

### Sequential Determinantal Point Processes (SeqDPPs):

Models, Algorithms, and Applications in Diverse and Sequential Subset Selection

**Boqing Gong** 

BoqingGo@outlook.com

#### Video summarization

#### Extractive video summarization



Subset Selection problem

Compositional video summarization

Limited to well-controlled videos



[Pritch et al.'09]

#### Two competing criteria

Extracting frames/shots

Individually important

Collectively diverse

[Wolf 1996, Vasconcelos and Lippman 1998, Aner and Kender 2002, Pal and Jojic 2005, Kang et al. 2006, Pritch et al. 2007, Jiang et al. 2009, Lee and Kwon 2012, Khosla et al. 2013, Kim et al. 2014, Song et al. 2015, Lee and Grauman 2015, ...]



1:00 pm 2:00 pm 3:00 pm 4:00 pm 5:00 pm 6:00 pm

Output: a storyboard summary

# Prior work (before 2014)

[Wolf 1996, Vasconcelos and Lippman 1998, Aner and Kender 2002, Pal and Jojic 2005, Kang et al. 2006, Pritch et al. 2007, Jiang et al. 2009, Lee and Kwon 2012, Khosla et al. 2013, Kim et al. 2014, Song et al. 2015, Lee and Grauman 2015, ...]

Measuring importance of frames/shots

Low-level visual cues, motion cues

Weakly supervised Web images, texts

Human labeled objects, attributes, etc.

#### Cons:

Indirect cues

System developers making decisions for users

### Our goal (2014): **Supervised** video summarization

Learn video summarizer from user summaries

What model constitutes a good video summarizer?

### Model selection for **Supervised** video summarization



### Why DPP?

Modeling subset selection

Modeling diversity & importance

A generative probabilistic model

Supervised video summarization

Maximum likelihood & large-margin estimation

Effective for document summarization

#### This talk

DPP SeqDPP Variations Lessons Learned

DPP

Large-margin DPP

BoqingGo@outlook.com

#### Discrete point process

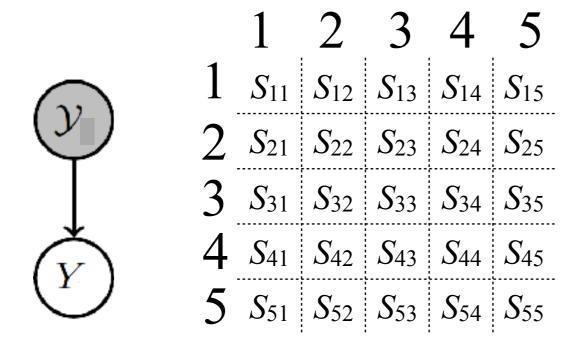
ullet N items (e.g., images or sentences):

$$\mathcal{Y} = \{1, 2, ..., N\}$$

- ullet 2 possible subsets
- ullet Probability measure  ${\mathcal P}$  over subsets  $Y\subseteq {\mathcal Y}$

Vanilla DPP is a discrete point process.

#### Determinantal point process (DPP)



$$P(Y = \{2, 4\})$$

$$\mathcal{Y} = \{1, 2, 3, 4, 5\}$$

 $Y \subseteq \mathcal{Y}$ : subset selection variable

Vanilla DPP is a discrete point process.

#### Determinantal point process (DPP)

$$P(Y = \{2, 4\})$$
 $\propto \det \begin{pmatrix} S_{22} & S_{24} \\ S_{42} & S_{44} \end{pmatrix}$ 

$$\mathcal{Y} = \{1, 2, 3, 4, 5\}$$

 $Y \subseteq \mathcal{Y}$ : subset selection variable

Vanilla DPP is a discrete point process.

#### DPP models diversity & importance

Items 2 and 4

diverse, larger probability

important, larger probability

$$P(Y = \{2, 4\})$$

$$\propto \det \begin{pmatrix} S_{22} & S_{24} \\ S_{42} & S_{44} \end{pmatrix}$$

$$= S_{22} \cdot S_{44} - S_{24} \cdot S_{42}$$

#### DPP models diversity & importance

$$P(Y = \{2, 4\})$$
 $\propto \det \begin{pmatrix} S_{22} & S_{24} \\ S_{42} & S_{44} \end{pmatrix}$ 
 $= S_{22} \cdot S_{44} - S_{24} \cdot S_{42}$ 

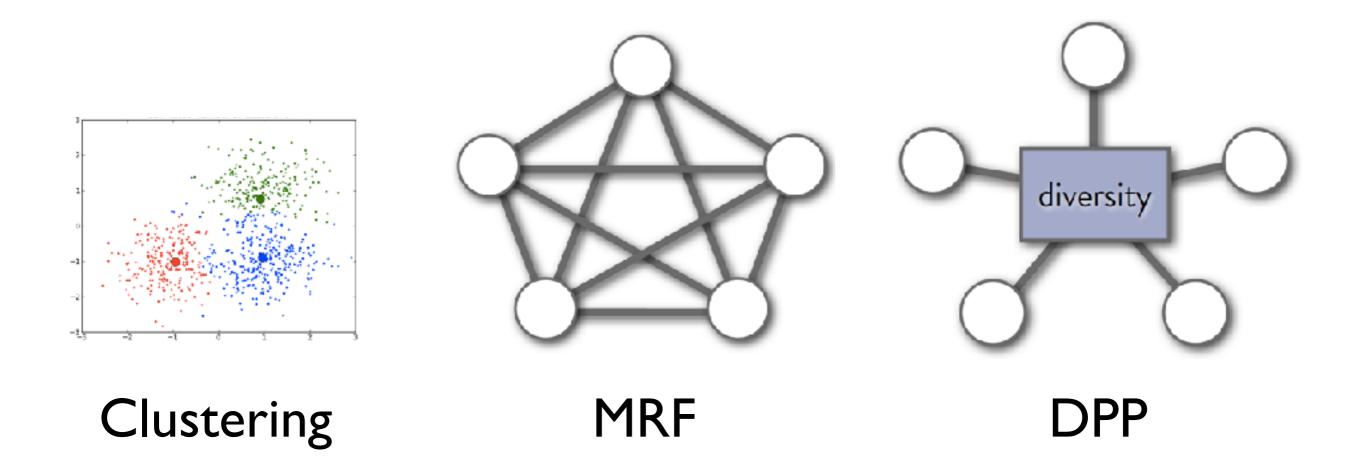
#### importance

#### DPP models diversity & importance

$$P(Y = \{2, 4\})$$
 $\propto \det \begin{pmatrix} S_{22} & S_{24} \\ S_{42} & S_{44} \end{pmatrix}$ 
 $= S_{22} \cdot S_{44} - S_{24} \cdot S_{42}$ 

#### Diversity

### Diversity



### Diversity

	MRF	DPP	
Inference	NP	Mostly tractable	
MAP inference	NP	NP	
Approx. MAP	Likewise NP	1/4	

#### DPP: some properties

Modeling subset selection, diversity, & importance

Log-submodular

MAP inference is NP-hard

1/4-approximation under some constraints

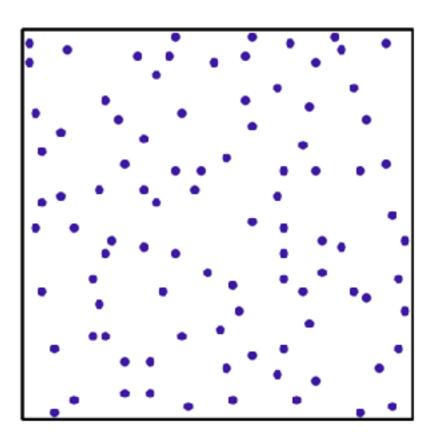
Efficient sampling

Two-stage sampling, MCMC sampling

Closed-form marginalization & conditioning

DPP

$$P(Y) \propto \det(L_Y)$$



• DPP

$$P(Y) \propto \det(L_Y)$$

• **k-DPP** [Kulesza & Taskar, 2011]  $_{\mathrm{S.t.}}$   $\mathrm{CARD}(Y)=k$ 

- DPP
- **k-DPP** [Kulesza & Taskar, 2011]
- Markov DPP [Affandi et al., 2012]

