Image and Video Phylogeny Reconstruction

Filipe de Oliveira Costa

Advisor: Prof. Anderson Rocha
Co-advisor: Prof. Zanoni Dias

RECOD reasoning for complex data

UNICAMP

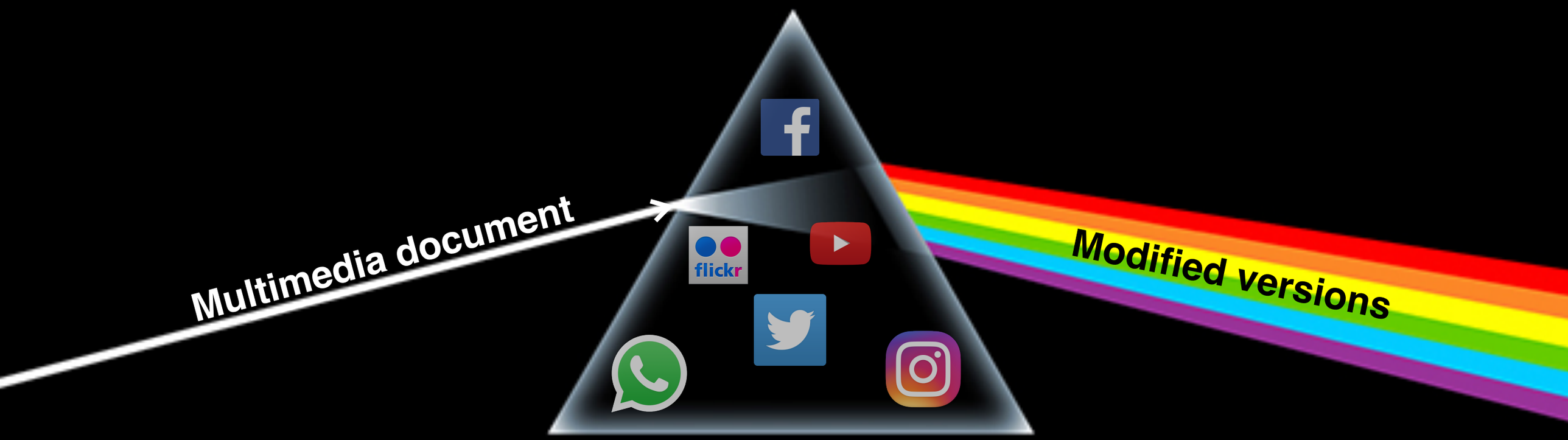
FAPESP

Outline

- Introduction
- Image Phylogeny
- Video Phylogeny
- Conclusions

Introduction

Dark side of the ~~moon~~ internet

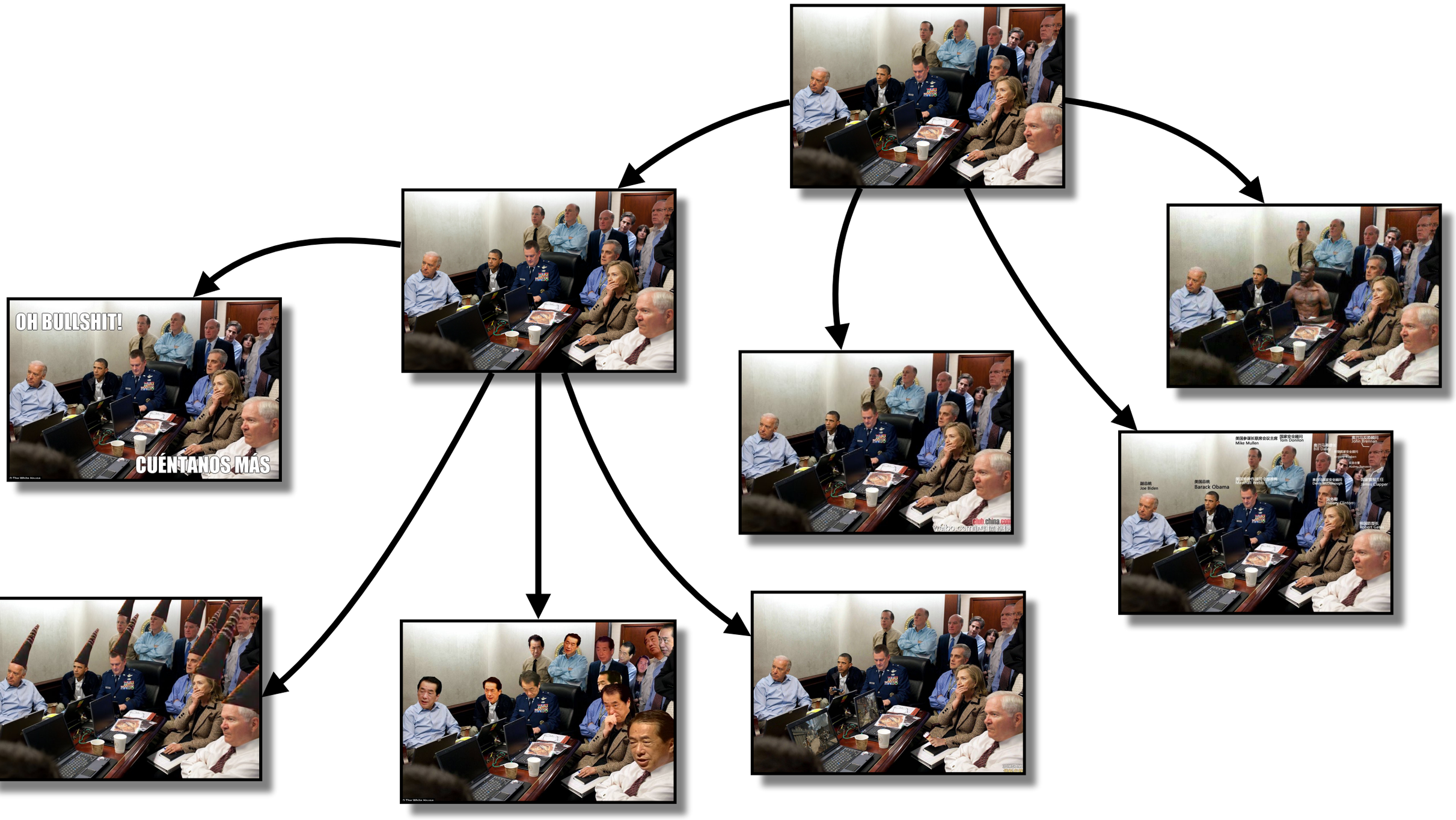


Source: Google Images

Near-duplicate detection

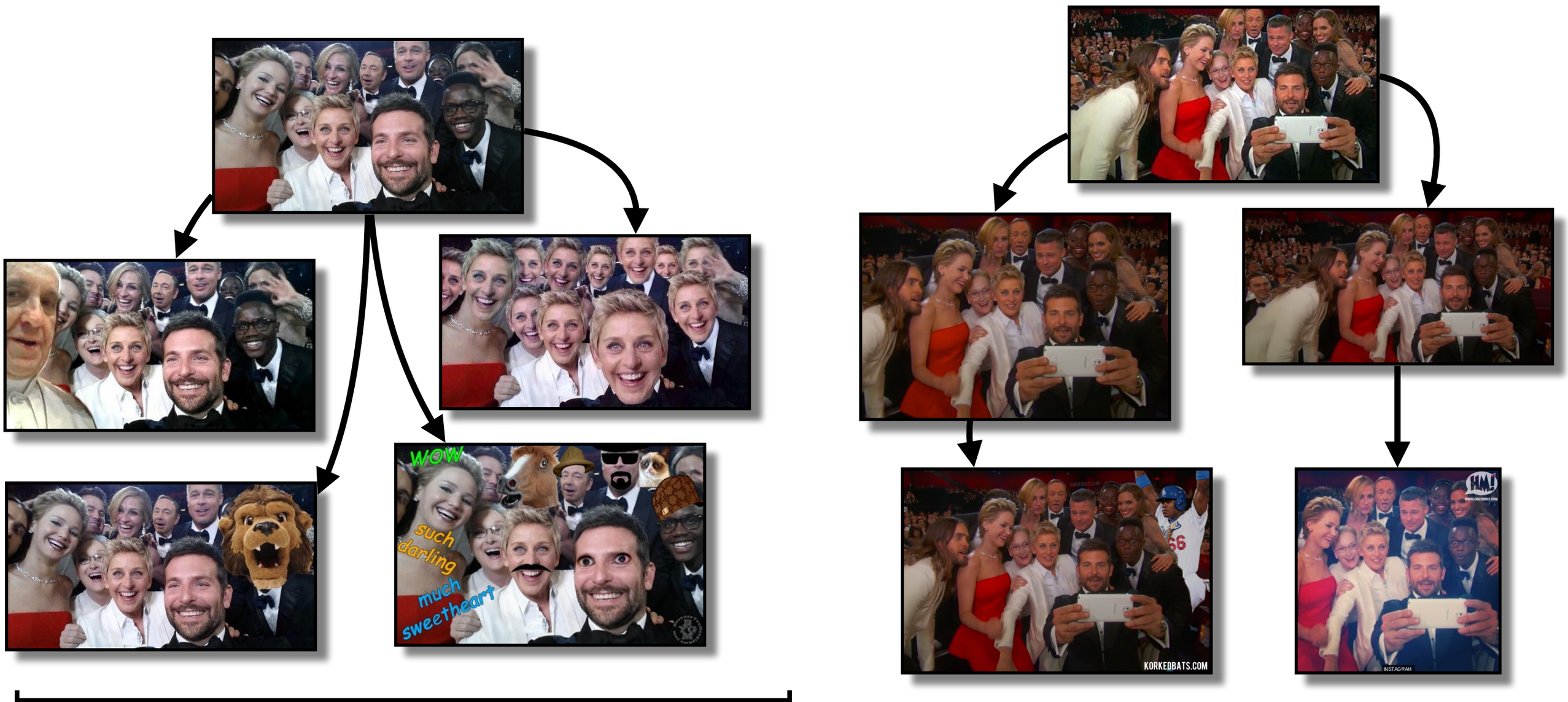
Near-duplicate detection

Multimedia Phylogeny



Multimedia Phylogeny

Multimedia Phylogeny



Main Objective

Design and develop solutions that allow us to identify the structure (phylogeny tree / forest) that represents the generation process overtime of images and videos

Contributions

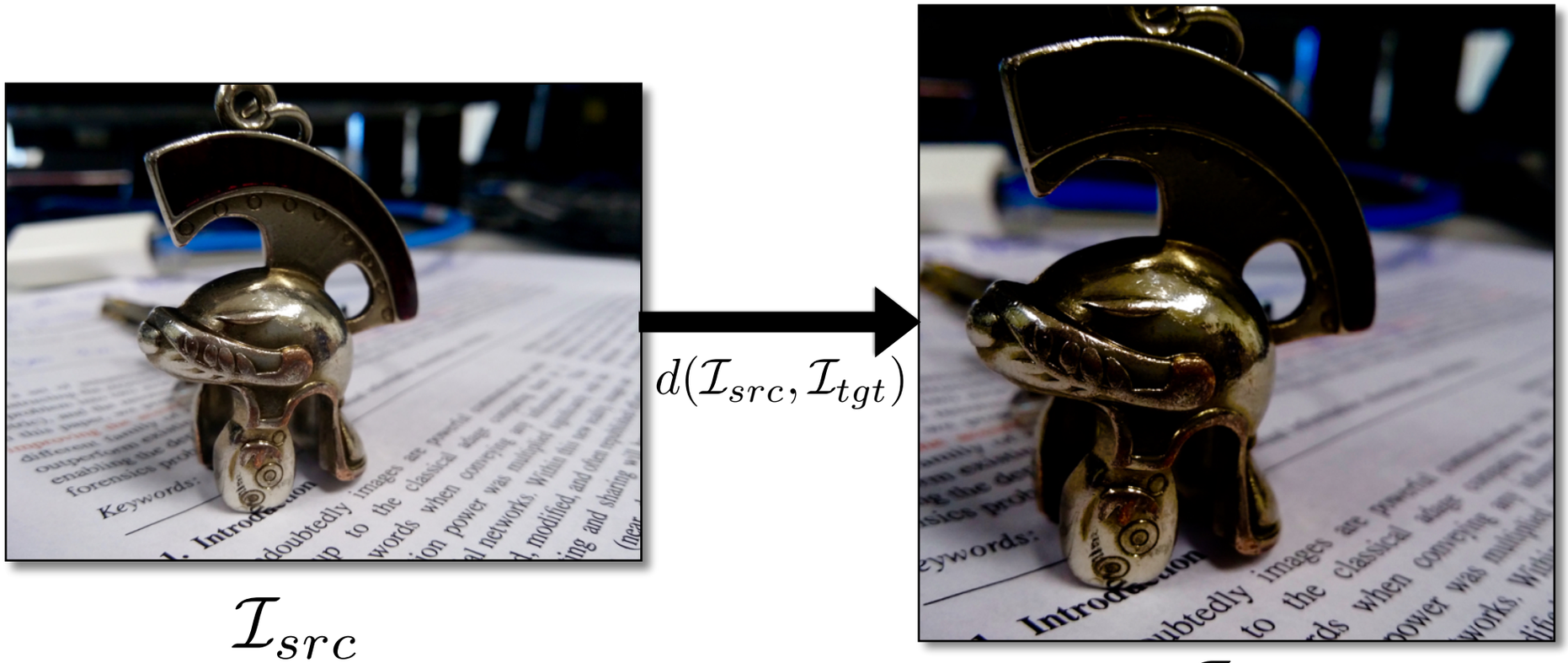
- Image phylogeny
  - Approaches to image phylogeny forest reconstruction
    - New dissimilarity measures for image phylogeny
- Video Phylogeny
  - Phylogeny reconstruction for misaligned and compressed video sequences

Image Phylogeny

Image Phylogeny

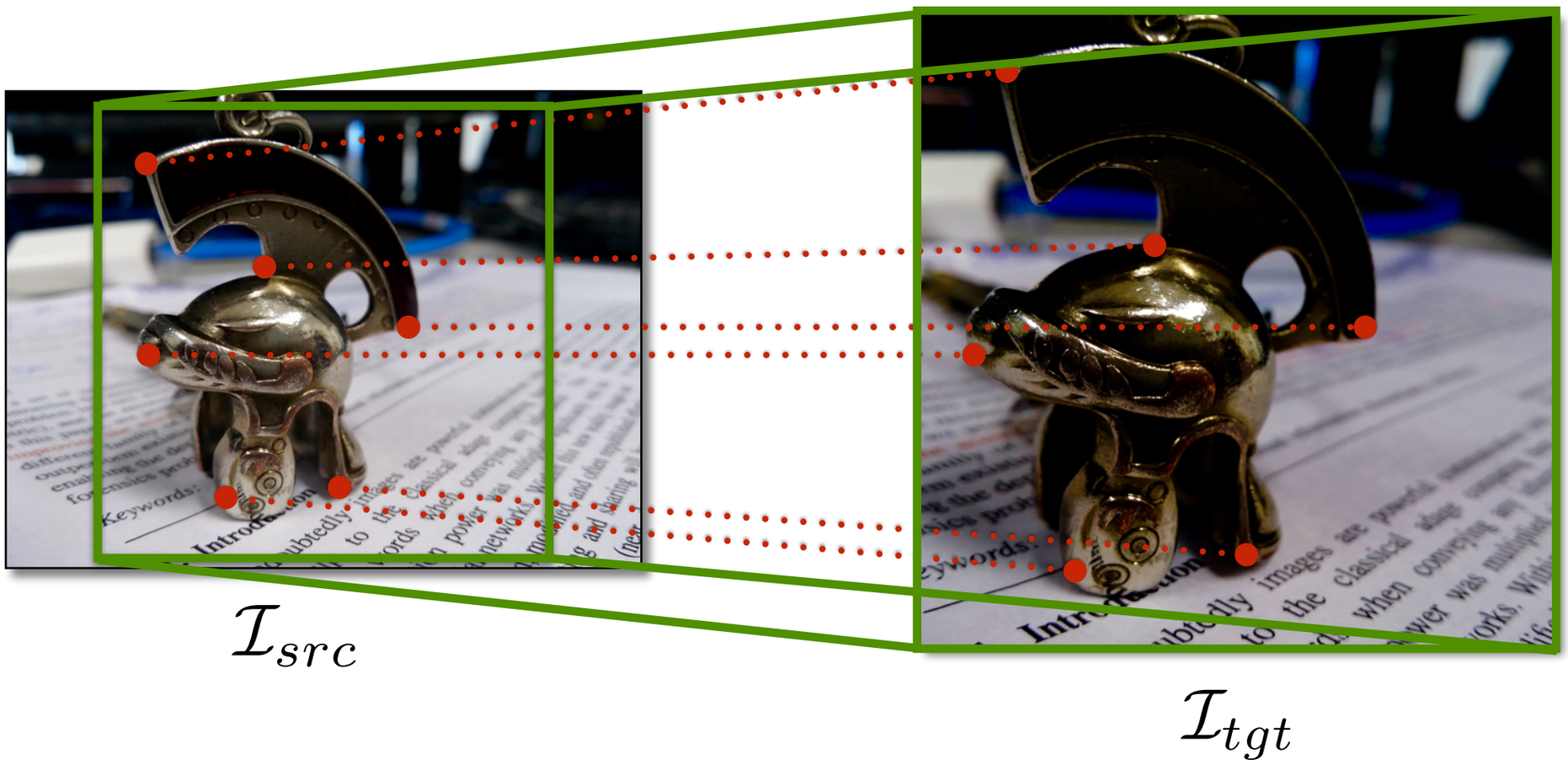
- Divided into two basic steps
  - Dissimilarity calculation
  - Phylogeny Reconstruction

Dissimilarity Calculation



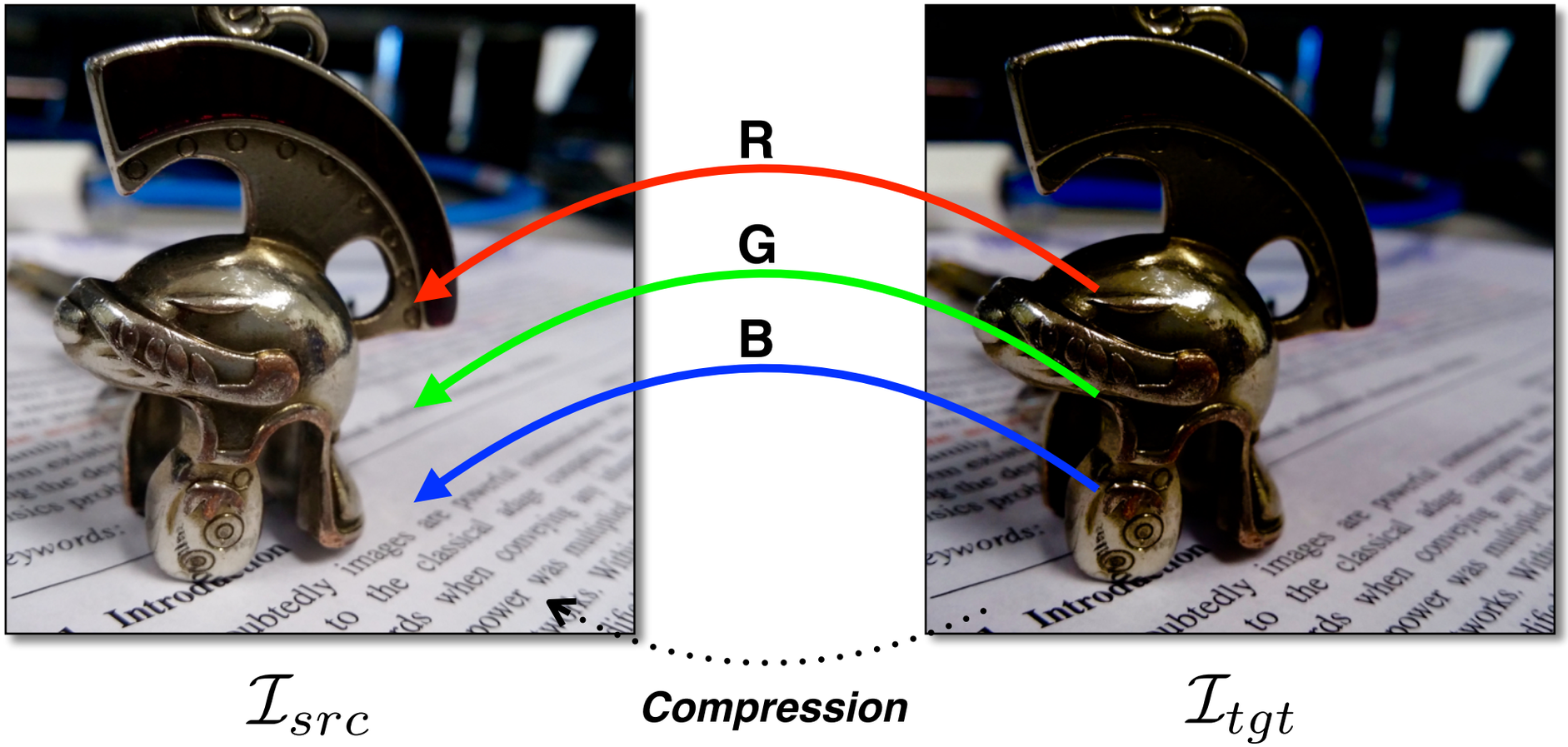
$$d(\mathcal{I}_{src}, \mathcal{I}_{tgt}) = \min_{T_{\overrightarrow{\beta}} \in \mathcal{T}} |\mathcal{I}_{tgt} - T_{\overrightarrow{\beta}}(\mathcal{I}_{src})| \text{ point-wise comparison } \mathcal{L}.$$

Dissimilarity Calculation



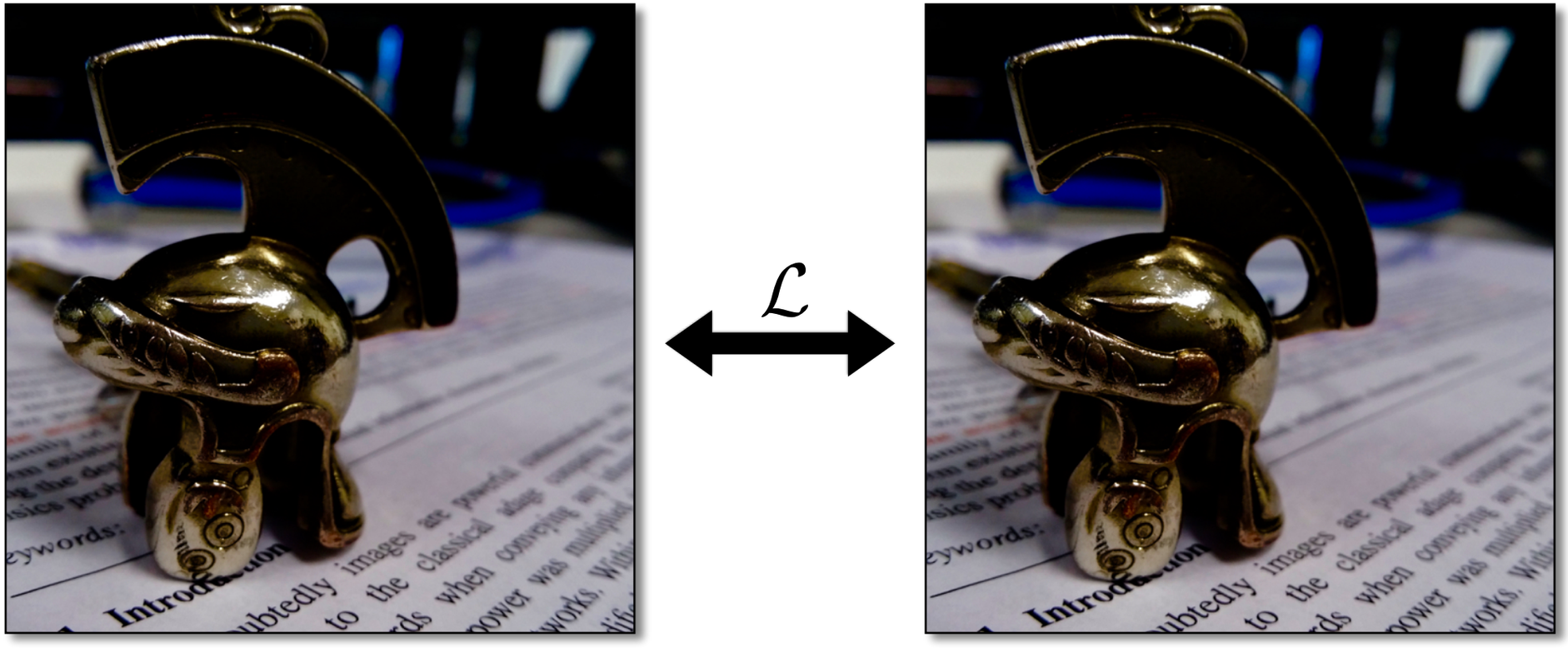
Geometric matching

Dissimilarity Calculation



Color and compression matching

Dissimilarity Calculation

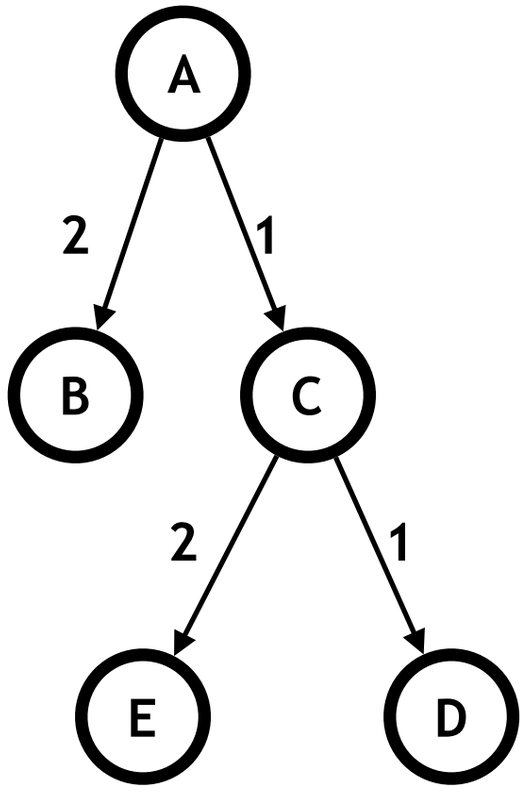


Point-wise comparison
(Mean Squared Error)

Phylogeny Reconstruction

	A	B	C	D	E
A	-	2	1	3	4
B	4	-	1	2	6
C	3	4	-	1	2
D	4	7	4	-	4
E	5	7	8	3	-

Dissimilarity matrix



Phylogeny tree

Phylogeny Reconstruction

- Oriented Kruskal (OK) [Dias et al. (2010)]
  - Based on the classic Kruskal's Minimum Spanning Tree algorithm
  - Adapted for oriented graphs
- Optimum Branching (OB) [Edmonds (1957)]
  - Always returns the optimum tree

Image Phylogeny Contributions

- Image Phylogeny Forests Reconstruction

F. O. Costa, M. Oikawa, Z. Dias, S. Goldenstein, and A. Rocha. *Image phylogeny forests reconstruction*. IEEE Transactions on Information Forensics and Security (TIFS), vol.9, no. 10, pp. 1533-1546, 2014

- New dissimilarity measures

F. O. Costa, A. Oliveira, P. Ferrara, Z. Dias, S. Goldenstein, and A. Rocha. *New dissimilarity measures for image phylogeny reconstruction*. Submitted to Springer Pattern Analysis and Applications, 2016

Image Phylogeny Forests Reconstruction

Image Phylogeny Forests Reconstruction

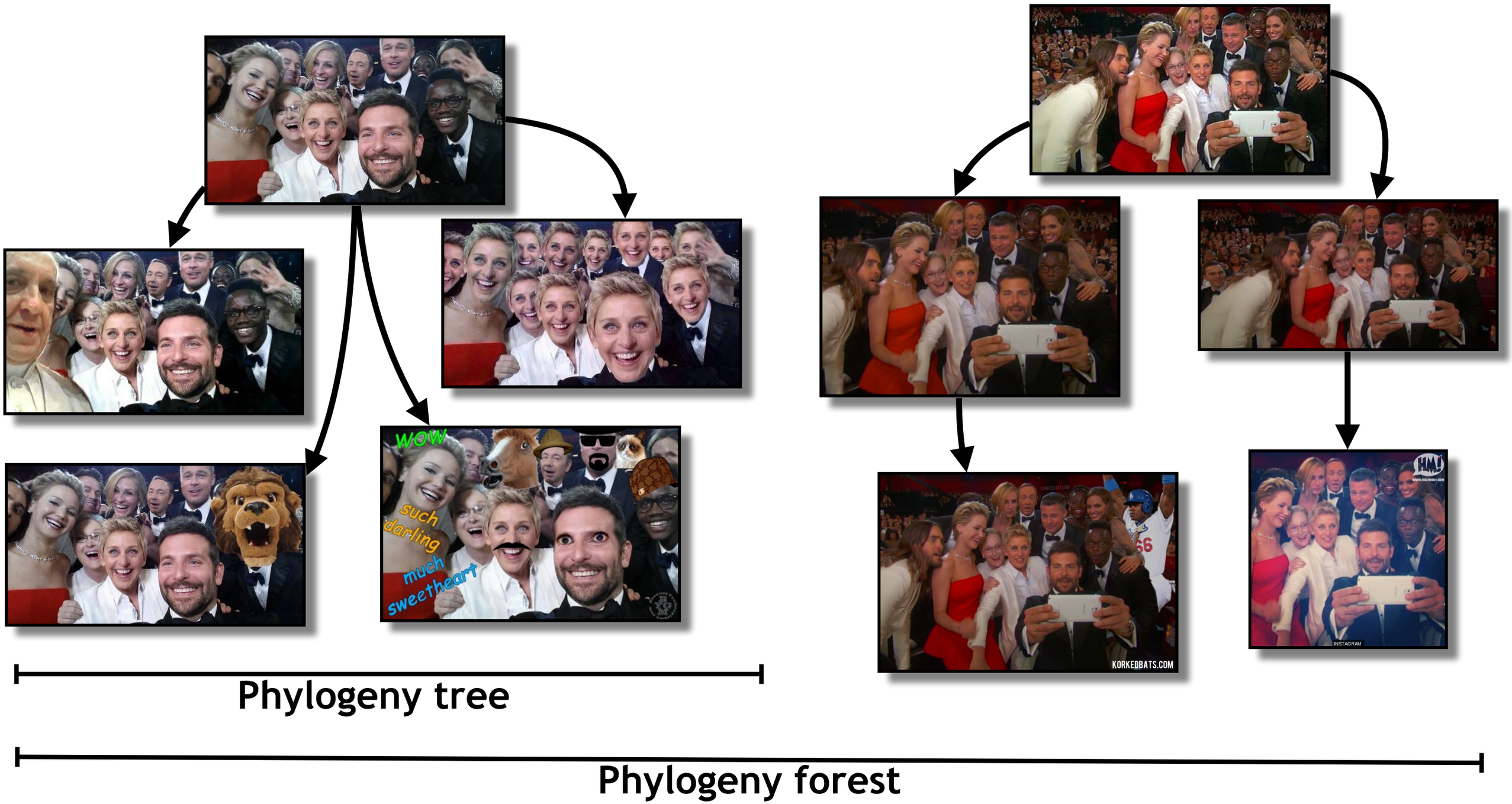


Image Phylogeny Forests Reconstruction

Dias et al.(2013b): Automatic Oriented Kruskal (AOK)

