accomplished by OCR tools



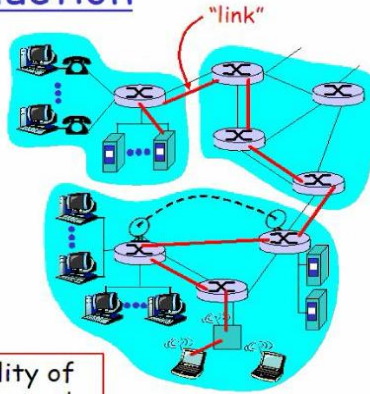
- Integration with ICS video framework?
- Accuracy on lecture video images?

Example images from lecture videos

Link Layer: Introduction

Some terminology:

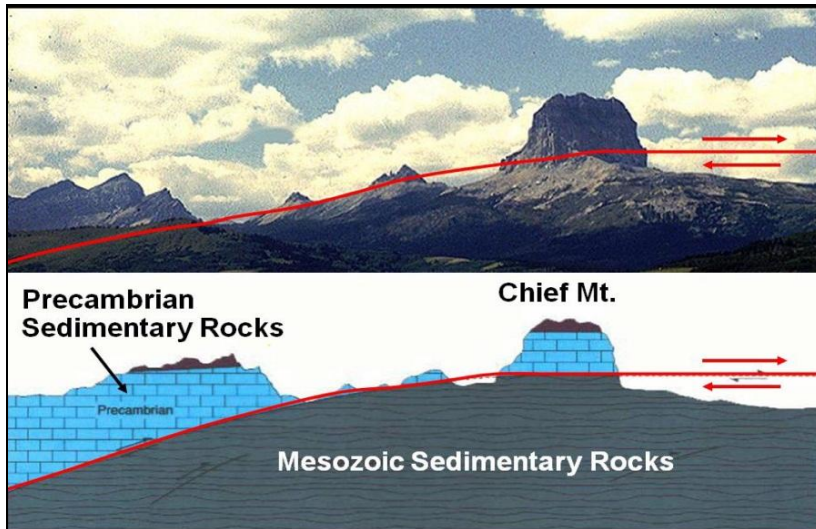
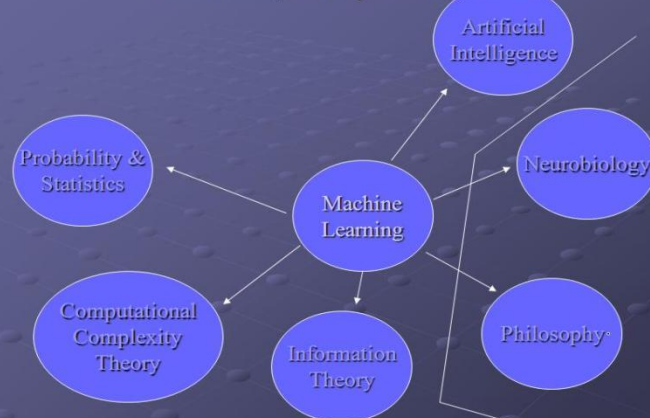
- hosts and routers are **nodes**
- communication channels that connect adjacent nodes along communication path are **links**
 - wired links
 - wireless links
 - LANs
- layer-2 packet is a **frame**, encapsulates datagram



data-link layer has responsibility of transferring datagram from one node to adjacent node over a link

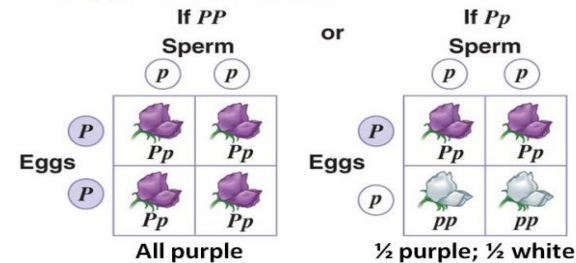
5: DataLink Layer 5-4

Multidisciplinary Field



Testcross

- How can you determine the genotype of a dominant phenotype?:
 - Purple-flowered pea plant is either **PP** or **Pp**: cross with a homozygous recessive (pp) white-flowered plant:



- By definition, the **testcross** is used to determine the genotype of an organism expressing a dominant phenotype by breeding with a recessive homozygote.