- DPP
- **k-DPP** [Kulesza & Taskar, 2011]
- Markov DPP [Affandi et al., 2012]
- Structured DPP [Kulesza & Taskar, 2010]
- Continuous DPP [Affandi et al., 2013]
- Sequential DPPs [Gong et al., NIPS'14, UAI'15] [ECCV'16, CVPR'17, ECCV'18ab]

#### This talk

DPP SeqDPP Variations Lessons Learned

Vanilla DPP for supervised video summarization

BoqingGo@outlook.com

### Video summarization by vanilla DPP

$$\mathcal{Y} = \{1, 2, 3, 4, 5\} = \left\{ \begin{array}{c} \\ \\ \\ \end{array} \right\}$$

$$Y = \{1, 2, 3, 4, 5\} = \left\{ \begin{array}{c} \\ \\ \end{array} \right\}$$

$$Y = \left\{ 1, 2, 3, 4, 5 \right\} = \left\{ \begin{array}{c} \\ \\ \end{array} \right\}$$

$$Y = \left\{ 1, 2, 3, 4, 5 \right\} = \left\{ \begin{array}{c} \\ \\ \end{array} \right\}$$

$$Y = \left\{ 1, 2, 3, 4, 5 \right\} = \left\{ \begin{array}{c} \\ \\ \end{array} \right\}$$

	1	2	3	4	)
1	$S_{11}$	$S_{12}$	$S_{13}$	$S_{14}$	$S_{15}$
2	$S_{21}$	4		$S_{24}$	$S_{25}$
3	$S_{31}$	$S_3$		$S_{34}$	$S_{35}$
4	$S_{41}$	$S_{42}$	$S_{43}$	$S_{44}$	$S_{45}$
5	$S_{51}$	$S_{52}$	$S_{53}$	$S_{54}$	$S_{55}$

# Parameterizing kernels for out-of-sample extension

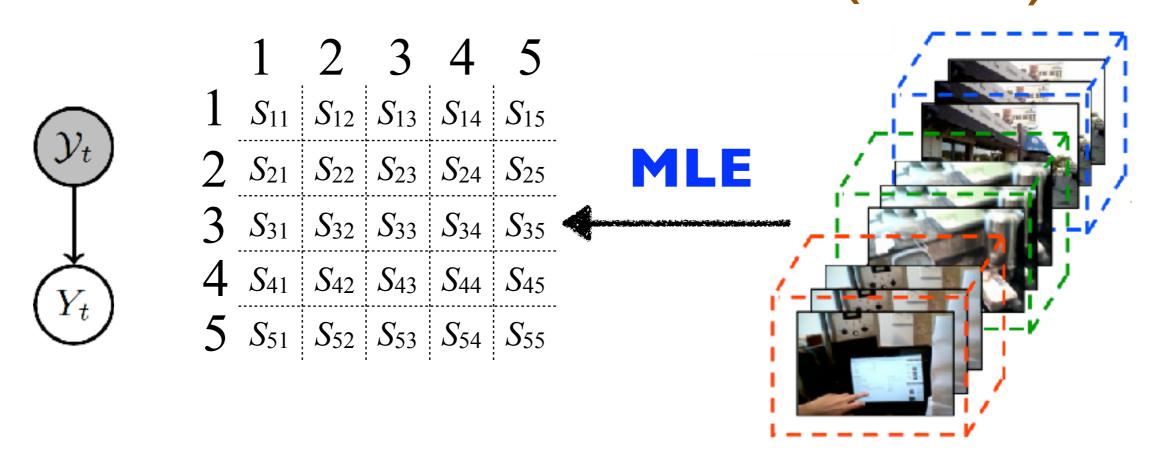
$$L_{ij} = \langle f(\mathbf{x}_i), f(\mathbf{x}_j) \rangle$$

1-layer neural network:  $f(\mathbf{x}) = W \tanh(U\mathbf{x})$ 

Linear:  $f(\mathbf{x}) = W\mathbf{x}$ 

	1	2	3	4	5
1	$S_{11}$	$S_{12}$	$S_{13}$	$S_{14}$	$S_{15}$
2	$S_{21}$	4		$S_{24}$	$S_{25}$
3	$S_{31}$	$S_3$		$S_{34}$	$S_{35}$
4	$S_{41}$	$S_{42}$	$S_{43}$	$S_{44}$	$S_{45}$
5	$S_{51}$	$S_{52}$	$S_{53}$	$S_{54}$	$S_{55}$

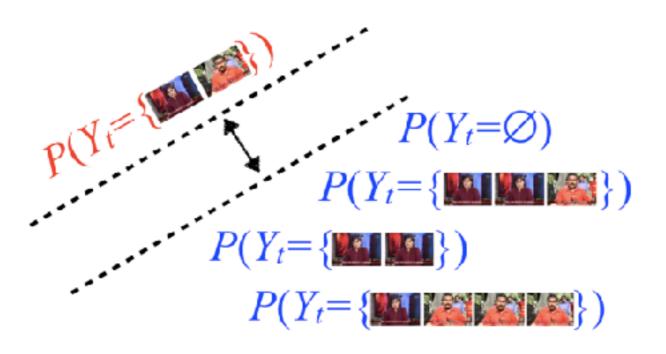
## Learning kernels by maximum likelihood estimation (MLE)



# Learning kernels by the large-margin principle [UAI'15]



Wei-Lun Chao



#### Advantages over MLE

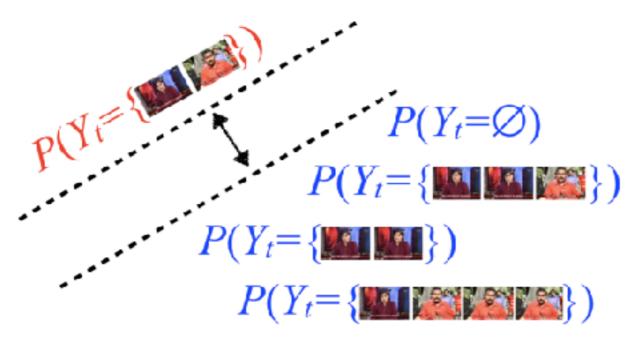
Tracking errors

Accepting various margins (e.g., trade-off precision & recall)

# Learning kernels by the large-margin principle [UAI'15]



Wei-Lun Chao



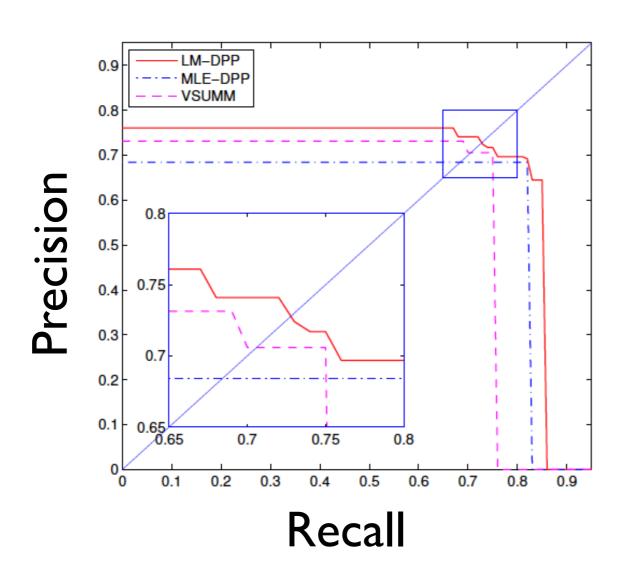
#### Main challenge:

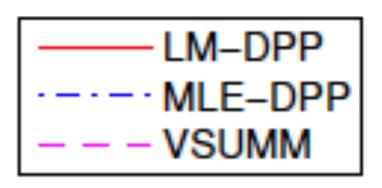
An exponential number of negative examples

#### Solution:

Multiplicative margin
Upper bound by softmax

# Large-margin DPP better balances precision & recall





### Video summarization by vanilla DPP: what's missing?

DPP fails to capture the **temporal structure** of videos









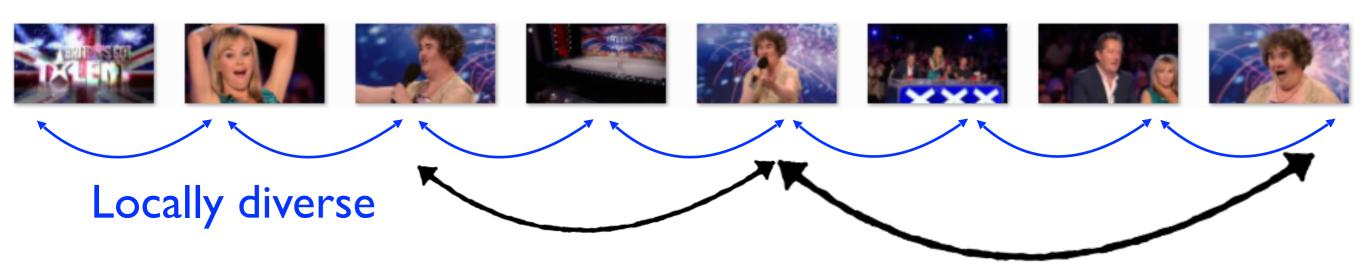




Susan Boyle performs in "Britain's Got Talent".

