Visual Search Engine for Handwritten and Typeset Math in Lecture Videos and LaTeX Notes

Kenny Davila and Richard Zanibbi

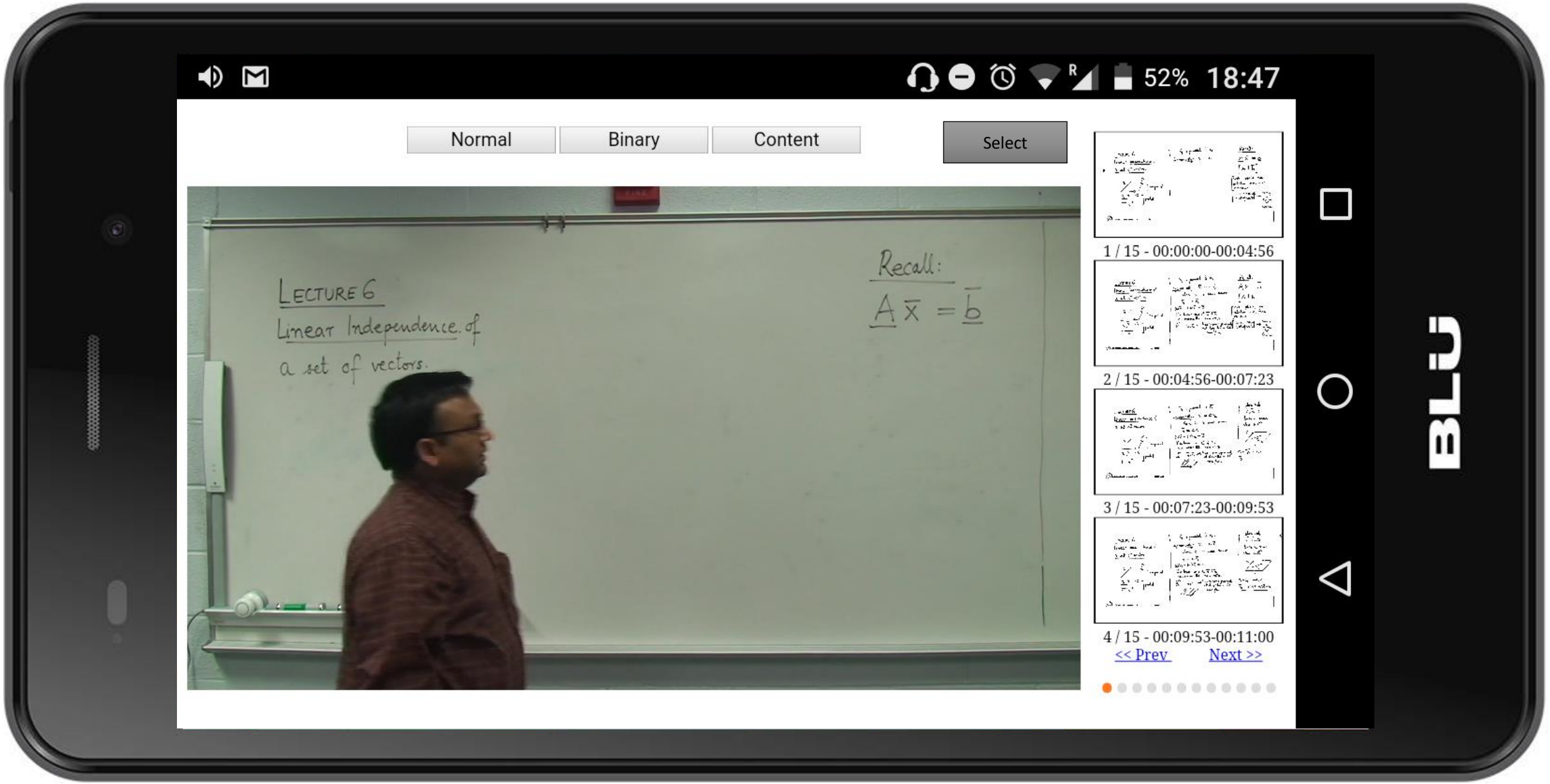
August 6, 2018

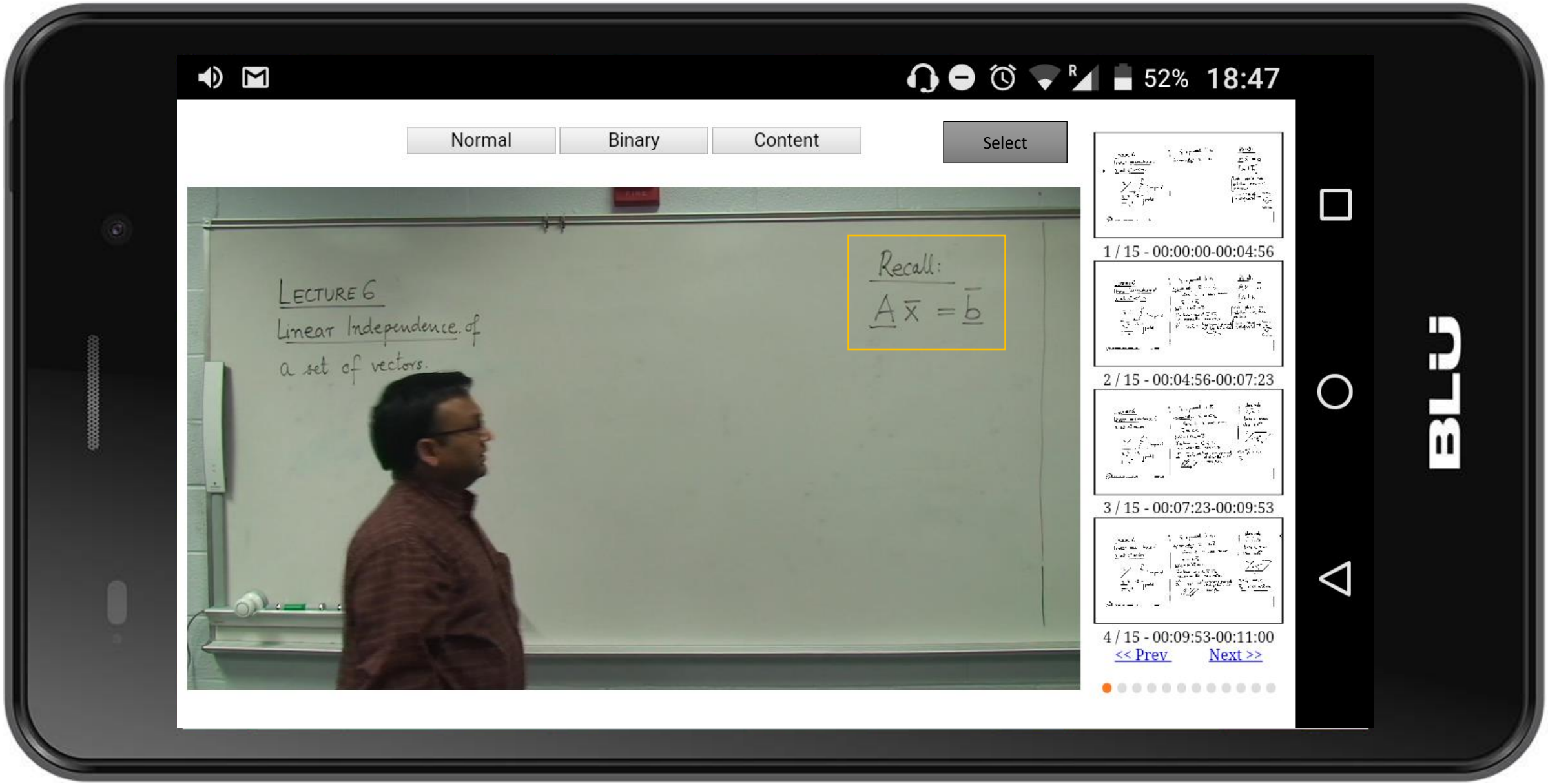
University at Buffalo

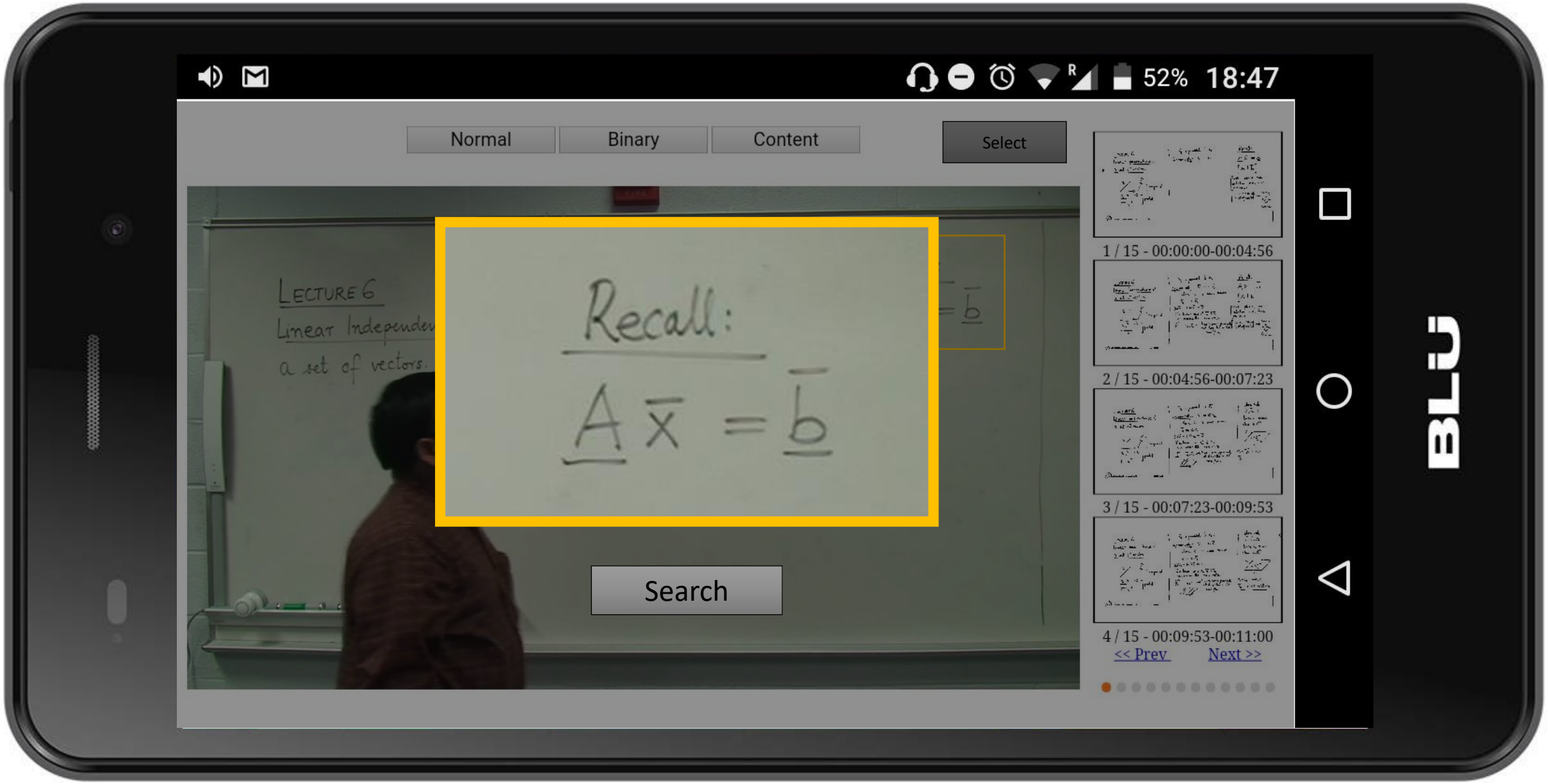
Center for Unified Biometrics and Sensors

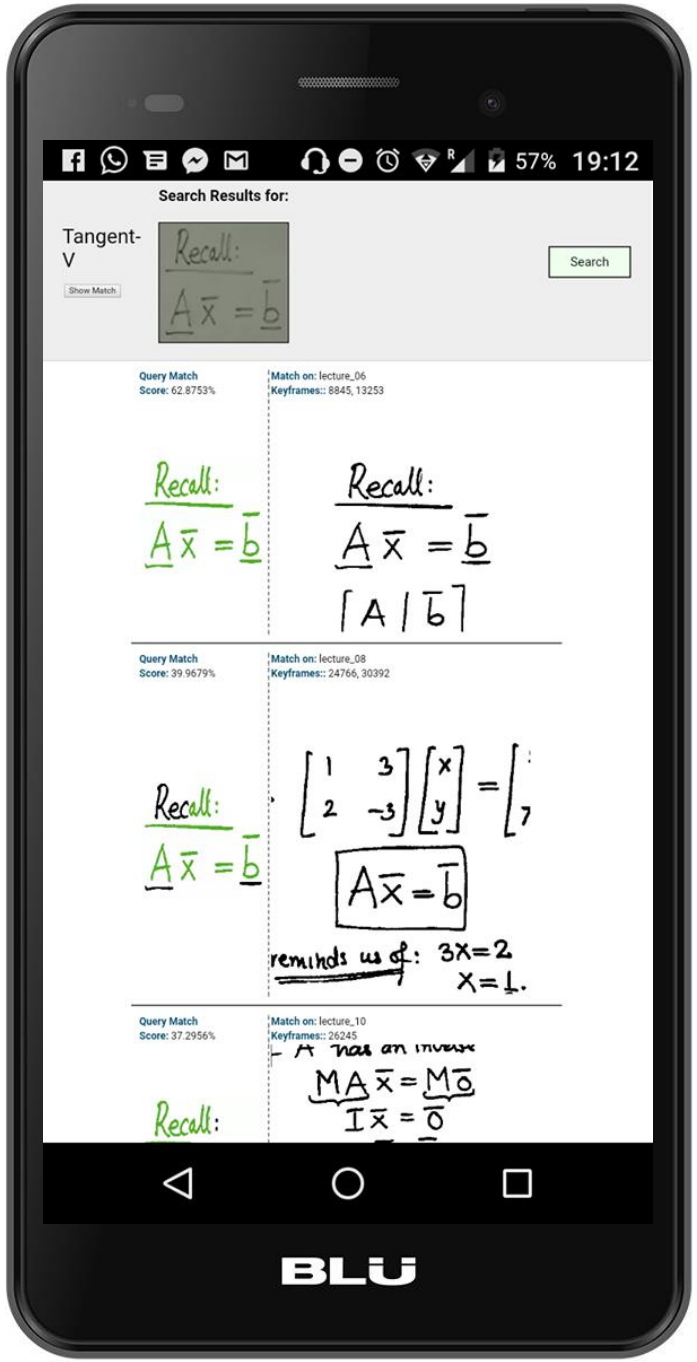
R·I·T

dprl [DOCUMENT AND PATTERN RECOGNITION LAB]


52% 18:47
Normal
Binary
Content
Select
LECTURE 6
Linear Independence of
a set of vectors.
Recall:
$A\bar{x} = \bar{b}$
1 / 15 - 00:00:00-00:04:56
2 / 15 - 00:04:56-00:07:23
3 / 15 - 00:07:23-00:09:53
4 / 15 - 00:09:53-00:11:00
<< Prev
Next >>
BLU


52% 18:47
Normal
Binary
Content
Select
LECTURE 6
Linear Independence of