- Adaptation of the OK algorithm for forests
- Select the edges in crescent order
- Stops selecting edges according to a threshold

Hypothesis #1

Using exact approaches for phylogeny forests reconstruction problem is at least as effective than using heuristic-based approaches

Image Phylogeny Forests Reconstruction

We designed and developed an algorithm for forests reconstruction

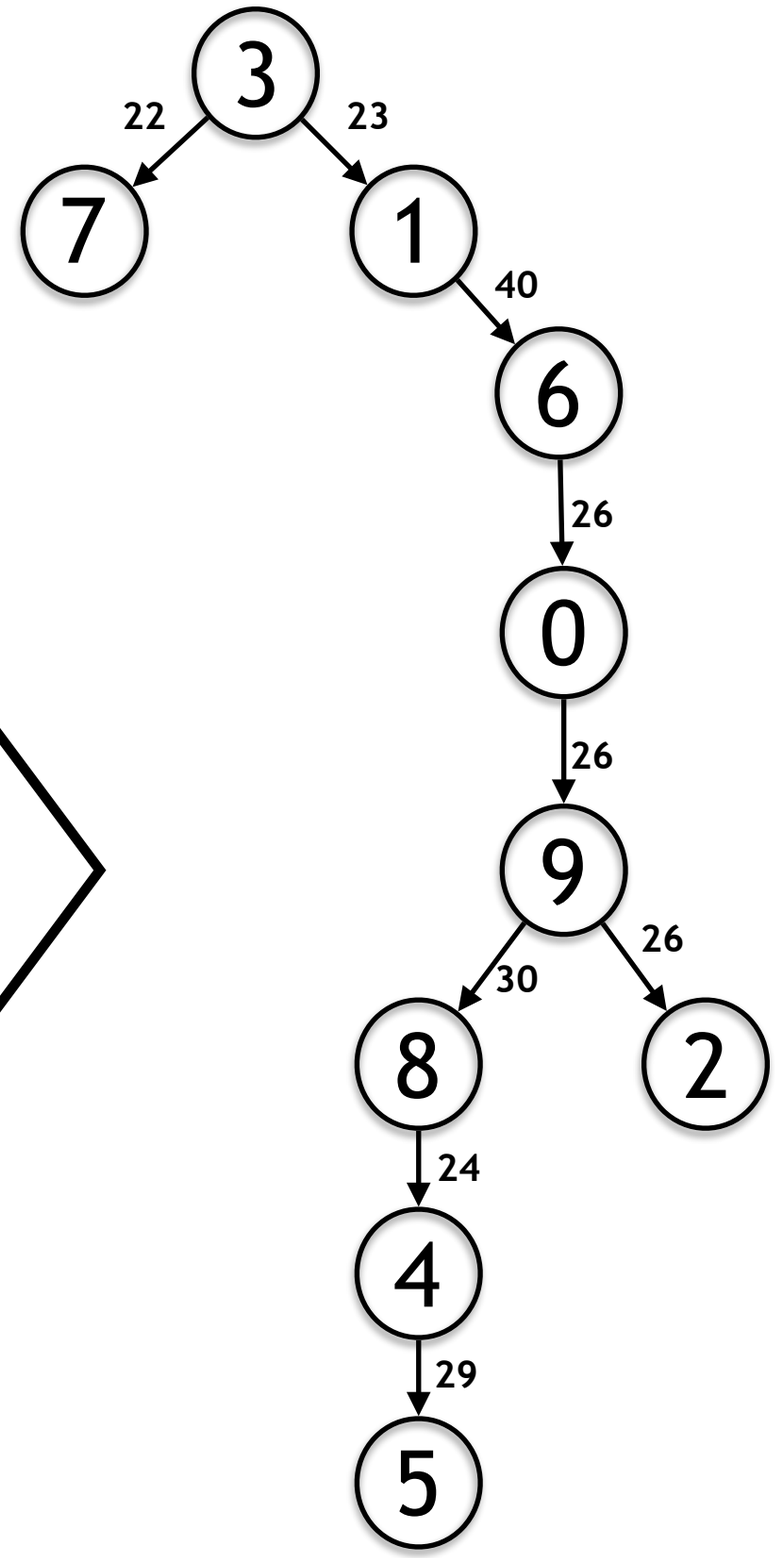
- Automatic Optimum Branching (AOB)
  - Adaptation of the OB algorithm

Automatic Optimum Branching (AOB)

Dissimilarity matrix

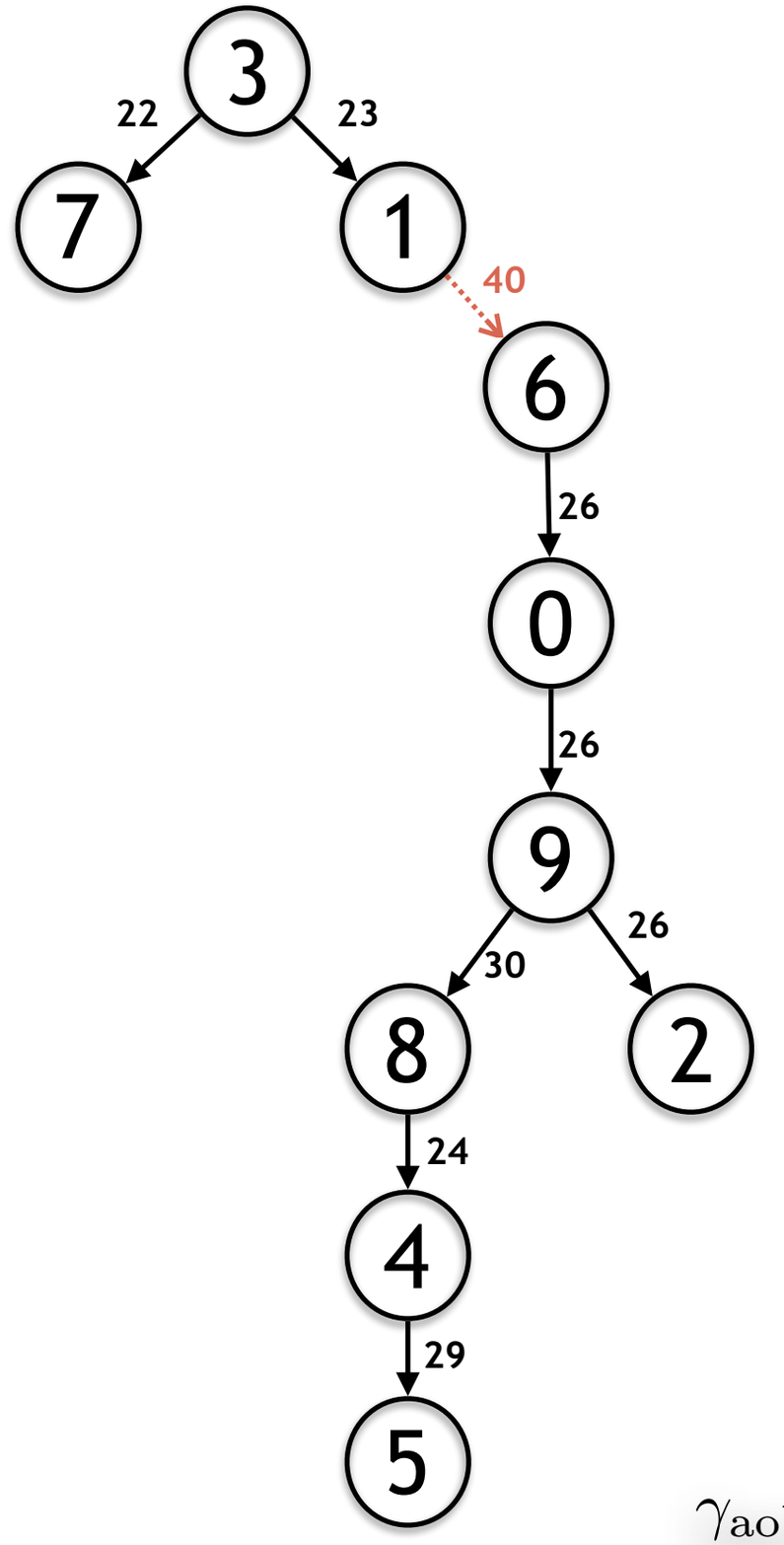
M	0	1	2	3	4	5	6	7	8	9
0	-	52	29	57	43	55	24	52	49	26
1	50	-	45	35	60	68	40	42	63	53
2	42	48	-	56	42	54	28	52	50	29
3	56	23	55	-	65	67	50	22	66	58
4	45	61	35	67	-	29	42	63	43	52
5	25	69	65	70	27	-	48	68	45	65
6	26	50	30	52	39	50	-	50	48	48
7	51	32	49	30	61	67	49	-	62	55
8	52	65	55	69	24	30	49	64	-	26
9	25	54	26	59	50	61	30	56	30	-

Optimum Branching



Automatic Optimum Branching (AOB)

Step	X	Y	d(X,Y)	Diff	$\gamma_{\mathrm{aob}} \times \sigma$
1	3	7	22	-	-
2	3	1	23	1	-
3	8	4	24	1	1,41
4	9	2	26	2	2,00
5	6	0	26	0	3,42
6	0	9	26	0	3,58
7	4	5	29	3	3,52
8	9	8	30	1	4,68
9	1	6	40	10	5,53



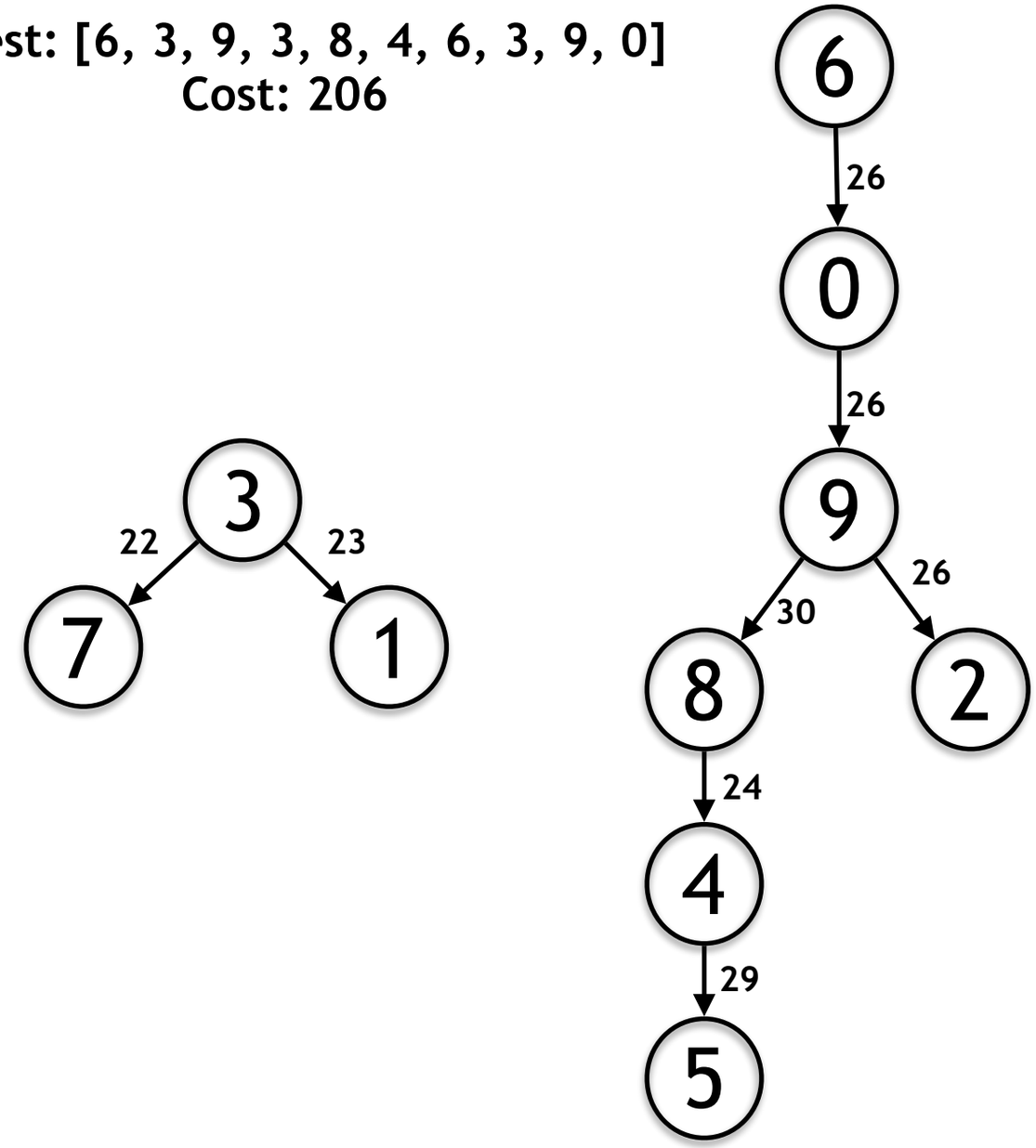
$\gamma_{\mathrm{aob}} = 2.0$

Automatic Optimum Branching (AOB)

Dissimilarity matrix

M	0	1	2	3	4	5	6	7	8	9
0	-	52	29	57	43	55	24	52	49	26
1	50	-	45	35	60	68	40	42	63	53
2	42	48	-	56	42	54	28	52	50	29
3	56	23	55	-	65	67	50	22	66	58
4	45	61	35	67	-	29	42	63	43	52
5	25	69	65	70	27	-	48	68	45	65
6	26	50	30	52	39	50	-	50	48	48
7	51	32	49	30	61	67	49	-	62	55
8	52	65	55	69	24	30	49	64	-	26
9	25	54	26	59	50	61	30	56	30	-

Forest: [6, 3, 9, 3, 8, 4, 6, 3, 9, 0]
Cost: 206

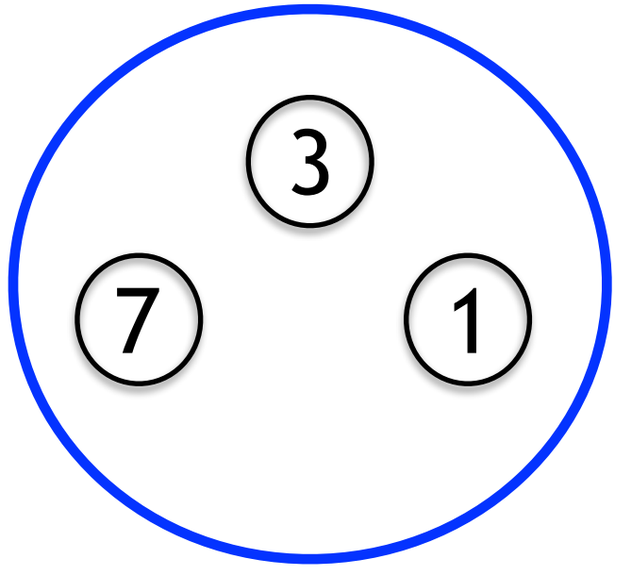


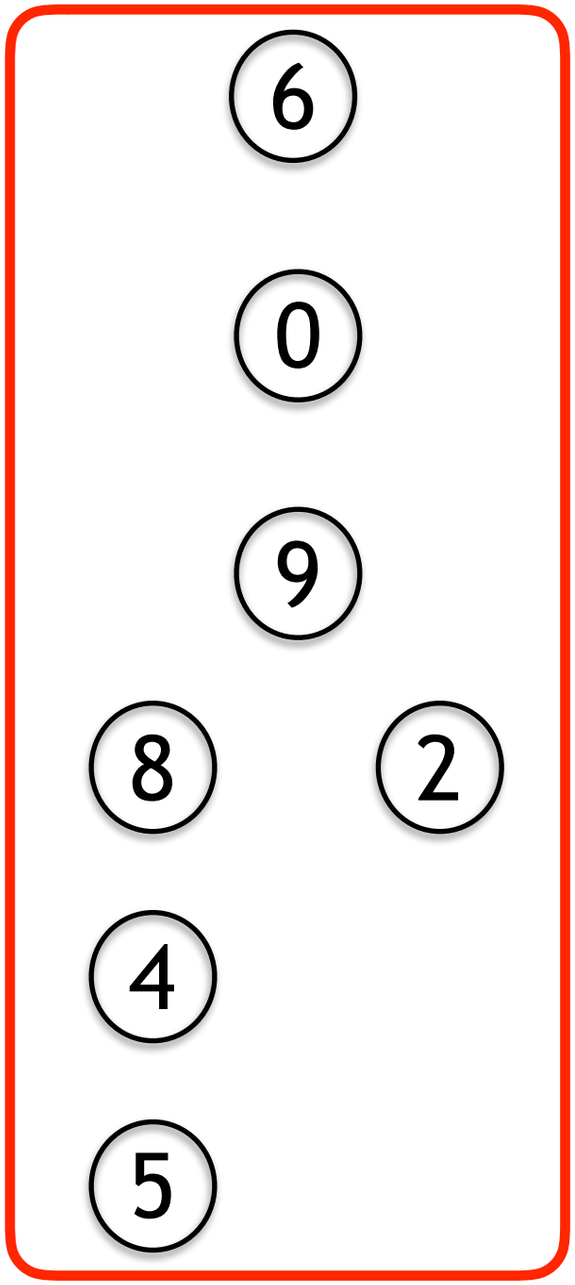
$\gamma_{\text{aob}} = 2.0$

Extended Automatic Optimum Branching (E-AOB)

Dissimilarity matrix

M	1	3	7	0	2	4	5	6	8	9
1	-	35	42	-	-	-	-	-	-	-
3	23	-	22	-	-	-	-	-	-	-
7	32	30	-	-	-	-	-	-	-	-
0	-	-	-	-	29	43	55	24	49	26
2	-	-	-	42	-	42	54	28	50	29
4	-	-	-	45	35	-	29	42	43	52
5	-	-	-	25	65	27	-	48	45	65
6	-	-	-	26	30	39	50	-	48	48
8	-	-	-	52	55	24	30	49	-	26
9	-	-	-	25	26	50	61	30	30	-





Extended Automatic Optimum Branching (E-AOB)

Dissimilarity matrix

M	1	3	7	0	2	4	5	6	8	9
1	-	35	42	-	-	-	-	-	-	-
3	23	-	22	-	-	-	-	-	-	-
7	32	30	-	-	-	-	-	-	-	-
0	-	-	-	-	29	43	55	24	49	26
2	-	-	-	42	-	42	54	28	50	29
4	-	-	-	45	35	-	29	42	43	52
5	-	-	-	25	65	27	-	48	45	65
6	-	-	-	26	30	39	50	-	48	48
8	-	-	-	52	55	24	30	49	-	26
9	-	-	-	25	26	50	61	30	30	-

Forest: [5, 3, 9, 3, 8, 4, 0, 3, 9, 0]
Cost: 204

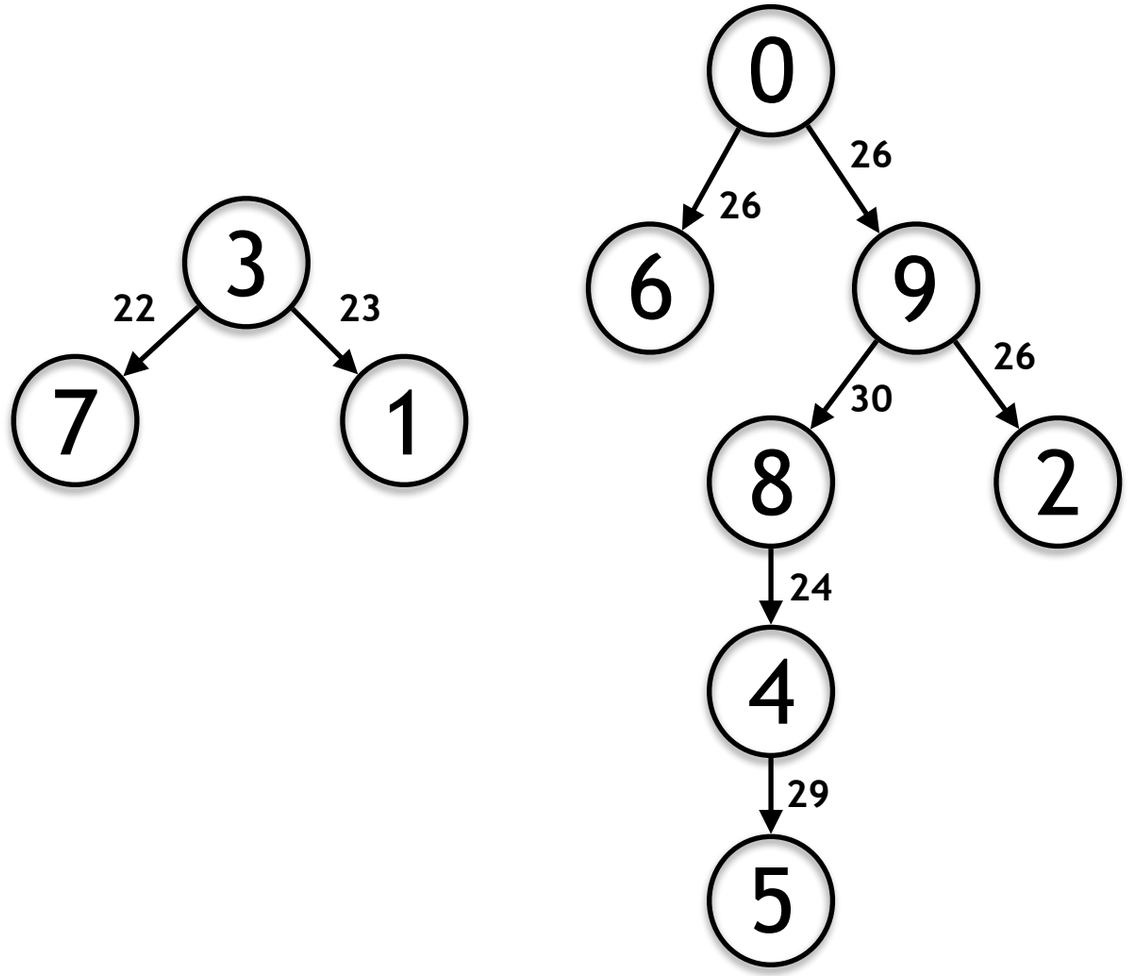


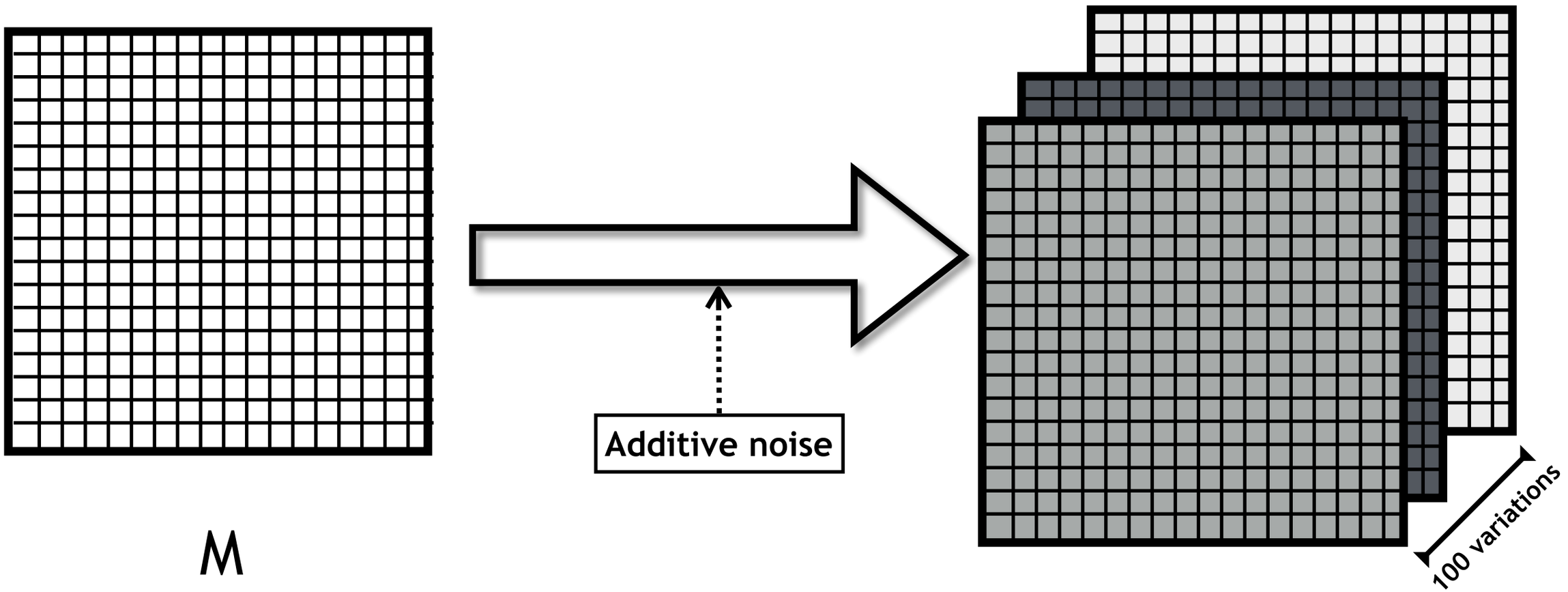
Image Phylogeny Forests Reconstruction

- Main approaches
  - AOK: Greedy approach
  - AOB/E-AOB: Exact approach

Hypothesis #2

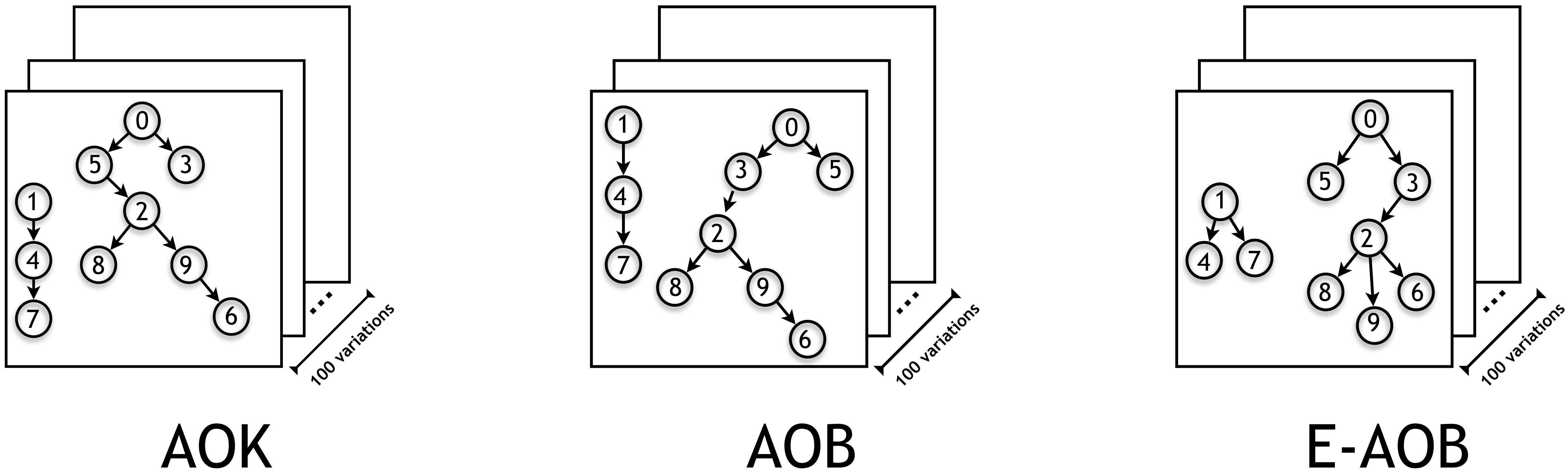
It is possible to lead to complementary properties of different forest reconstruction algorithm and combine them for improving the quality of the phylogeny forest reconstruction