"Britain's Got Talent" ... surprises a lady.

### Need of a "sequential" DPP



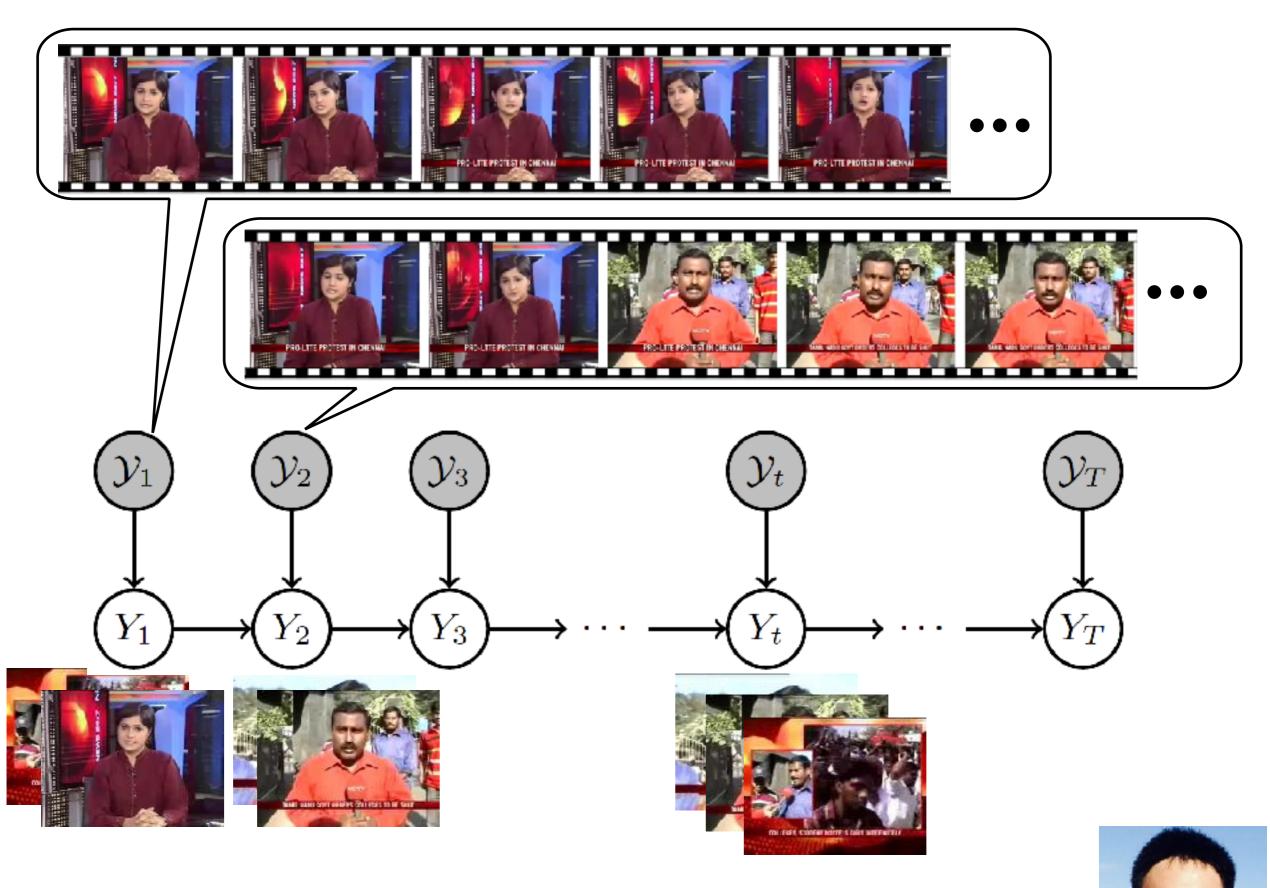
Globally not as diverse as locally

#### This talk

DPP SeqDPP Variations Lessons Learned

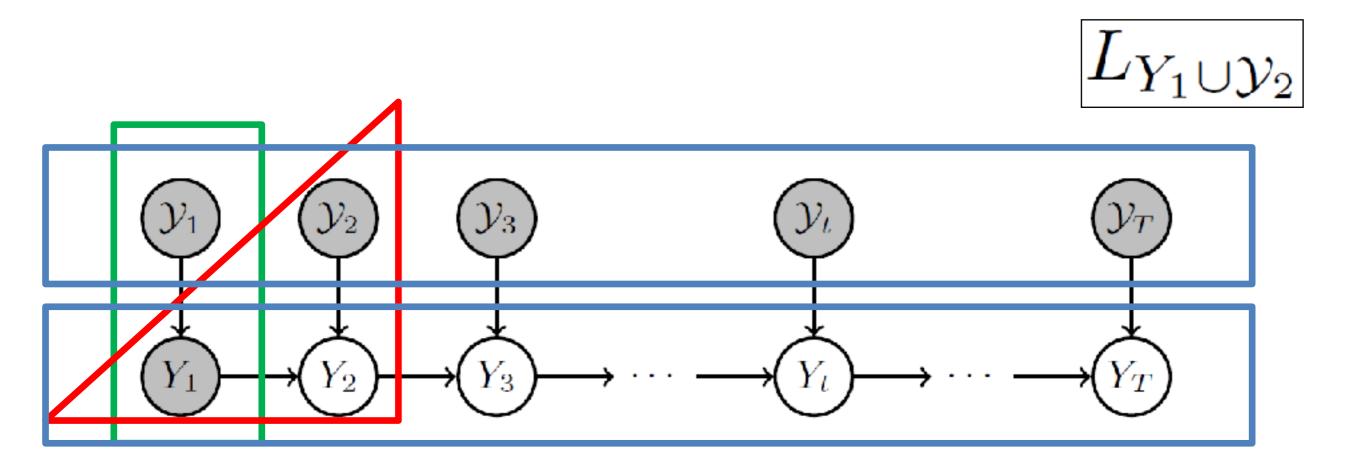
Sequential DPP for supervised video summarization

BoqingGo@outlook.com



[NIPS'14]

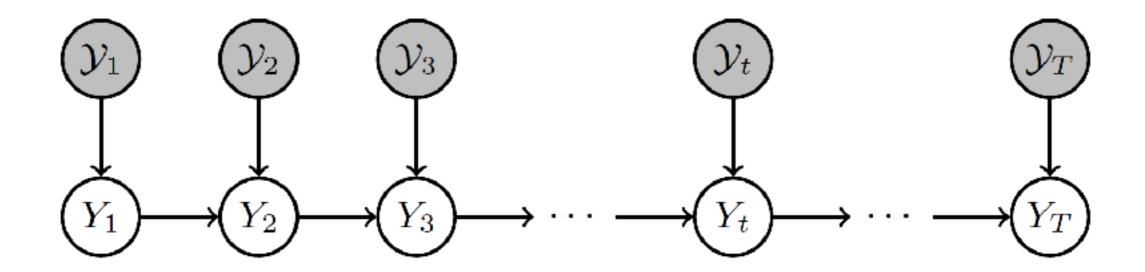
### Sequential DPP (seqDPP)



$$P(Y_1 = \mathbf{y}_1, Y_2 = \mathbf{y}_2, \dots, Y_T = \mathbf{y}_T) = P(Y_1 = \mathbf{y}_1) \prod_{t=2}^{T} P(Y_t = \mathbf{y}_t | Y_{t-1} = \mathbf{y}_{t-1})$$

Conditional probability: still a DPP!