a set of vectors.
Recall:
$A\bar{x} = \bar{b}$
1 / 15 - 00:00:00-00:04:56
2 / 15 - 00:04:56-00:07:23
3 / 15 - 00:07:23-00:09:53
4 / 15 - 00:09:53-00:11:00
<< Prev
Next >>
BLU


52% 18:47
Normal
Binary
Content
Select
LECTURE 6
Linear Independe
a set of vectors.
Recall:
$A\bar{x} = \bar{b}$
Search
1 / 15 - 00:00:00-00:04:56
2 / 15 - 00:04:56-00:07:23
3 / 15 - 00:07:23-00:09:53
4 / 15 - 00:09:53-00:11:00
<< Prev
Next >>
BLU



SEARCH RESULTS

Found in Lecture Videos

1. Linear Algebra – Lecture 06
2. Linear Algebra – Lecture 08
3. Linear Algebra – Lecture 10

…

Related Topics

1. Systems of Equations
2. Matrix Reduction
3. Linear Algebra

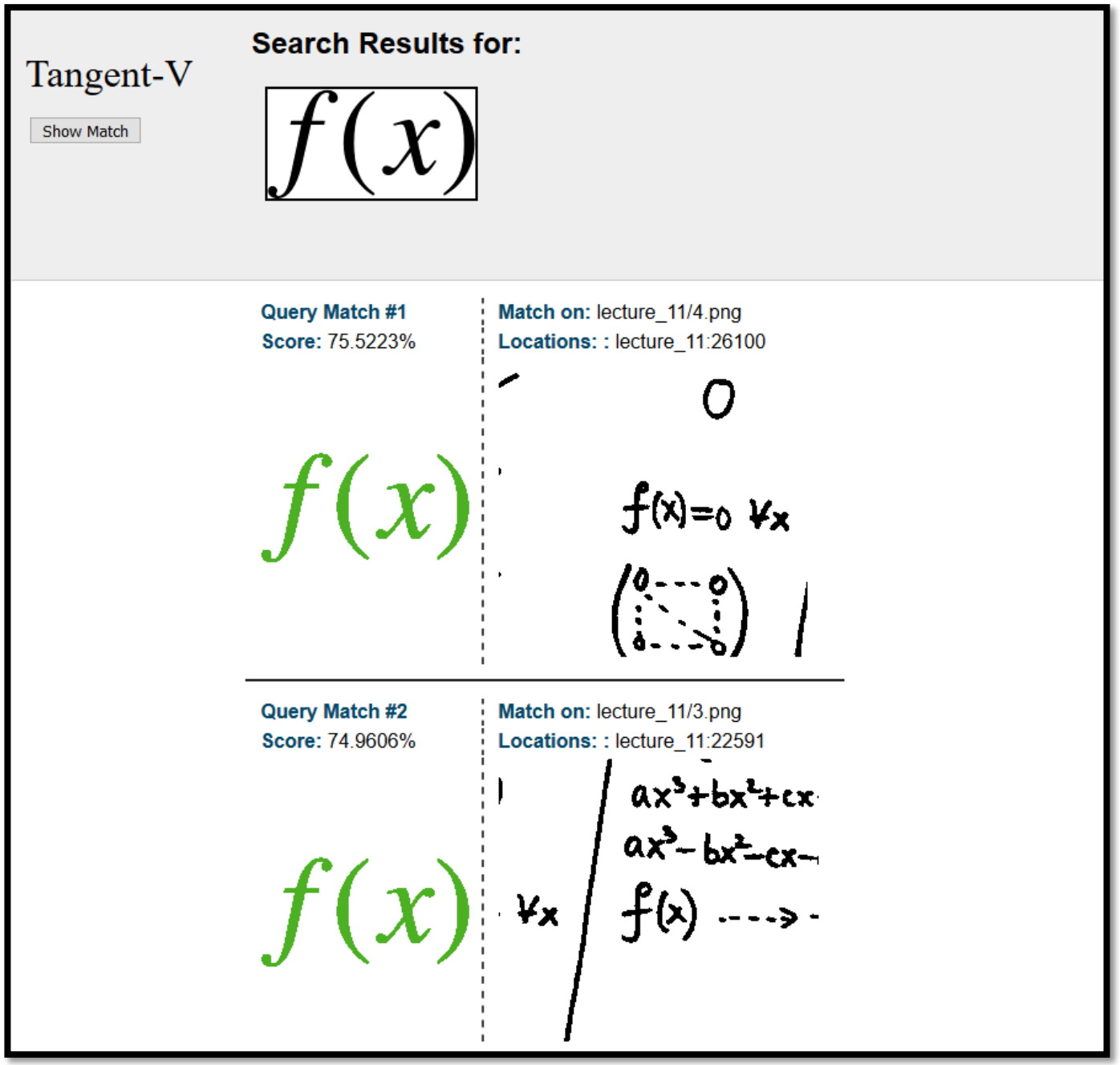
What about other
Mathematical Expressions?

