Motivation

➢ Video Summarization attempts to provide a highlight of the most critical and important events in a video.
➢ Different summarization models capture different aspects of a video. Types of characteristics captured by summarization models:
   ➢ Diversity
   ➢ Representation
   ➢ Importance
   ➢ Coverage
➢ What comprises of the key aspects of a video, depends largely on the domain.
➢ Different domains, require different summarization models.
➢ Which summarization model to use for a domain?

Our Contributions

➔ Study the role and characteristics of various summarization models.
➔ Investigate several diversity, coverage and representation models, and demonstrate how different models are applicable in different kinds of summarization tasks.
➔ Empirically and quantitatively prove the behaviour of the functions on various videos across domains.
➔ Discuss computational scalability of the optimization algorithms and the usage of computational tricks, such as lazy evaluations and memoization.
➔ Explore various video summarization variants and study the details on how to design a video summarization system.

Submodular Summarization Framework

➔ Our framework studies the following variants of Video Summarization:
   ➢ Extractive Summarization
   ➢ Query Summarization
   ➢ Entity Summarization
➔ In all variants, we use Convolutional Neural Networks as feature extractors, before performing submodular optimization to generate the summary.

The following table illustrates the various submodular functions studied by us, as a part of our framework.

<table>
<thead>
<tr>
<th>Name</th>
<th>Facility Location</th>
<th>Saturated Coverage</th>
<th>Weighted Set Cover</th>
<th>Feature Based</th>
<th>Prob. Set Cover</th>
<th>Temporal Sum</th>
<th>Displacement Min</th>
<th>Displacement Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>p_i(x)</td>
<td>( \sum_i p_i(x_i) ), ( a_i(x_i) )</td>
<td>( \sum_i a_i(x_i) )</td>
<td>( \sum_i \mathcal{V}(x_i) )</td>
<td>( \Omega(x) )</td>
<td>( \Omega(x) )</td>
<td>( \mathcal{S}(x) )</td>
<td>( \mathcal{S}(x) )</td>
<td>( \mathcal{S}(x) )</td>
</tr>
</tbody>
</table>

Experiments and Results

➔ In the illustrated results, we show the following:
   ➢ Extractive summarization on TV shows.
   ➢ Entity summarization on TV shows.
   ➢ Query summarization for a query "skyscraper".
➢ In each case, we compare Representation, Diversity and Coverage models.

Conclusions

➔ We present a unified picture of multi-faceted video summarization for extractive, query based and entity based summarization.
➔ Take a closer look at different summarization models and argue the benefits of these models in different domains by comparing qualitative and quantitative results.
➔ Implementation tricks, like memoization, can drastically improve the summary generation time, as compared to computational gains, such as usage of GPUs.