#### SeqDPP vs. DPP



Modeling importance, diversity, and sequential structure

More efficient inference:  $O(2^N) \rightarrow O(M \cdot 2^{N/M})$ 

Summarizing streaming videos on the fly

#### Experimental study

Three benchmark datasets:

Open video project, Youtube (50), Kodak

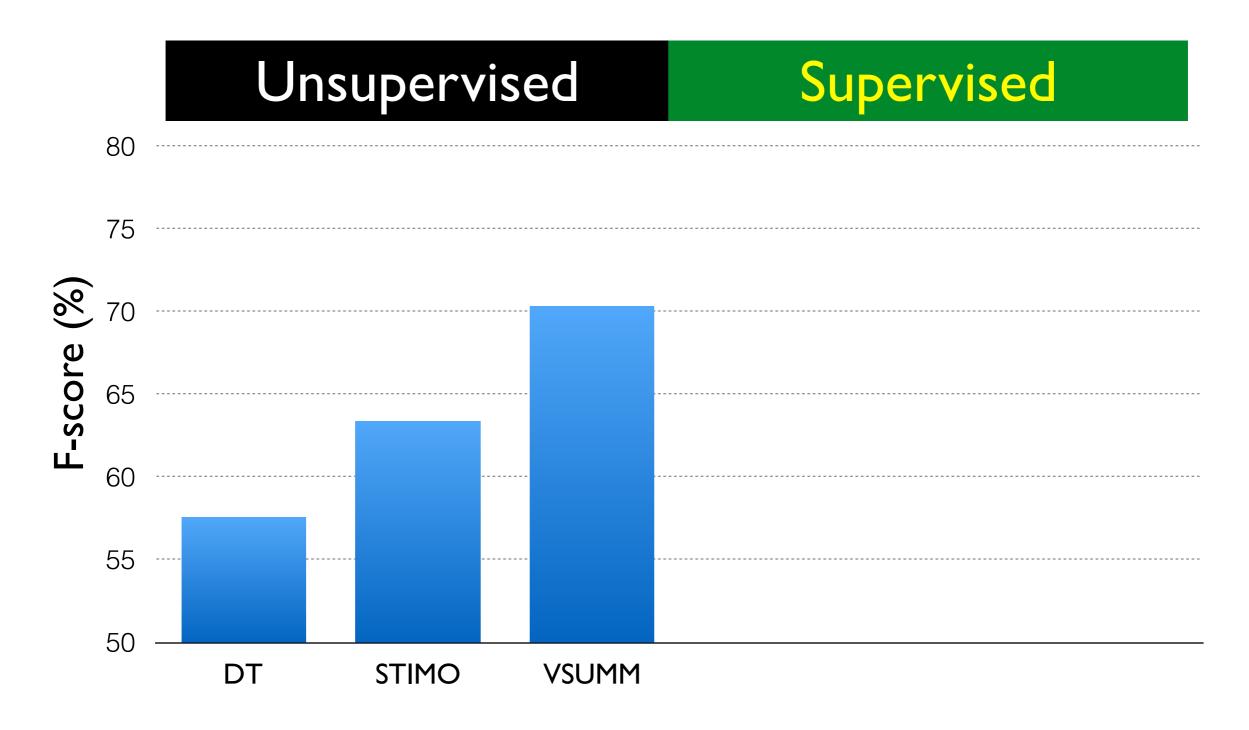
Preprocessing: down-sampling I frame/sec

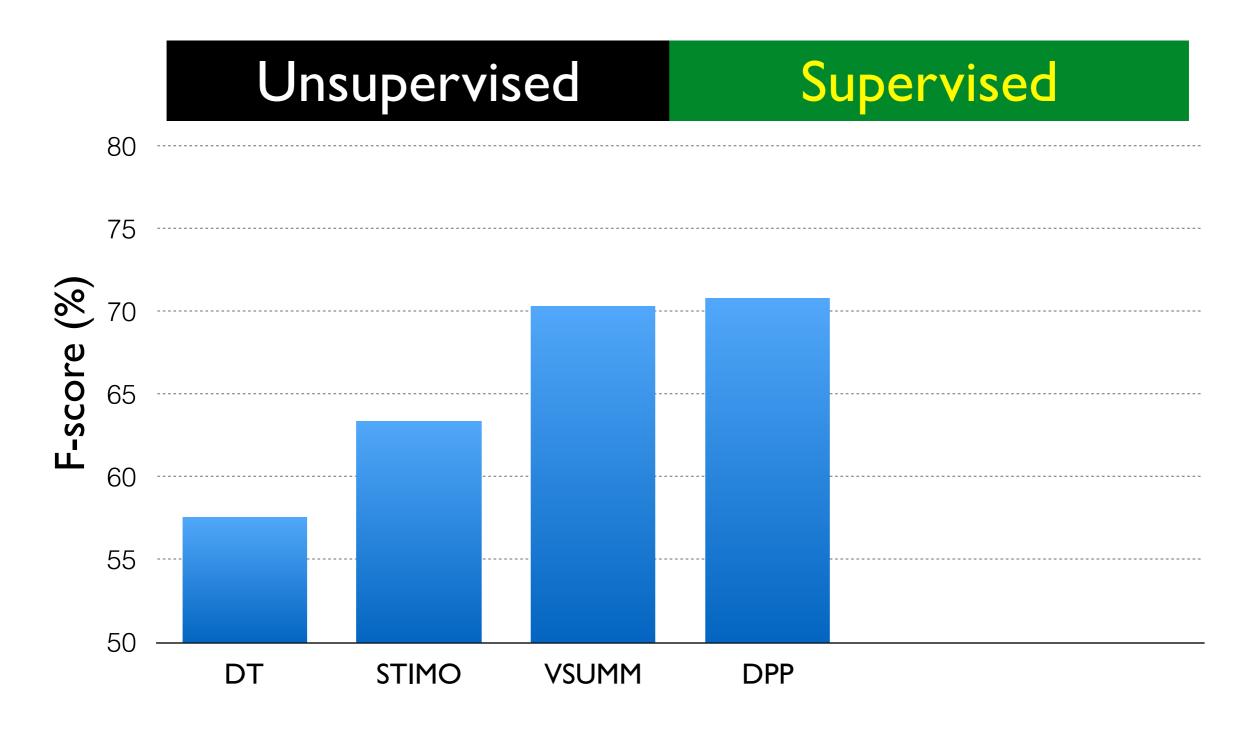
Features: saliency, Fisher vectors, context