Outputs of OCR Tools

Question 3

Where did the story say that there was a statue raised in Mrs. Bethune's honor?

Washington, D.C.

Miami

Mayesville
South Carolina

EXTRACT

GOOCR

MODI

Where did the story say that there was a statue raised in Mrs. Bethune's honor?

_B-Nik is

Where did the story say that there was a statue raised in Mrs. Bethune's honor?

Washington D.C., Miami, Mayesville, South Carolina

Where did the story say that there was a statue raised in Mrs. Bethune's honor? Washington, D.C. Miami, Florida Mayesville, South Carolina

Segmentation of Text Regions

Enlargement with interpolation

Color Inversion

Segmentation of Text Region

Image Example



Noisy, $\sigma=10$ (MSE=100)

denoised (T=3, MSE=56)

39

Segmentation of Text Region and Enlargement

Image Example



Noisy, $\sigma=10$ (MSE=100)

denoised (T=3, MSE=56)

a) Original image

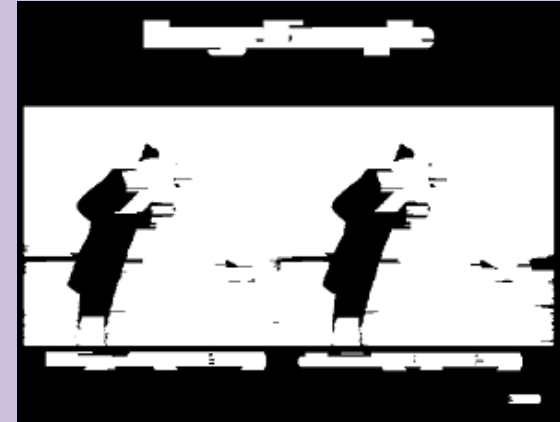
Image Example



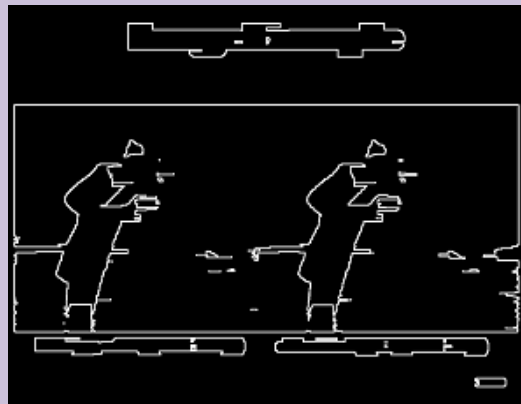
Noisy, $\sigma=10$ (MSE=100)

denoised (T=3, MSE=56)

b) Binarization



c) Dilation effect



d) Edge detection

Image Example

Noisy, $\sigma=10$ (MSE=100)

denoised (T=3, MSE=56)

e) Blob extraction

Image Example

Noisy, $\sigma=10$ (MSE=100)

denoised (T=3, MSE=56)