Fusion approach



M

Fusion approach

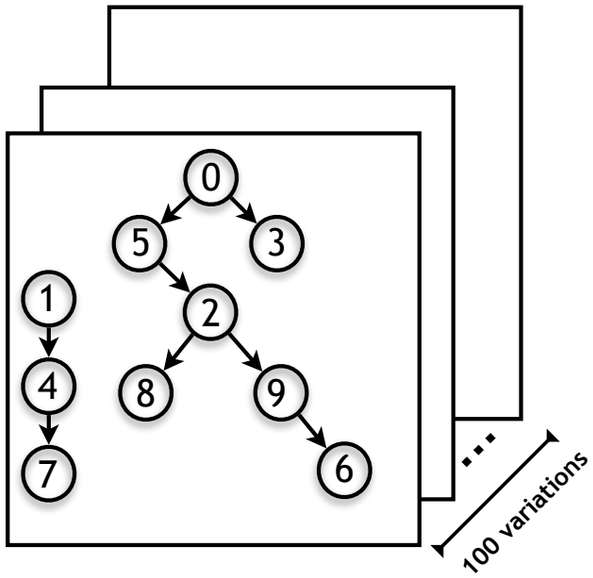


Tree count

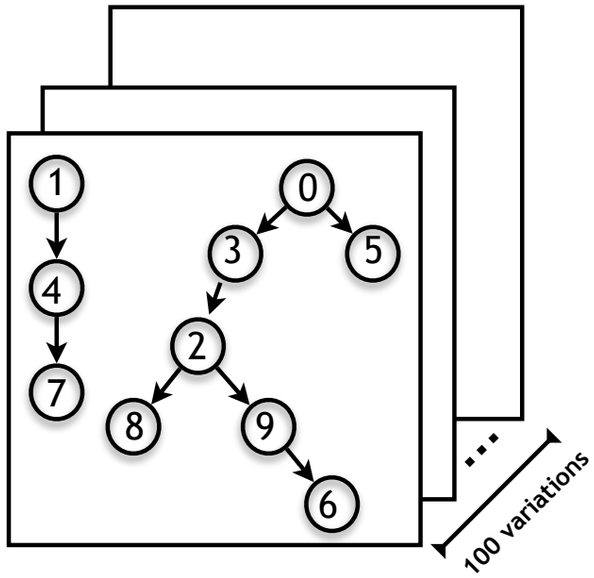
V = [1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, **2, 2**, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3,
3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5]

Median(V) = (2 + 2)/2 = 2

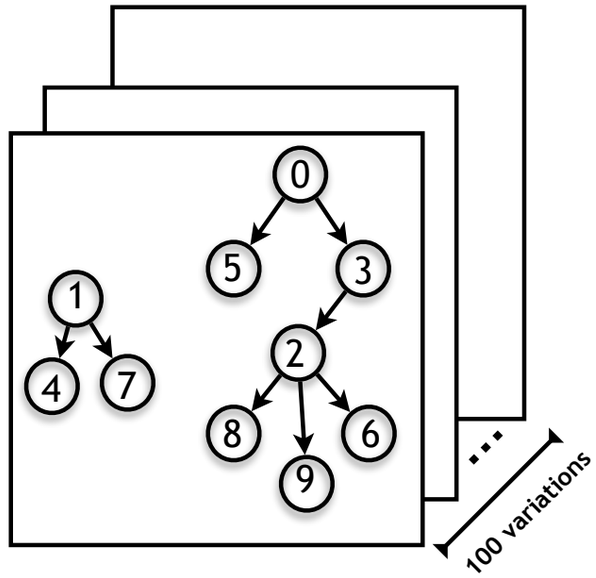
Fusion approach



AOK



AOB

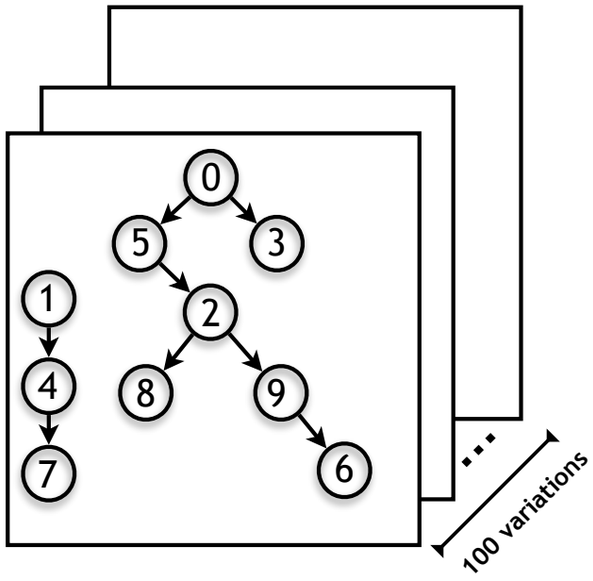


E-AOB

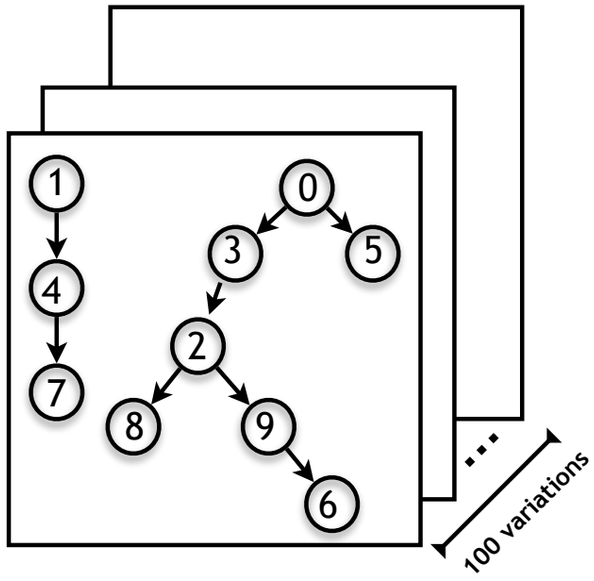
Sum of votes of each node
[**260, 250**, 90, 30, 35, 24, 0, 0, 0, 0]

Roots {0, 1}

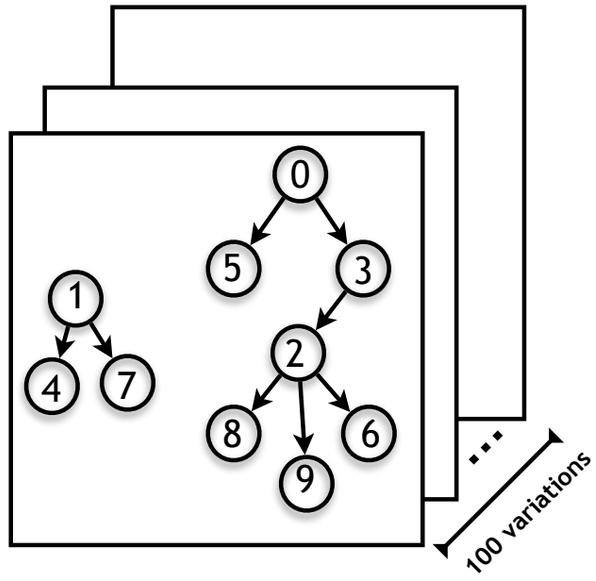
Fusion approach



AOK

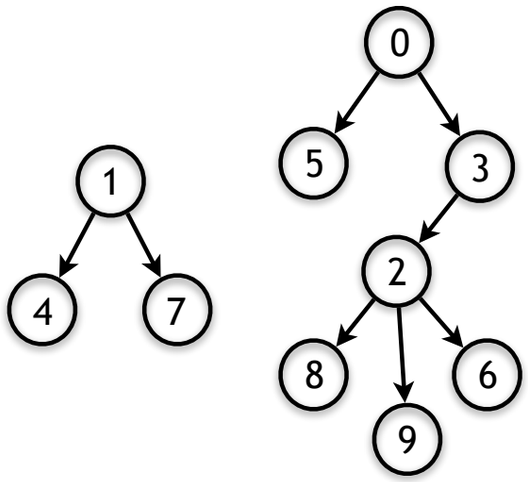


AOB



E-AOB

M	0	1	2	3	4	5	6	7	8	9
0	-	-	0	270	0	286	0	0	0	0
1	-	-	0	0	255	0	0	280	0	0
2	-	-	-	0	0	0	130	0	270	300
3	-	-	280	-	0	0	0	0	30	0
4	-	-	0	0	-	0	0	20	0	0
5	-	-	20	0	0	-	0	0	0	0
6	-	-	0	0	0	0	-	0	0	0
7	-	-	0	0	10	0	0	-	0	0
8	-	-	0	0	0	0	0	0	-	0
9	-	-	0	0	0	0	170	0	0	-



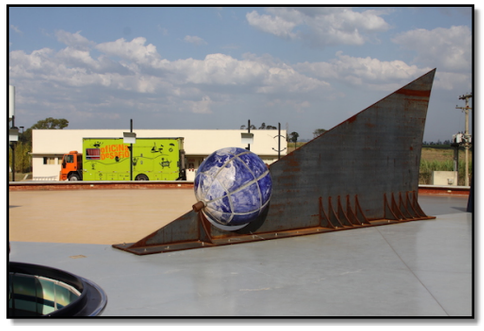
Final forest
[0, 1, 3, 0, 1, 0, 2, 1, 2, 2]

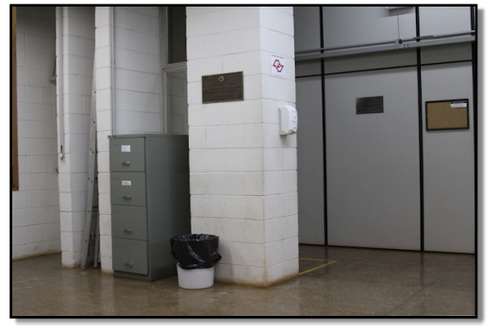
Experiments

Datasets

- 36,000 test cases
  - [2..10] trees
- Ground truth is available
- One Camera and Multiple Cameras



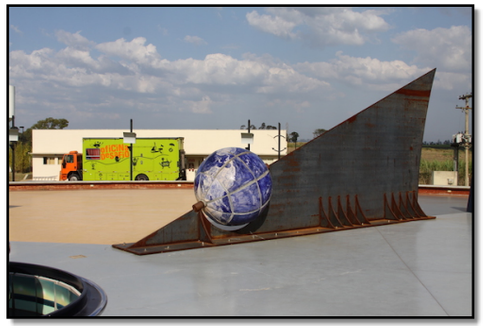


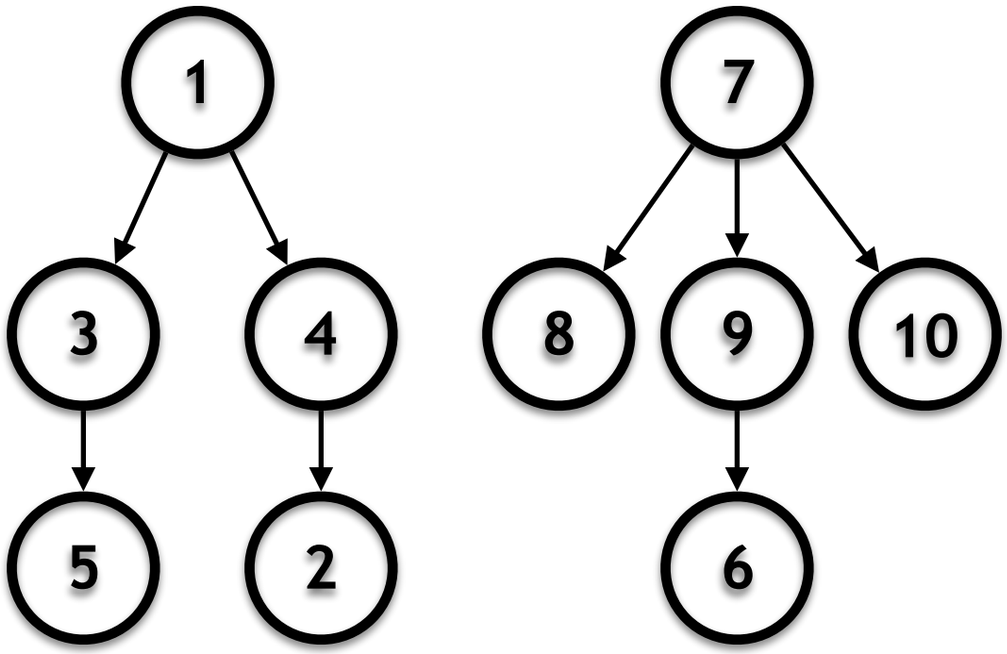
Datasets

- Transformations
  - Crop
  - Resize
  - Brightness, contrast and gamma correction
  - Rotation
  - JPEG Compression