Could I write my **queries** instead
of using Images?

What about other Mathematical Expressions?

Could I write my **queries** instead of using Images?

Yes, using LaTeX

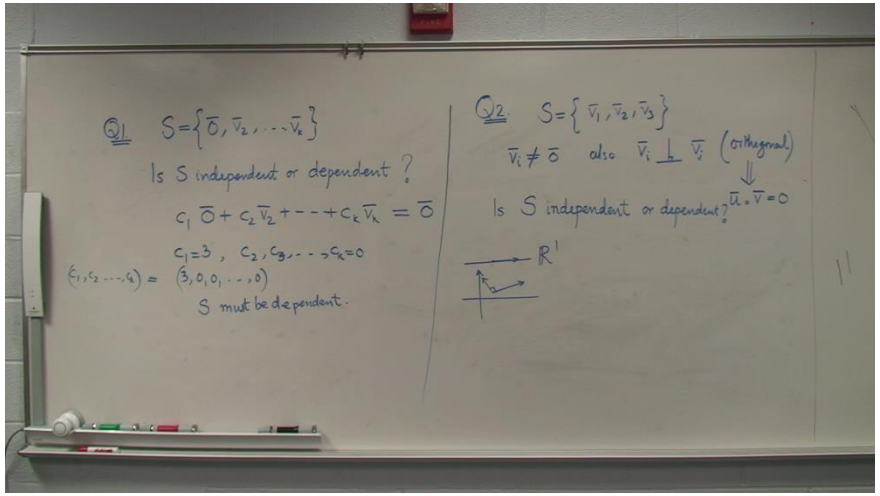


Potential Search Modes

LaTeX → Whiteboard

LaTeX → LaTeX

Lecture Video



Linear Independence

Question: Consider the set of vectors $S = \{\vec{0}, \vec{v}_2, ..., \vec{v}_k\}$. Is S independent or dependent? To answer this, we must ask whether the following homogeneous equation has a non-zero solution:

$$c_1\vec{0} + c_2\vec{v}_2 + ... + c_k\vec{v}_k = \vec{0}$$

Consider for instance the solution $c_1 = 3, c_2, c_3, ..., c_k = 0$. This is a non-zero or non-trivial solution, hence S must be dependent.

Question: Let $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Suppose none of the vectors in S are zero vectors, and each is perpendicular to one another. That is,

$$\vec{v}_i \neq \vec{0} \qquad \vec{v}_i \perp \vec{v}_j \qquad \text{for all } i \neq j$$

Lecture Notes

Whiteboard → Whiteboard

Whiteboard → LaTeX

Tangent-V Visual Search Engine

Applied to **Indexing** and **Retrieval** of formulae from **Lecture** materials

Based on Matching **Symbol Pairs** from **Line of Sight Graphs (LOS)**

Domain knowledge is given by Recognition Module

- Currently: **Mathematical Symbol Recognition**

Source **code** released: https://cs.rit.edu/~dprl/Software.html

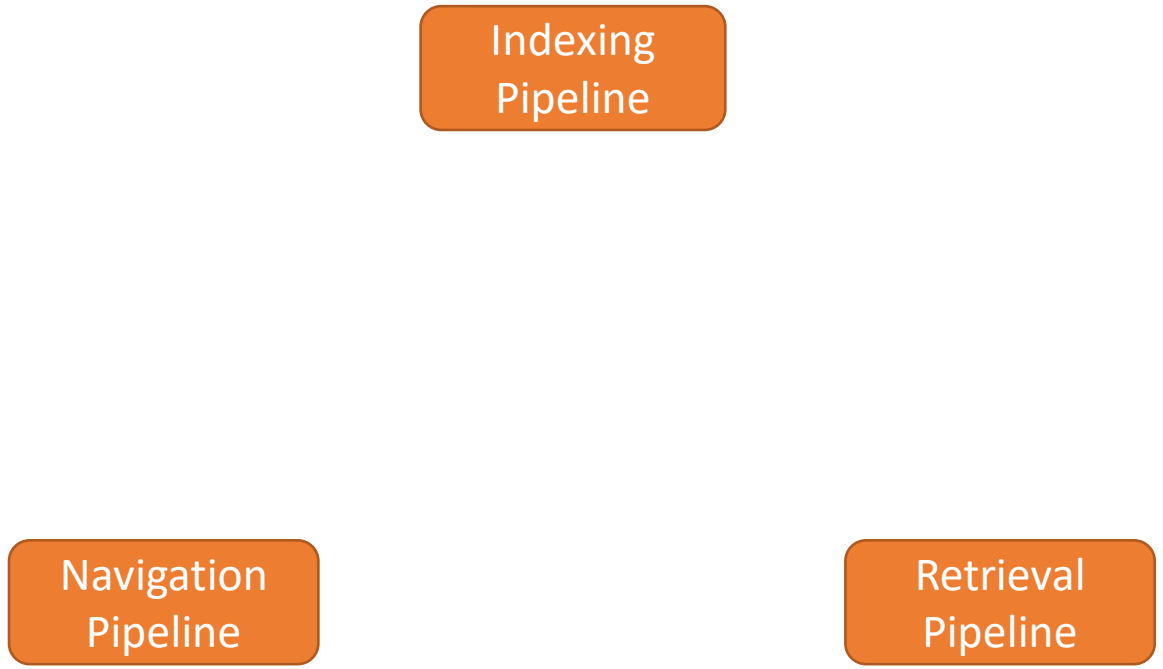
Related Work

Related fields:

- Content-Based Image Retrieval [1]
- Word Spotting [2]
- **Mathematical Information Retrieval** [3]
  - Formula Representation: Semantic vs **Appearance**
  - Retrieval Modality: Symbol vs **Image-based**
  - **Tangent-V** generalizes the **Tangent-S** formula retrieval model [4]

[1] J. Sivic & A. Zisserman, “Video Google: A text retrieval approach to object matching in videos,” in ICCV 2003
[2] S. Sudholt & G. A. Fink, “Phocnet: A deep convolutional neural network for word spotting in handwritten documents,” in ICFHR 2016
[3] R. Zanibbi & D. Blostein, “Recognition and retrieval of mathematical expressions,” IJDAR, vol. 15, no. 4, 2012.
[4] K. Davila & R. Zanibbi, “Layout and semantics: Combining representations for mathematical formula search,” SIGIR, 2017

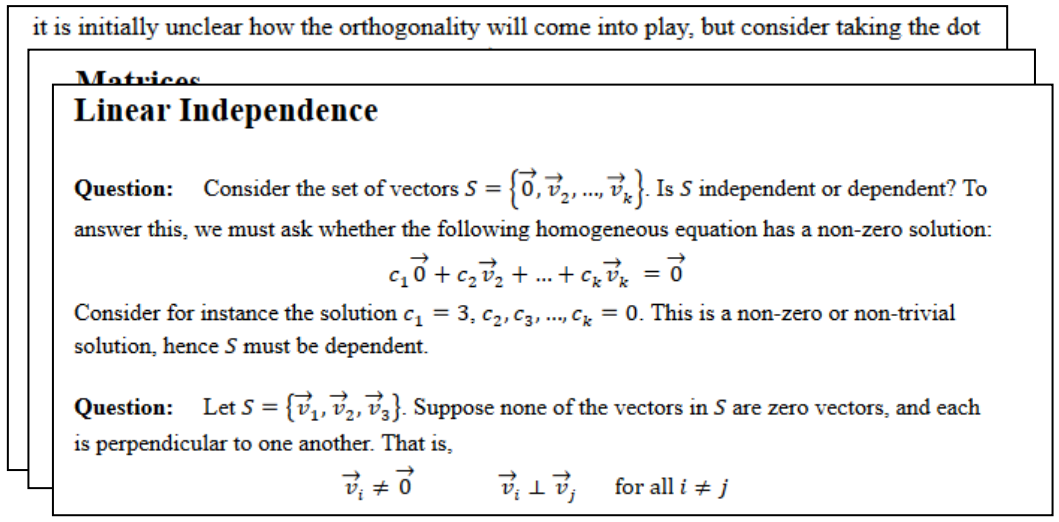
Tangent-V Overview



Supplementary Lecture Notes (LaTeX)

Input

Lecture Notes



it is initially unclear how the orthogonality will come into play, but consider taking the dot

Matrices

Linear Independence

Question: Consider the set of vectors $S = \{\vec{0}, \vec{v}_2, ..., \vec{v}_k\}$. Is S independent or dependent? To answer this, we must ask whether the following homogeneous equation has a non-zero solution:

$$c_1\vec{0} + c_2\vec{v}_2 + ... + c_k\vec{v}_k = \vec{0}$$

Consider for instance the solution $c_1 = 3, c_2, c_3, ..., c_k = 0$. This is a non-zero or non-trivial solution, hence S must be dependent.