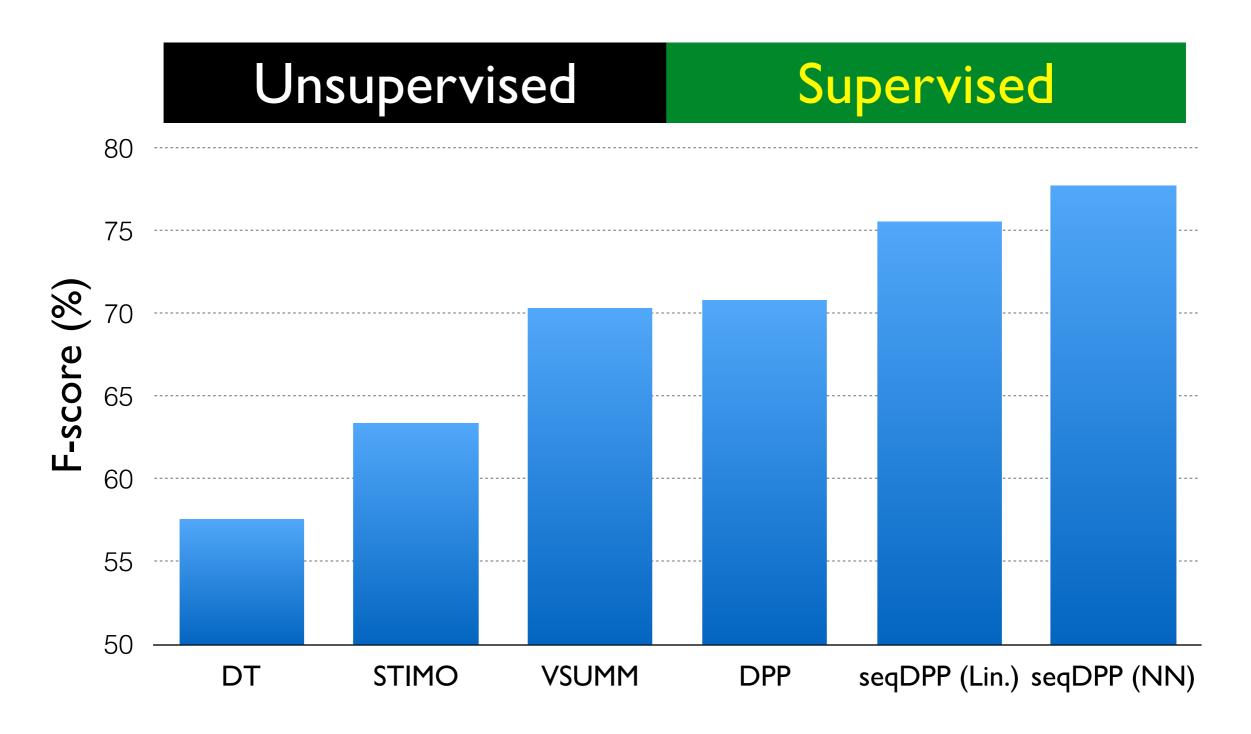
**Evaluation:** 

Precision, recall, F-score by the VSUMM package

#### Experimental results







## **SeqDPP**

Code: <a href="https://github.com/pujols/Video-summarization">https://github.com/pujols/Video-summarization</a>

#### Large-Margin Determinantal Point Processes

Wei-Lun Chao\*

U. of Southern California Los Angeles, CA 90089 Boqing Gong\*

U. of Southern California Los Angeles, CA 90089 Kristen Grauman

U. of Texas at Austin Austin, TX 78701 Fei Sha

U. of Southern California Los Angeles, CA 90089

#### Diverse Sequential Subset Selection for Supervised Video Summarization

#### Boqing Gong\*

Department of Computer Science University of Southern California Los Angeles, CA 90089 boqinggo@usc.edu

#### Kristen Grauman

Department of Computer Science University of Texas at Austin Austin, TX 78701 grauman@cs.utexas.edu

#### Wei-Lun Chao\*

Department of Computer Science University of Southern California Los Angeles, CA 90089 weilunc@usc.edu

#### Fei Sha

Department of Computer Science University of Southern California Los Angeles, CA 90089 feisha@usc.edu [UAI 2015]

[NIPS 2014]

## Thus far,

**Supervised** video summarization

DPP: MLE & large-margin

Sequential DPP

Experimental results & analysis

## Lessons learned

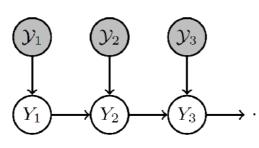
Video summarization is **subjective** 

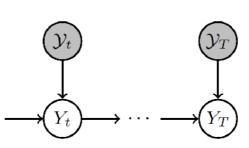
I. <u>Personalization</u>
System needs a channel to infer user's preference

#### 2. Evaluation is hard





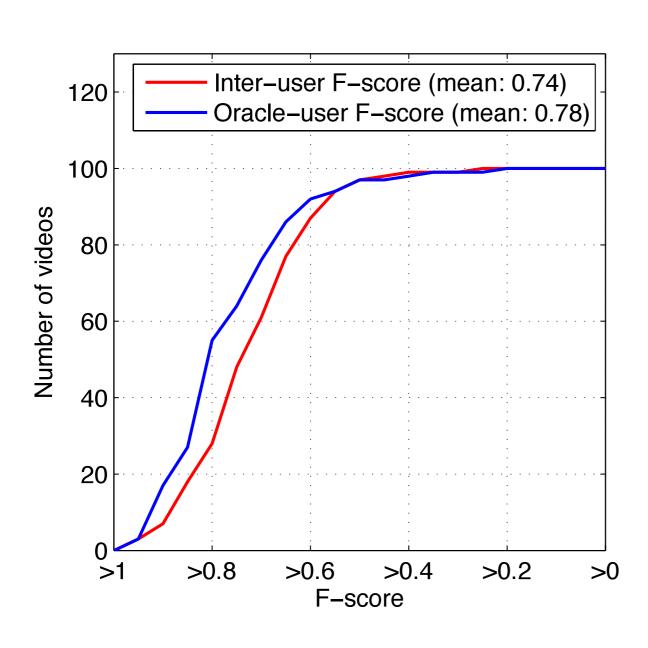








## Inter-user agreement



100 videos

Five summaries per video

No "groundtruth" summary

Fairly high inter-user agreement

## This talk

DPP

SeqDPP

**Variations** 

Lessons Learned

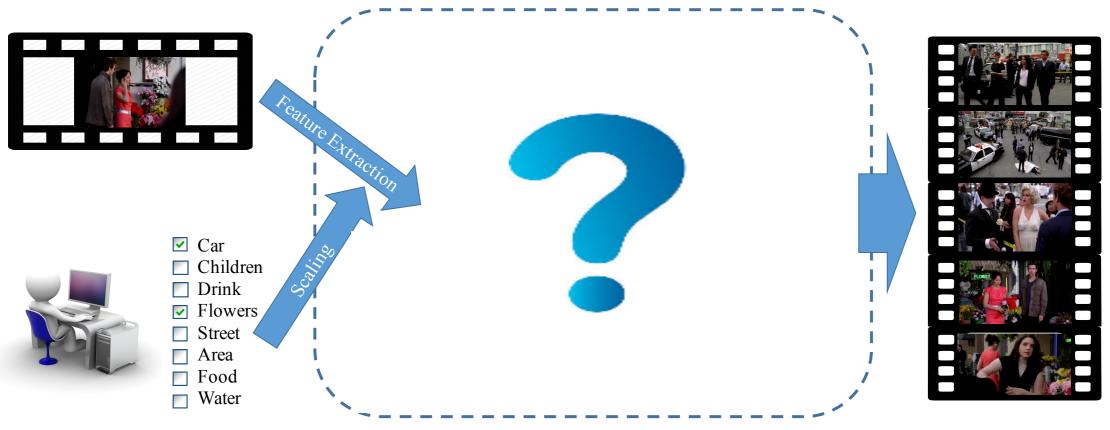
## User-subjectivity

- I. Personalizing video summarizers
- 2. An improved evaluation metric



# Query-focused video summarization

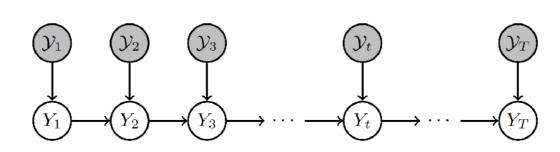








(a) Input: Video & Query



(b) Algorithm: Sequential & Hierarchical Determinantal Point Process (SH-DPP) (c) Output: Summary





# Query-focused video sumarization



Decision to include a frame/short in summary

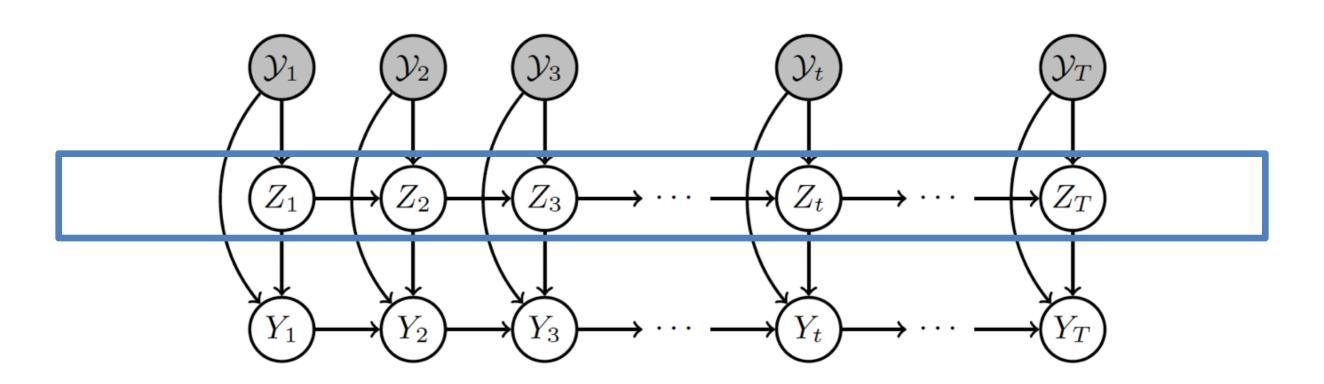
Relevance to query (be responsive to user input)

Importance in the context (maintain story flow)

Collective diversity

Two levels of summarization granularity.