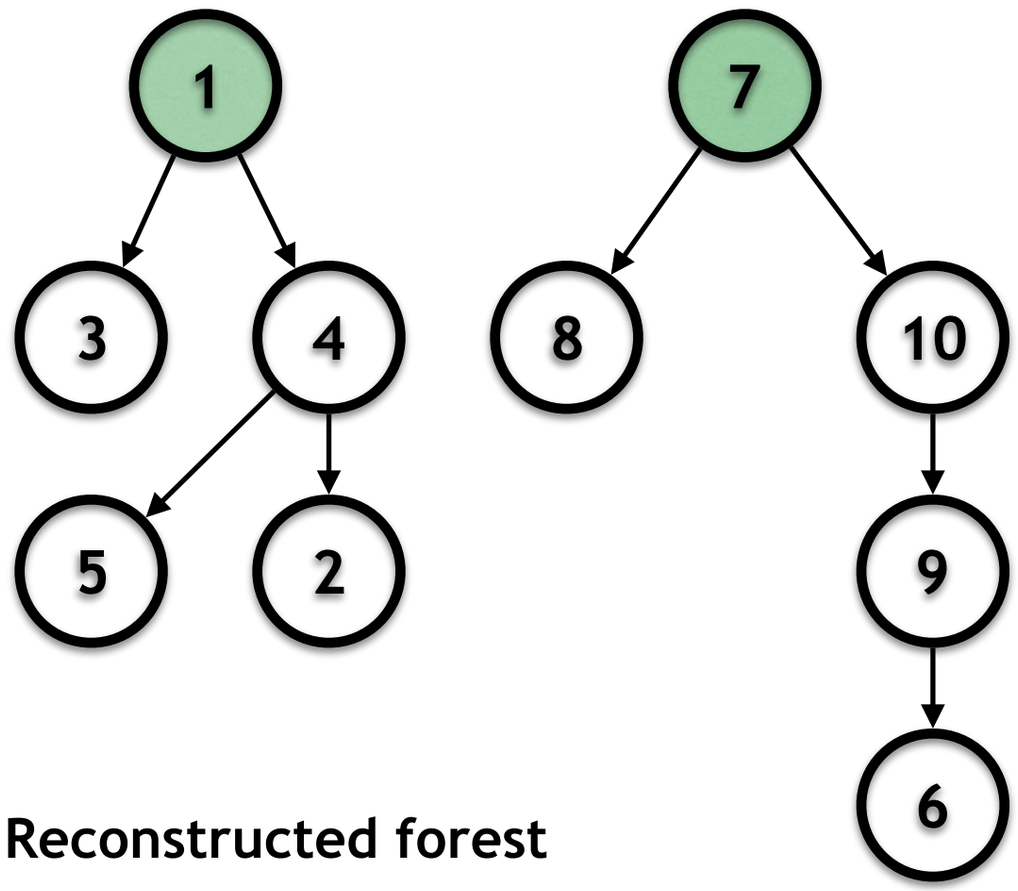



Evaluation



Ground truth

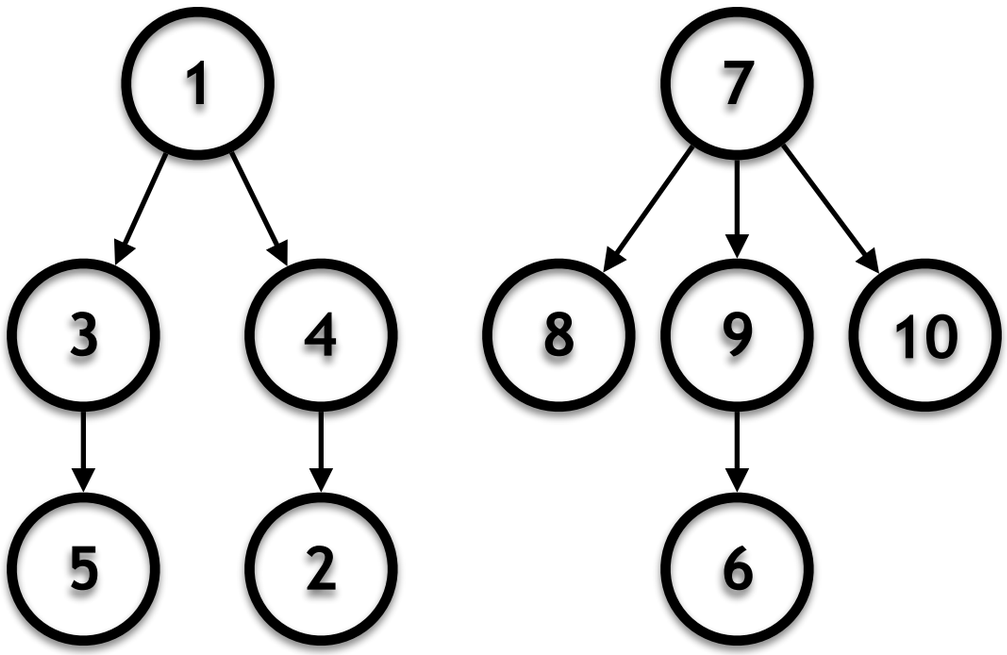
Roots



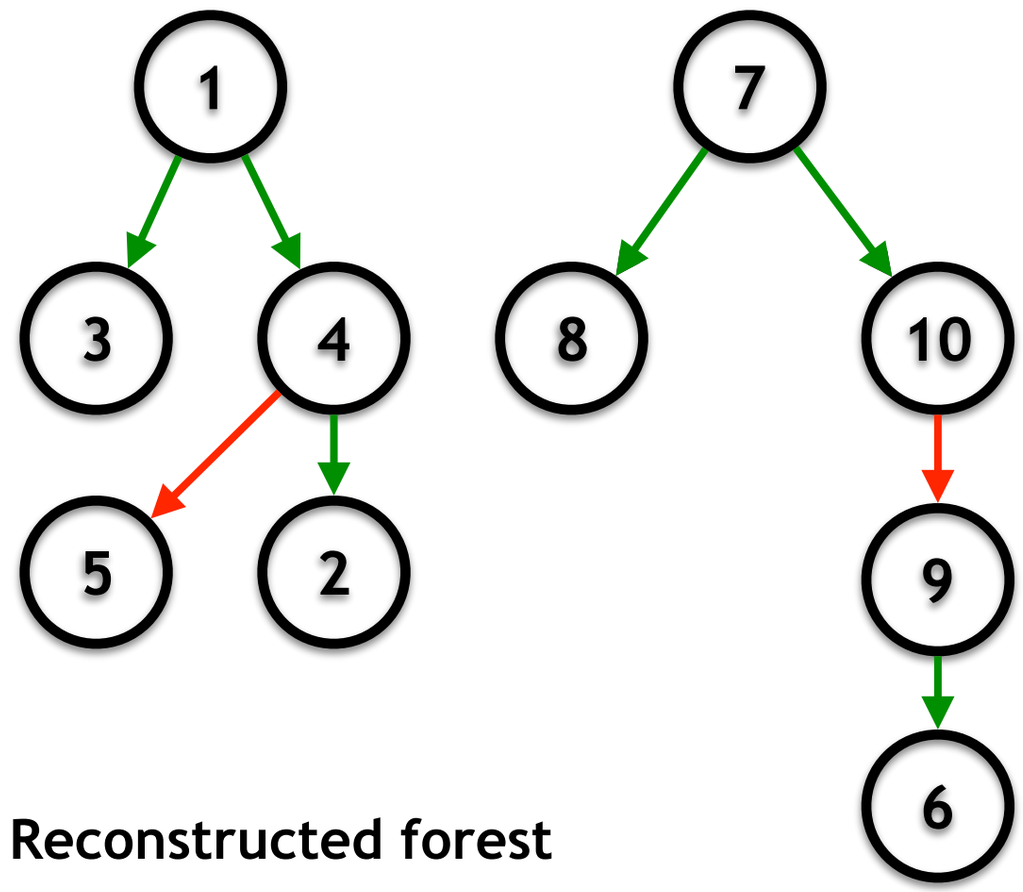
Reconstructed forest

Evaluation

Edges

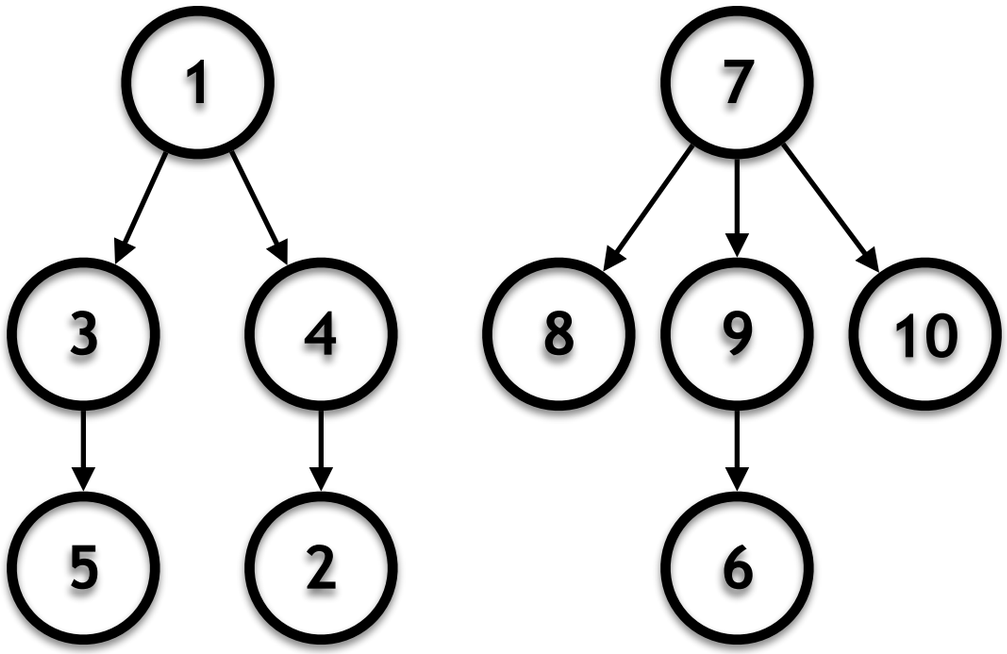


Ground truth



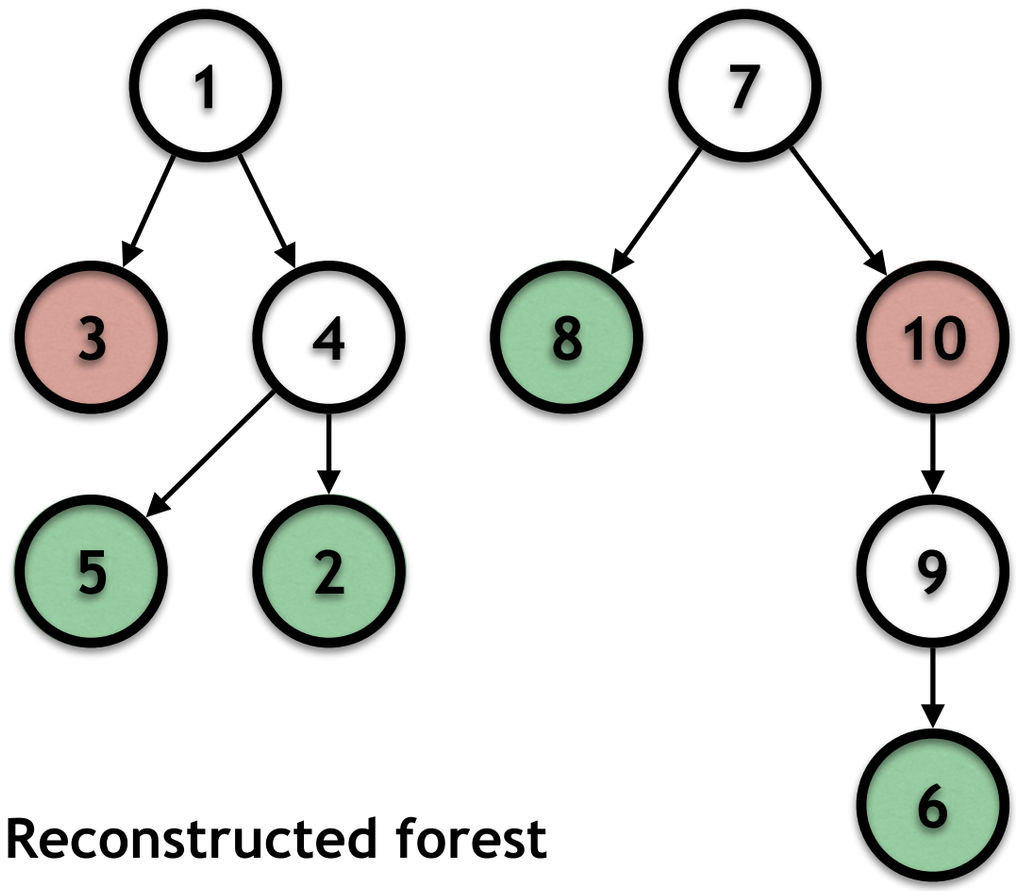
Reconstructed forest

Evaluation



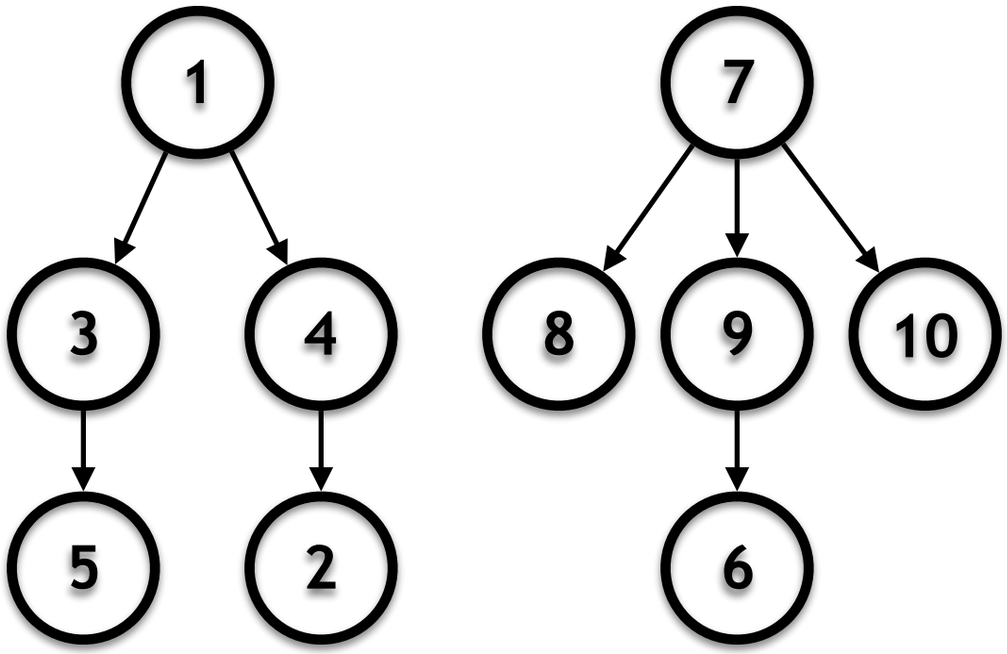
Ground truth

Leaves



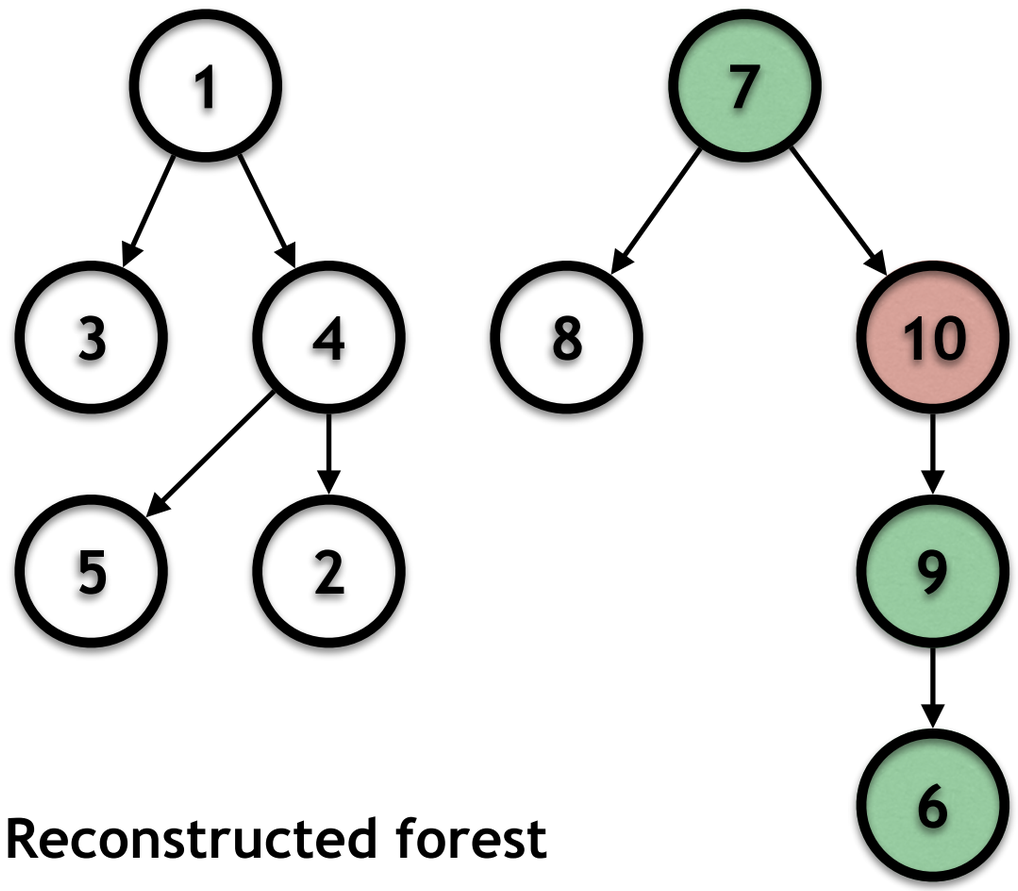
Reconstructed forest

Evaluation



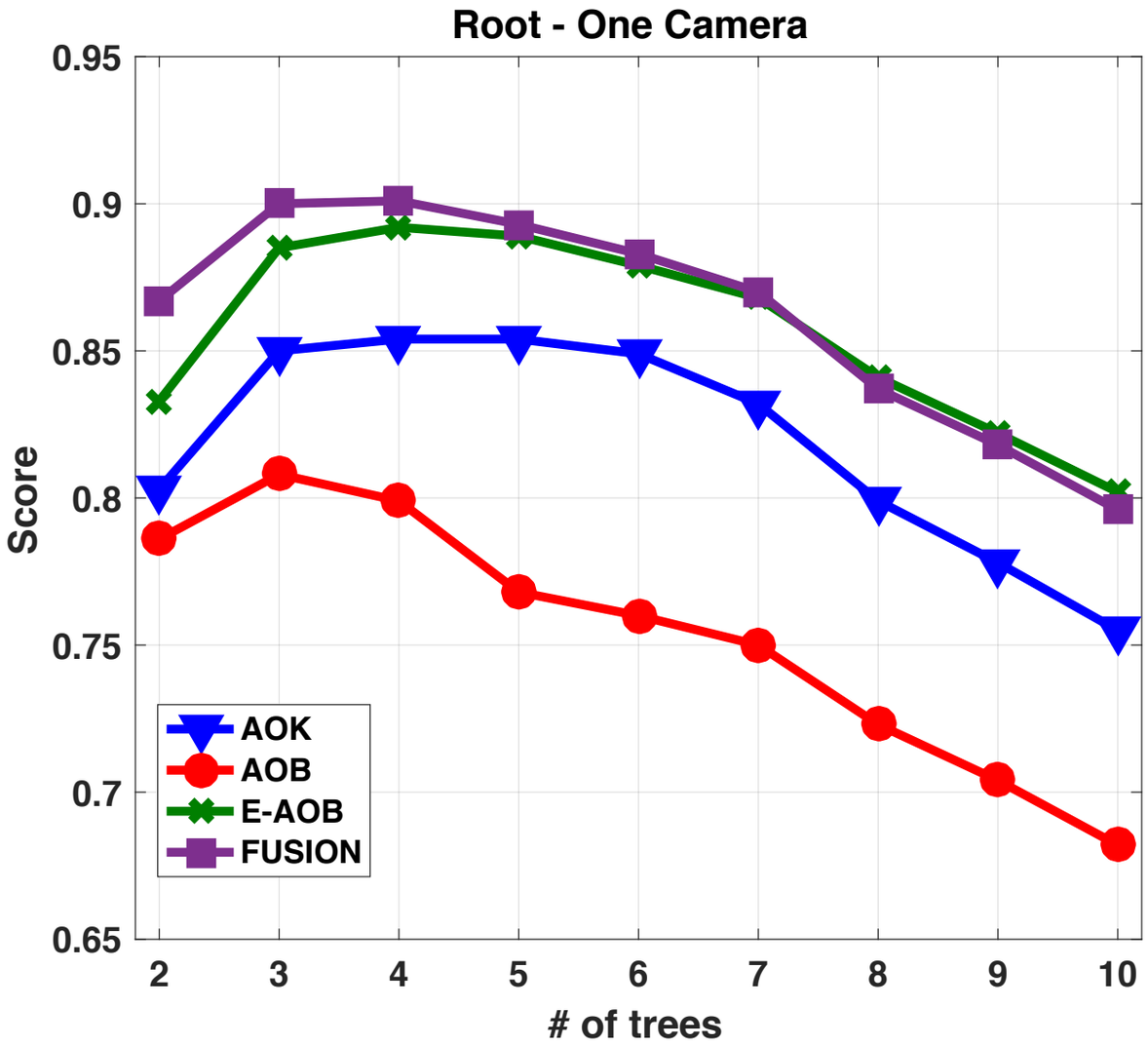
Ground truth

Ancestry

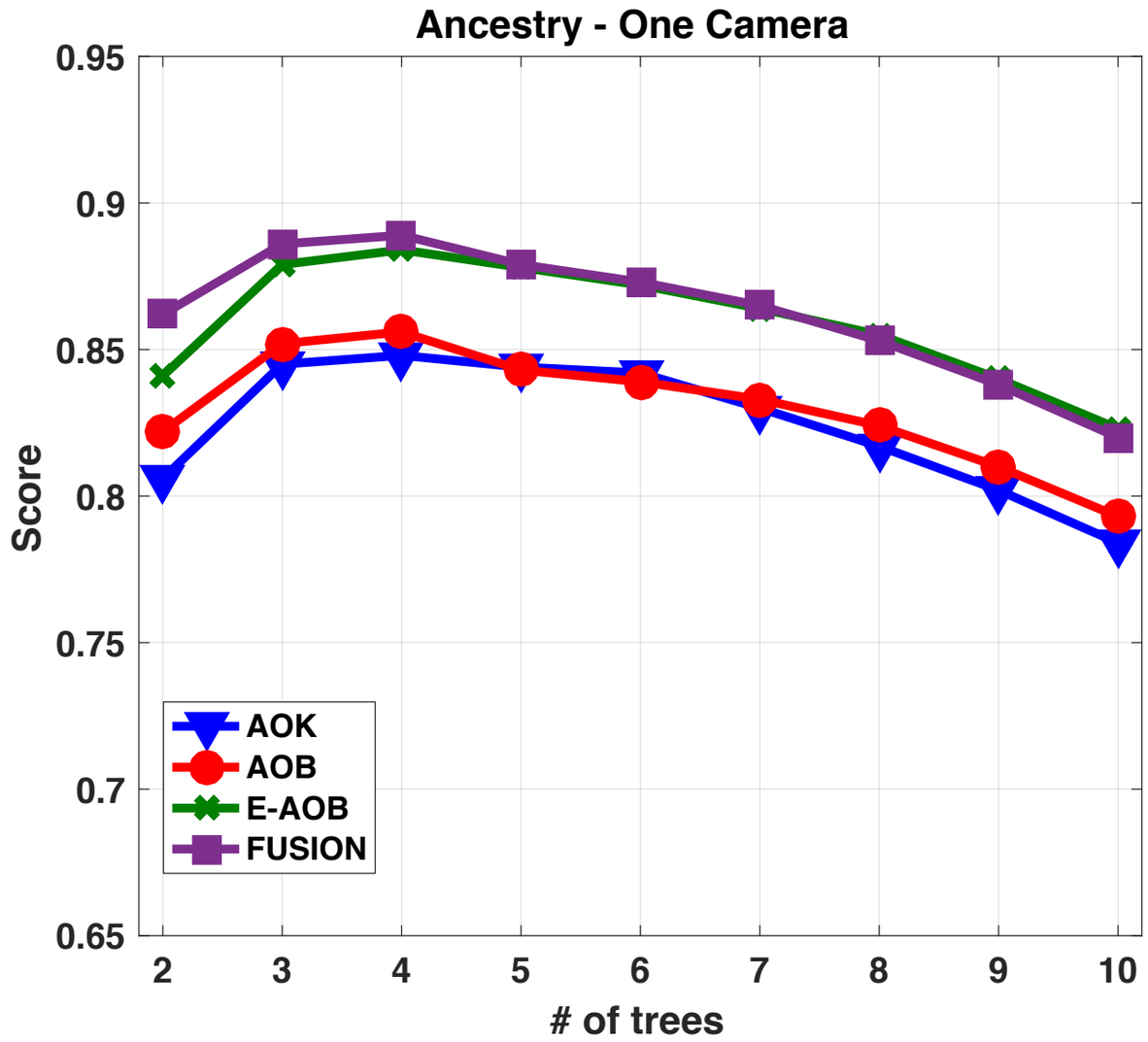


Reconstructed forest

Results - One Camera

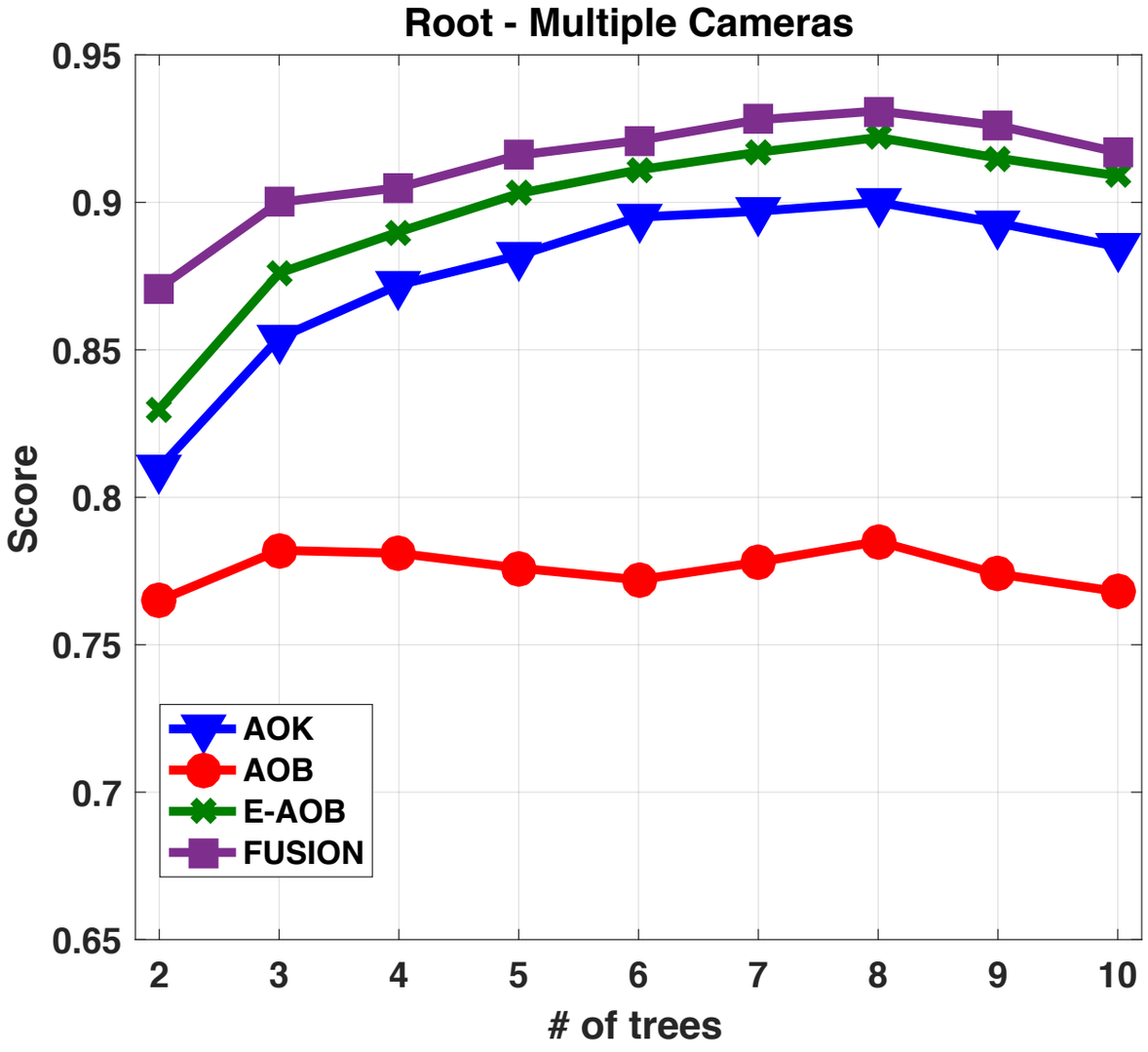


Error Reduction (E-AOB vs. AOK): 19%
Error Reduction (FUSION vs. AOK): 17%

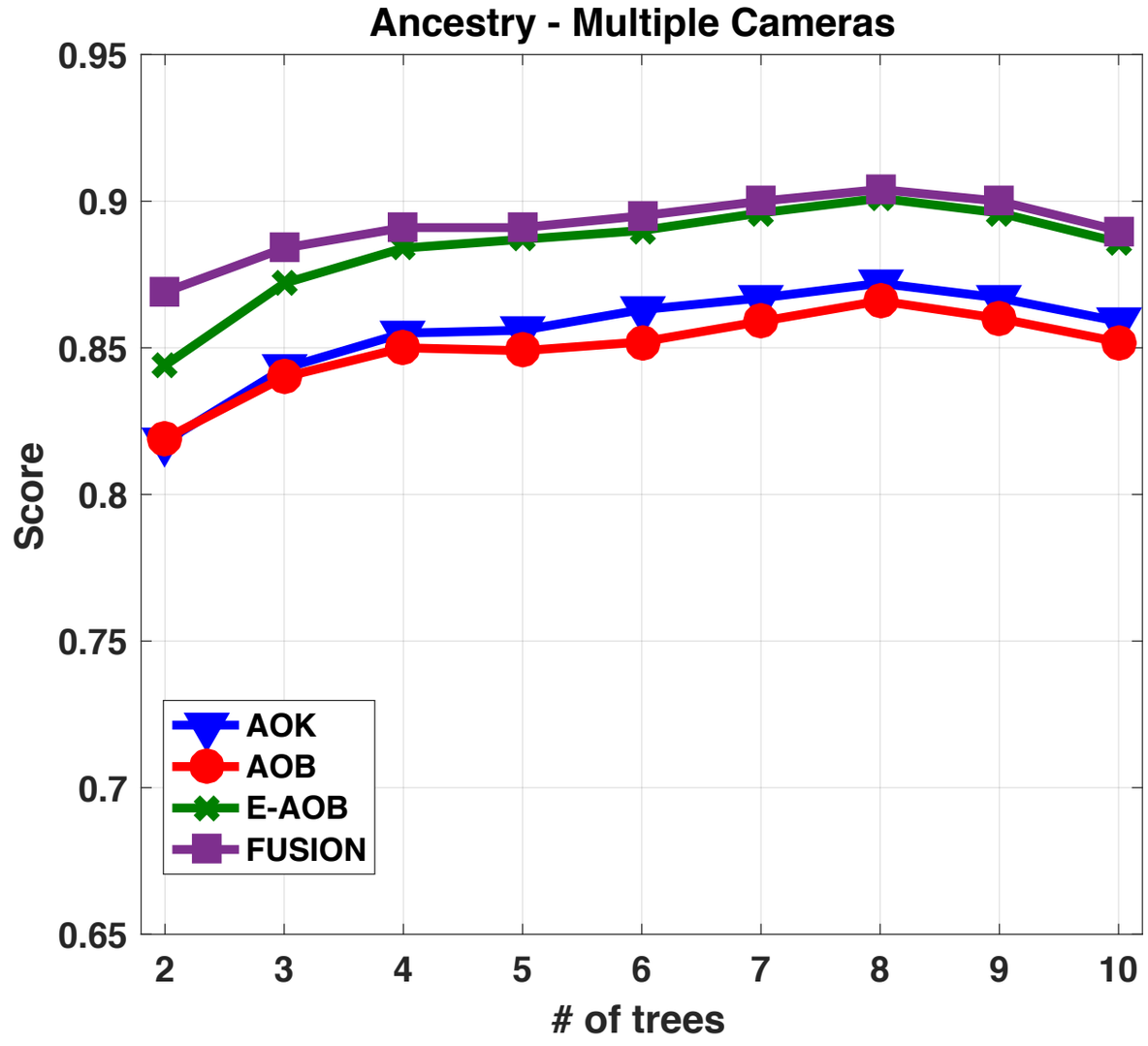


Error Reduction (E-AOB vs. AOK): 18%
Error Reduction (Fusion vs. AOK): 17%

Results - Multiple Cameras



Error Reduction (E-AOB vs. AOK): 20%
Error Reduction (FUSION vs. AOK): 28%



Error Reduction (E-AOB vs. AOK): 19%
Error Reduction (Fusion vs. AOK): 21%

New dissimilarity measures for image phylogeny reconstruction

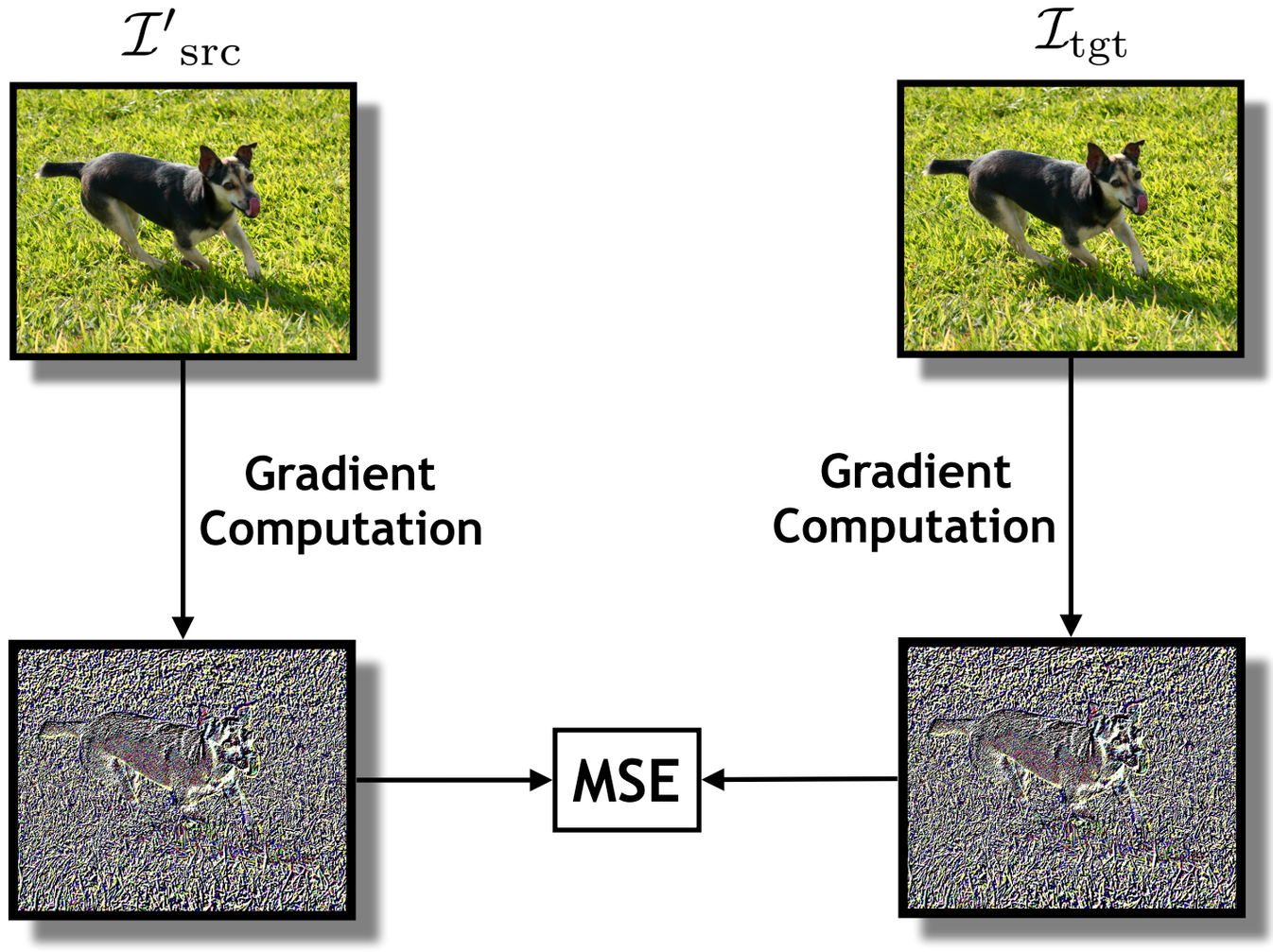
New dissimilarity measures for image phylogeny reconstruction

- Previous work: Mean Squared Error (MSE)
- Problems
  - Few information about the content of the images
  - Small misaligned caused on mapping can affect the results

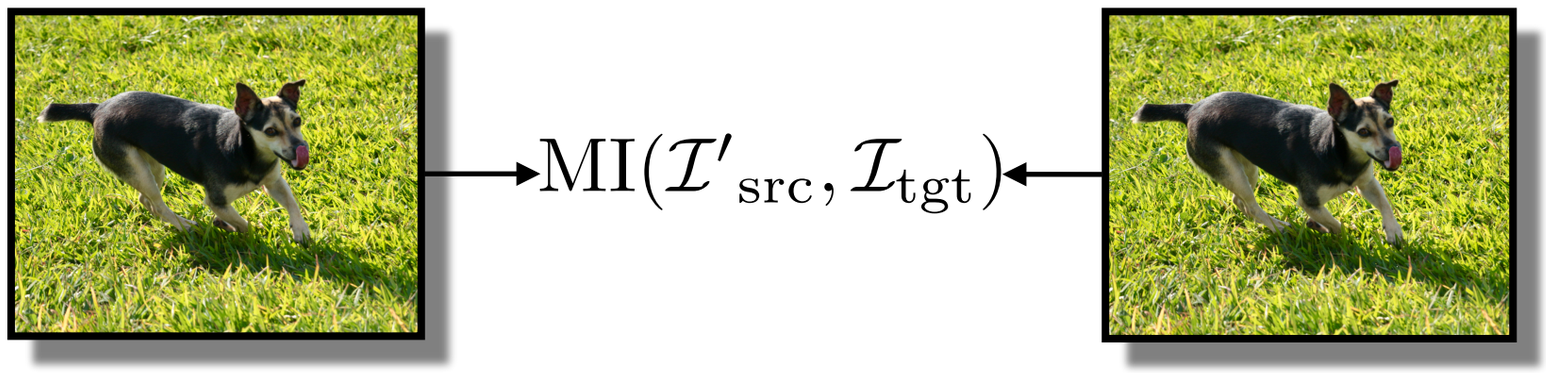
Hypothesis #3

The comparison of the distributions of the pixel values of two near-duplicate images is more effective for the dissimilarity calculation than their point-wise comparison

Gradient point-wise comparison

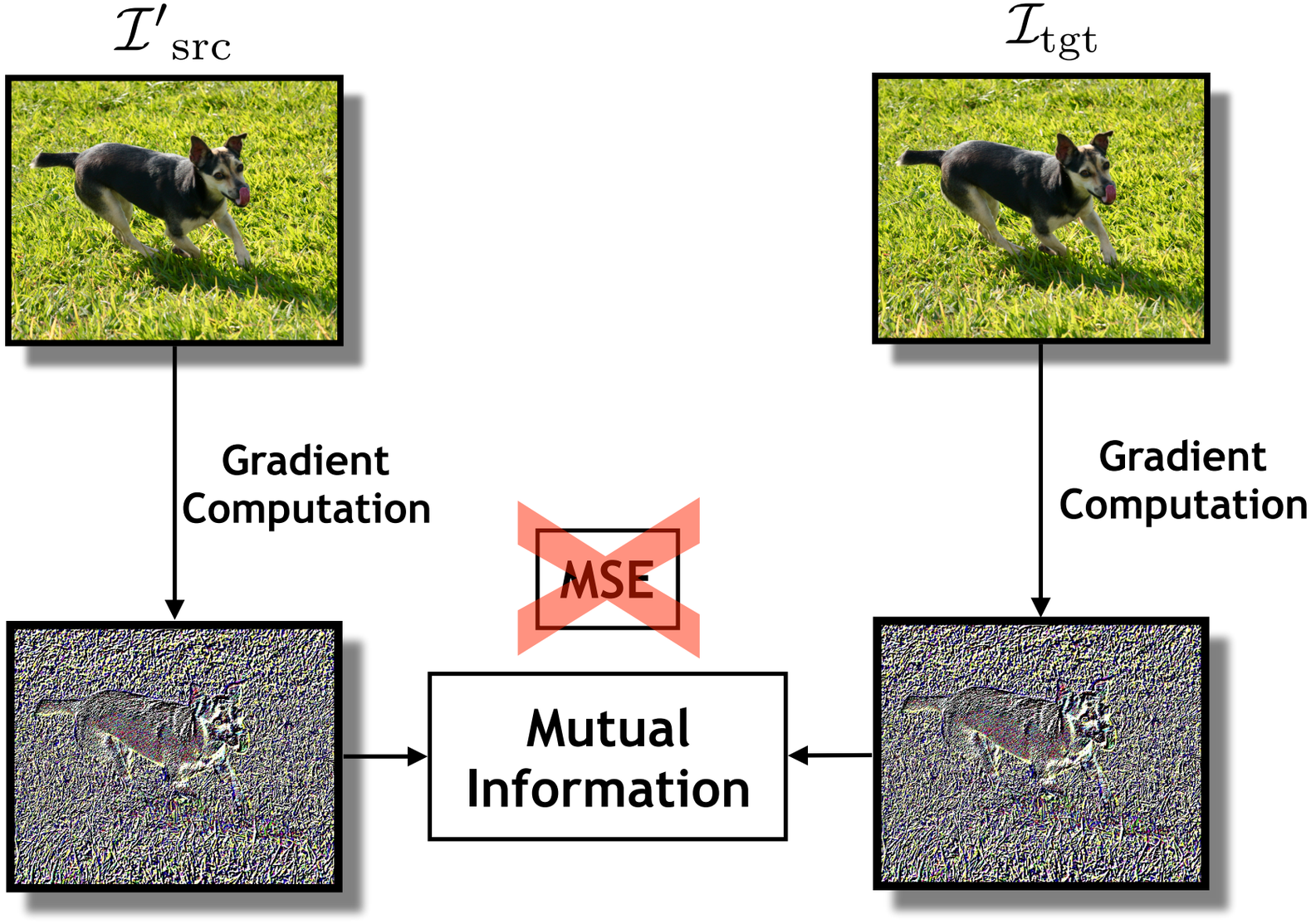


Mutual Information comparison



$$\text{MI}(\mathcal{I}'_{src}, \mathcal{I}_{tgt}) = H(\mathcal{I}_{tgt}) - H(\mathcal{I}_{tgt}|\mathcal{I}'_{src})$$
$$H(X) = -E_x[log(P(X))]$$

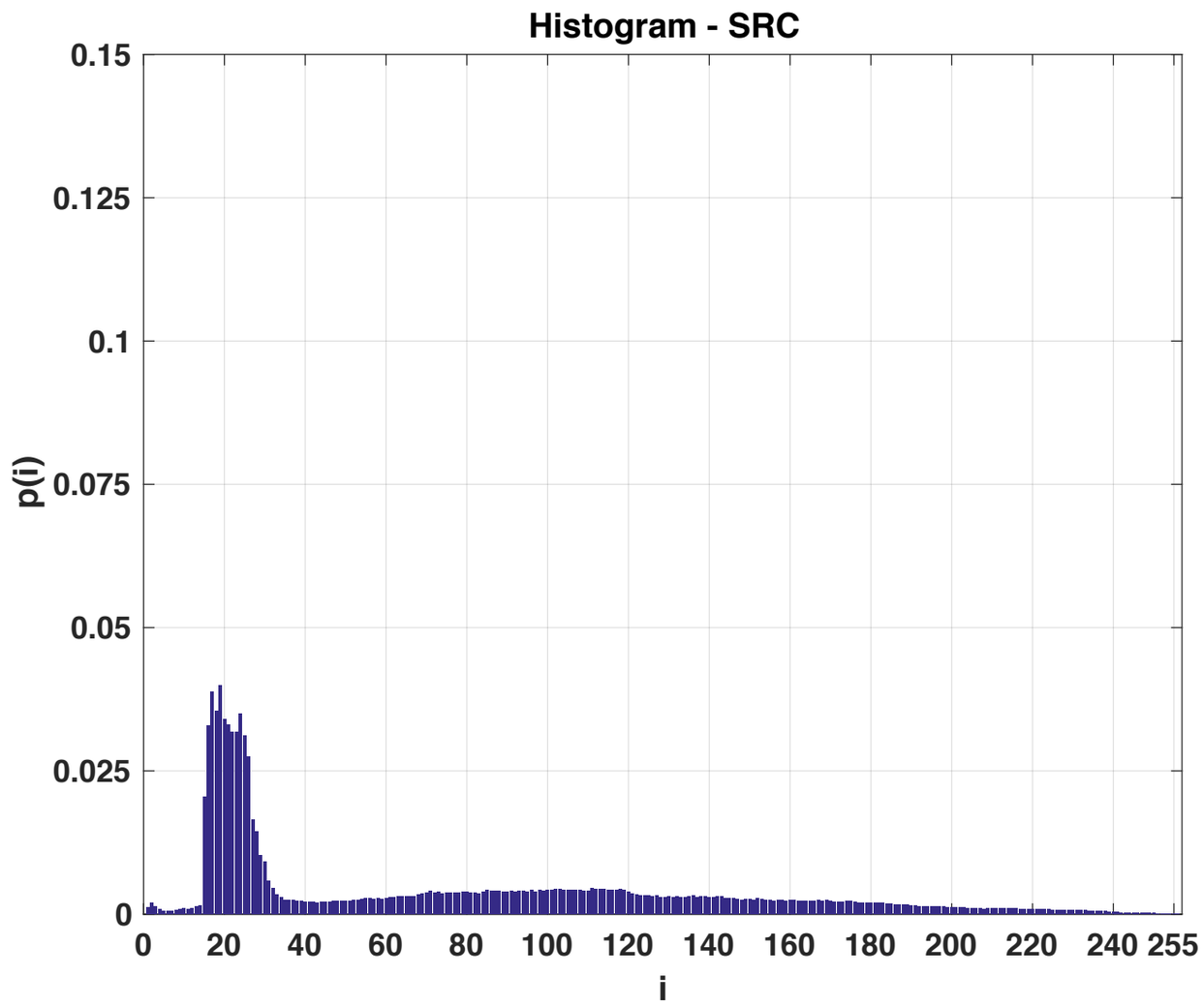
Comparing gradients with mutual information

