Video Summarization summary

C. Aslanidou, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr

Version 1.2

Date: September 2021

Artificial Intelligence & Information Analysis Lab

Contents

- Video Summarization
  - Video Summarization usage
  - Video Summarization Aim
  - Video Summarization purposes
  - Video Summarization Techniques
  - Video Summarization Methods
  - Video Summarization Models
  - Video Summarization Applications
  - Automatic video summarization applications

Contents

- Video Summarization with Neural Networks (NN)
- Deep Architectures for Video Summarization
- Generative Adversarial Networks (GANs)
- Techniques for video summarization with GANs
- GAN Models for video summarization
- Refences

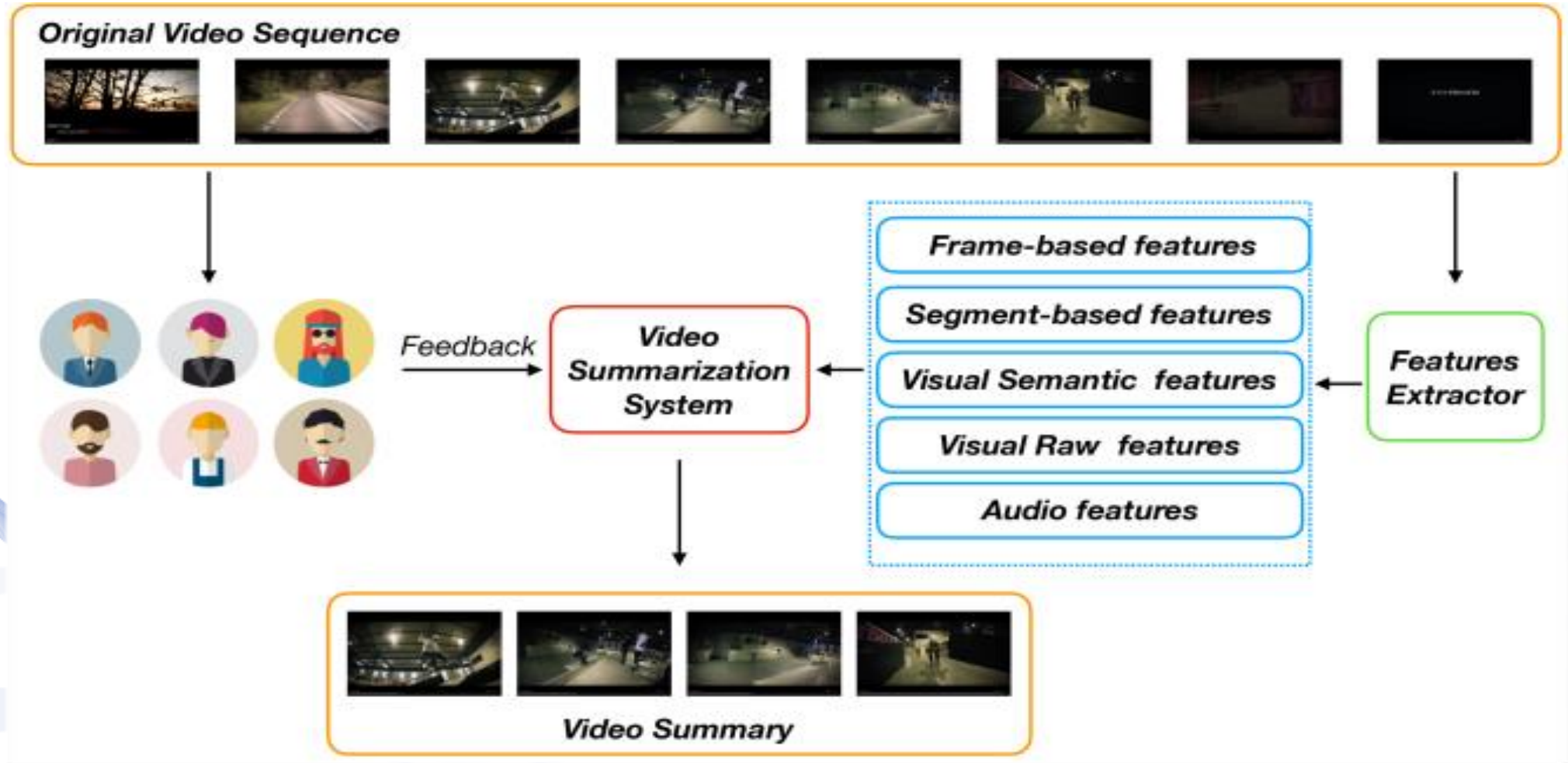
Video Summarization

Video summarization has recently become an active area of research due to a myriad of possible applications, such as in the entertainment industry, sports, and surveillance.