# Sequential and hierarchical DPP (SH-DPP)



Z-layer summarizes query-relevant video shots/frames.

# Z-layer: responsive to user query q

≅ SeqDPP: Markov process with DPP

Summarizes shots/frames relevant to query

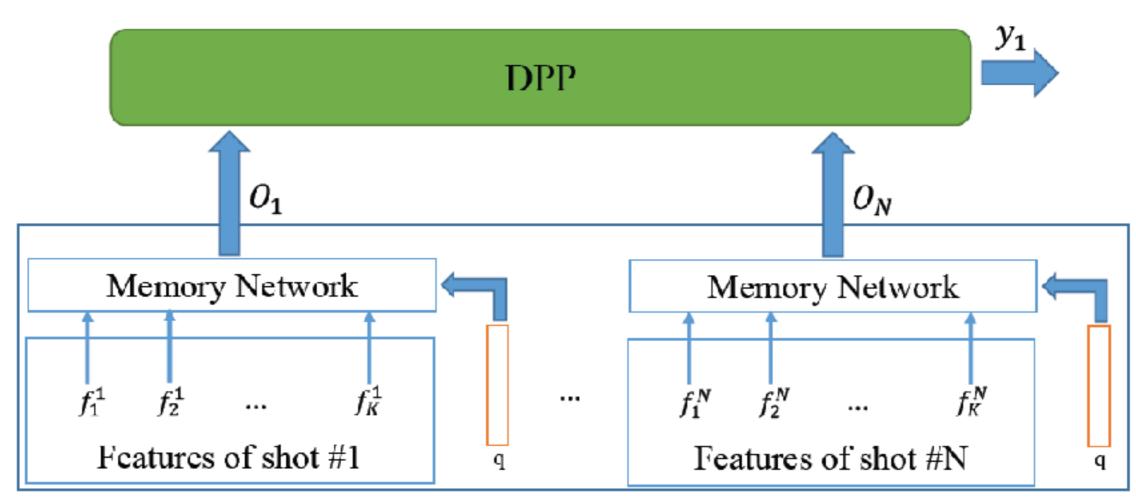
The DPP kernel is thus query-dependent

$$\Omega_{ij} = [\boldsymbol{f}_i(\boldsymbol{q})]^T W^T W[\boldsymbol{f}_j(\boldsymbol{q})]$$

Z-layer summarizes query-relevant video shots/frames.

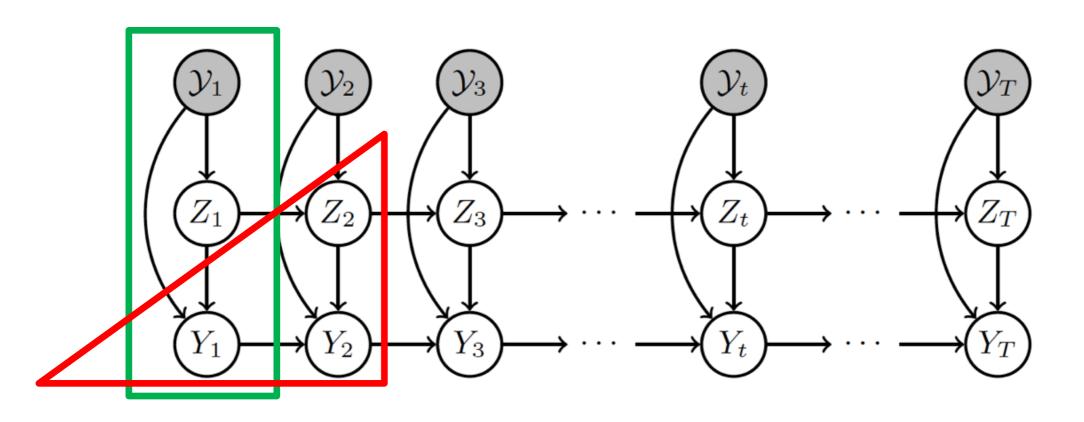


# Z-layer: responsive to user query q

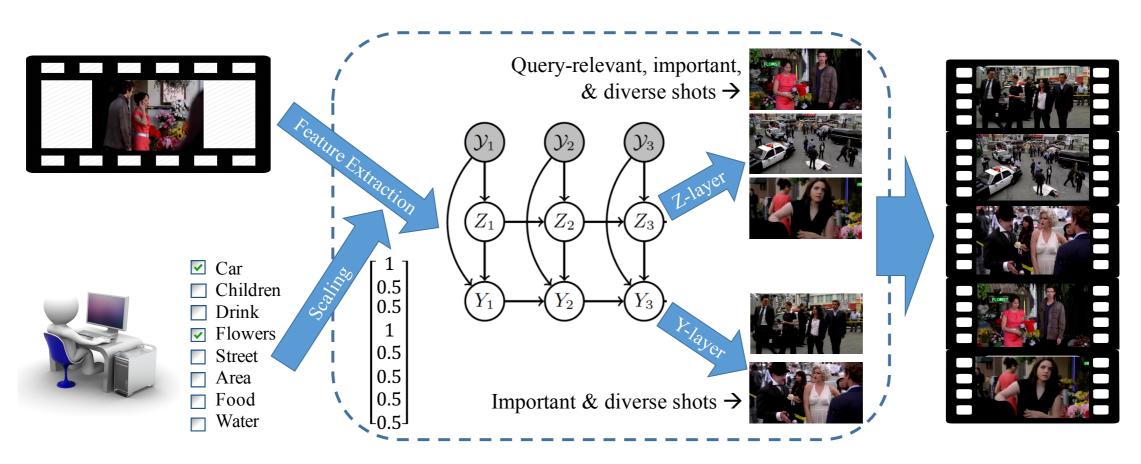




# Y-layer: summ. remaining video (maintain story flow)



# Query-focused video summarization



(a) Input: Video & Query (b) Algorithm: Sequential & Hierarchical Determinantal Point Process (SH-DPP) (c) Output: Summary

## Query: CAR+PHONE

Cho and Lisbon examine Hanson's CAR



Lisbon and Rigsby speak on the PHONE.



## Relevant to query

Felicia Scott speaks to Sydney on the PHONE, while the movie is being filmed.



### **Query: CAR+PHONE**

Cho and Lisbon examine Hanson's CAR



Lisbon and Rigsby speak on the PHONE.



### Relevant to query

Felicia Scott speaks to Sydney on the PHONE, while the movie is being filmed.





Jane finishes his conversation with the policeman.



Mitch Cavenaugh enters the RV, and explains the drugs are his



Jane speaks to Felicia Scott about how well she is acting.

Important in context

(maintain story flow)

### Query: FLOWER+WALL

My friend drove the car, and I sat in the passenger seat.



I looked at FLOWERS at the booth.



I watched the TV on the WALL.



I looked across the room.



My friend drove the car, and I sat in the passenger seat.

Detection Scores: FLOWER: 0.86 WALL: 0.65



Y-layer

Z-layer



I walked toward the tents



I watched the lady ask my friend a question.



My friend and I walked down the street on the sidewalk.



I walked down the street on the sidewalk.

#### **Ground-truth Summary**

My friend drove the car, and I sat in the passenger seat. I got out of the car. I walked toward the tents. I looked at the fruit at the booth. My friend and I walked through the market. My friend and I looked at FLOWERS at the booth. My friend drove the car, and I sat in the passenger seat.