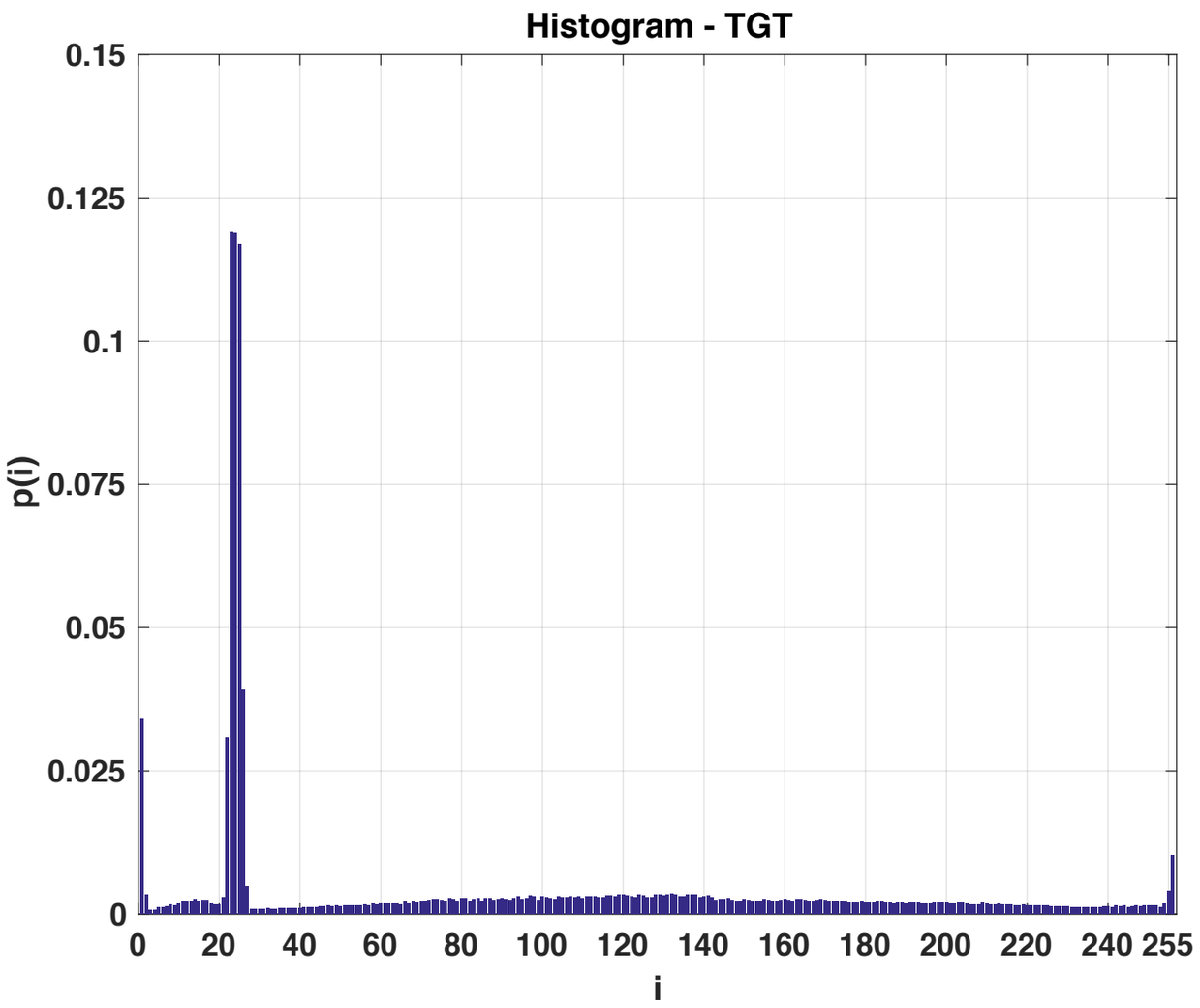


Color matching

- Previous work
  - Normalization based on mean and standard deviation of another image
- Problem
  - Unexpected Artifacts
- New color matching approach
  - Histogram matching [Gonzales (2007)]

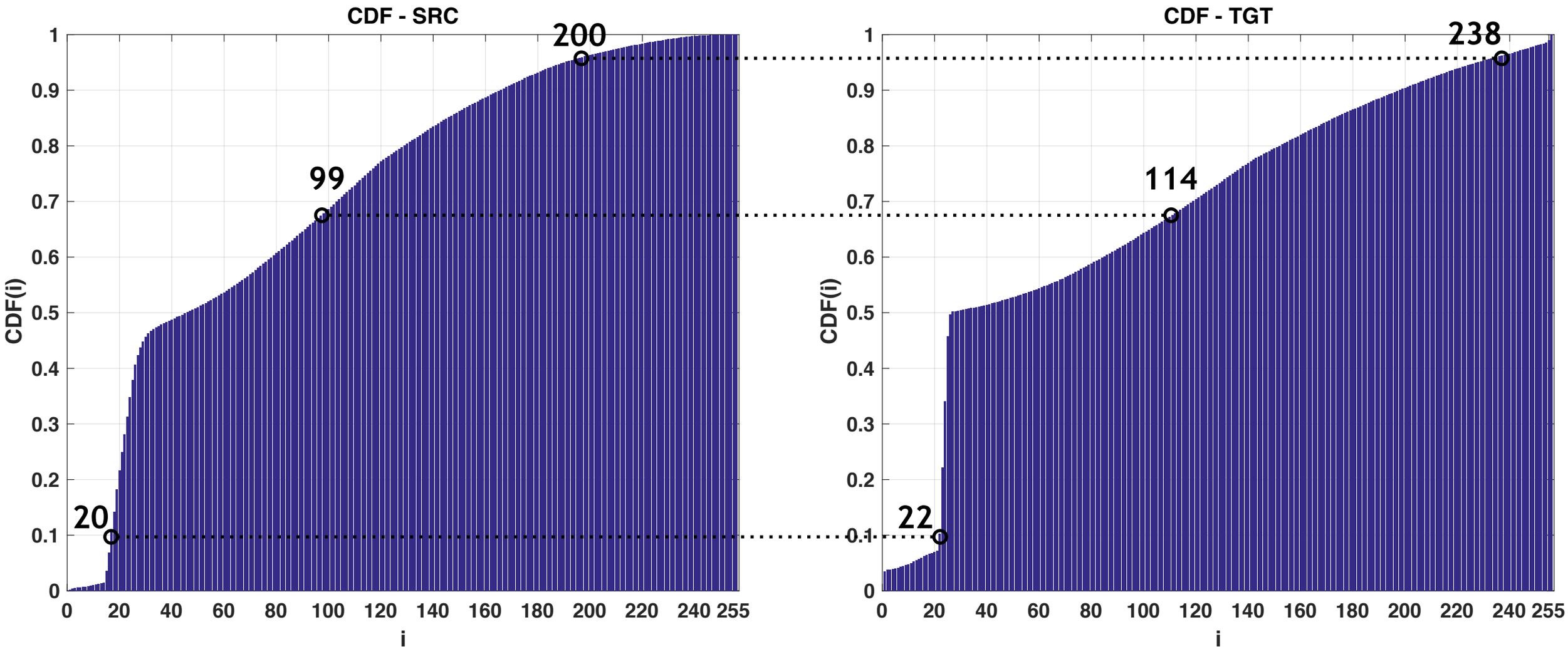
Histogram color matching





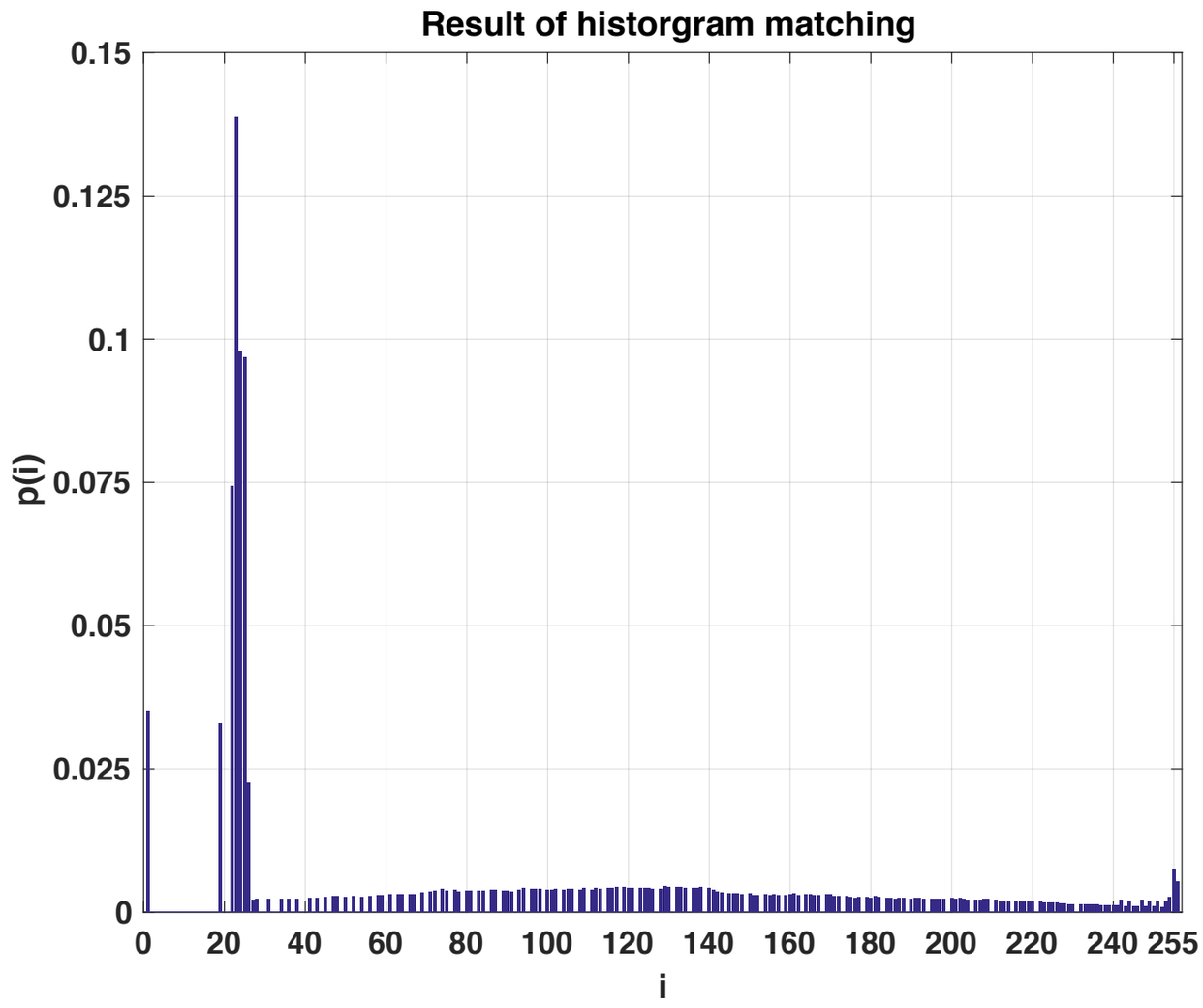
Histogram

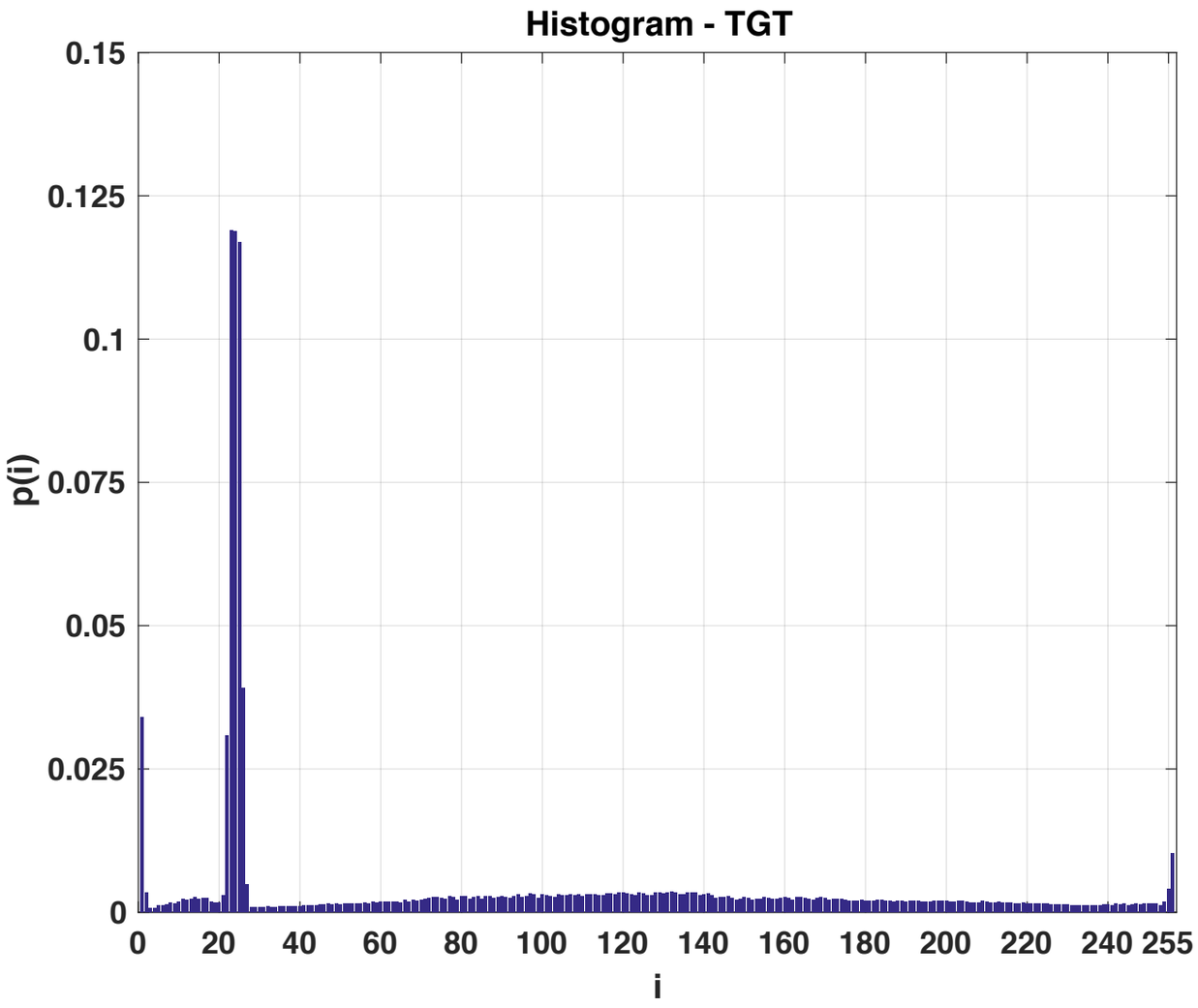
Histogram color matching



Cumulative Distribution Function

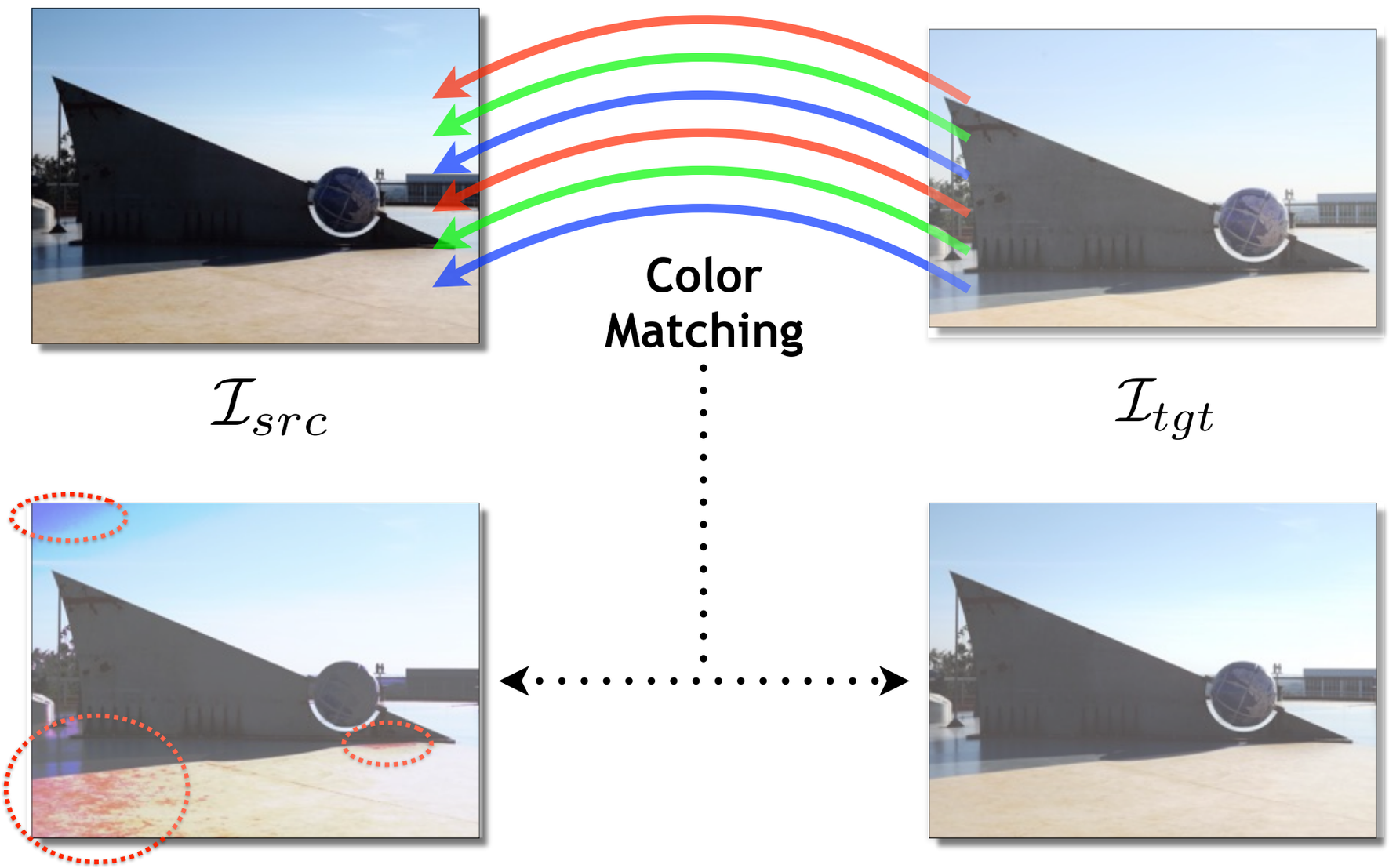
Histogram color matching





Resultant histogram

Color matching approaches

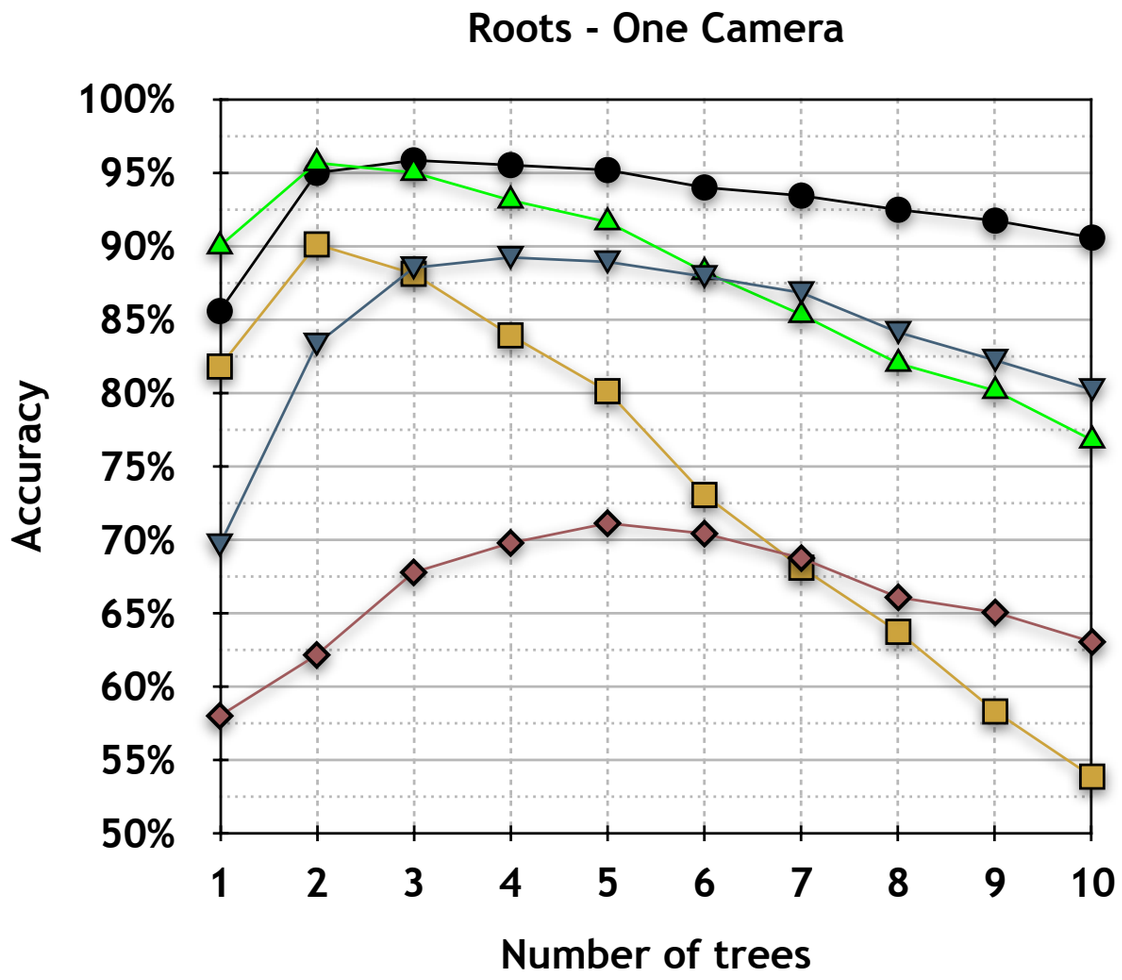


Experiments

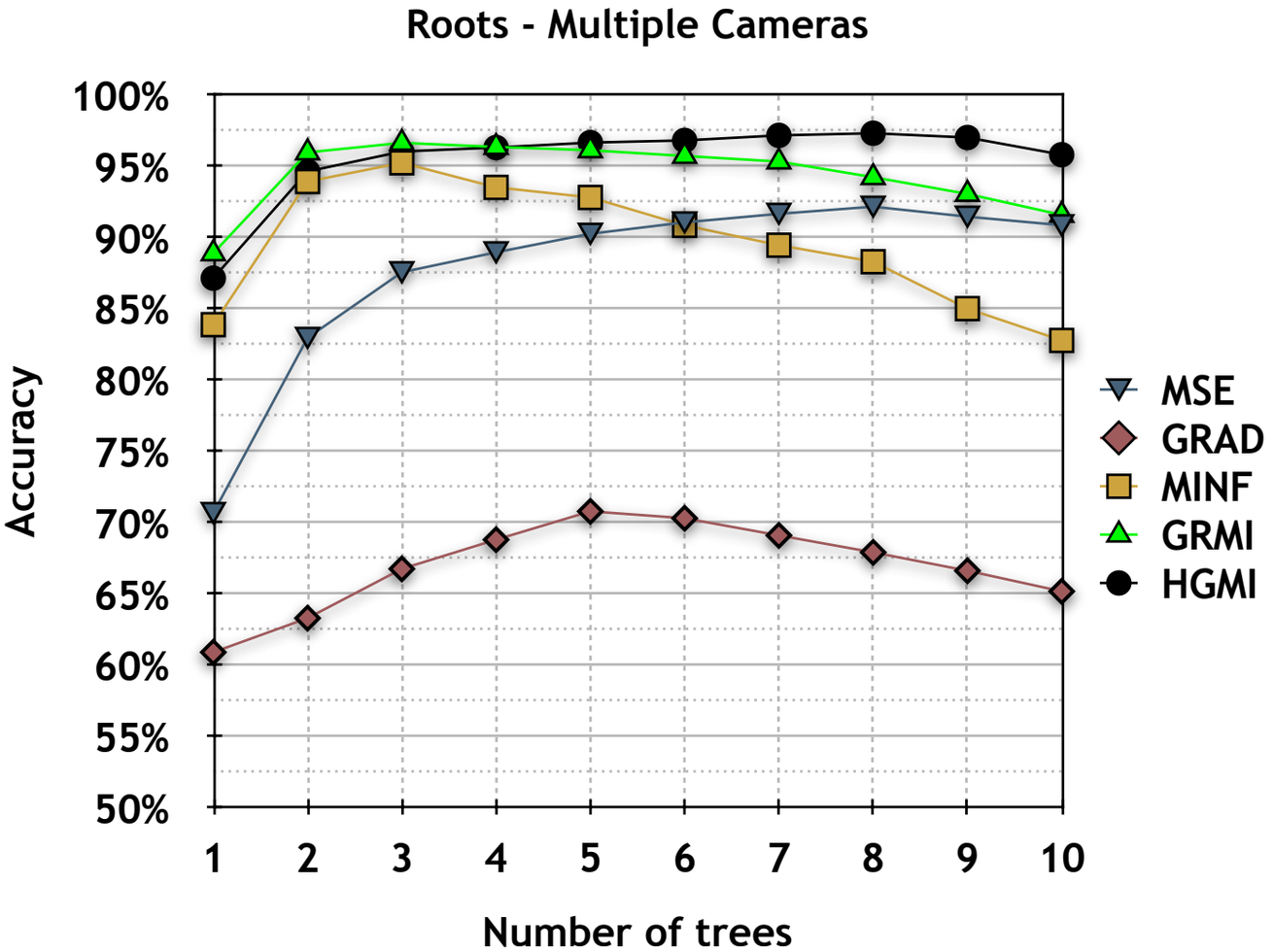
Experimental setup

- Reconstruction
  - Extended AOB
- Evaluation
  - Roots, edges, leaves and ancestry