Question: Let $S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Suppose none of the vectors in S are zero vectors, and each is perpendicular to one another. That is,

$$\vec{v}_i \neq \vec{0} \qquad \vec{v}_i \perp \vec{v}_j \quad \text{for all } i \neq j$$

LaTeX

Output

Math Expressions

$$S = \{\vec{0}, \vec{v}_2, \ldots, \vec{v}_k\} \quad S$$

$$c_1\vec{0} + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k \quad S$$

$$c_1 = 3 \qquad c_2, c_3, \ldots, c_k = 0$$

$$S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$$

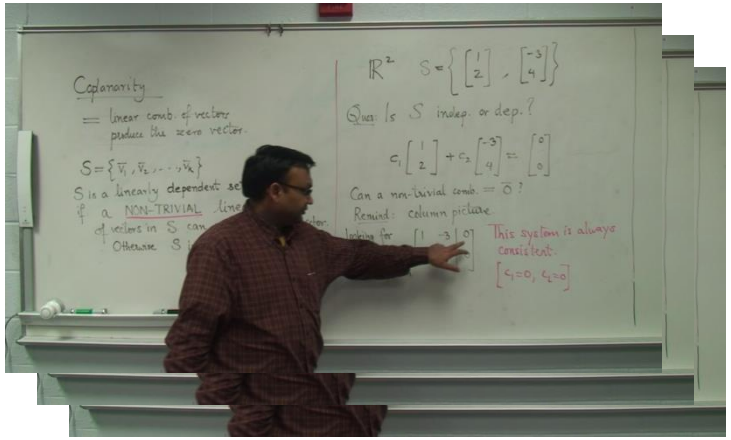
$$\vec{v}_i \neq \vec{0} \qquad \vec{v}_i \perp \vec{v}_j \quad \text{for all } i \neq j$$

Binary Images

Preprocessing Lecture Video Summarization [1]

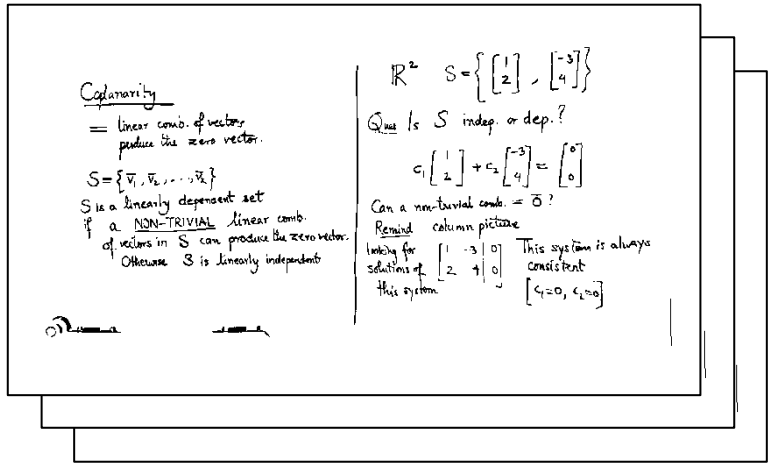
Input

Lecture Video



Output

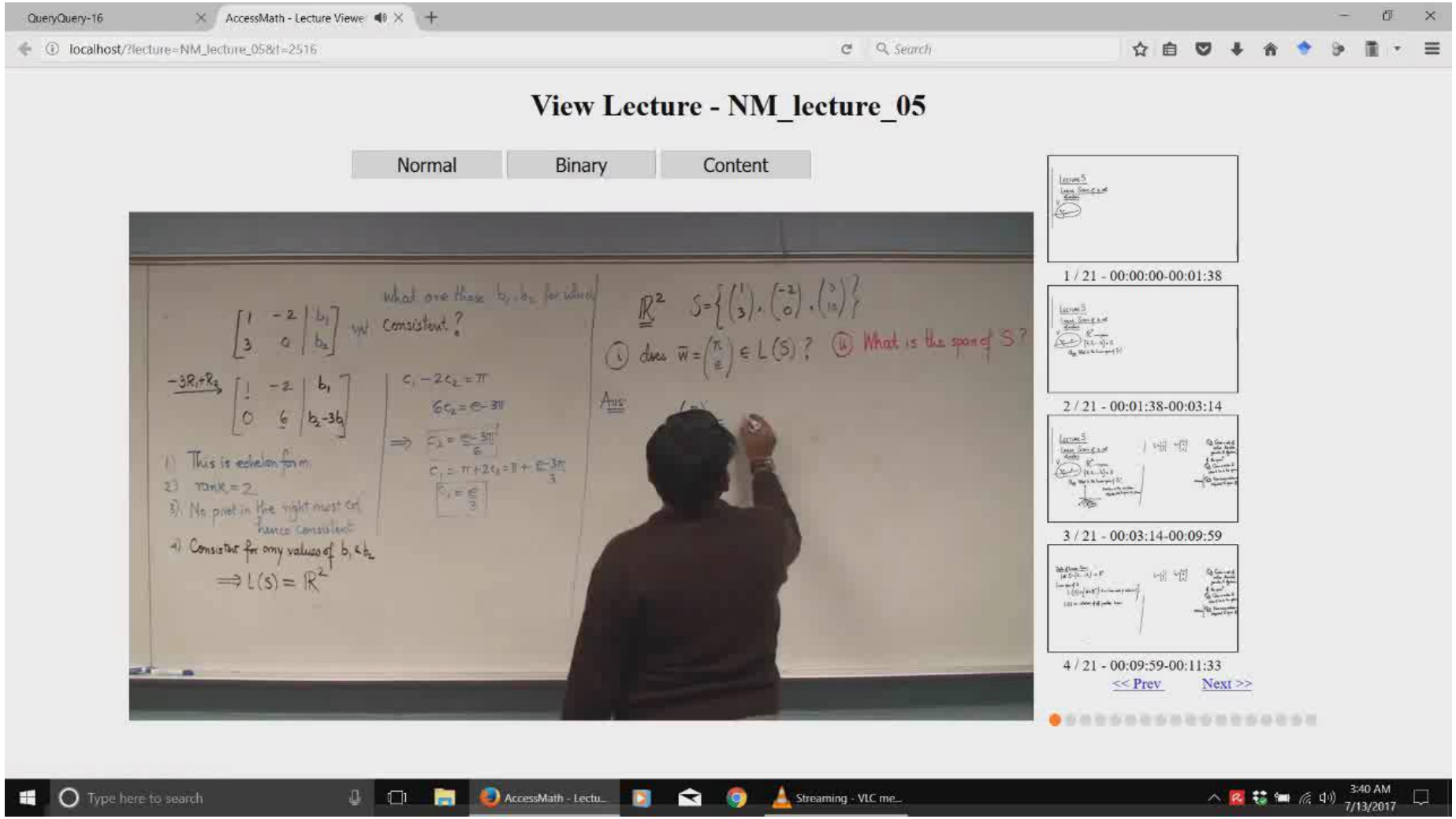
Whiteboard Contents Keyframes



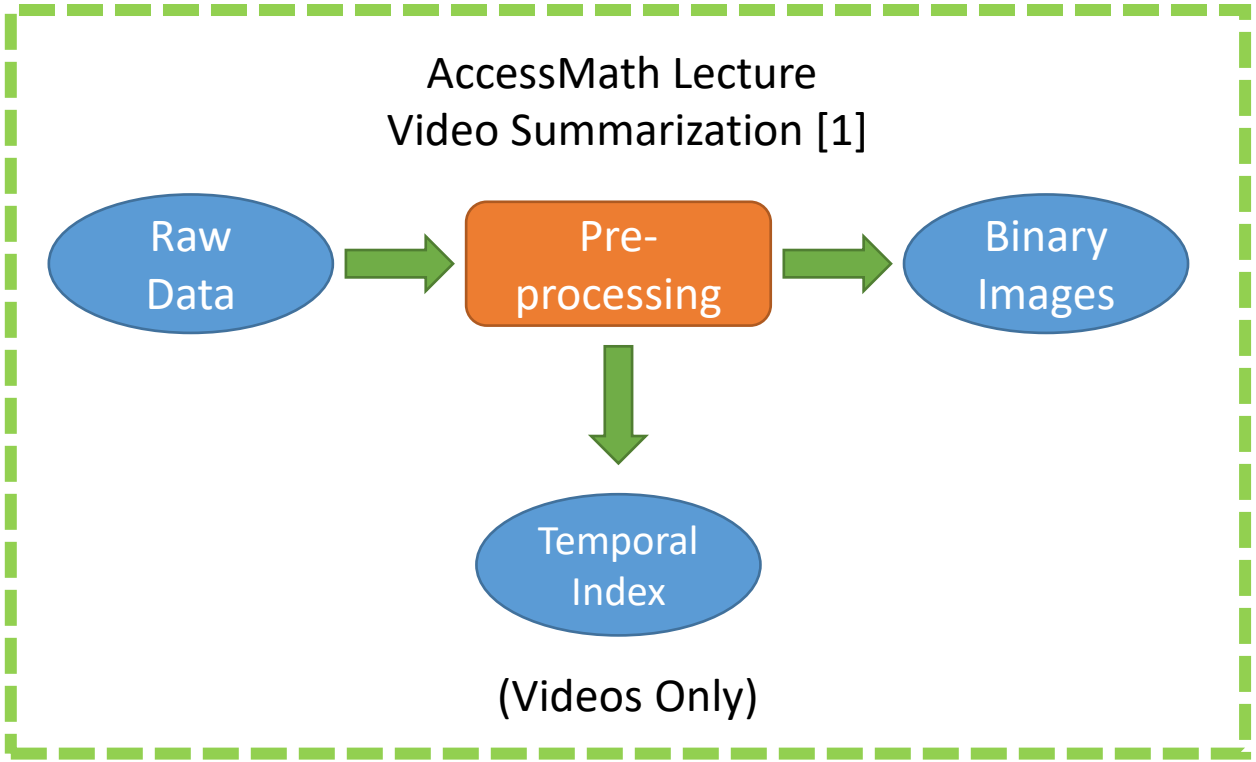
MTS/ MP4 → Content Extraction → Spatio-temporal Analysis → Temporal Index → Temporal Segmentation → Binary Images