Users would ideally like to browse through videos quickly to get an idea of the content.

This will enable faster browsing of large video datasets, as well as better grouping and access to these videos.

Video Summarization



(Image from Heartbeat Fritz AI)

Video Summarization usage

The video summarization is a brief presentation of a video. It presents some points presented in the video. It does not have to be too long. In fact, it may be some frames of the whole video.

Video summarization is used to generate a short summary of the content of a longer video by selecting and presenting the most informative or interesting materials for potential users.

Video Summarization Aim

The aim of video summarization is to speed up browsing of a large collection of video data, and achieve efficient access and representation of the video content.

By watching the summary, users can make quick decisions on the usefulness of the video.

Video Summarization purposes

- **Summarization of personal videos**

Image from Bing Blogs

Video Summarization purposes

- **Sport highlights**

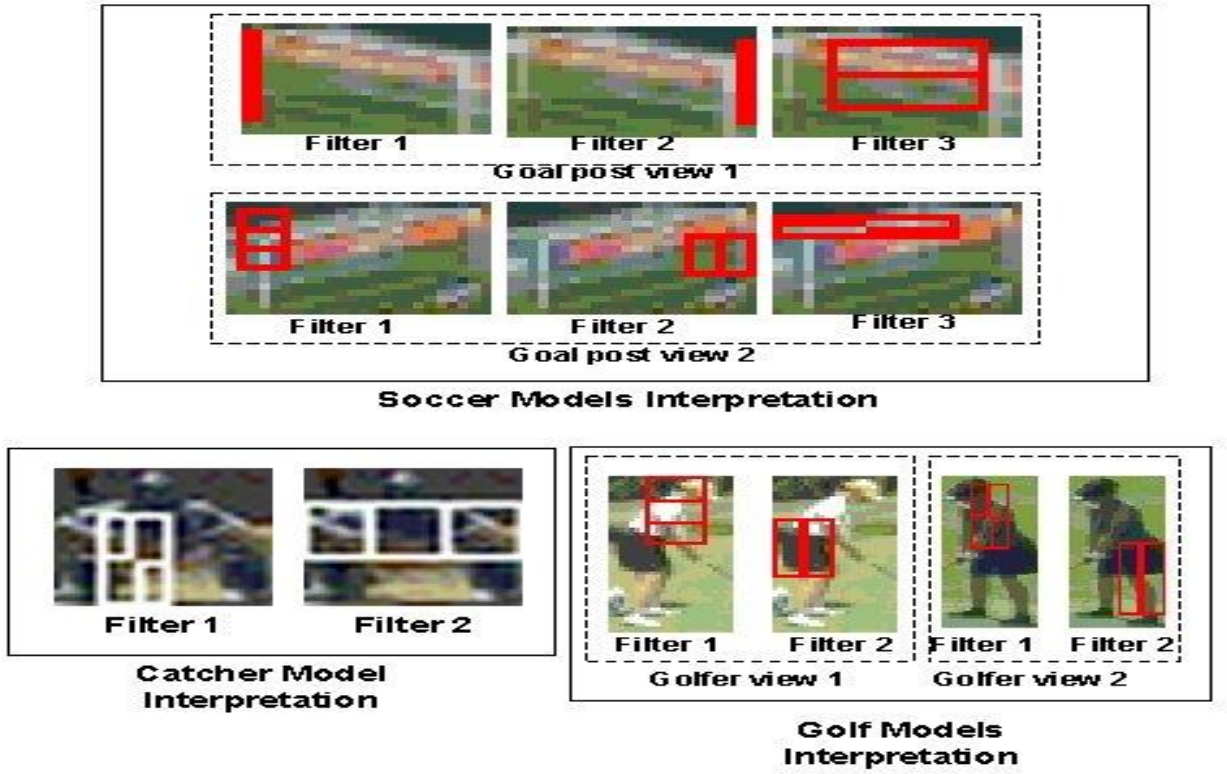


Image from slidetodoc.com

Video Summarization purposes

- **The trailer of automated movies and serial**

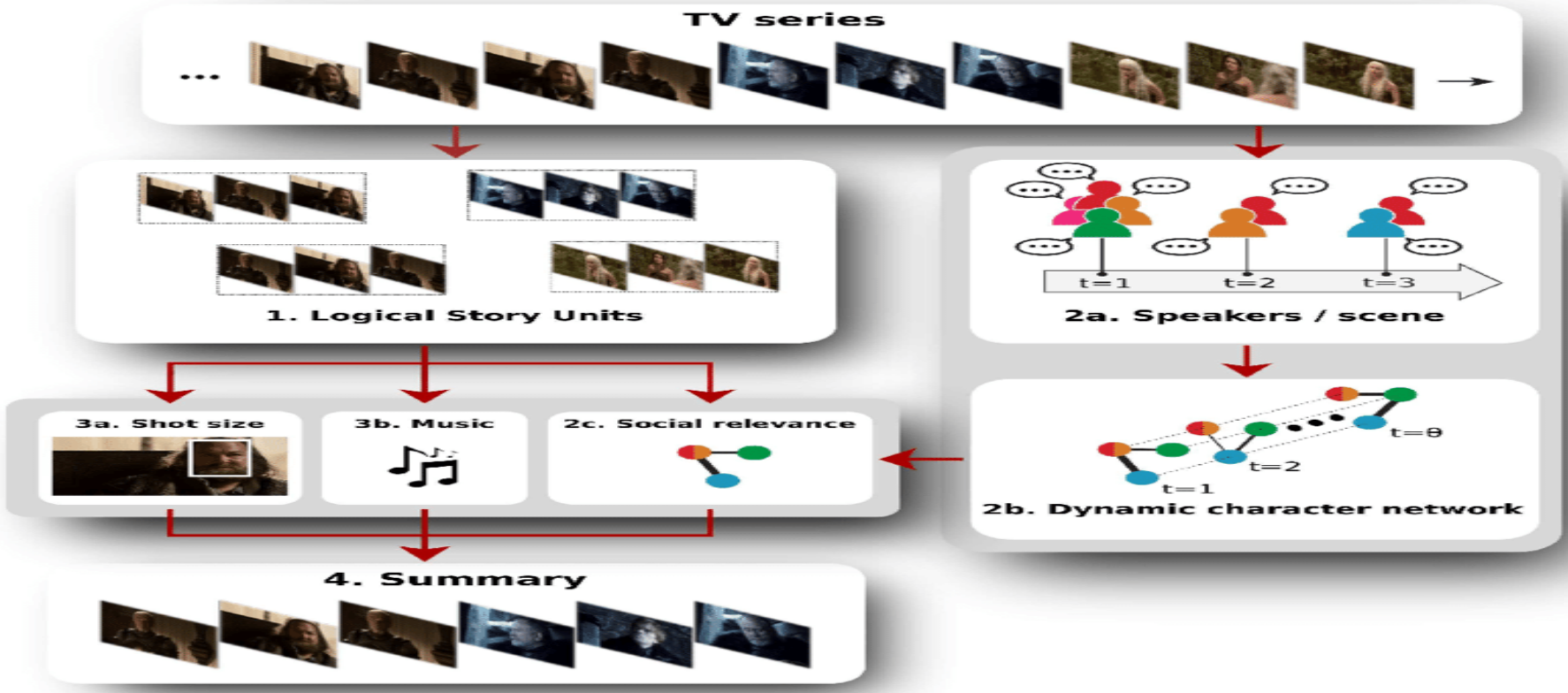
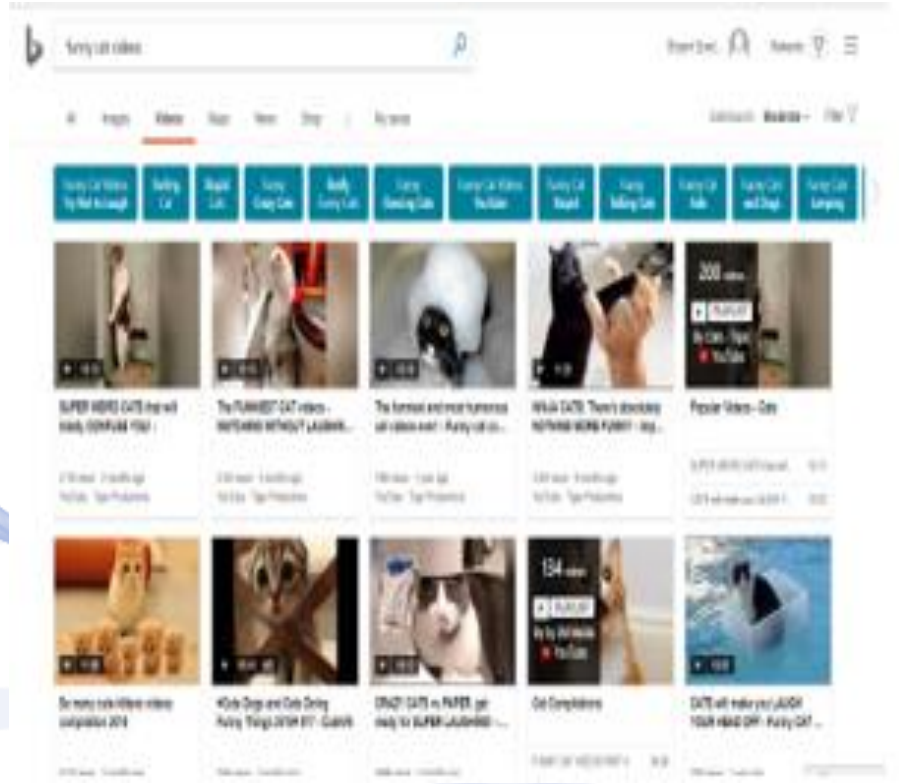