Results

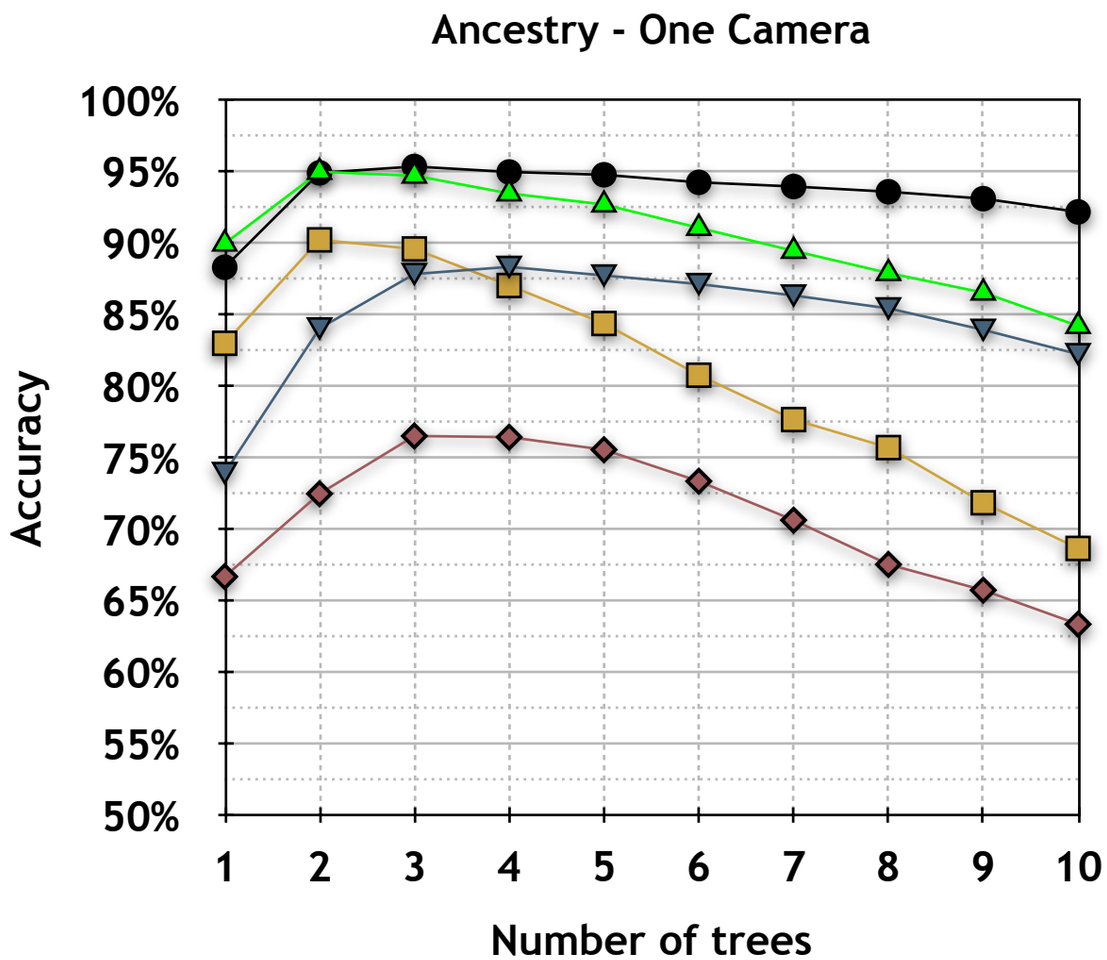


Error Reduction (HGMI vs. MSE): 53%

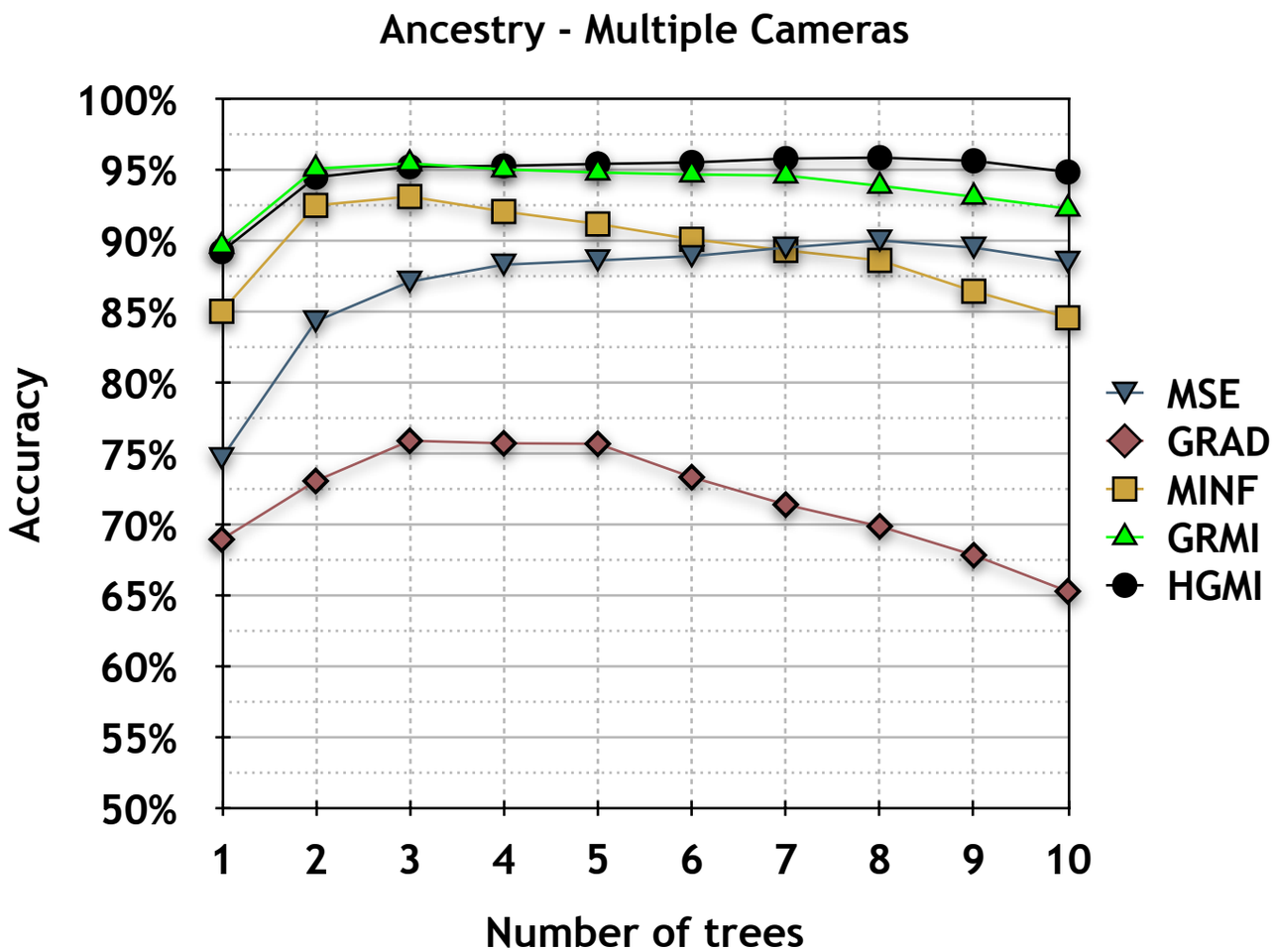


Error Reduction (HGMI vs. MSE): 47%

Results



Error Reduction (HGMI vs. MSE): 60%



Error Reduction (HGMI vs. MSE): 58%

Qualitative results
Ellen DeGeneres' Selfie

Group a

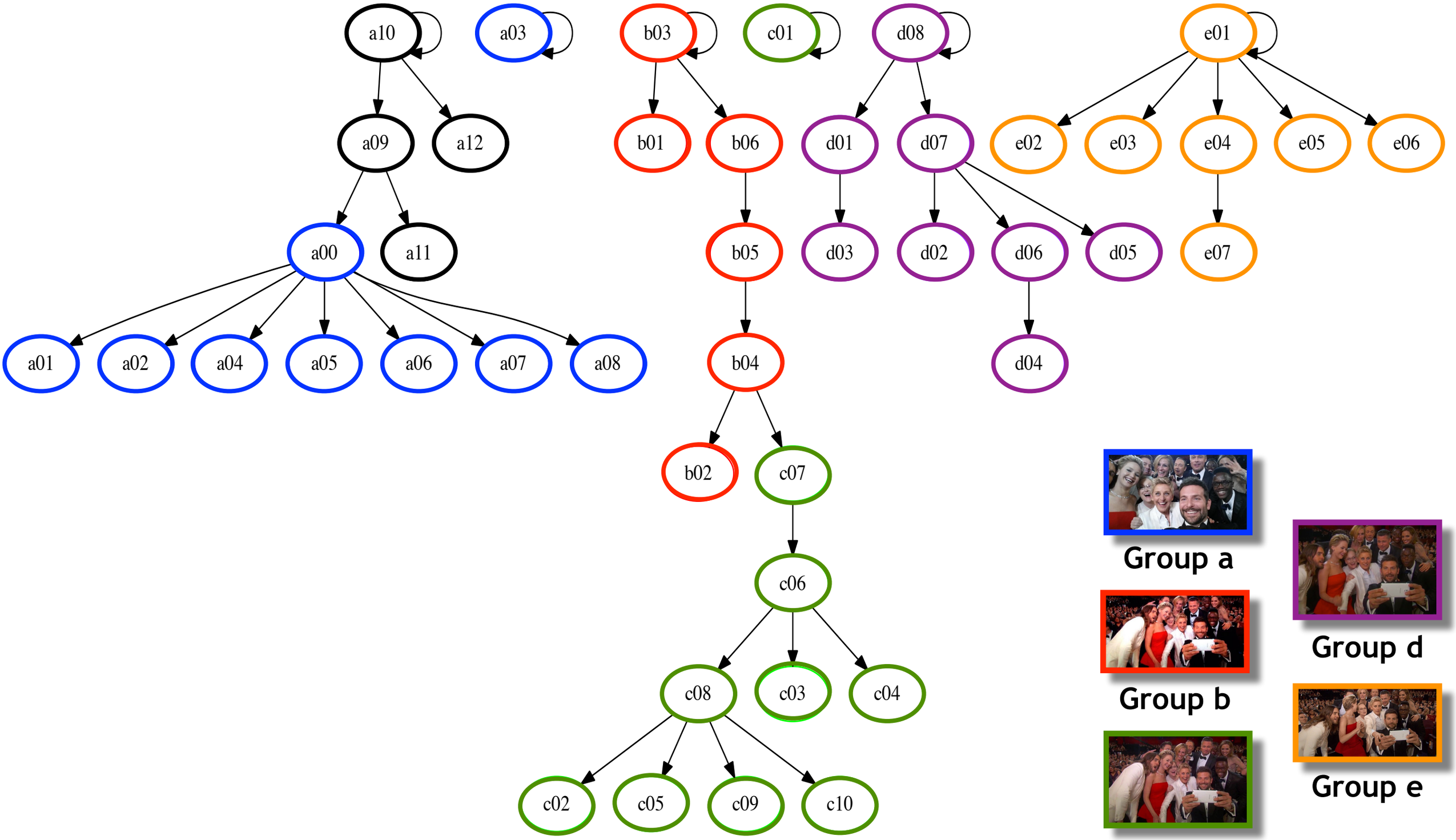
Group b

Group c

Group d

Group e

Qualitative results
Ellen DeGeneres' Selfie



Group a

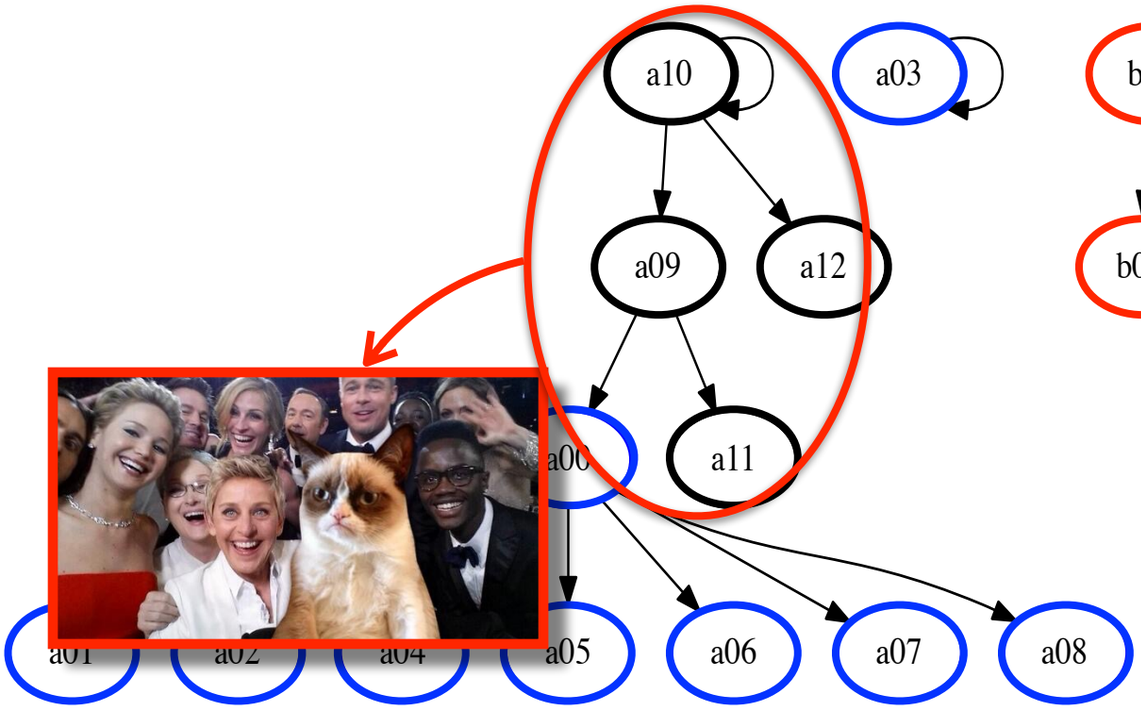
Group d

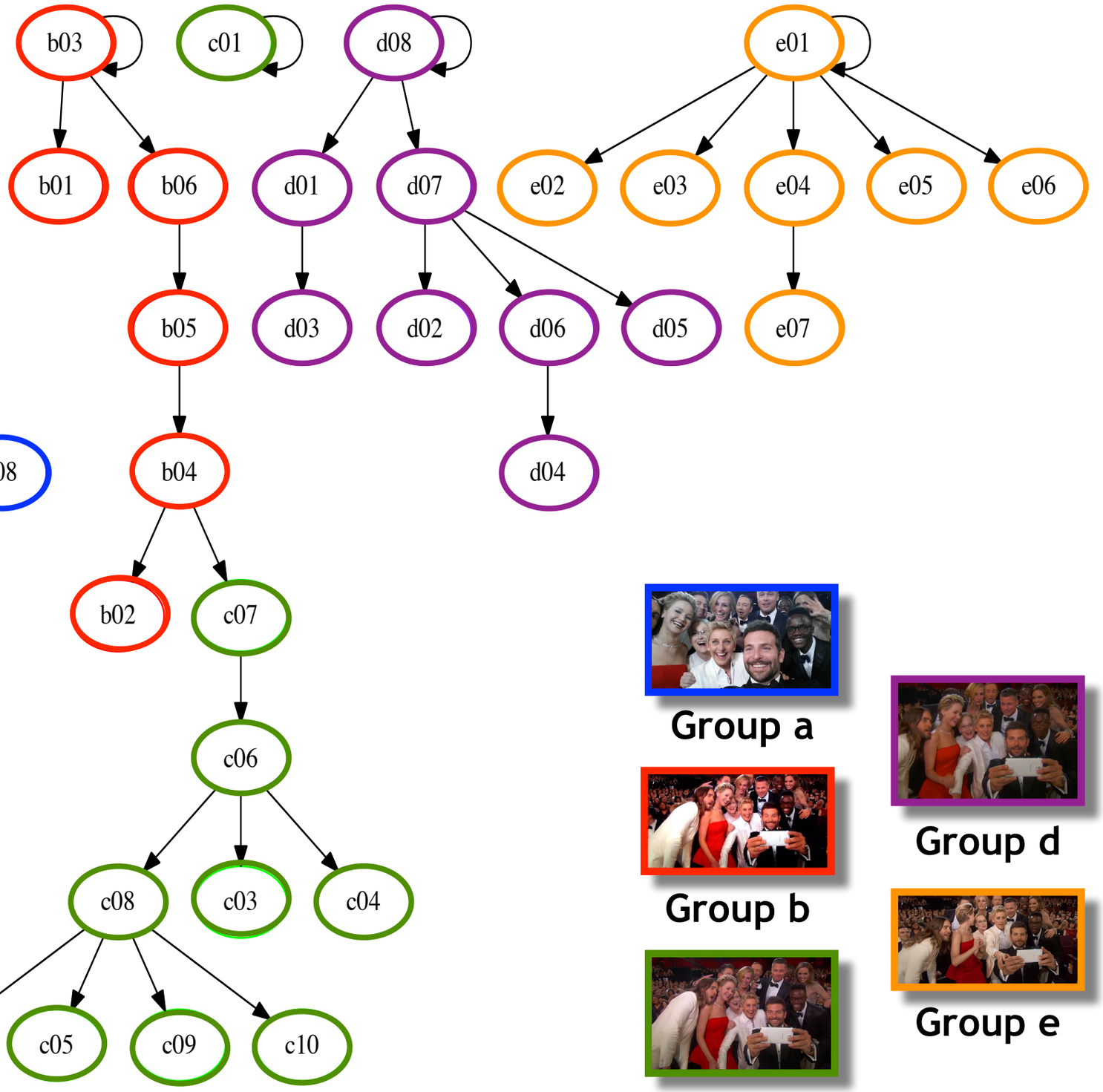
Group b

Group e

Group c

Qualitative results
Ellen DeGeneres' Selfie





Group a

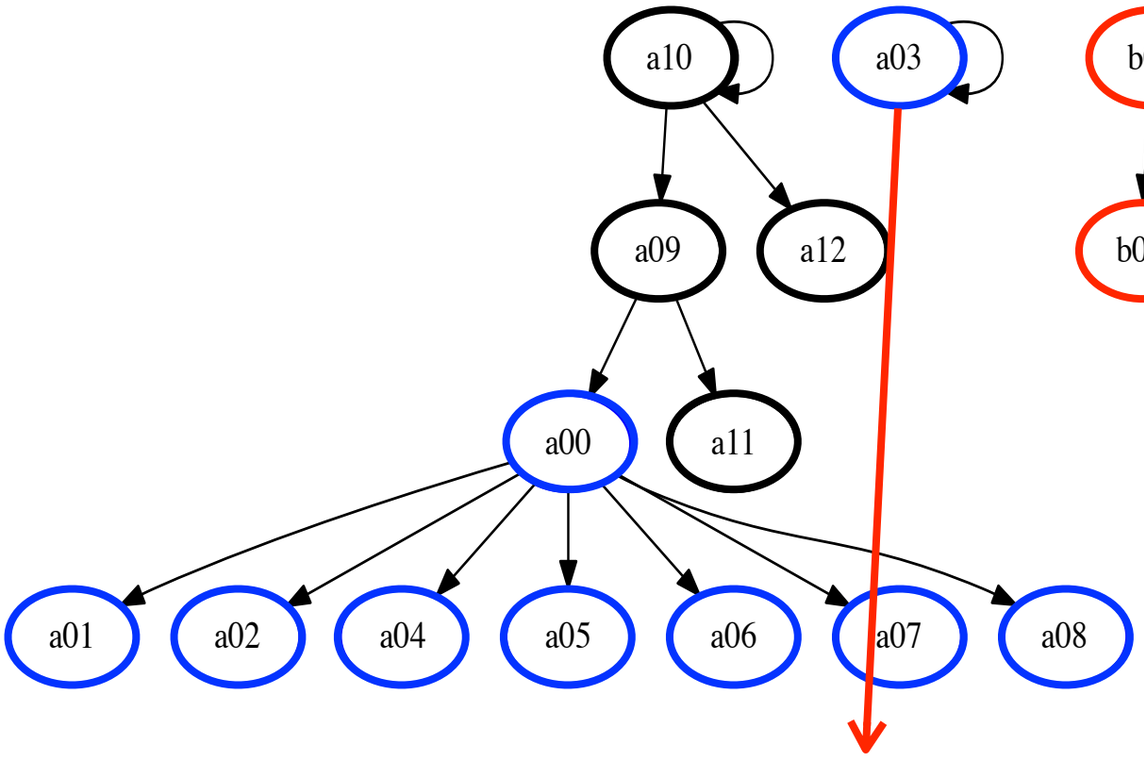
Group d

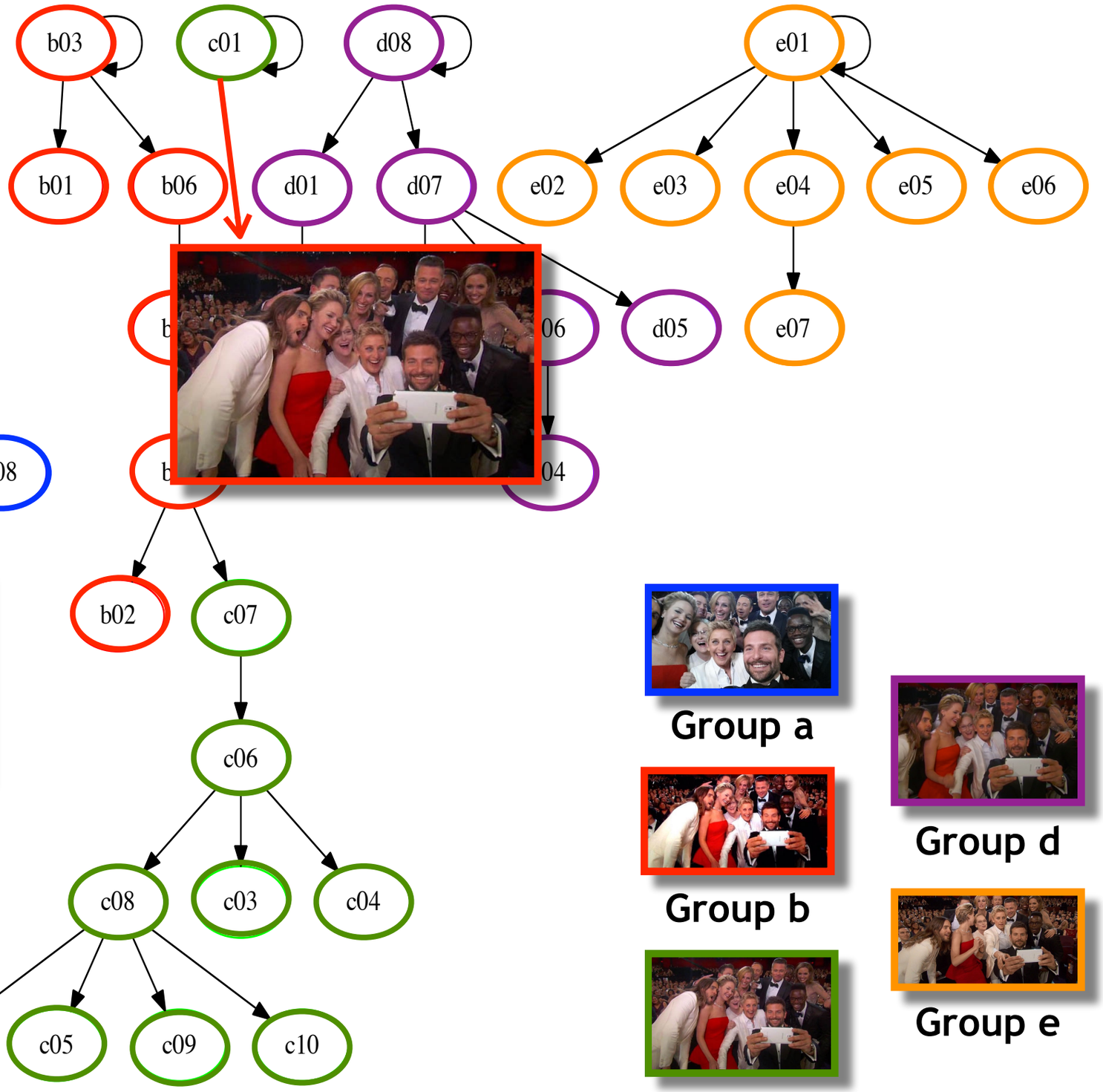
Group b

Group e

Group c

Qualitative results
Ellen DeGeneres' Selfie





Group a

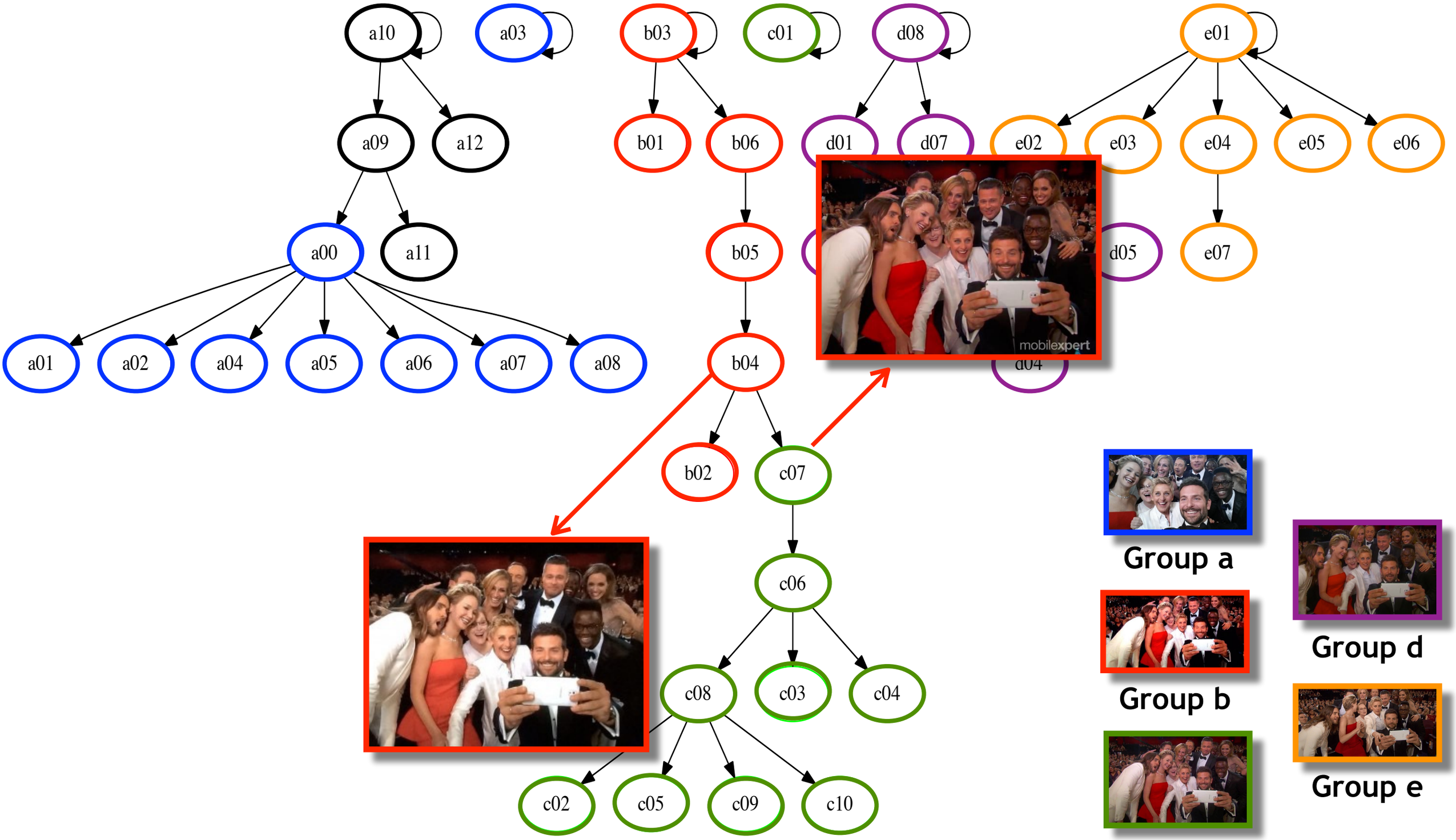
Group d

Group b

Group e

Group c

Qualitative results
Ellen DeGeneres' Selfie



Group a

Group d

Group b

Group e

Group c

Video Phylogeny

Video Phylogeny

- More challenging problem than Image Phylogeny
- Now we have a new dimension: time
- Videos can be misaligned
- Videos can have different compression parameters

Hypothesis #4

Temporal alignment is paramount for a proper video phylogeny reconstruction process

Video Phylogeny Contributions

- Phylogeny reconstruction for misaligned and compressed video sequences

F. O. Costa, S. Lameri, P. Bestagini, Z. Dias, A. Rocha, M. Tagliassachi, S. Tubaro. *Phylogeny reconstruction for misaligned and compressed video sequences*. IEEE International Conference on Image Processing (ICIP), 2015

- Hashing-based Temporal Alignment for Video Phylogeny Reconstruction

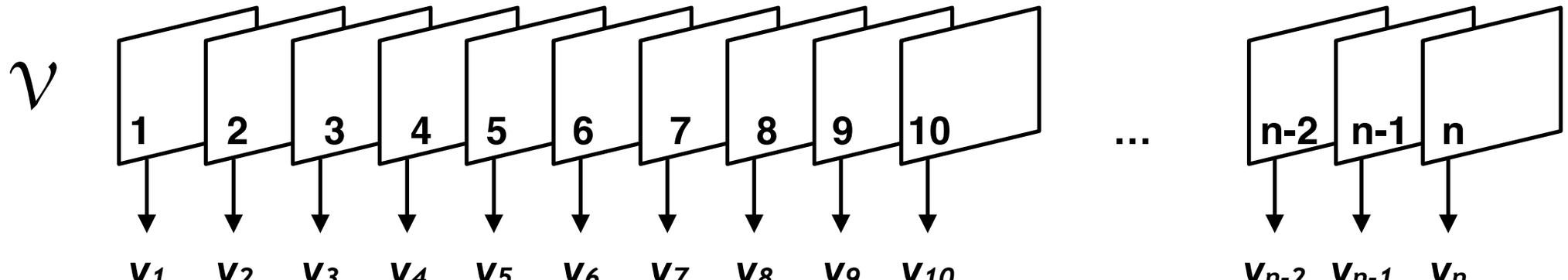
F. O. Costa, P. Bestagini, S. Tubaro, Z. Dias, A. Rocha. *Hashing-based Temporal Alignment for Video Phylogeny Reconstruction*. Submitted to IEEE WIFS, 2016

Phylogeny reconstruction for misaligned and compressed video sequences

Video dissimilarity calculation

- Temporal alignment
- Frame registration
- Color matching
- Coding matching (controlled)
- Comparison (MSE)

Temporal alignment



y_i = **Average of luminance of $\mathcal{V}$(i)**

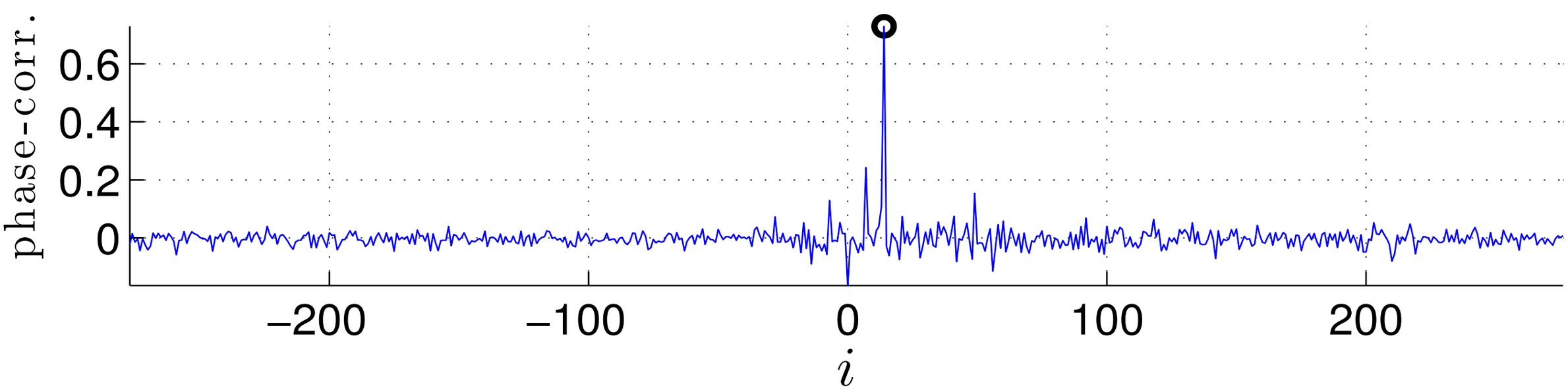
Difference of luminance

$$LD_{\mathcal{V}} = [y_2 - y_1, y_3 - y_2, \ldots, y_{n-1} - y_{n-2}, y_n - y_{n-1}]$$

Video Alignment

Phase-correlation:

