Fast Graph Sampling for Short Video Summarization using Gershgorin Disc Alignment

Paper id: 5050

Sadid Sahami [1] Gene Cheung [2] Chia-Wen Lin [1]

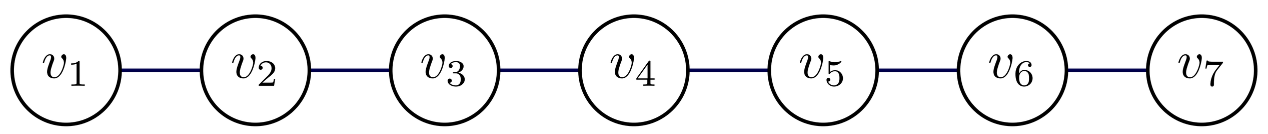
[1]National Tsing Hua University, Hsinchu, Taiwan

[2]York University, Toronto, Canada

Video Summarization

- **Problem**: An unsupervised linear-time complexity key-frame selector for short videos
- **Prior Work**:
  - Group 1: Heuristic or random based methods, sometimes fast but no optimization objective, poorer performance [1]
  - Group 2: Optimization based with usually complex/approximate solution [2]
  - Our goal: Devise an optimization-based but linear-complexity solution with advantages of both groups

Video model: Similarity Path Graph (SPG)



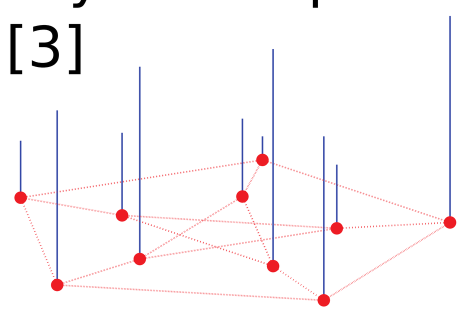
The similarity metric is defined as:

$$W_{i,i+1} := \frac{\|\mathbf{f}_i - \cos(\theta_{i,i+1})\mathbf{f}_{i+1}\|_2 + \|\mathbf{f}_{i+1} - \cos(\theta_{i,i+1})\mathbf{f}_i\|_2}{\|\mathbf{f}_i\|_2 + \|\mathbf{f}_{i+1}\|_2}$$

where $\theta_{i,i+1} = \angle(\mathbf{f}_i, \mathbf{f}_{i+1})$.

Graph Signal Processing (GSP)

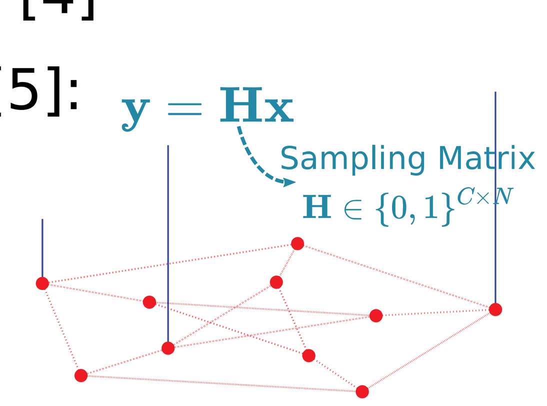
- GSP: Study of how to analyze and process data associated with graphs [3]
- Graph: $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$
- Graph Signal: $\mathbf{x} \in \mathbb{R}^N$
  - set of scalers associated to vertices



- Combinatorial graph Laplacian matrix: $\mathbf{L} = \mathbf{D} - \mathbf{W}$ where $\mathbf{D}$ is the degree matrix
- GLR: $\mathbf{x}^T\mathbf{L}\mathbf{x}$ captures signal smoothness w.r.t graph
  - Graph Laplacian Regularizer (a smoothness prior)