I sat with my friend and looked over at the TV on the WALL. I sat at the table while my friend drank. I ate pizza with my friend and we looked at the TV. I looked at the TV on the WALL and then looked back at my friend. I watched the TV on the WALL's at the restaurant.

I walked out the shop with my friend. My friend and I walked down the street on the sidewalk. I walked on the side walk.

## This talk

DPP

SeqDPP

**Variations** 

**Lessons Learned** 

## User-subjectivity

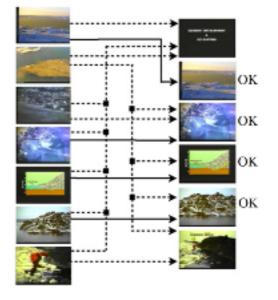
- I. Personalizing video summarizers
- 2. An improved evaluation metric



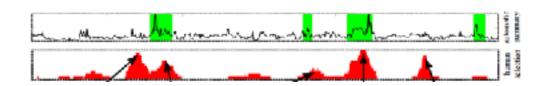
## What makes a good evaluation for video summarization?



A/B test



**Bipartite** matching [Avila et al. 2011]



Time overlap [Gygli et al. 2014]



My friends and I rode on a train.



My friends and I talked with the Pooh mascot.

walked up to the counter in the rate. I gave my rules to the back a tokuch my tea. I wente more entered. We bread and I walked out of the rate t from my car outside. I walked into the real, My dend and I walked around the mail. Heaked at me phone while standing in my bitches. Fured the rice context. I added the shapped vegetables to the colding pet. Indirect the inspections in the pet. placed the cooking pot onto the colline table. I corried my most two the filling rooms, I watched talesisten while eating my ment. I washed the debte in the sink.

at the table and are a need together. I walked down the street with my blood, I walked through th the the acces with my friend. I walked through the perking garage. I drove the car. I walked into the rafe. I put my things down on the table. Hooked down at my log top. I paid for items at the registers sat at a table with my friend and bushed at notes Wyfried and total the table and talent, I walked through the store with my friend, I drove the car. I period the car. I walked into the real. My Intent and I walked around the mail. I washed the claber Iffiled the pet with water from the sink and place. ton the manter. Ich opport to onlors with a lettle stimed the ingredient into the cooking pot. added some food to my bow livith the chearticle. eathed the dishes in the sink.

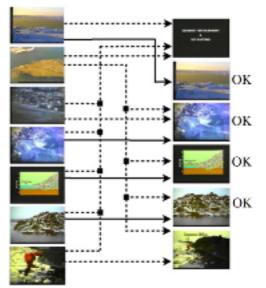
wated in line with my friend. My friend and I sat

Video → text [Yeung et al. 2014]

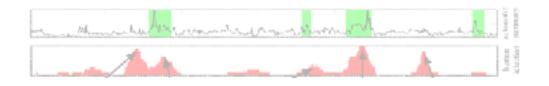
# What makes a good evaluation for video summarization?



A/B test



Bipartite matching [Avila et al. 2011]



Time overlap [Gygli et al. 2014]





My friends and I talked with the Pook mascot.

softed up to the counter in the rate. Lyce manufact in the hards thank may be I worken more unique, the formal and two literature. I worken manufacture of the hard soft would all that the late of the entry or outside. I walked into the mail. My bised and I walked account the mail. Hosterde manufacture is labeled to the counter. I added the chapped segments to the context. I added the chapped segments in the post, I place this cooking pot into the colling but onto the colling table. I confed my mail into the like proon. I wanthed to be soften as while walking say mail. I walked the debelor the into.

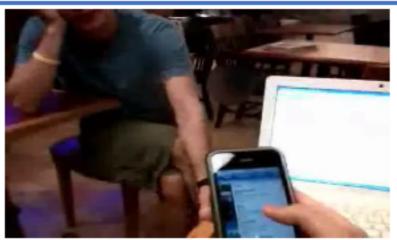
th the

I worted in the with my finest. My finest and that at the table and attrial month (pipellor. I washed down the direct with my friend. I wished through the store with my friend. I wished through the store with my friend. I wished through the cash. I put my things down on the table. I looked down at well up the lips that the table through the store with my friend and better thin the vittle my fire of the latest and the store with my friend, indeed the store with my friend, indeed the cash. I washed the store with my friend, indeed the click. I washed the store with my friend, indeed the click manner of the cash of the control of t

Video → text [Yeung et al. 2014]

# Captions per video shot

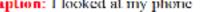
# Dense concepts



Caption: I looked at my phone

#### Dense Tags:

Face Computer Men Phone Hands Chair Room Desk Hall



Caption: I walked around my bedroom

Caption: I waited in line with my friend

#### Dense Tags:

Lady Food Men Drink Hands Hat Computer Market Building Desk



Caption: I drove the car in traffic

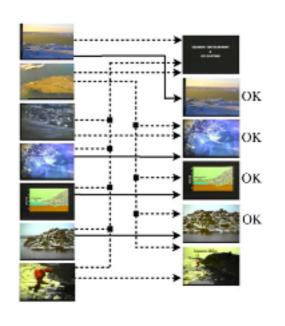
#### Dense Tags:

Chair Computer Room Desk Office

#### Dense Tags:

Sky Street Building Hands Car Tree Window

# What makes a good evaluation for video summarization?



Bipartite matching [Avila et al. 2011]

Bipartite matching of concept vectors



[Lady, Man, Phone, Cab, Street, Building, Restaurants, ...]

## This talk

**DPP** 

SeqDPP

**Variations** 

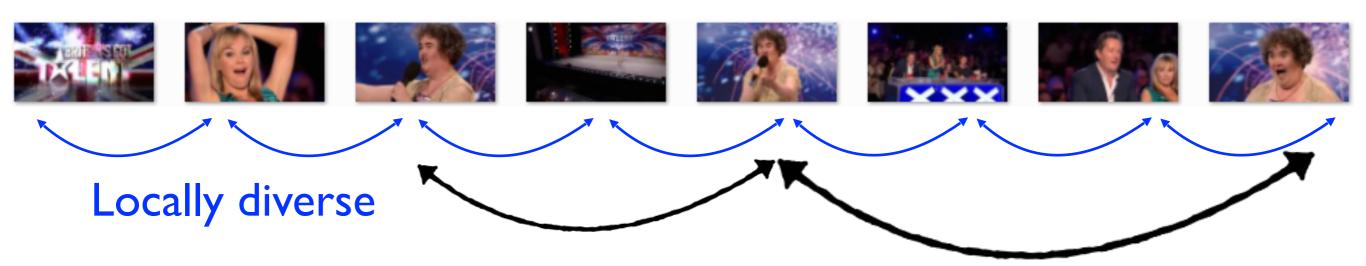
Lessons Learned

## Improving seqDPP

- I. Reinforcing seqDPP
- 2. Large-margin seqDPP



# How local is the local diversity?



Globally not as diverse as locally

Adaptively infer "locality" on the fly

The locality is hidden → Infer it by a latent variable

Direct MLE training incurs an involved EM algorithm

Instead, learn by reinforcement learning

# How local is the local diversity?

How Local is the Local Diversity? Reinforcing Sequential Determinantal Point Processes with Dynamic Ground Sets for Supervised Video Summarization

Yandong Li<sup>1</sup>0000000320051334, Liqiang Wang<sup>1</sup>0000000212654656, Tianbao Yang<sup>2</sup>0000000278585438, and Boqing Gong<sup>3</sup>0000000339155977

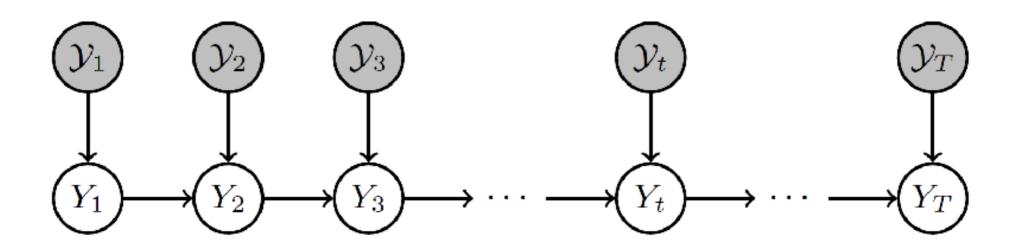


[ECCV 2018b]

Adaptively infer "locality" on the fly Learn by reinforcement learning

- → Avoiding exposure bias
- → Optimizing for the evaluation metrics, vs. surrogate loss

# How to control the summary length?



SeqDPPs automatically determine summary lengths Most competing methods need user-supplied lengths How to make the summary lengths controllable in seqDPPs?