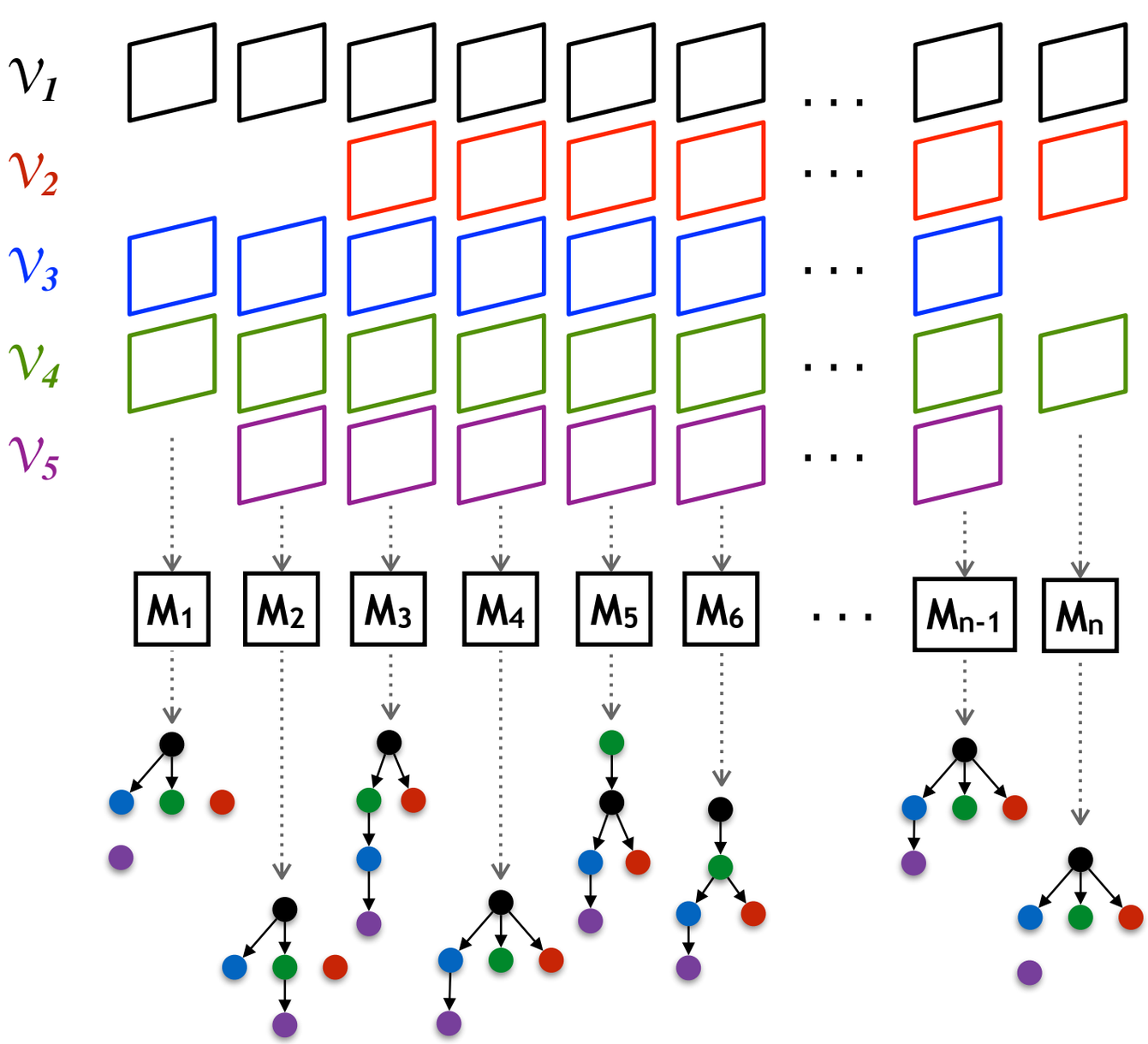
$$\hat{i} = \arg\max_i \mathcal{F}^{-1}\left[\frac{\mathcal{F}[LD\mathcal{V}_{\text{src}}]\cdot\mathcal{F}[LD\mathcal{V}_{\text{tgt}}]^*}{|\mathcal{F}[LD\mathcal{V}_{\text{src}}]\cdot\mathcal{F}[LD\mathcal{V}_{\text{tgt}}]^*|}\right](i)$$

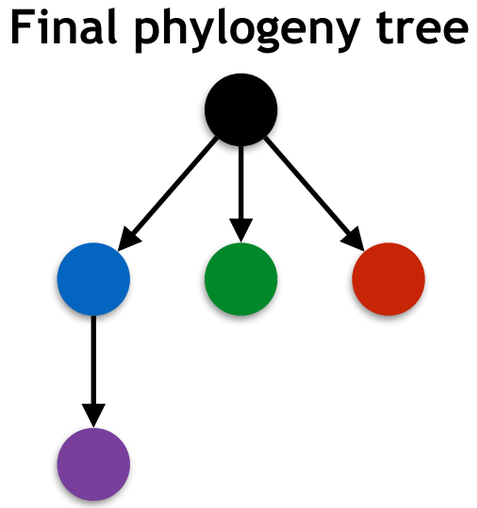


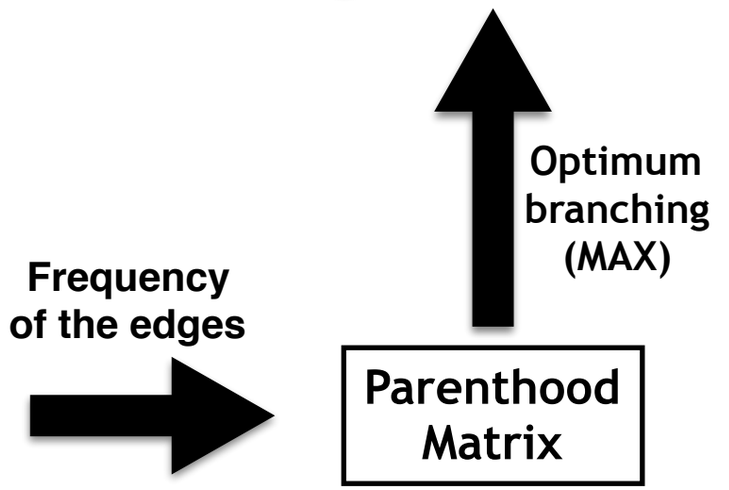
Comparison

- We perform two kinds of comparison
  - Average of MSE for the correspondent frames of two videos
  - Parenthood matrix

Parenthood matrix







Experiments

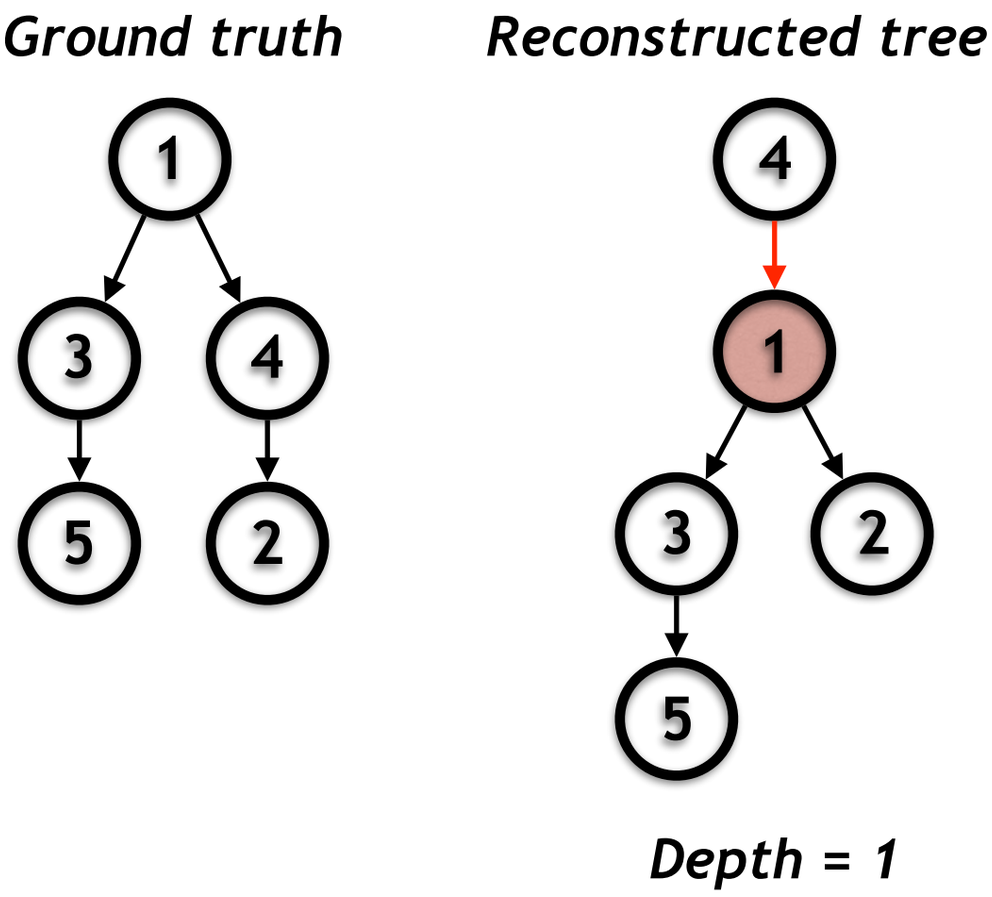
Dataset

- 200 trees with 10 videos
  - 100 - No clip
  - 100 - Clip at the Beginning/end of the stream
- Probability of clip: 50%
- CODEC: MPEG2, MPEG4, H264
- Crop, resize, bright, contrast, coding and temporal clipping

Available at https://media.xiph.org/video/derf/

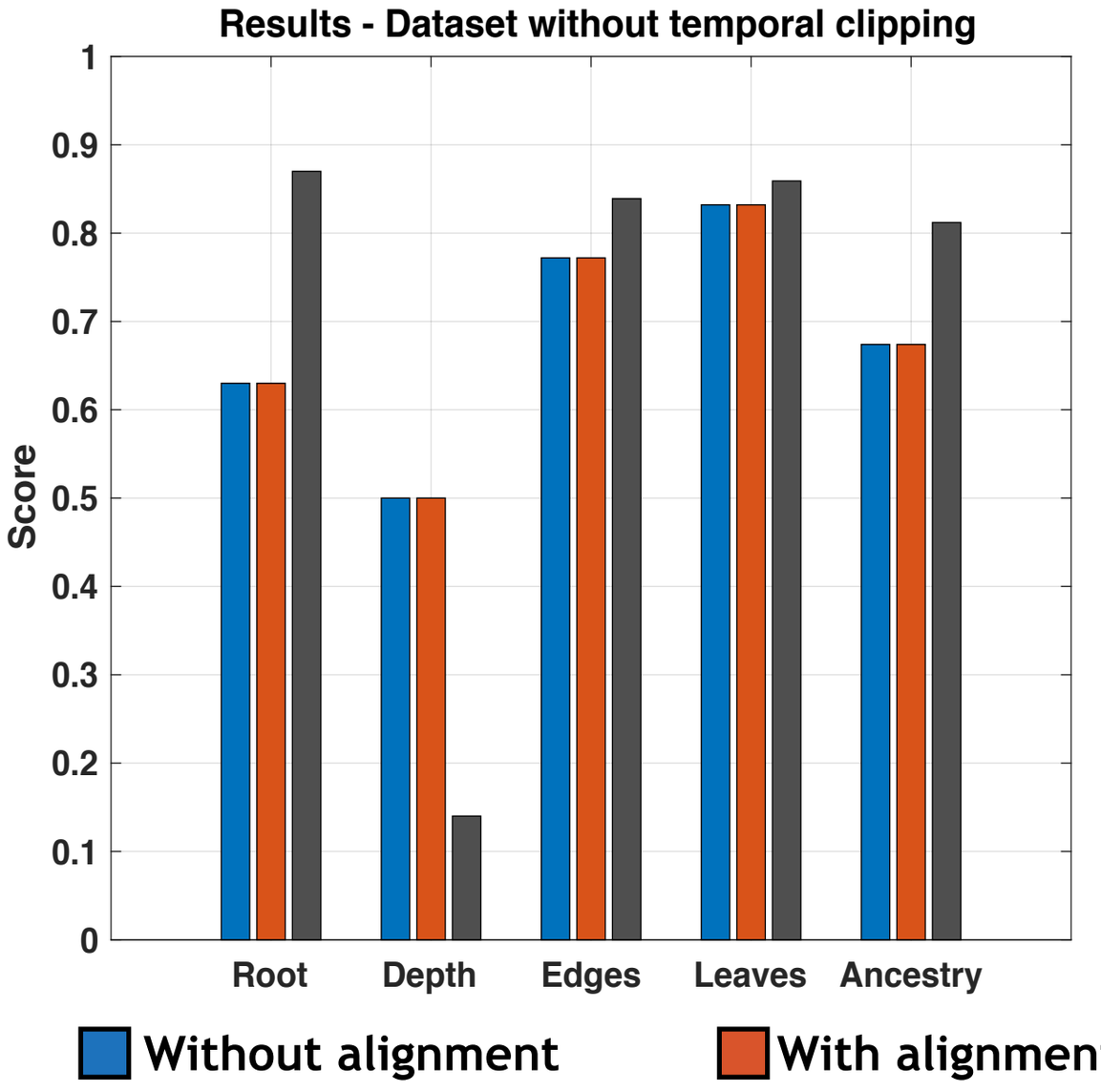
Experimental Setup

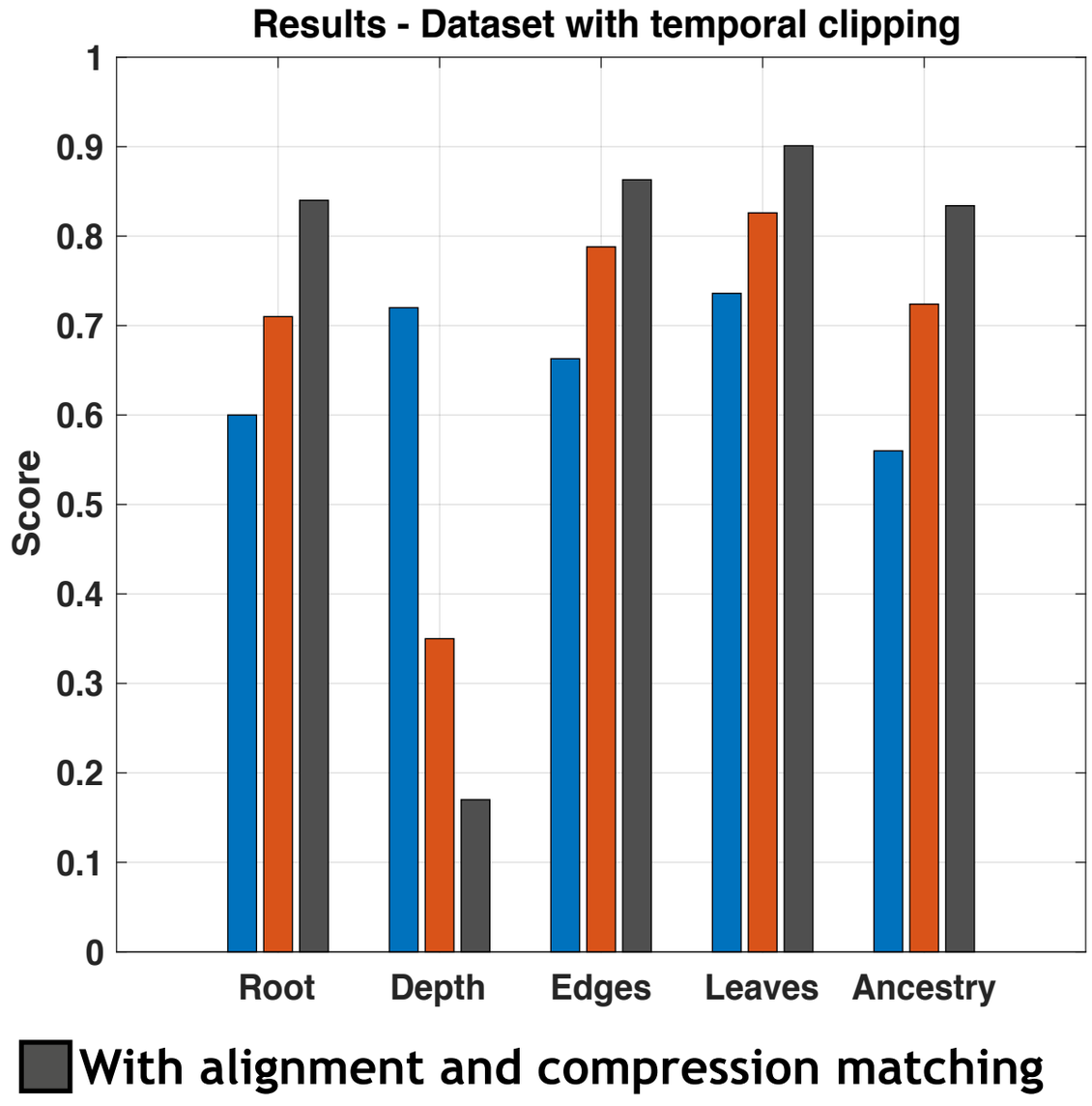
- Phylogeny Reconstruction
  - OB algorithm
- Evaluation metrics
  - Root, edges, leaves, ancestry and ***depth***



Results #1

Verifying alignment and coding matching





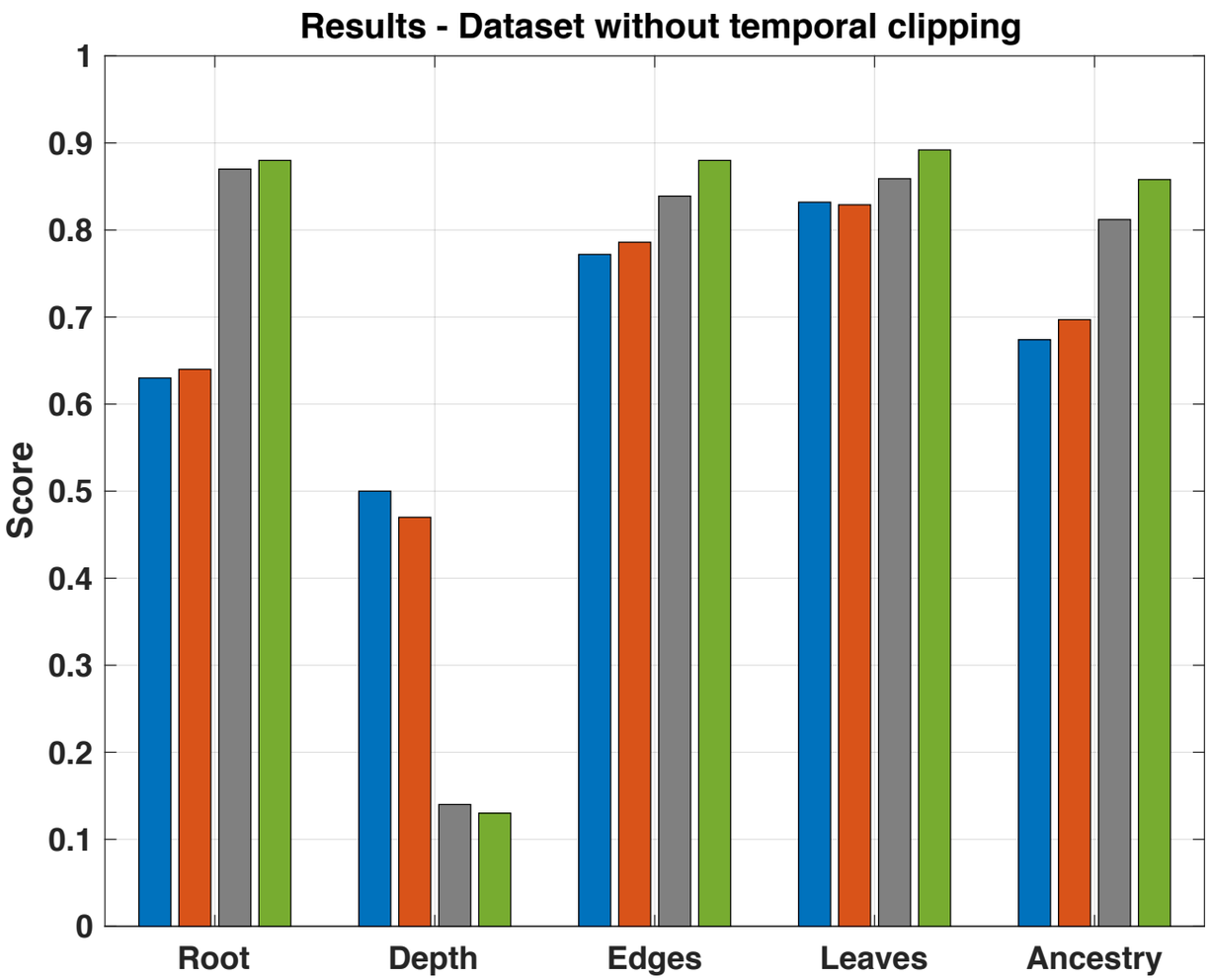
Without alignment

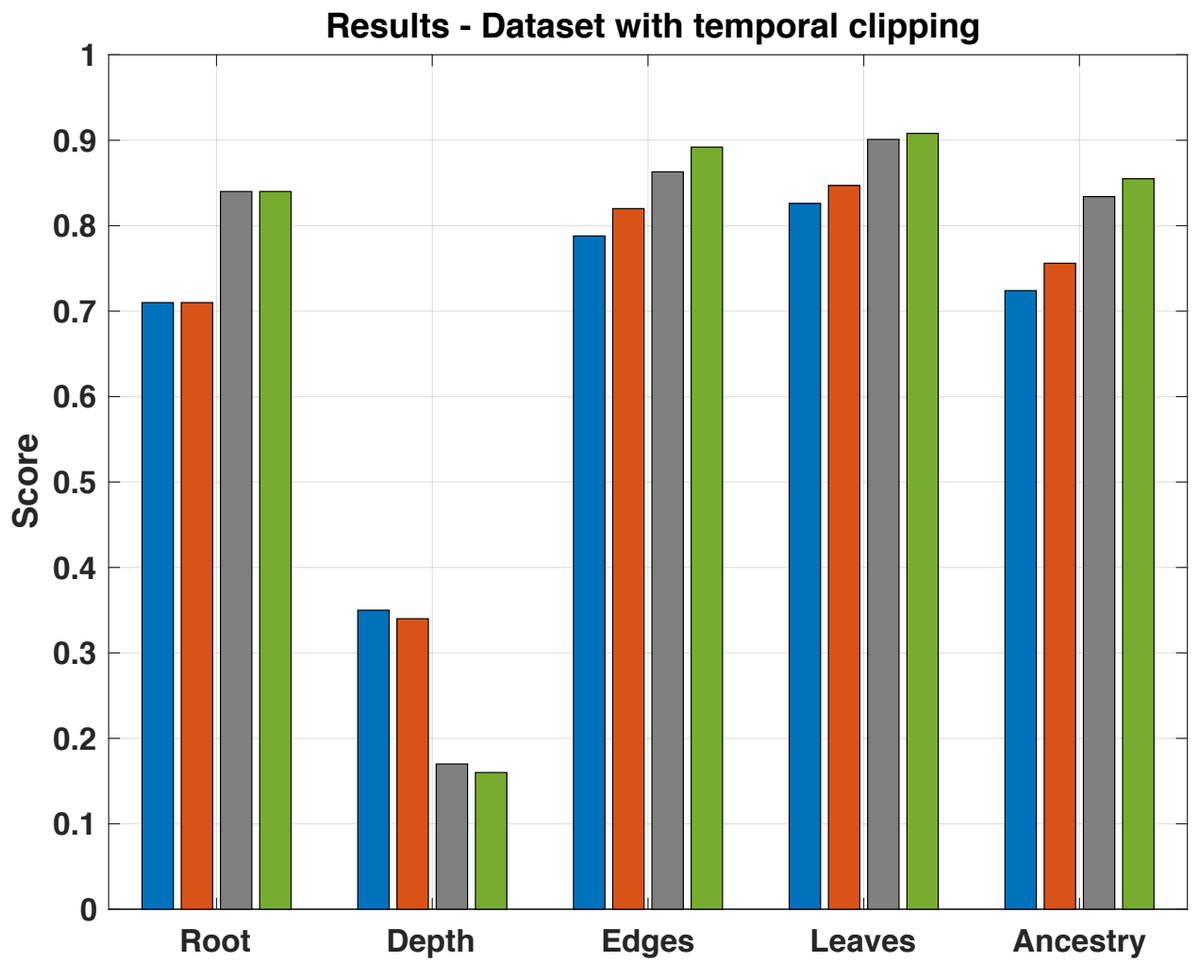
With alignment

With alignment and compression matching

Results #2

Comparison: Average of MSE (A) x Parenthood matrix (P)





A - NO COMPRESSION MATCHING

P - NO COMPRESSION MATCHING

A - WITH COMPRESSION MATCHING

P - WITH COMPRESSION MATCHING

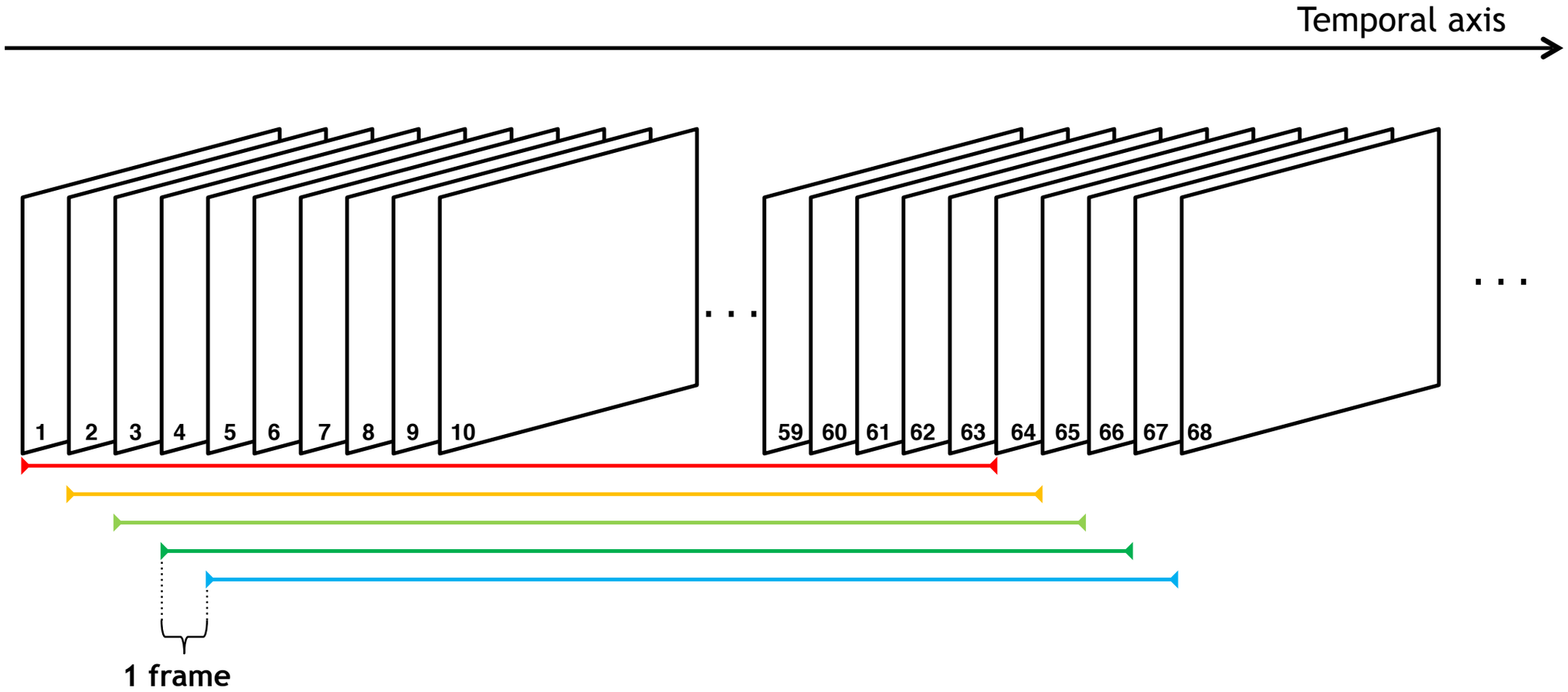
Hashing-based temporal alignment for video phylogeny