f) Enlargement

Inversions of Colors

Original Image R / G / B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

Inversion1

255-R / G / B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

Inversion2

R/ 255-G / B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

Inversion3

R/ G / 255-B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

Inversion4

255-R/ 255-G / B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

Inversion5

R/ 255-G / 255-B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

Inversion6 255-R/ G / 255-B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

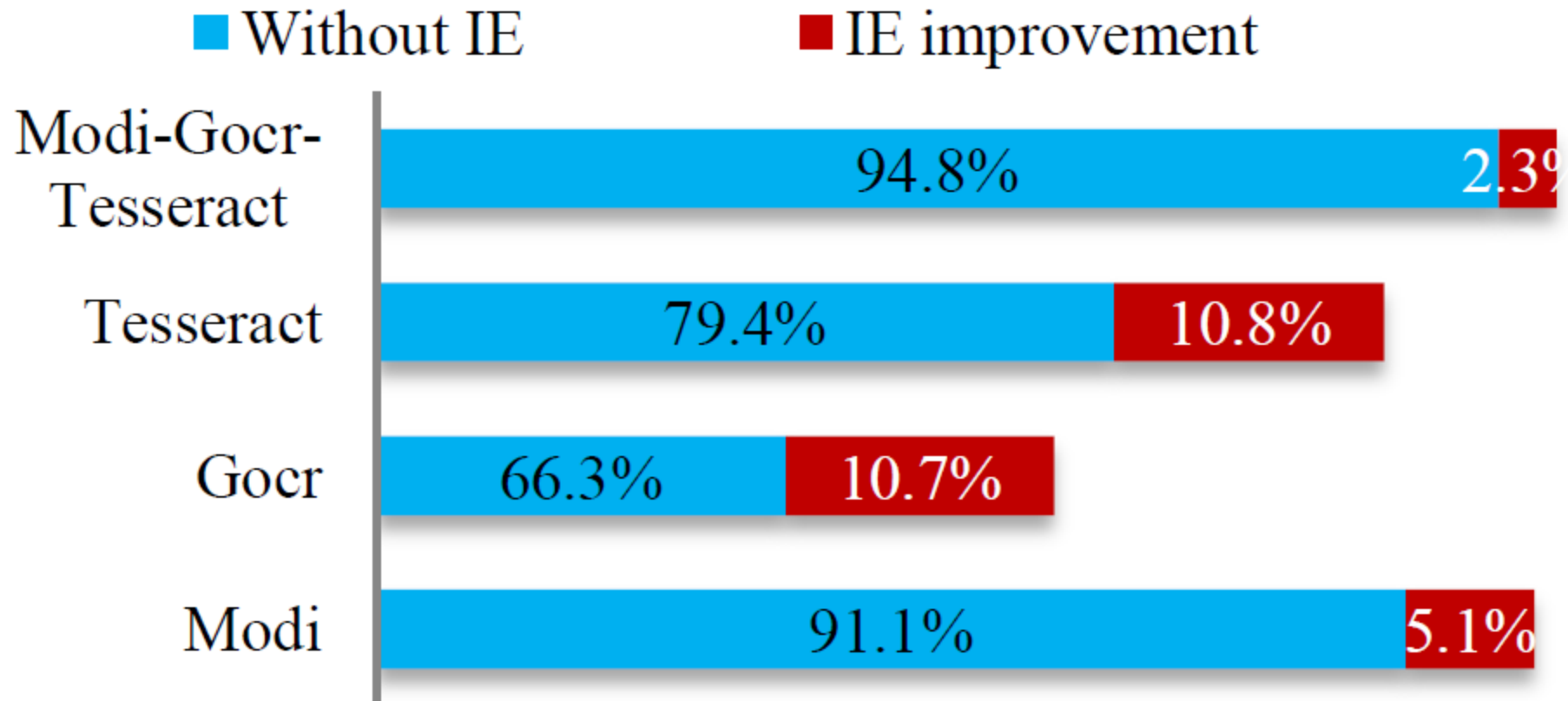
Inversion7

255-R / 255-G / 255-B



Input (source code): program in a computer programming language
The code is translated to object code and immediately executed.
Some programming languages are interpreted rather than compiled

Recognition accuracy with OCR Tools and Image Enhancement(IE)