Graph Sampling

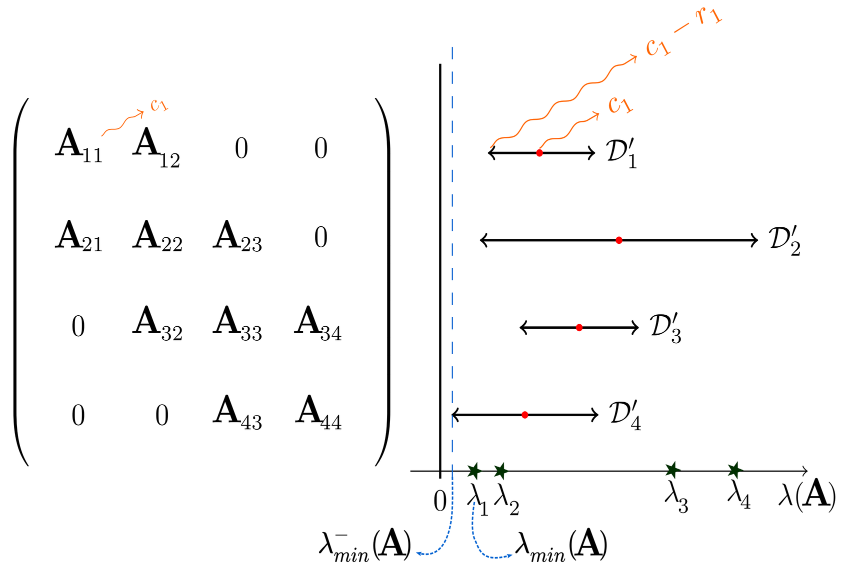
- Problem of finding best subset of samples, $\mathbf{y} \in \mathbb{R}^C$, such that it cost least reconstruction error [4]
- One particular objective of interest: [5]:
  $\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \mu\mathbf{x}^\top\mathbf{L}\mathbf{x} \quad (1)$
- The solution $\mathbf{x}^*$: $\underbrace{(\mathbf{H}^\top\mathbf{H} + \mu\mathbf{L})}_{\mathbf{B}}\mathbf{x}^* = \mathbf{H}^\top\mathbf{y}$
  - $\mathbf{H}$ is matrix of one-hot vectors then $\mathbf{H}^T\mathbf{H} := \mathrm{diag}(\mathbf{a})$



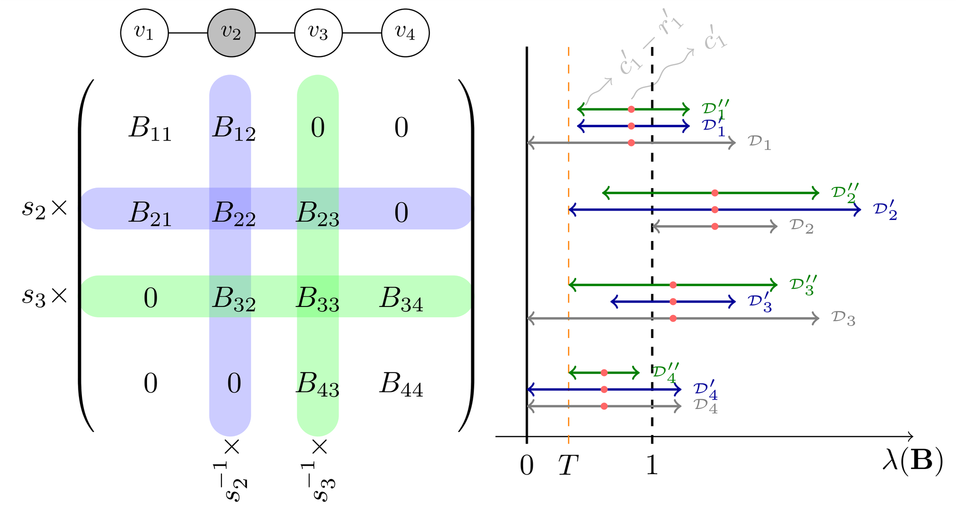
- Equivalently find $\mathbf{H}$ to $\max\ \lambda_{\min}(\mathbf{B})$ [6]

Gershgorin Circle Theorem (GCT)

- GCT: each eigenvalue $\lambda(\mathbf{A})$ resides in at least one Gershgorin disc [7]
- Discs center and radius: $c_i = A_{i,i}$, $r_i = \Sigma_{j\neq i} A_{i,j}$
  - As a corollary: Gershgorin disc lower bound $\lambda^-_{\min}(\mathbf{A}) \leq \lambda_{\min}(\mathbf{A})$.