[1] Davila, K., Zanibbi, R. Whiteboard Content Summarization via Spatio-Temporal Conflict Minimization in Lecture Videos. ICDAR 2017

Lecture Video Navigation from Keyframes

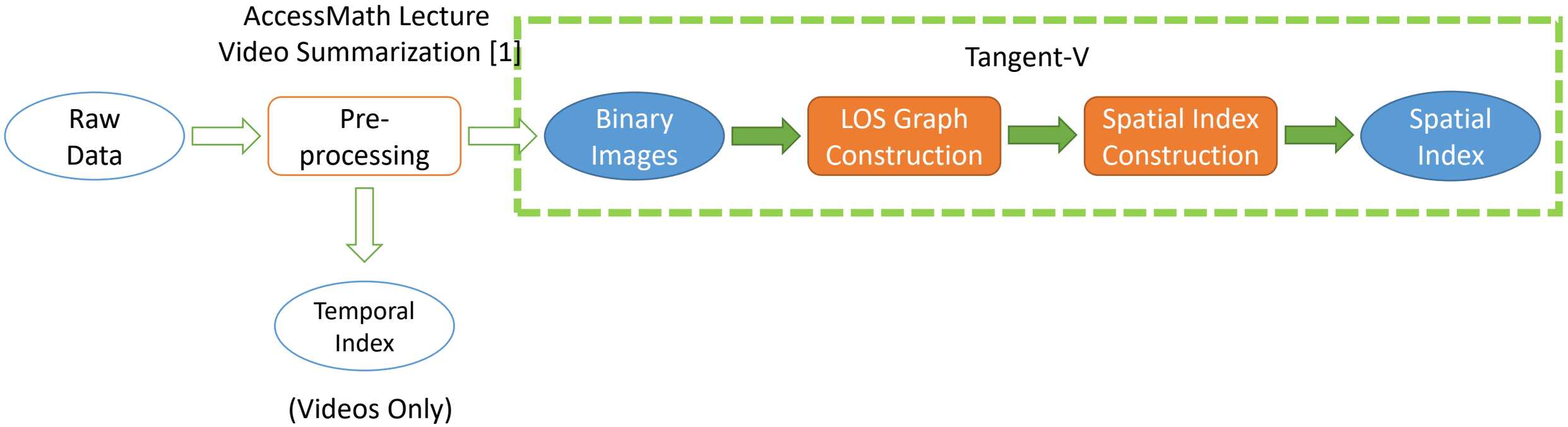


Indexing Pipeline (Overview)



[1] Davila, K., Zanibbi, R. Whiteboard Content Summarization via Spatio-Temporal Conflict Minimization in Lecture Videos. ICDAR 2017

Indexing Pipeline (Overview)



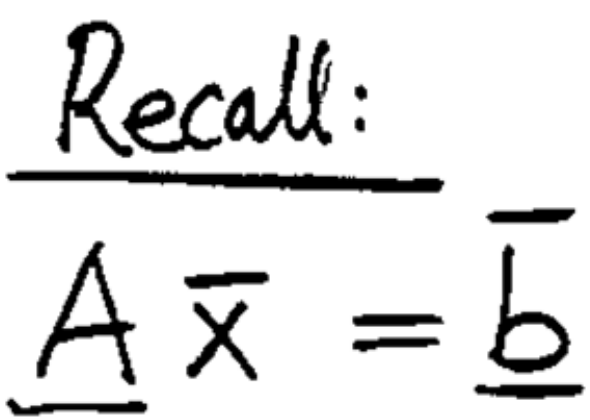
[1] Davila, K., Zanibbi, R. Whiteboard Content Summarization via Spatio-Temporal Conflict Minimization in Lecture Videos. ICDAR 2017

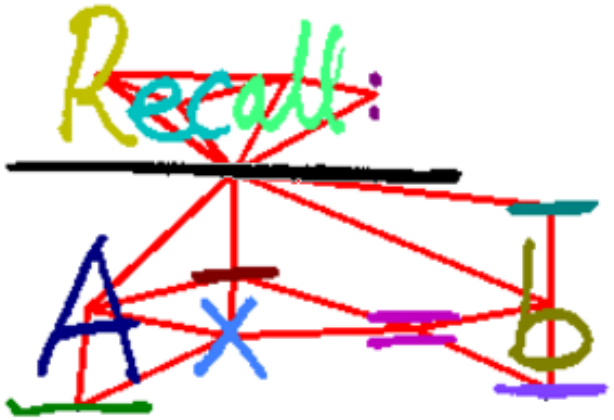
Line of Sight (LOS) Graphs

Uses **Connected Components (CC)** as **Nodes**

Two **nodes** are **connected** if

- One can **see** the other
- Max. distance factor considered for whiteboard content (2 times median size)





Line of Sight (LOS) Graphs

True **Node Labels/Relationships** are **unknown**

- After Symbol Recognition, each **Node** has top **k labels** with **probabilities**
  - $\sum_{\omega \in \Omega} p(\omega | s_x) \geq 80\% \quad k \leq 10$
- **Edges** have **3D unit vectors** indicating direction

2^x

(0.707, 0.707, 0.000)

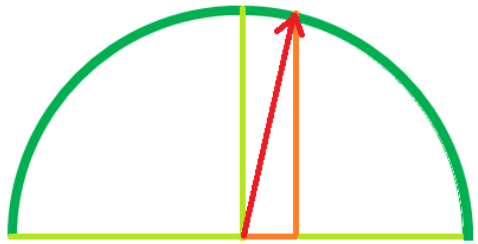
$2x$

(1.000, 0.000, 0.000)

x^2

(-0.707, -0.707, 0.000)

$\sqrt{x}$



(0.146, -0.146, 0.978)

Spatial Indexing using Symbol Pairs

Inverted Index for Symbol Pairs

Entries: Pairs of symbol labels $(\boldsymbol{\omega_1}, \boldsymbol{\omega_2})$

Posting lists: Pair locations in images with $\left(\boldsymbol{ID}, \boldsymbol{p_1}, \boldsymbol{p_2}, \vec{\boldsymbol{c}}, \boldsymbol{s_p}(\boldsymbol{c_1}, \boldsymbol{c_2})\right)$

Top k-labels per node $(\boldsymbol{\Omega})$

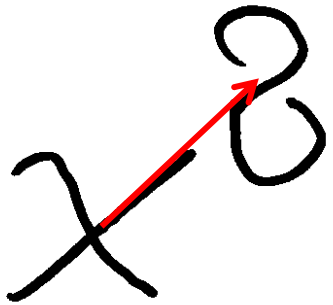
Tuples Generated $(\boldsymbol{\Omega_1 \times \Omega_2})$

$$\left(\boldsymbol{\omega_1}, \boldsymbol{\omega_2}, \boldsymbol{p_1}, \boldsymbol{p_2}, \vec{\boldsymbol{c}}, \boldsymbol{s_p}\right)$$

$\boldsymbol{p_x}$ - $\boldsymbol{p}(\boldsymbol{\omega_x}|\boldsymbol{s_x})$

$\vec{\boldsymbol{c}}$ - 3D Unit Vector from $\boldsymbol{s_1}$ to $\boldsymbol{s_2}$

$\boldsymbol{s_p}$ - Size Ratio between $\boldsymbol{s_1}$ and $\boldsymbol{s_2}$