Image from ResearchGate

Video Summarization purposes

- **Video search engines**



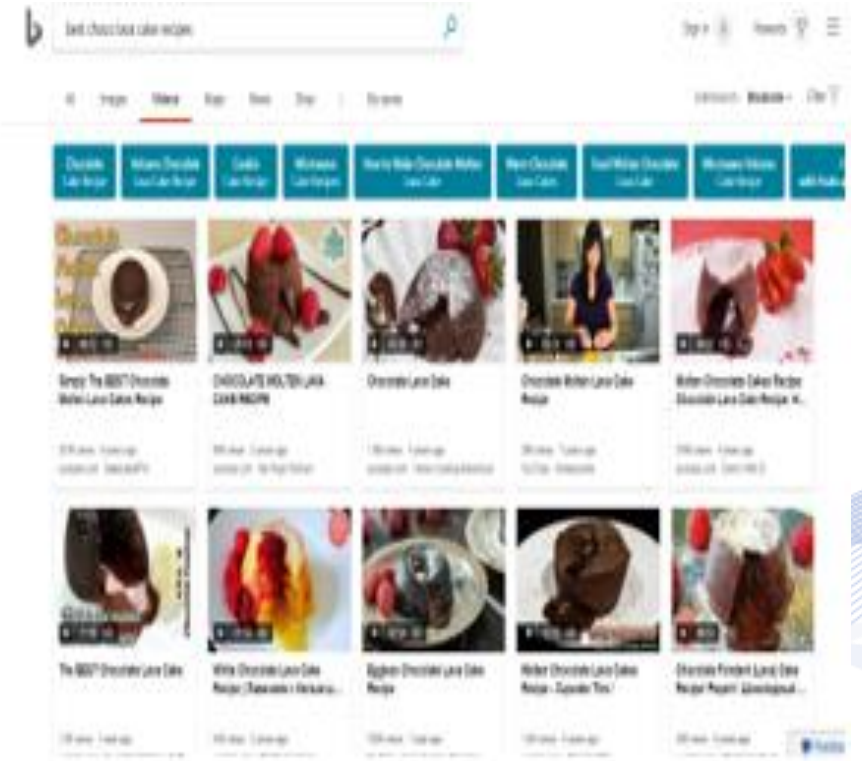


Image from Bing Blogs

Video Summarization Techniques

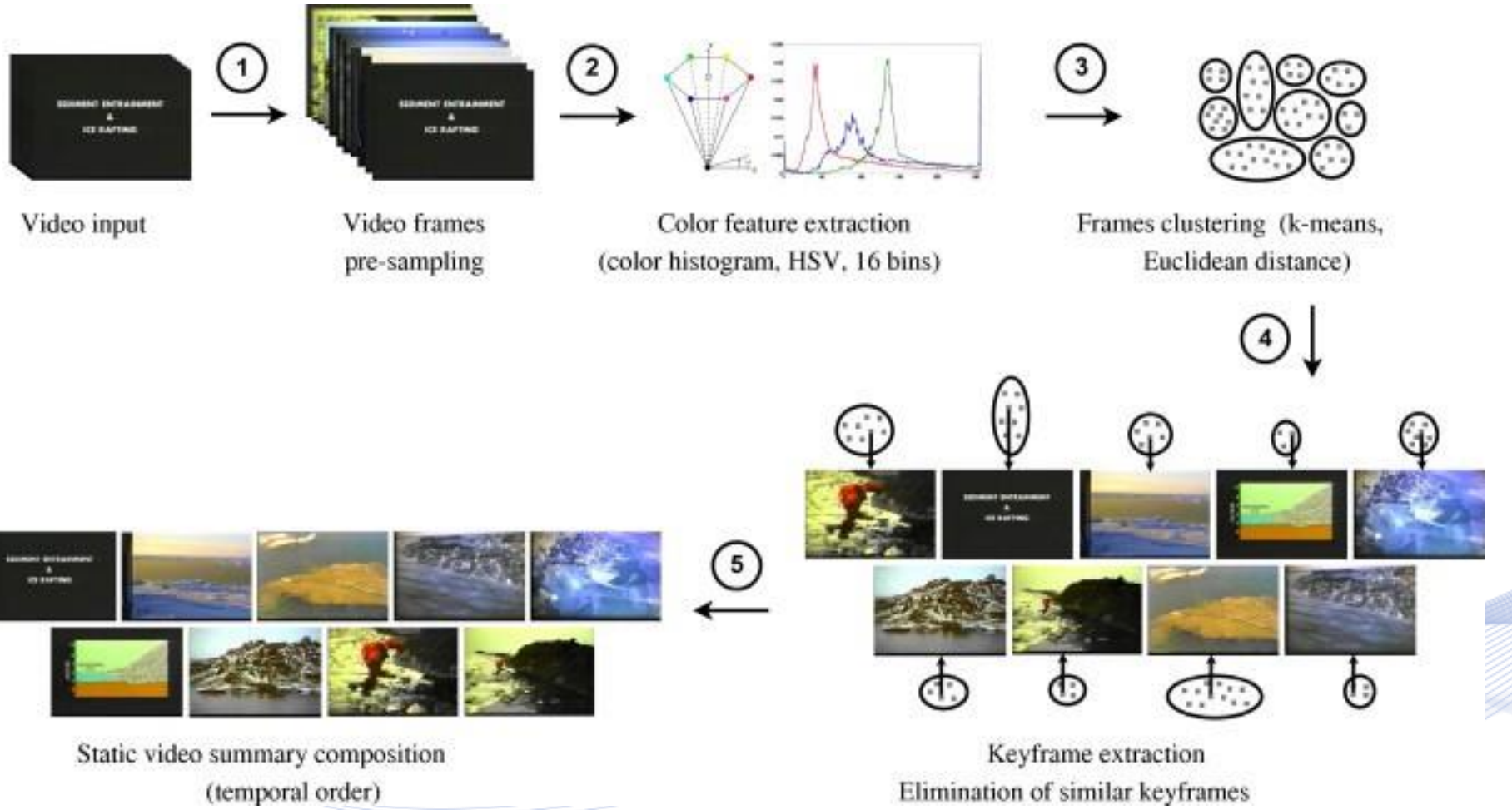


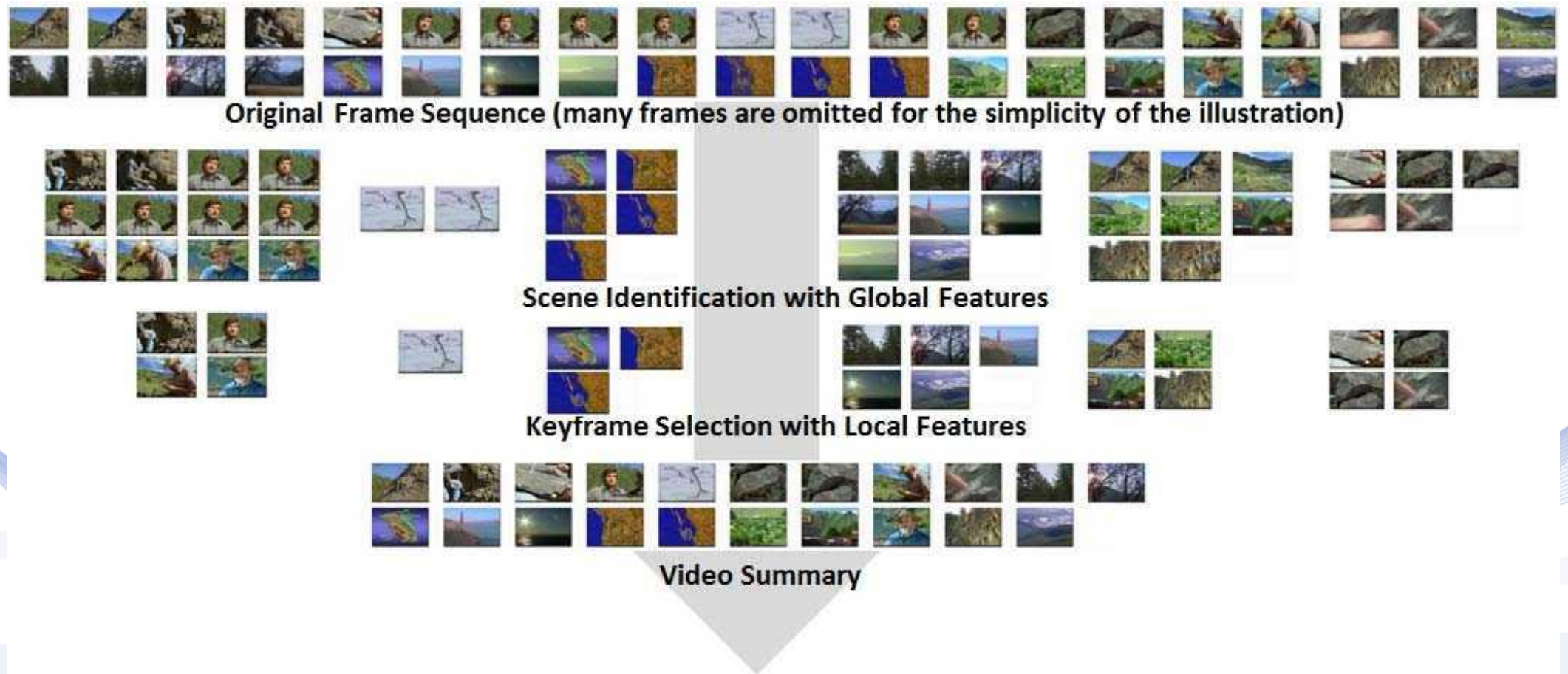
Image from Heartbeat-Fritz AI

Video Summarization Techniques

Multiple techniques/approaches have been developed which have the main concern to precise the video contents and generate a video summary.

These techniques are classified into **four major categories** based on their properties and characteristics. [BUR2020]

Video Summarization Techniques



Video Summarization with Global and Local Features (Image from ResearchGate)

Video Summarization Techniques

Video Summarization with Global and Local Features (Image from ResearchGate)