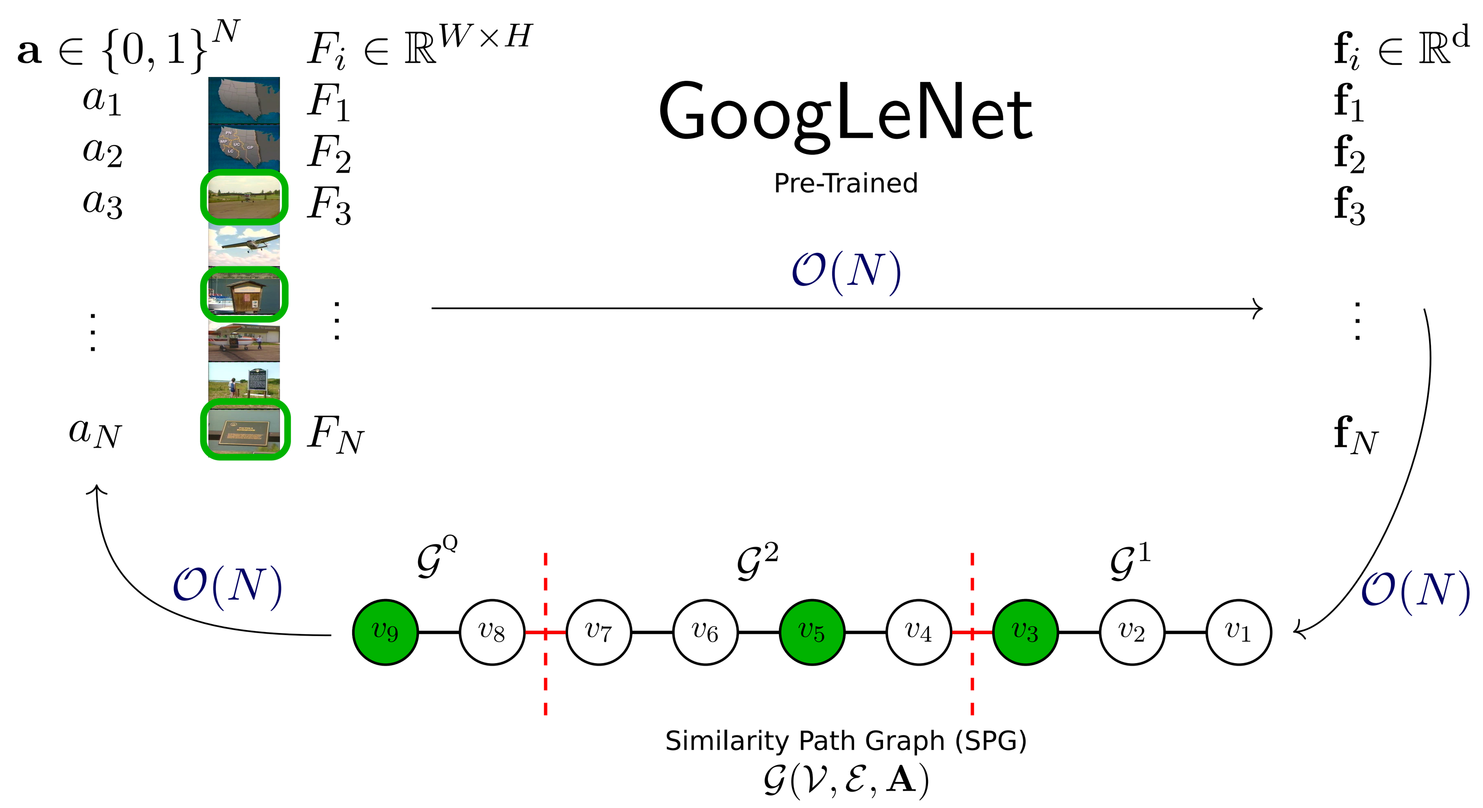


Gershgorin Disc Alignment based Sampling



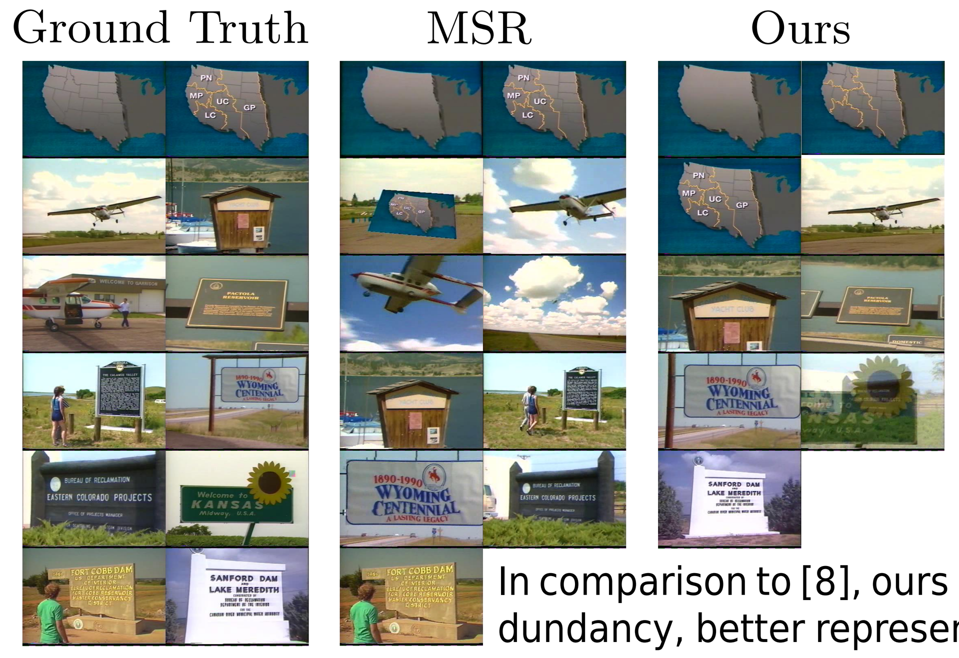
- Paradigm: Solve (1) by maximizing $\lambda^-_{\min}(\mathbf{B})$
- Sampling a node i, shift D_i by 1 ($D_i \rightarrow D'_i$)
- With $\lambda(\mathbf{B}) = \lambda(\mathbf{SBS}^{-1})$, scalers can expand/contracts discs[7]
- Using shift and $\mathbf{S}$ operators, align $\lambda^-_{\min}(\mathbf{B})$ beyond T ($D'_i \rightarrow D''_i$)
- Binary search largest T for the desired sampling budget