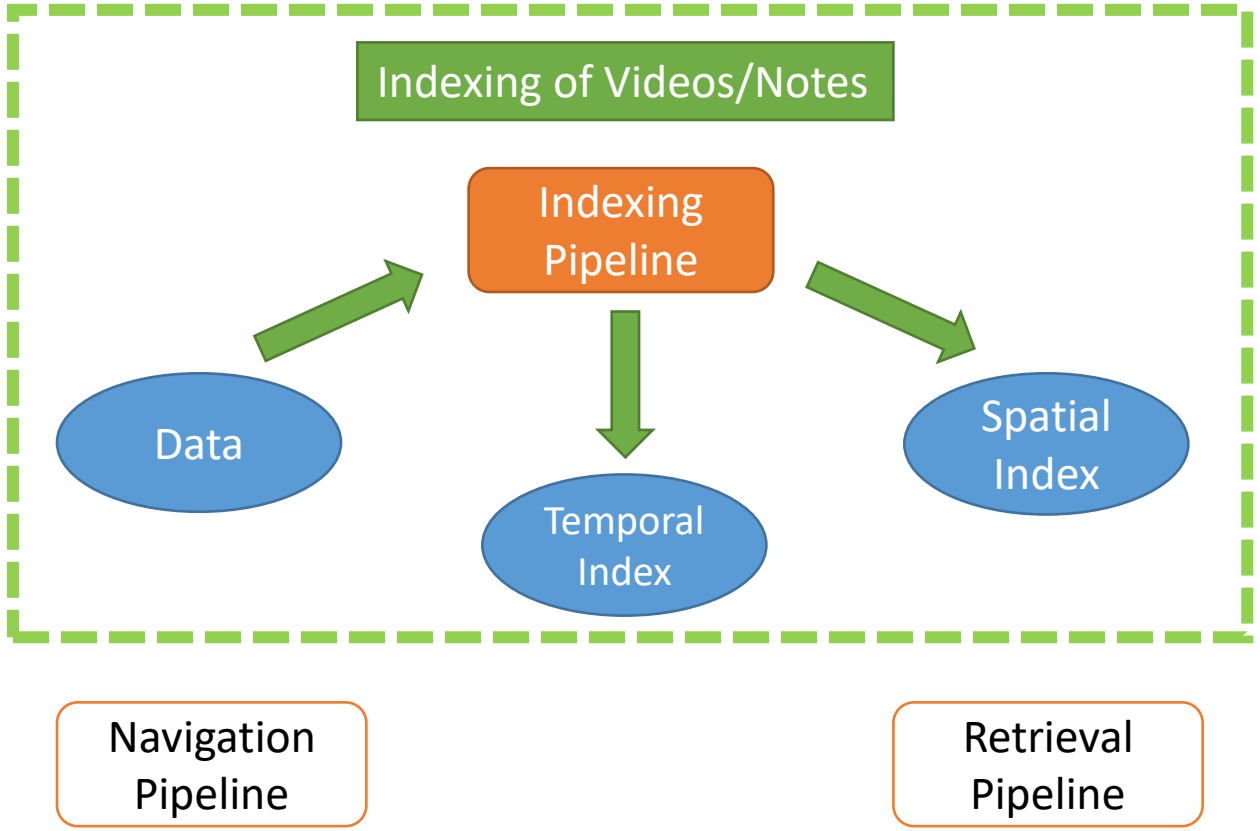
$$S_1 = x \qquad S_2 = 8$$

$$\Omega_1 = \{(x, 0.8), (X, 0.2)\} \qquad \Omega_2 = \{(8, 0.6), (\&, 0.3)\}$$

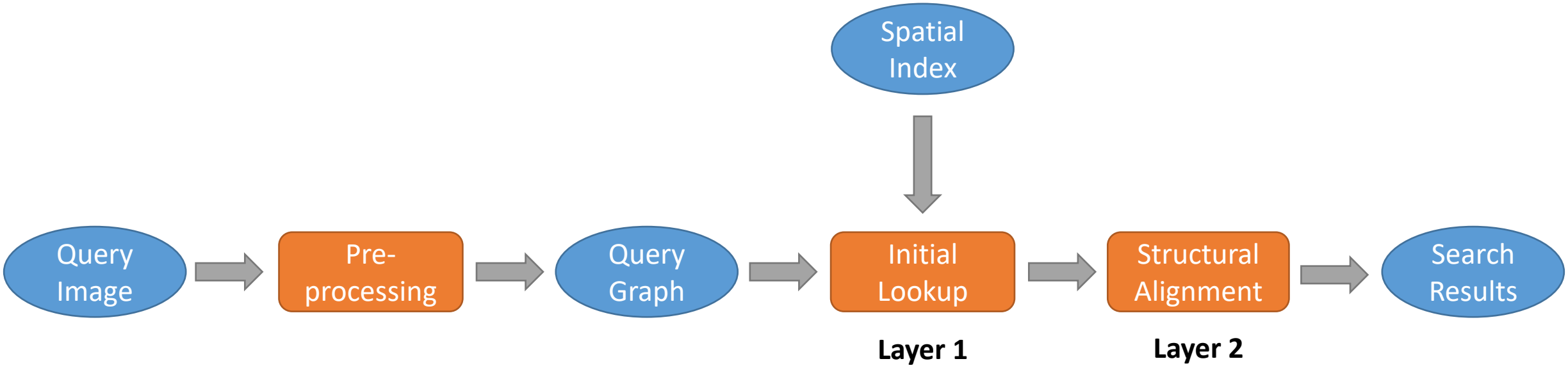
$$\vec{\boldsymbol{c}} = \langle \boldsymbol{0.71}, \boldsymbol{-0.71}, \boldsymbol{0.00} \rangle$$

$$\boldsymbol{s_p} = 1.26$$

Tangent-V Overview



Tangent-V Retrieval Model



Layer 1: Initial Lookup

Query **symbol pairs** are used to find matches on their corresponding **entries** on the **inverted index** structure

A **match** between index symbol pair $P^c = (c_1, c_2)$ and query pair $P^q = (q_1, q_2)$ will be accepted as **valid** if and only if:

1 - They are **spatially consistent**:

$$\vec{\boldsymbol{c}} \cdot \vec{\boldsymbol{q}} \geq \cos(45^\circ)$$

2 - Optionally, if they have **consistent size ratios** (not too small/large)

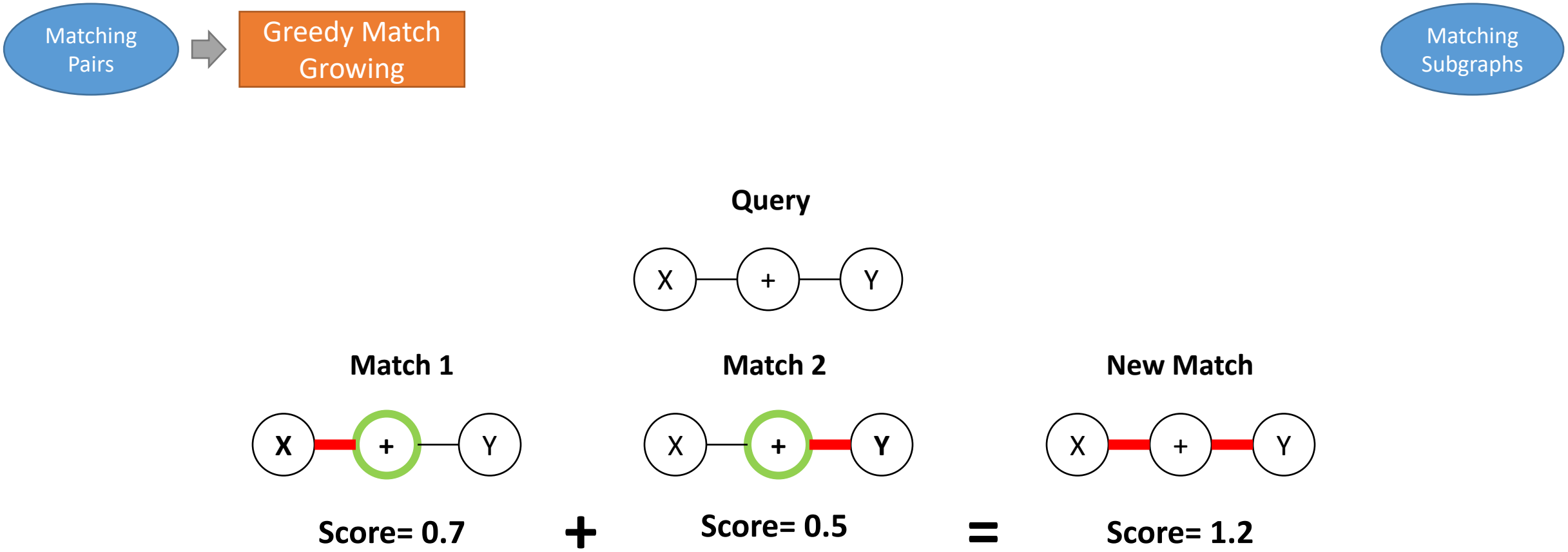
Matching Pairs Scores are then **aggregated** by unique **Graph Pair IDs**