Hashing-based temporal alignment

- Near-duplicate matching
- Near-duplicate extraction
- Near-duplicate alignment

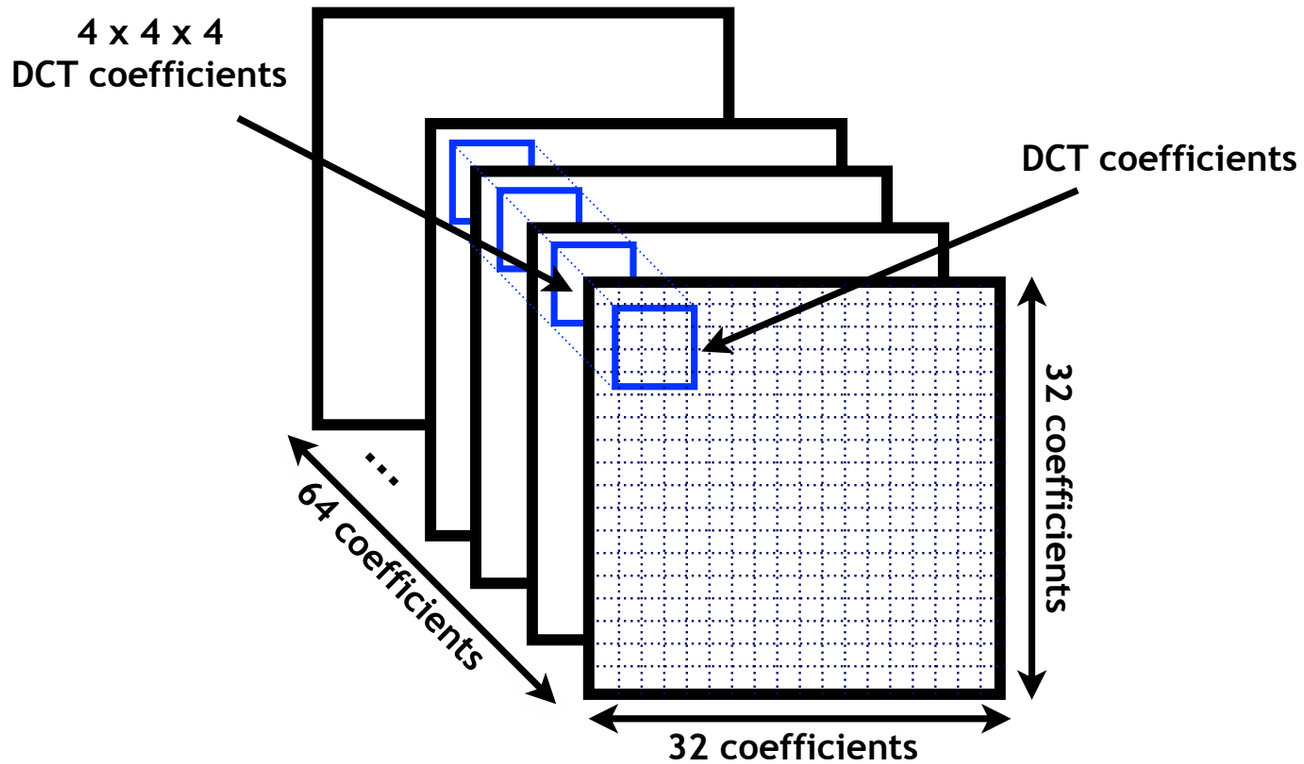
Step 1
Near-duplicate matching

- Groups of 64 frames
- All frames are resized to 32 x 32 pixel



Step 1
Near-duplicate matching

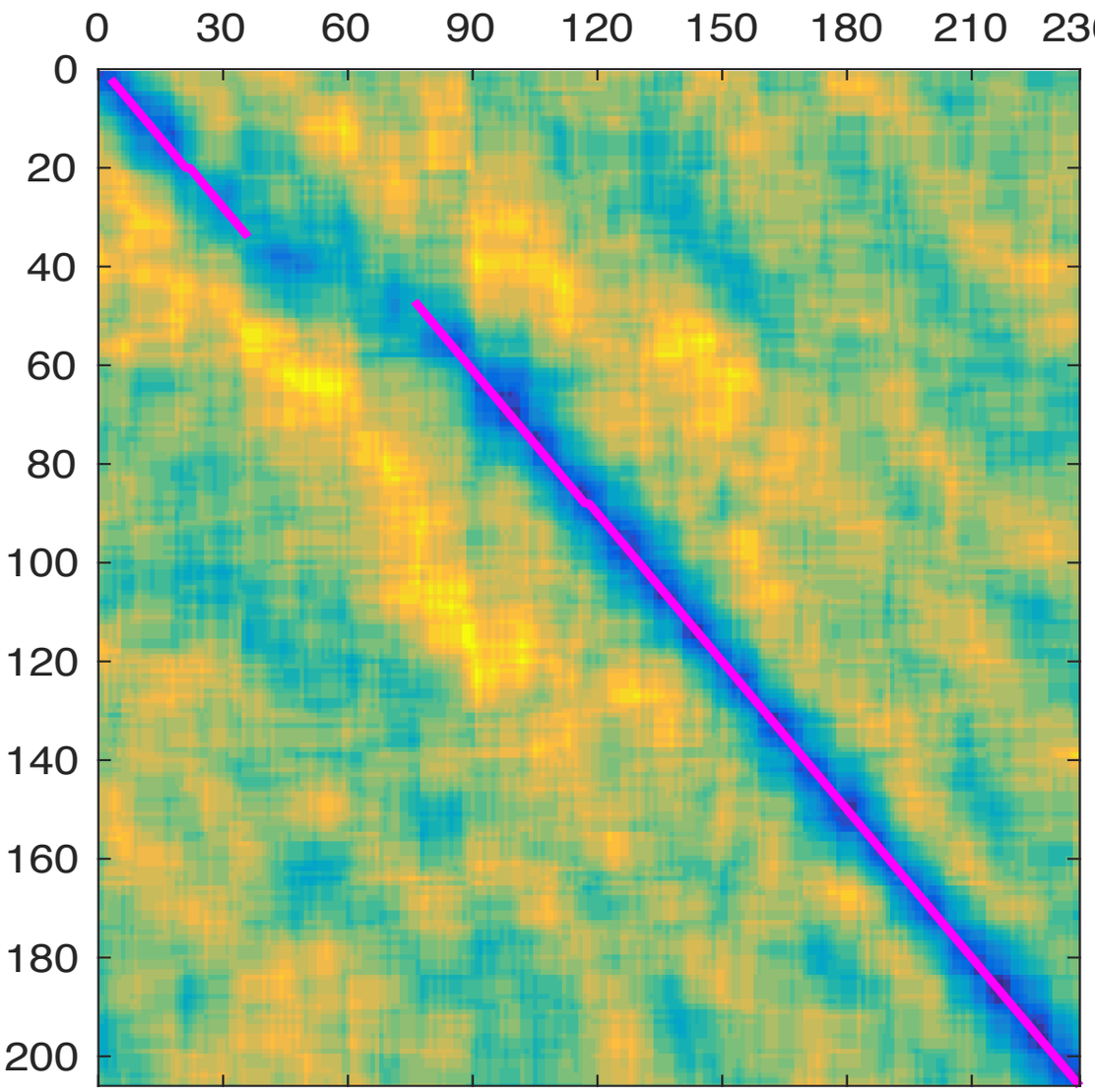
3D Discrete Cosine Transform



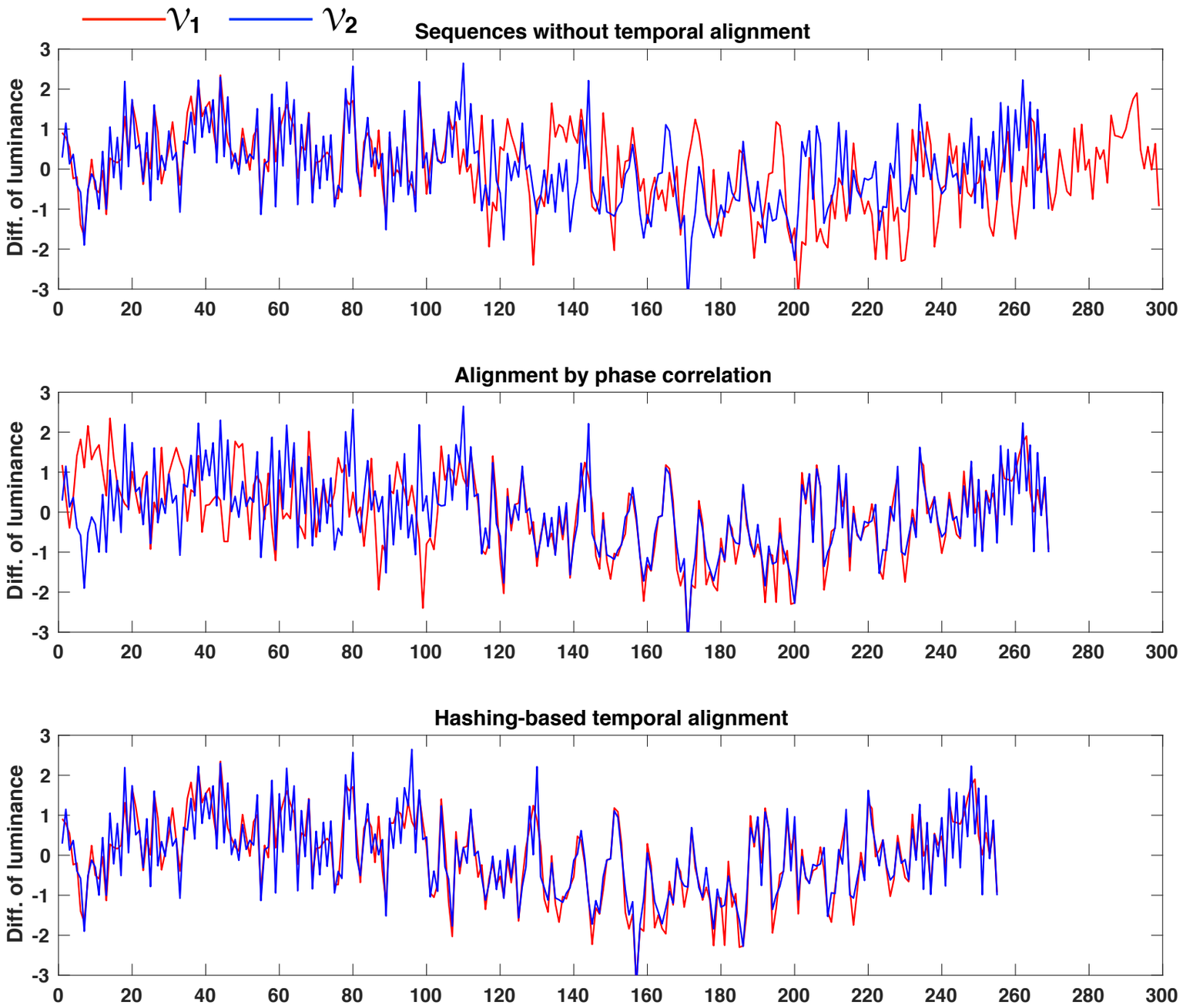
Step 1
Near-duplicate matching

- Binary hash
  - According to the median of DCT
- Distance matrix
  - Hamming distance of two binary hashes of different videos

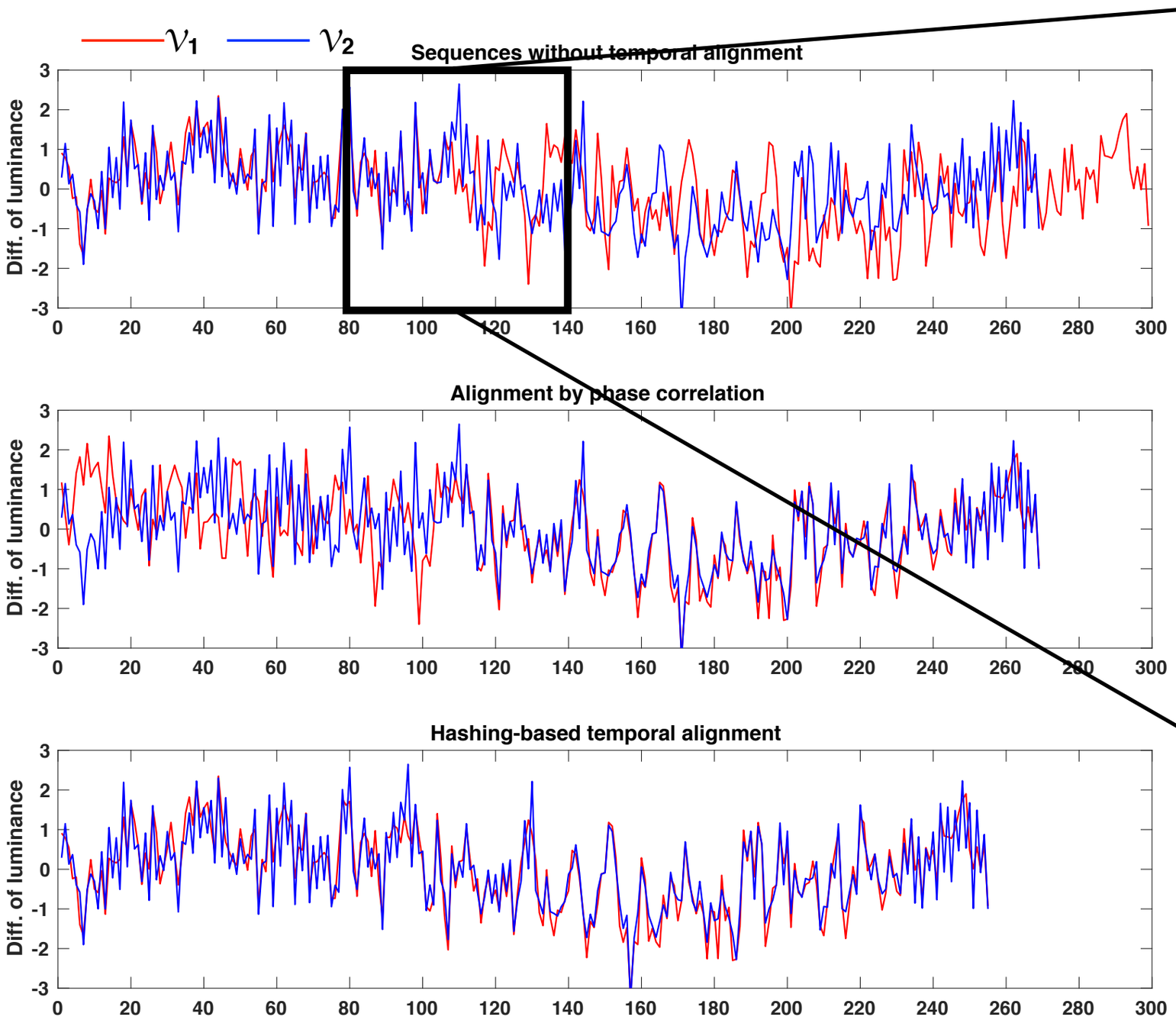
Step 2
Near-duplicate extraction

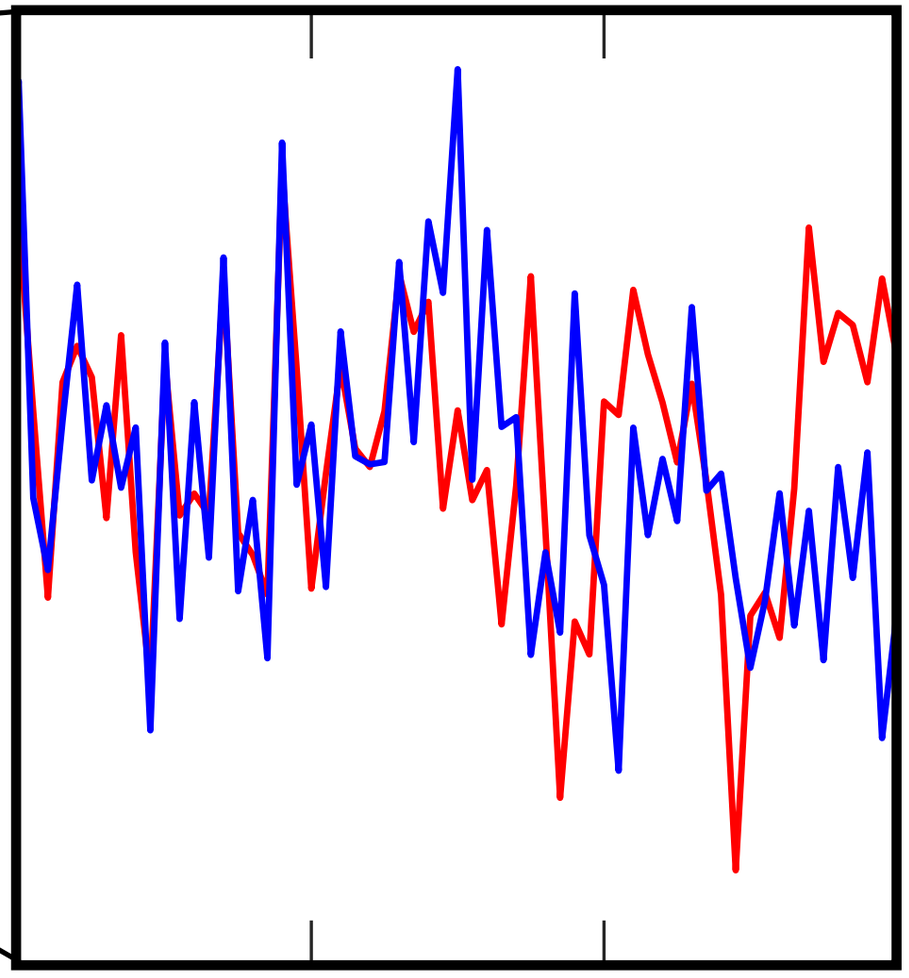


Step 3
Near-duplicate alignment

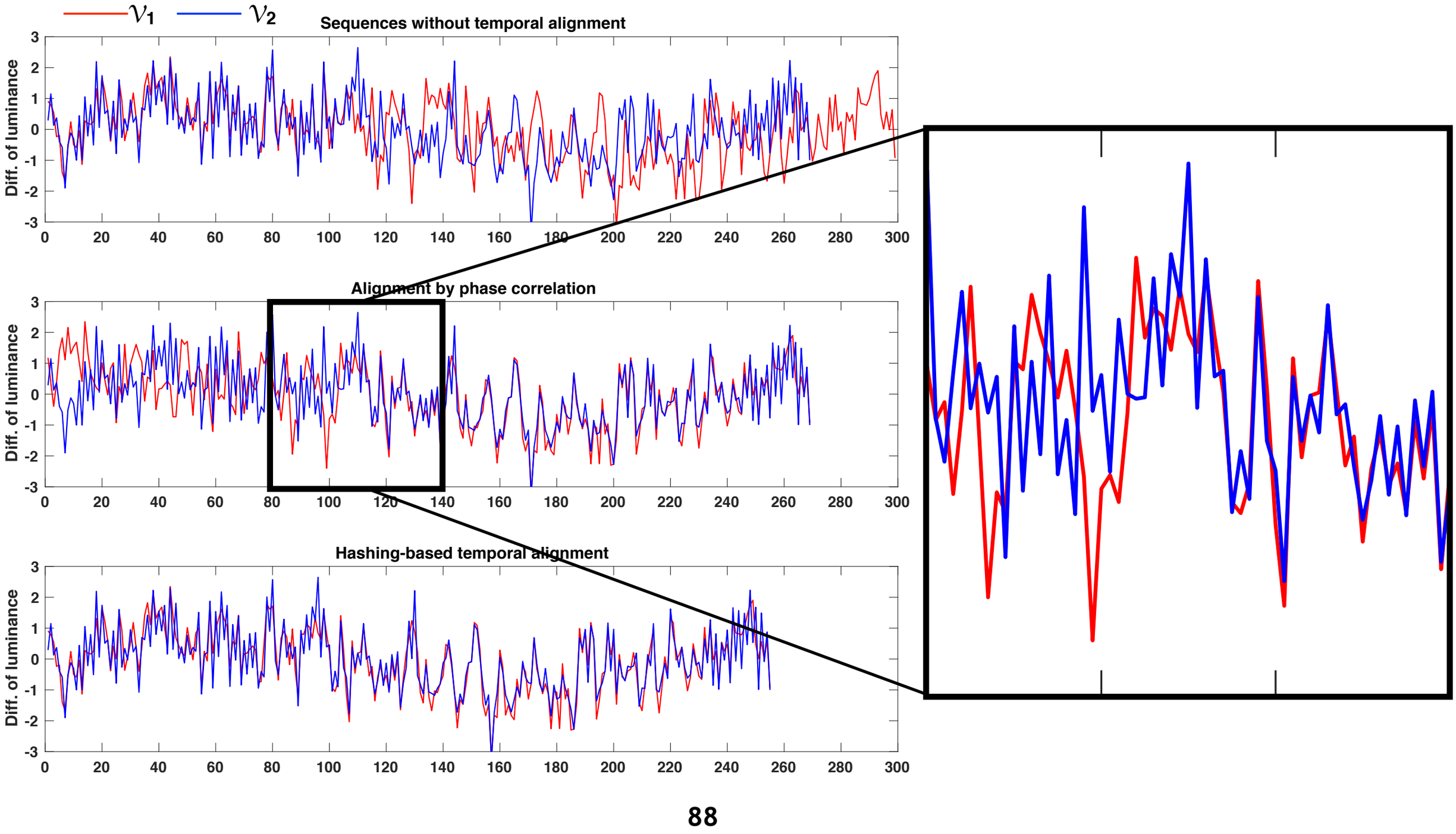


Step 3
Near-duplicate alignment

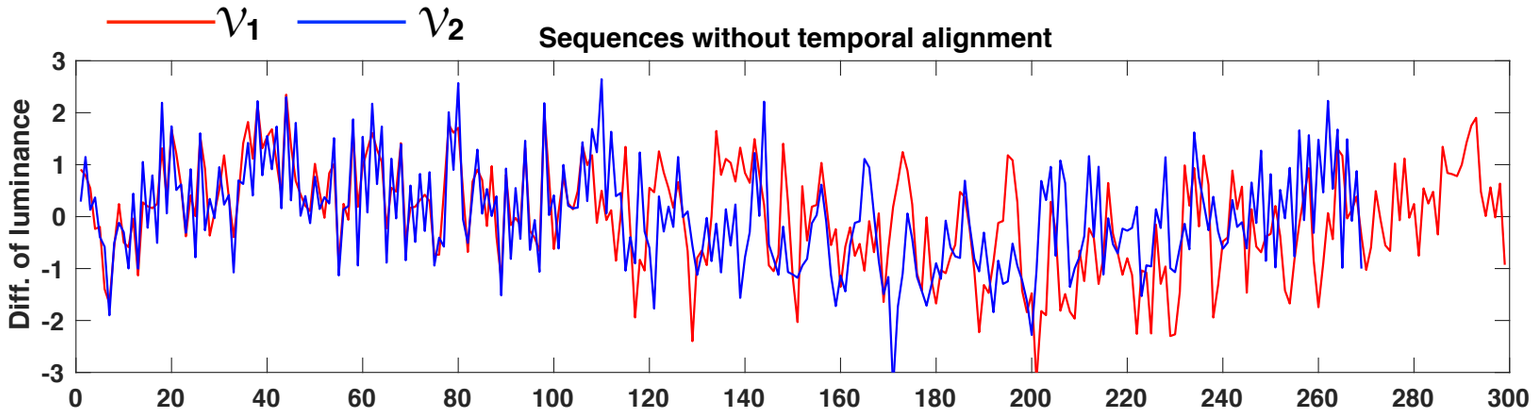


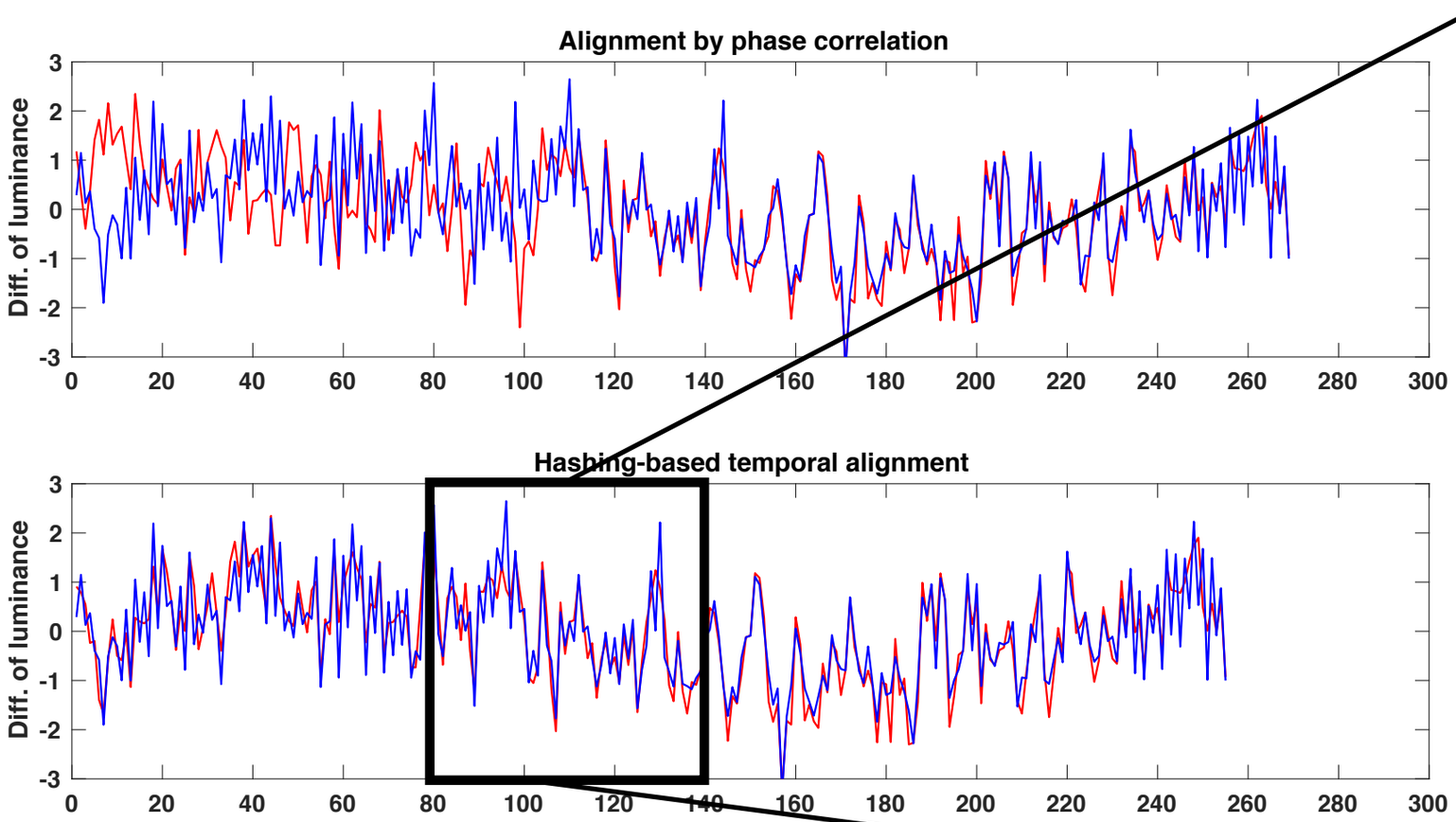


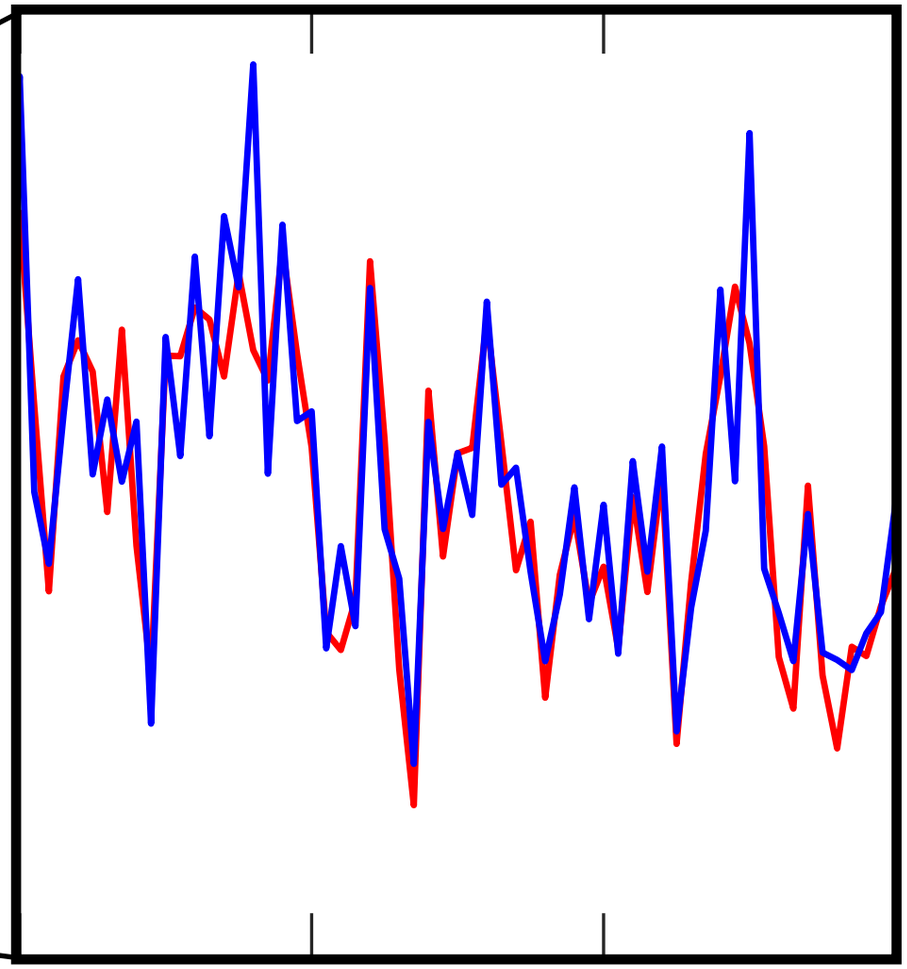
Step 3
Near-duplicate alignment



Step 3
Near-duplicate alignment







Experiments

Experiments

- Dataset
  - 300 trees with 10 videos
  - 100 - No clip
  - 100 - Clip at the Beginning/end of the stream
  - 100 - Clip anywhere

Results

		Difference of luminance					Hashing-based				
		Root	Depth	Edges	Leaves	Ancestry	Root	Depth	Edges	Leaves	Ancestry
No coding matching	No clip	0,630	0,500	0,772	**0,832**	0,674	**0,650**	**0,440**	**0,777**	0,828	**0,702**
	Clip border	**0,710**	0,35	0,788	0,826	0,724	**0,710**	**0,330**	**0,807**	**0,846**	**0,740**
	Clip any	0,620	0,560	0,724	0,758	0,628	**0,670**	**0,450**	**0,755**	**0,789**	**0,683**
With coding matching	No clip	**0,870**	**0,140**	**0,839**	**0,859**	**0,812**	0,690	0,400	0,778	0,831	0,711
	Clip border	**0,840**	**0,170**	**0,863**	**0,901**	**0,834**	0,660	0,400	0,780	0,824	0,722
	Clip any	**0,720**	**0,390**	**0,801**	**0,828**	**0,696**	0,660	0,470	0,765	0,810	0,696

Diff. of luminance vs. Hashing-based alignment

Conclusion

Conclusion

- New solutions for multimedia phylogeny
- Robust methods for phylogeny forest reconstruction
- More robust dissimilarity measures
  - Gradient + mutual information
- Good results for video phylogeny when dealing with misaligned and compressed videos

Future work

- Other transformations
  - Blurring, sharpness, content insertion…
- Videos with different frame rates
- Video dissimilarity calculation using time information
- Robust methods to reliably estimate the compression parameters directly from the video

Acknowledgements

- Examiners
- Prof. Anderson Rocha and Prof. Zanoni Dias
- RECOD - IC/Unicamp
- ISPG - Politecnico di Milano
- FAPESP
- CAPES

Thanks!