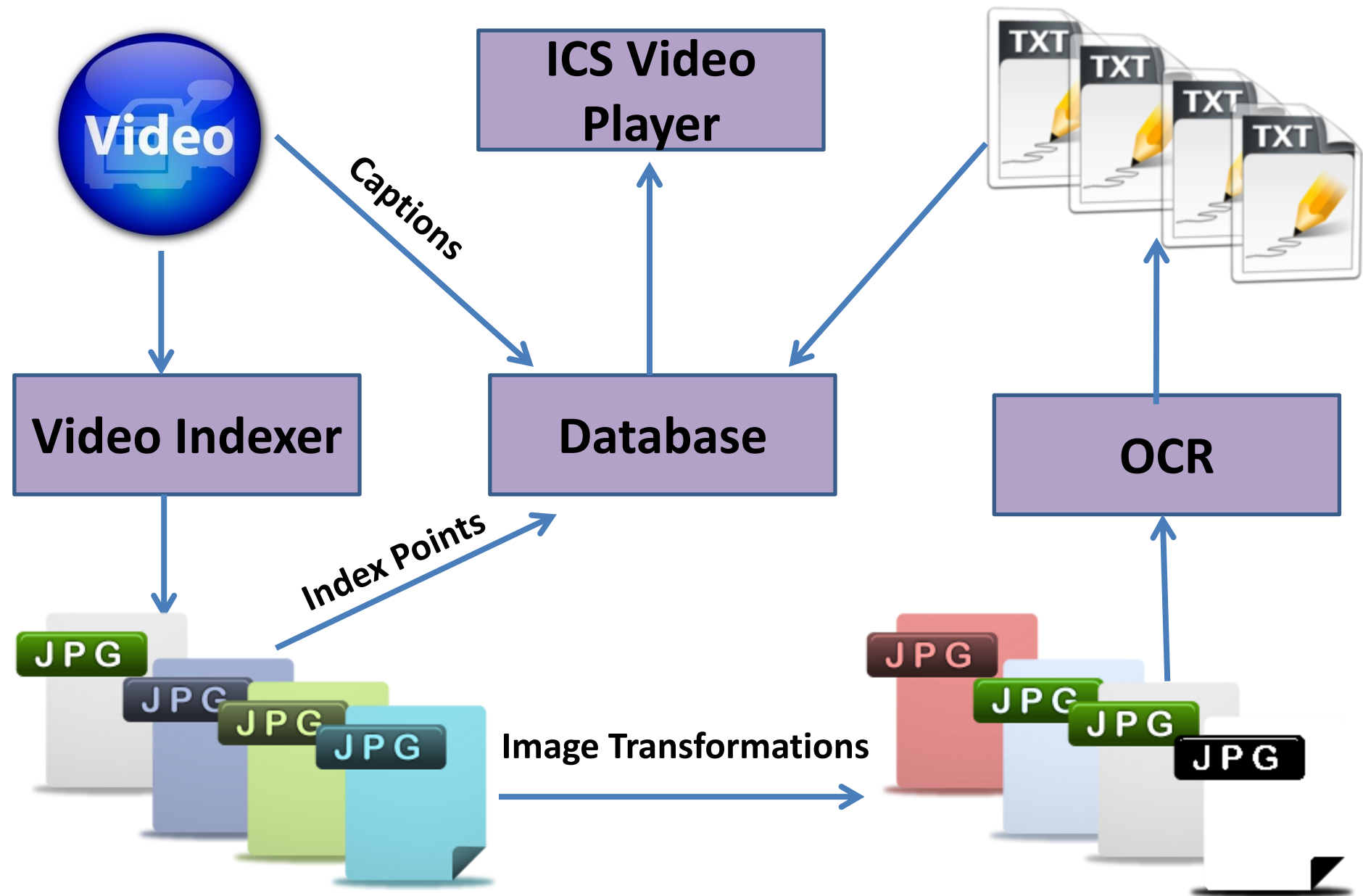


Maximum accuracy with all OCR engines with IE : **97.1%**.

CAPTIONING

Captioning

- Motivation is to make video lectures available to the deaf students
- Improves the experience of hearing students
- Non-native English speakers
- ICS player supports **captions**;
currently captions created **manually**



DEMO

ASSESSMENT OF STUDENT USE, EXPERIENCE, PERCEPTIONS

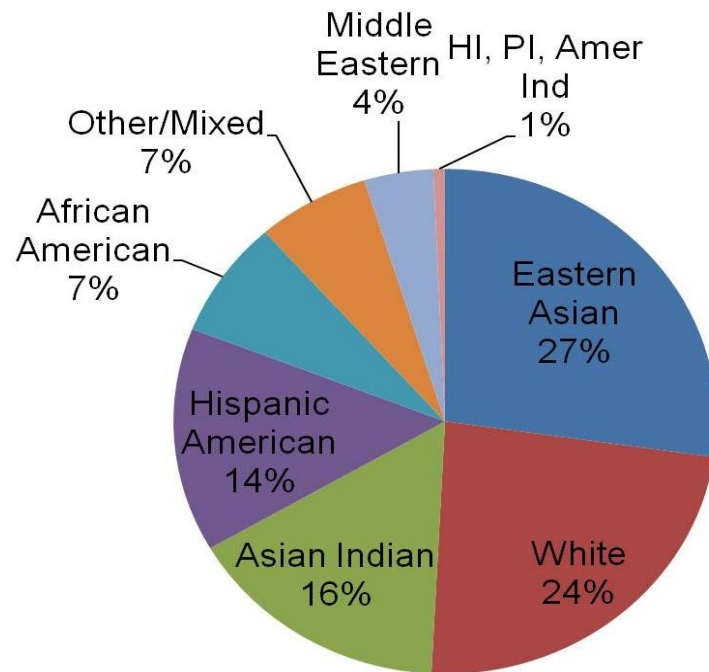
Assessment Questions

- What do students perceive as the value of ICS videos?
- What is the perceived value of the index?
- What is the perceived value of the search tool?