Layer 2: Structural Alignment

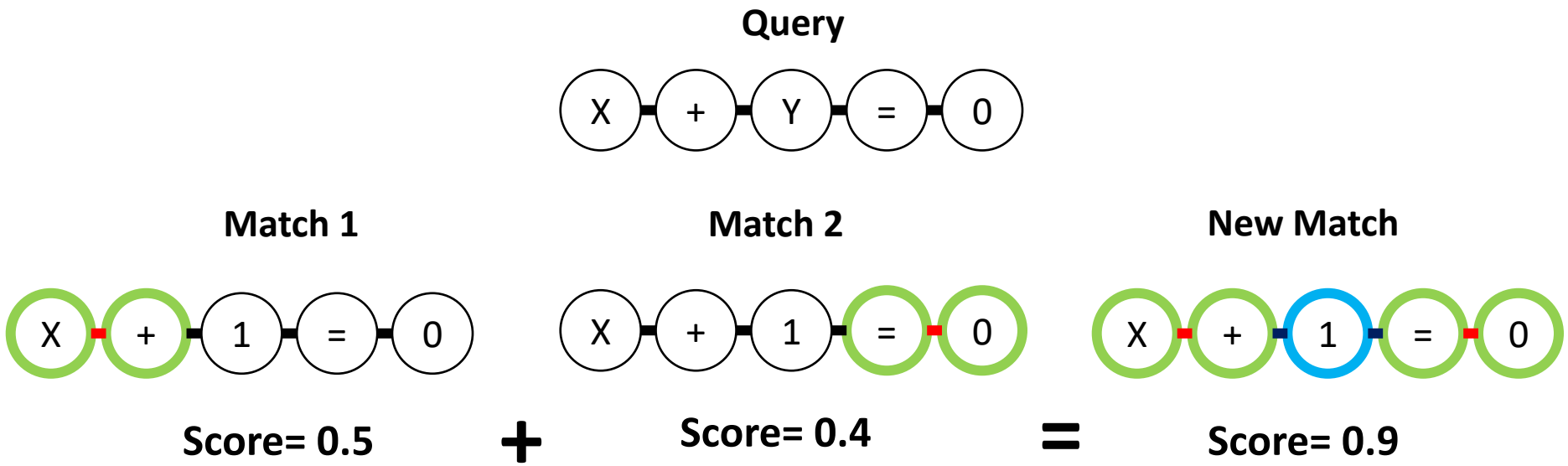
Matching Pairs

Matching Subgraphs

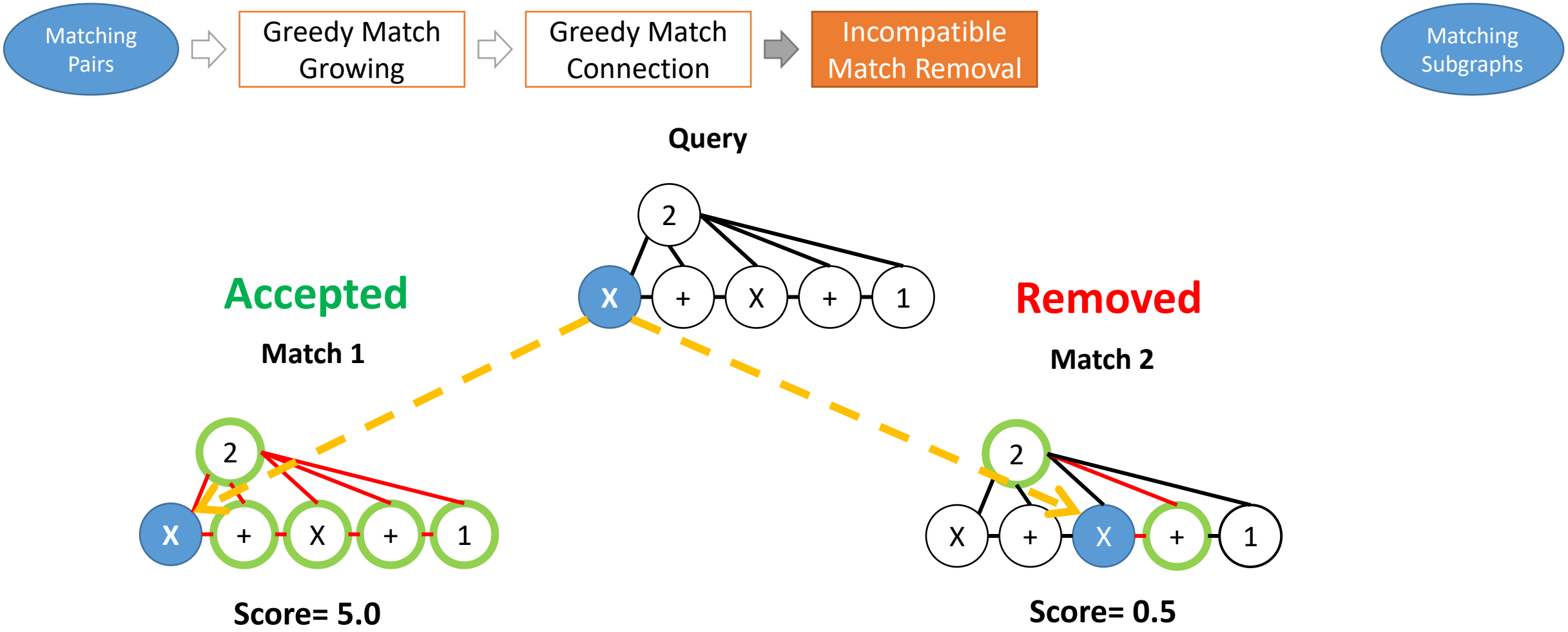
Layer 2: Structural Alignment



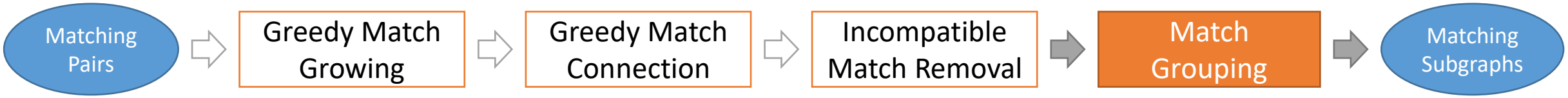
Layer 2: Structural Alignment



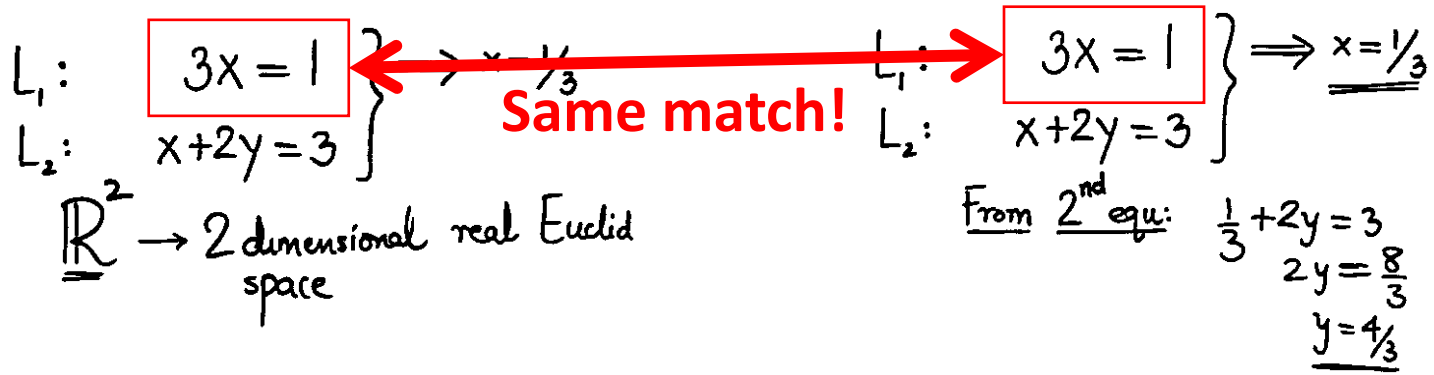
Layer 2: Structural Alignment



Layer 2: Structural Alignment



Query: $3x = 1$



Lecture 01 – KF #5 **Lecture 01 – KF #6**

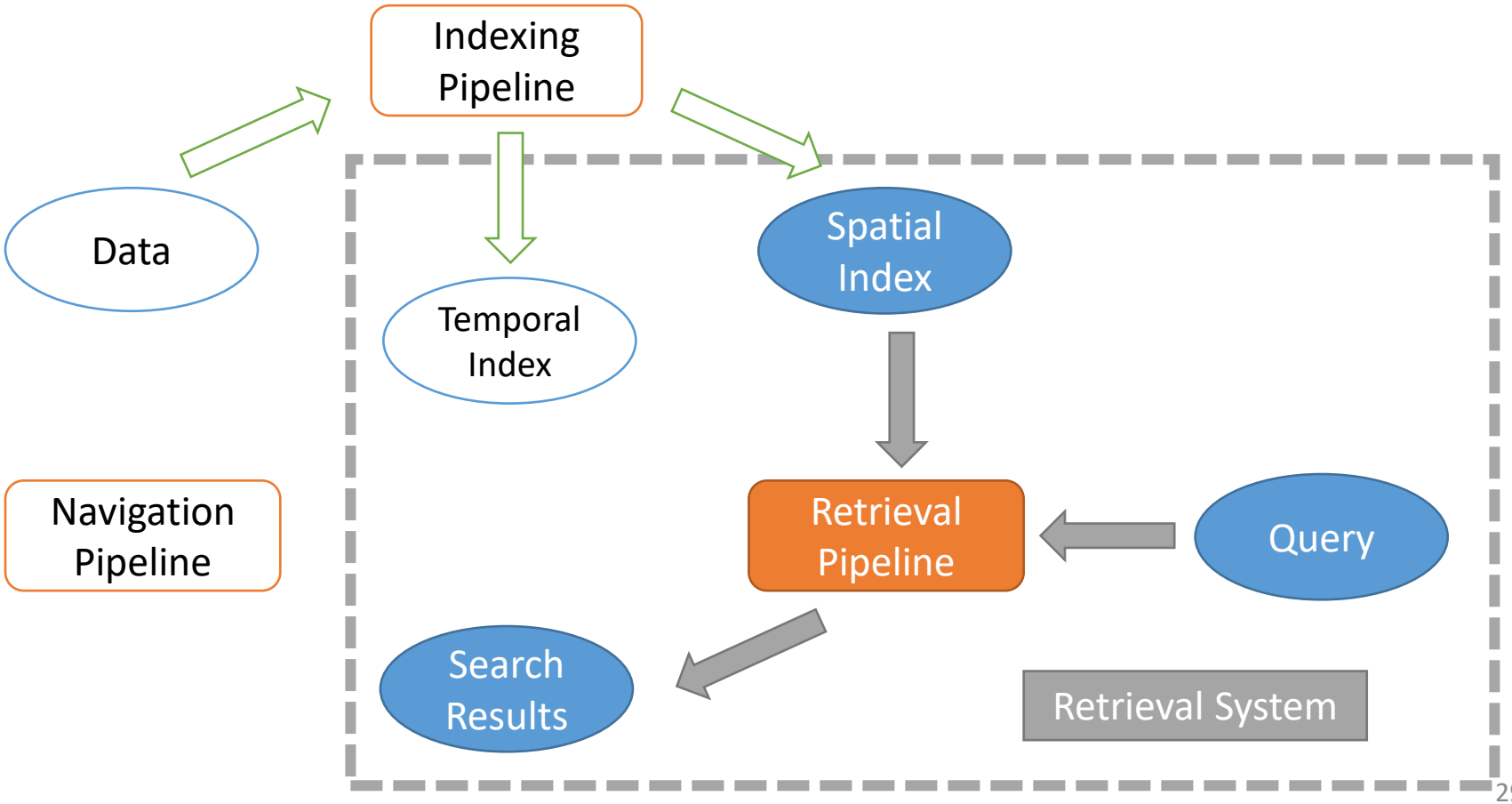
Match Scoring and Ranking

We introduce two **scoring schemes**: **α** and **h**

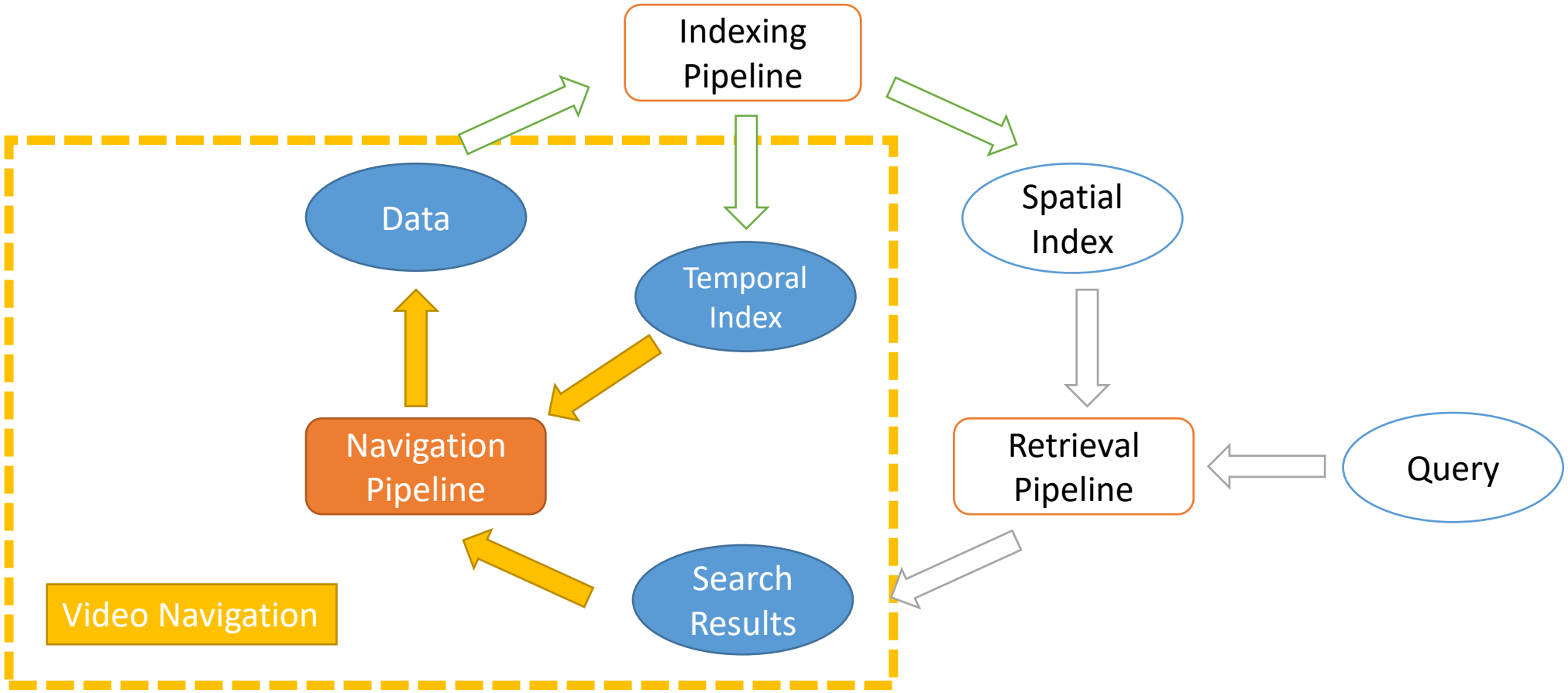
Item	$\alpha(M)$	$h(M)$
Description	A weighted **edge recall**	Harmonic mean of weighted **edge recall** and **node recall**
Edge weighting	**pair-wise** symbol alignments and **scaled cosine similarity**	**scaled cosine similarity**
Node weighting	-	**Individual** symbol alignments
Based on	-	**Maximum Subtree Similarity (MSS)** [1]
Execution Times	Faster	Slower

[1] R. Zanibbi, K. Davila, A. Kane, & F. Tompa, "Multi-stage math formula search: Using appearance-based similarity metrics at scale," SIGIR, 2016

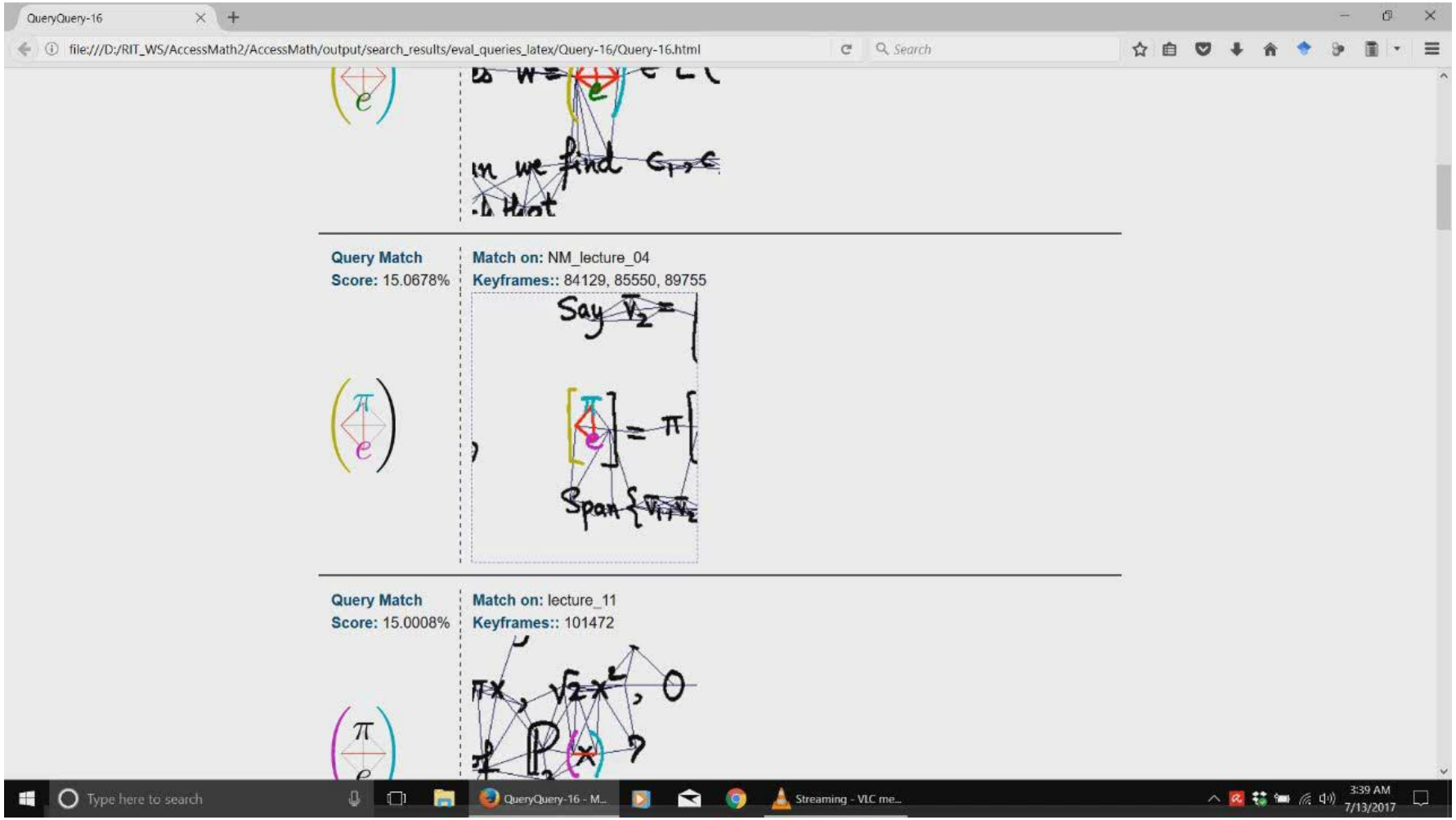
Tangent-V Overview



Tangent-V Overview



Lecture Video Navigation from Search Results



Check our demo at: https://youtu.be/gn24qo1MLN0

Experiments

AccessMath **Dataset**

- 13 Lecture videos with supplementary notes

A total of 20 evaluation **queries** were chosen with rejection sampling

A total of 4 combinations of Query-vs-Index **modalities**

- Handwritten expressions
- Typeset expressions

For a given query, the **target** is to find a math expression that **contains** the **whole query graph**

- query is **same expression**
- query is **sub-expression**

Evaluation Metrics

Two **metrics** are considered

- Recall @ 10: Target found @ rank ≤ 10