## Generalized DPPs

#### Improving Sequential Determinantal Point Processes for Supervised Video Summarization

Aidean Sharghi $^{1[0000000320051334]}$ , Ali Borji $^{1}$ , Chengtao Li $^{2[0000000323462753]}$ , Tianbao Yang $^{3[0000000278585438]}$ , and Boqing Gong $^{4[0000000339155977]}$ 



[ECCV 2018a]

### Disentangling size and content in subset selection

$$P_L(Y;L) = \frac{1}{\det(L+I)} \sum_{J \subseteq \mathcal{Y}} P_E(Y;J) \prod_{n \in J} \lambda_n,$$
$$\propto \sum_{k=0}^{N} \sum_{J \subseteq \mathcal{Y}, |J|=k} P_E(Y;J) \prod_{n \in J} \lambda_n$$

# Large-margin learning of seqDPPs

### Improving Sequential Determinantal Point Processes for Supervised Video Summarization

Aidean Sharghi $^{1}$ [0000000320051334], Ali Borji $^{1}$ , Chengtao Li $^{2}$ [0000000323462753], Tianbao Yang $^{3}$ [0000000278585438], and Boqing Gong $^{4}$ [0000000339155977]



[ECCV 2018a]

Define the margins by using evaluation metrics

## This talk

DPP

SeqDPP

**Variations** 

Lessons Learned

# What makes a good video summarizer?

Video summarization: a subjective process

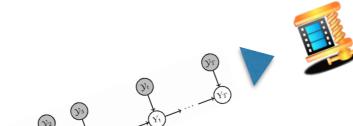


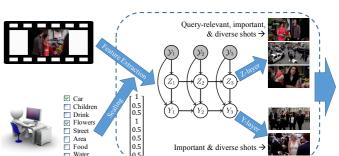




Prior: unsupervised SeqDPP: average user SH-DPP: "the" user

[Wolf 1996, Vasconcelos and Lippman 1998, Aner and Kender 2002, Pal and Jojic 2005, Kang et al. 2006, Pritch et al. 2007, Jiang et al. 2009, Lee and Kwon 2012, Khosla et al. 2013, Kim et al. 2014, Song et al. 2015, Lee and Grauman 2015, ... ]





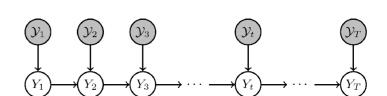


(a) Input: Video & Ouery (b) Algorithm: Sequential & Hierarchical Determinantal Point Process (SH-DPP) (c) Output:

## SeqDPPs: models

### Sequential DPPs (seqDPPs)

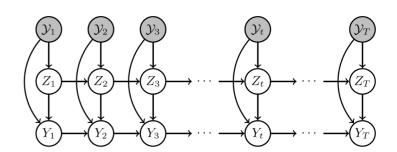
Diverse sequential subset selection



### Hierarchical seqDPPs (SH-DPPs)

Multi-granularity subset selection

Query-focused, user-tailored



## Generalized seqDPPs (seqGDPP)

Disentangling size & content

User-controllable summary lengths

## SeqDPPs: algorithms

Maximum likelihood estimation (MLE)

## Reinforcement learning

### Large-margin learning

Adaptively infers the "locality"

Avoids exposure bias

Accounts for evaluation metrics

#### Diverse Sequential Subset Selection for Supervised Video Summarization

#### Boqing Gong\*

Department of Computer Science University of Southern California Los Angeles, CA 90089 boqinggo@usc.edu

#### Kristen Grauman

Department of Computer Science University of Texas at Austin Austin, TX 78701 grauman@cs.utoxas.edu

#### Wei-Lun Chao\*

Department of Computer Science University of Southern California Los Angeles, CA 90089 weilunc@usc.edu

#### Fei Sha

Department of Computer Science University of Southern California Los Angeles, CA 90089 feisha@usc.edu

#### How Local is the Local Diversity? Reinforcing Sequential Determinantal Point Processes with Dynamic Ground Sets for Supervised Video Summarization

Yandong Li<sup>1</sup>0000000320051334, Liqiang Wang<sup>1</sup>0000000212654656, Tianbao Yang<sup>2</sup>0000000278585438, and Boqing Gong<sup>3</sup>0000000339155977

#### Improving Sequential Determinantal Point Processes for Supervised Video Summarization

 $\label{eq:Aidean Sharghi} Aidean Sharghi^{1[0000000320051334]}, Ali Borji^{1}, Chengtao \ Li^{2[0000000323462753]}, Tianbao \ Yang^{3[0000000278585438]}, and Boqing Gong^{4[0000000339155977]}$ 

## **SeqDPP**

Code: <a href="https://github.com/pujols/Video-summarization">https://github.com/pujols/Video-summarization</a>

#### Large-Margin Determinantal Point Processes

Wei-Lun Chao\*

U. of Southern California Los Angeles, CA 90089 Boqing Gong\*

U. of Southern California Los Angeles, CA 90089 Kristen Grauman

U. of Texas at Austin Austin, TX 78701 Fei Sha

U. of Southern California Los Angeles, CA 90089

#### Diverse Sequential Subset Selection for Supervised Video Summarization

#### Boqing Gong\*

Department of Computer Science University of Southern California Los Angeles, CA 90089 boqinggo@usc.edu

#### Kristen Grauman

Department of Computer Science University of Texas at Austin Austin, TX 78701 grauman@cs.utexas.edu

#### Wei-Lun Chao\*

Department of Computer Science University of Southern California Los Angeles, CA 90089 weilunc@usc.edu

#### Fei Sha

Department of Computer Science University of Southern California Los Angeles, CA 90089 feisha@usc.edu [UAI 2015]

[NIPS 2014]

### **SH-DPP**

Code & data: <a href="https://www.aidean-sharghi.com/cvpr2017">https://www.aidean-sharghi.com/cvpr2017</a>

#### Query-Focused Extractive Video Summarization

Aidean Sharghi, Boqing Gong, Mubarak Shah

Center for Research in Computer Vision, University of Central Florida aidean.sharghi@knights.ucf.edu,bgong@crcv.ucf.edu,shah@crcv.ucf.edu

[ECCV 2016]

Query-Focused Video Summarization: Dataset, Evaluation, and A Memory Network Based Approach

[CVPR 2017]

Aidean Sharghi<sup>†</sup>, Jacob Laurel<sup>‡</sup>, and Boqing Gong<sup>†</sup>

## **Seq-GDPP** & large-margin training

Data: <a href="https://www.aidean-sharghi.com/eccv2018">https://www.aidean-sharghi.com/eccv2018</a>

### Improving Sequential Determinantal Point Processes for Supervised Video Summarization

[ECCV 2018a]

Aidean Sharghi $^{1[000000320051334]}$ , Ali Borji $^{1}$ , Chengtao Li $^{2[0000000323462753]}$ , Tianbao Yang $^{3[0000000278585438]}$ , and Boqing Gong $^{4[0000000339155977]}$ 

## Reinforcing SeqDPP

How Local is the Local Diversity? Reinforcing Sequential Determinantal Point Processes with Dynamic Ground Sets for Supervised Video Summarization

[ECCV 2018b]

Yandong Li<sup>1</sup>0000000320051334, Liqiang Wang<sup>1</sup>0000000212654656, Tianbao Yang<sup>2</sup>0000000278585438, and Boqing Gong<sup>3</sup>0000000339155977

## Acknowledgements

**U. Southern California** 

Fei Sha, Wei-Lun Chao

U. Texas at Austin: Kristen Grauman

**U.** Central Florida

Aidean Sharghi, Yandong Li, Liqiang Wang

MIT: Chengtao Li

U. Iowa: Tianbao Yang