Overview of our method



- Sampling **(not fast)**: Sample SPG by maximizing $\lambda_{\min}(\mathbf{B})$ where $\mathbf{B} = \mathrm{diag}(\mathbf{a}) + \mu\mathbf{L}$ and $\mathbf{a}$ the sampling vector
- GDA based sampling **(fast)**: Avoid eigenvalue decomposition by maximizing $\lambda^-_{\min}(\mathbf{B})$ instead based on GDAS [6]
- Specialized GDA-based sampling **(faster)**: We prove that, by partitioning $\mathcal{G}$ into sub-graphs $\{\mathcal{G}^q\}_{q=1}^{Q}$, then $\min_q \lambda^-_{\min}(\mathbf{B}^q)$ is a lower bound for $\lambda^-_{\min}(\mathbf{B})$ which enables even faster sampling for SPG

Numerical & Qualitative Results



In comparison to [8], ours has less redundancy, better representation and better sparsity

Table 1. Results on VSUMM benchmark

Algorithm	P (%)	R (%)	F_1 (%)
DT	35.51	26.71	29.43
STIMO	34.73	40.03	35.75
VSUMM	**47.26**	42.34	43.52
MSR	36.94	57.61	43.39
AGDS	37.57	64.60	45.52
SBOMP [2]	39.28	62.28	46.68
SBOMPn [2]	**41.23**	**68.47**	**49.70**
Ours	39.67	**71.48**	**48.92**

- Precision, Recall
- $P_u = \frac{|\mathcal{A}\cap\mathcal{U}_u|}{|\mathcal{A}|}$,
- $R_u = \frac{|\mathcal{A}\cap\mathcal{U}_u|}{|\mathcal{U}_u|}$,
- $F_{1,u} = \frac{2P_uR_u}{P_u+R_u}$
- u users id
- $F_1 = \frac{2PR}{P+R}$

Algorithm	Complexity ($\mathcal{O}$)
SBOMP[2]	$\mathcal{O}(\mathrm{d}N^2m + \mathrm{d}^2Nm^3)$
Ours	$\mathcal{O}(ND^2\log 1/\epsilon)$

N: Number of frames
d: feature vector dimension
m: Number of keyframes
D: Maximum recursion ($D \ll N$)
ϵ: Binary search precision

Conclusion

- New class of key-frame selector based on *Graph Sampling*
- Scalable key-frame selector with comparable results
- Devise specialized GDAS[6] based graph sampling for SPG

References

[1] S. E. F. de Avila, A. P. B. Lopes, A. da Luz, and A. de Albuquerque Araújo, "VSUMM: A mechanism designed to produce static video summaries and a novel evaluation method," en, *Pattern Recognition Letters*, vol. 32, no. 1, pp. 56–68, Jan. 2011, ISSN: 0167-8655.

[2] S. Mei, M. Ma, S. Wan, J. Hou, Z. Wang, and D. D. Feng, "Patch Based Video Summarization With Block Sparse Representation," *IEEE Transactions on Multimedia*, vol. 23, pp. 732–747, 2021, ISSN: 1941-0077.

[3] A. Ortega, *Introduction to Graph Signal Processing*. Cambridge University Press, 2022.

[4] Y. Tanaka, Y. C. Eldar, A. Ortega, and G. Cheung, "Sampling signals on graphs: From theory to applications," *IEEE Signal Processing Magazine*, vol. 37, no. 6, pp. 14–30, Nov. 2020, ISSN: 1558-0792.

[5] J. Pang and G. Cheung, "Graph laplacian regularization for image denoising: Analysis in the continuous domain," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1770–1785, 2017.

[6] Y. Bai, F. Wang, G. Cheung, Y. Nakatsukasa, and W. Gao, "Fast graph sampling set selection using Gershgorin disc alignment," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2419–2434, 2020, ISSN: 1941-0476.

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge university press, 2012.

[8] S. Mei, G. Guan, Z. Wang, S. Wan, M. He, and D. Dagan Feng, "Video summarization via minimum sparse reconstruction," en, *Pattern Recognition*, vol. 48, no. 2, pp. 522–533, Feb. 2015, ISSN: 0031-3203.

s.sahami@ieee.org genec@yorku.ca cwlin@ee.nthu.edu.tw