- MRR @ 10: Mean of Reciprocal Rank (RR), with

$$RR = \begin{cases} \frac{1}{r} & 1 \leq r \leq 10 \\ 0 & otherwise \end{cases}$$

Results: Recall @ 10

		Weighted Edge Recall α			Harmonic Mean h		
Query	Index	$\boldsymbol{\alpha}$	$\boldsymbol{\alpha_{\wedge}}$	$\boldsymbol{\alpha_{\wedge s}}$	$\boldsymbol{h}$	$\boldsymbol{h_{\wedge}}$	$\boldsymbol{h_{\wedge s}}$
LaTeX		**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Whiteboard		0.95	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
LaTeX	Whiteboard	0.95	0.95	0.90	0.95	**1.00**	0.95
Whiteboard	LaTeX	0.80	0.85	0.85	**0.90**	**0.90**	**0.90**

Results: MRR @ 10

		Weighted Edge Recall α			Harmonic Mean h		
Query	Index	$\boldsymbol{\alpha}$	$\boldsymbol{\alpha_{\wedge}}$	$\boldsymbol{\alpha_{\wedge s}}$	$\boldsymbol{h}$	$\boldsymbol{h_{\wedge}}$	$\boldsymbol{h_{\wedge s}}$
LaTeX		0.98	**1.00**	**1.00**	0.98	**1.00**	**1.00**
Whiteboard		0.93	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
LaTeX	Whiteboard	0.66	0.69	0.71	**0.89**	0.84	0.86
Whiteboard	LaTeX	0.63	0.71	0.74	0.74	0.78	**0.84**

Conclusions

Tangent-V is effective for search between Typeset and Handwriting
- **Multiple labels help** finding targets when recognition accuracy is low

Tangent-V can also be used to create **navigational tools**

New symbol recognizers can be used for indexing of **new domains**
- **Code** is released for others to try on new domains (http://cs.rit.edu/~dprl/Software.html)

Future work:
- Test unsupervised symbol classification
- Explore Vector formats
- Speed-up search

Thank You!

Source code: www.cs.rit.edu/~dprl/Software.html

This material is based upon work supported by the **National Science Foundation** (USA) under Grants No. IIS-1016815 and HCC-1218801.

We also thank **Anurag Agarwal** for helping in the creation of the lecture videos used to evaluate our system.