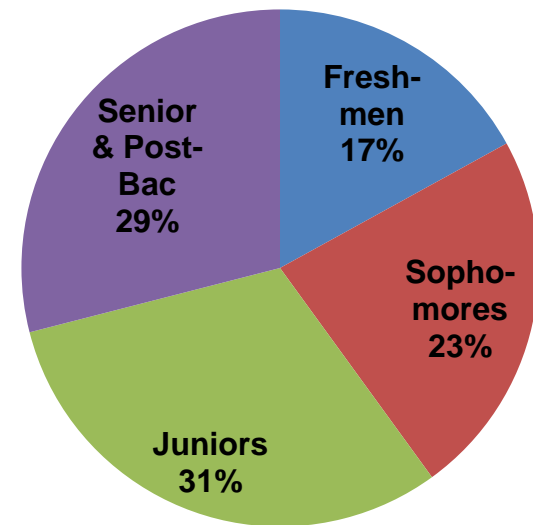
Survey Administration

- 1,167 student surveys
 - Spring 2010 (N=612) and Spring 2011 (N=555)
- 18 courses
 - Biology, computer science, chemistry, geology, and mathematics
 - Some courses have *very high* enrollments
- Last week of semester
 - Faculty sent out link, reminders

Sample Profile



Race/Ethnicity



Student Level

60% female, 40 male%

Mean one-way commute time: 30-45 minutes

57% of students work to support themselves (7% >36 hours/week)

Strong commitment to education (class attendance, course grade, GPA)

Reasons for Using Videos (n=1804)

To hear a lecture that I had missed because I had not gone to class.	73%
To review before a test or quiz.	70%
To review concepts that were particularly difficulty for me.	64%
To review concepts I didn't understand in the class-based lecture.	32%
To review concepts I could not hear in the class-based lecture.	64%
To preview a lecture before going to class.	7%
To review a lecture later on the same day that the lecture was presented in class.	17%
To review a lecture later in the same week that the lecture was presented in class.	32%

ICS Videos Strongly Valued for Learning and Grades

Survey Item	N	Mean (6-pt scale)	Std. Dev.
Lecture videos help me to clarify material that was not clear in class.	825	5.45	.826
Lecture videos are useful for reviewing.	841	5.63	.693
Having access to lecture videos for this class is important to me.	837	5.61	.774
The lecture videos helped me to study for quizzes or tests.	829	5.51	.842
Importance of video lectures for this class for getting the grade you wanted.	885	3.51 / 4-pt scale	.751

Perceptions of Index

- 97%: index was helpful
- 89%: intuitive
- 96%: easy to use
- 90%: time intervals appropriate

“sometimes I would have to pause the lecture to take care of other responsibilities that I had to attend to, and when I was ready to come back to the lecture I'd pick up exactly where I was at, it was great!”

Index Use Correlated with Increased Perceived Value of Videos

- “Overall Attitude Toward Video” (scale of six items) (Cronbach’s $\alpha = .872$)
- Index users valued videos more
 - ($t(616) = -3.284, p < .001$)
 - Small to moderate effect ($d = 0.265$)

Perceptions of Search Tool (N=50)

- 96%: easy to use
- 98%: results easy to use
- 95%: easier to navigate video
- 88%: helpful either most or all the time
- 82%: search results relevant to search most or all of the time

How Did Availability of ICS Videos Affect Class Attendance?

- 86% strongly agreed that it was important for them to go to class, whether or not the video lectures were available
- 67%: attended 75-100% of classes
- **Class attendance positively and significantly correlated with video use ($r(1019) = .234, p < .000$)**

CONCLUSION

Conclusion

- Reported : ICS videos and their usage for STEM coursework
- Indexing and search features :
very helpful and easy to use
- New and innovative direction for effective use of videos in STEM coursework.



- Large scale usage and assessments
- Captioning with speech recognition tools
- Text and semantic based indexing and search

THANK YOU

<http://icsvideos.cs.uh.edu/>

The framework is freely available to educational institutions.



National Science Foundation
WHERE DISCOVERIES BEGIN