Video Summarization Techniques

- **Event-based video summarization**

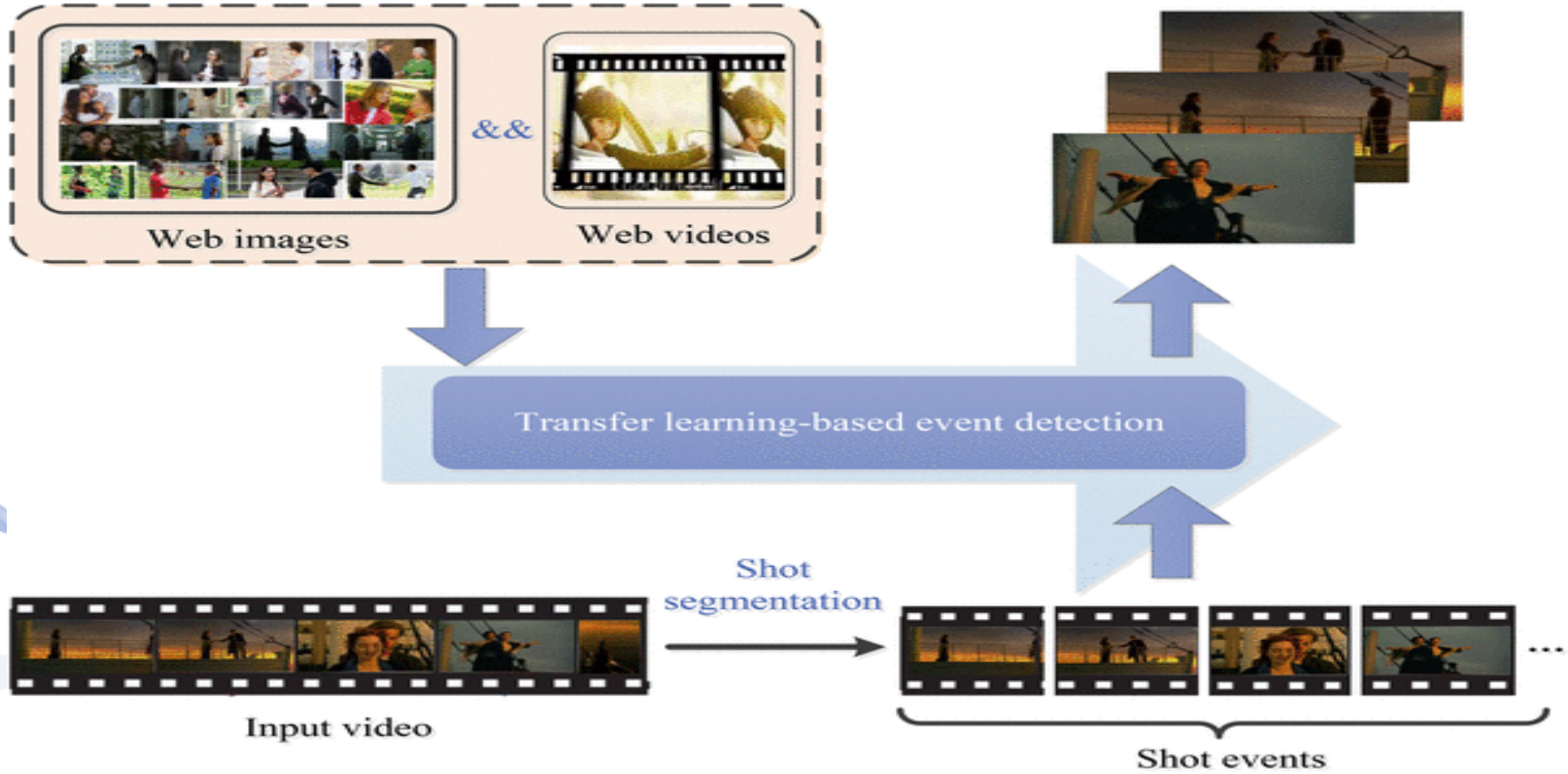


Image from ResearchGate

Video Summarization Techniques

- **Motion Based video Summarization**

Image from ScienceDirect

Video Summarization Techniques

- **Color Based video Summarization**

(a)

(b)

Image from Semantic Scholar

Video Summarization Techniques

- **Object-Based video summarization**

Image from ScienceDirect

Video Summarization Techniques

- **Attention-based video summarization**

Ngo et al. presented the approach for summarization that relies on both the perceptual quality and content balance of the video summary. Therefore, a clustering method is applied to edit/cut the video, and a motion attention model is used to measure the shot's quality and clusters. [BUR2020]

Video Summarization Techniques

- **Attention-based video summarization**

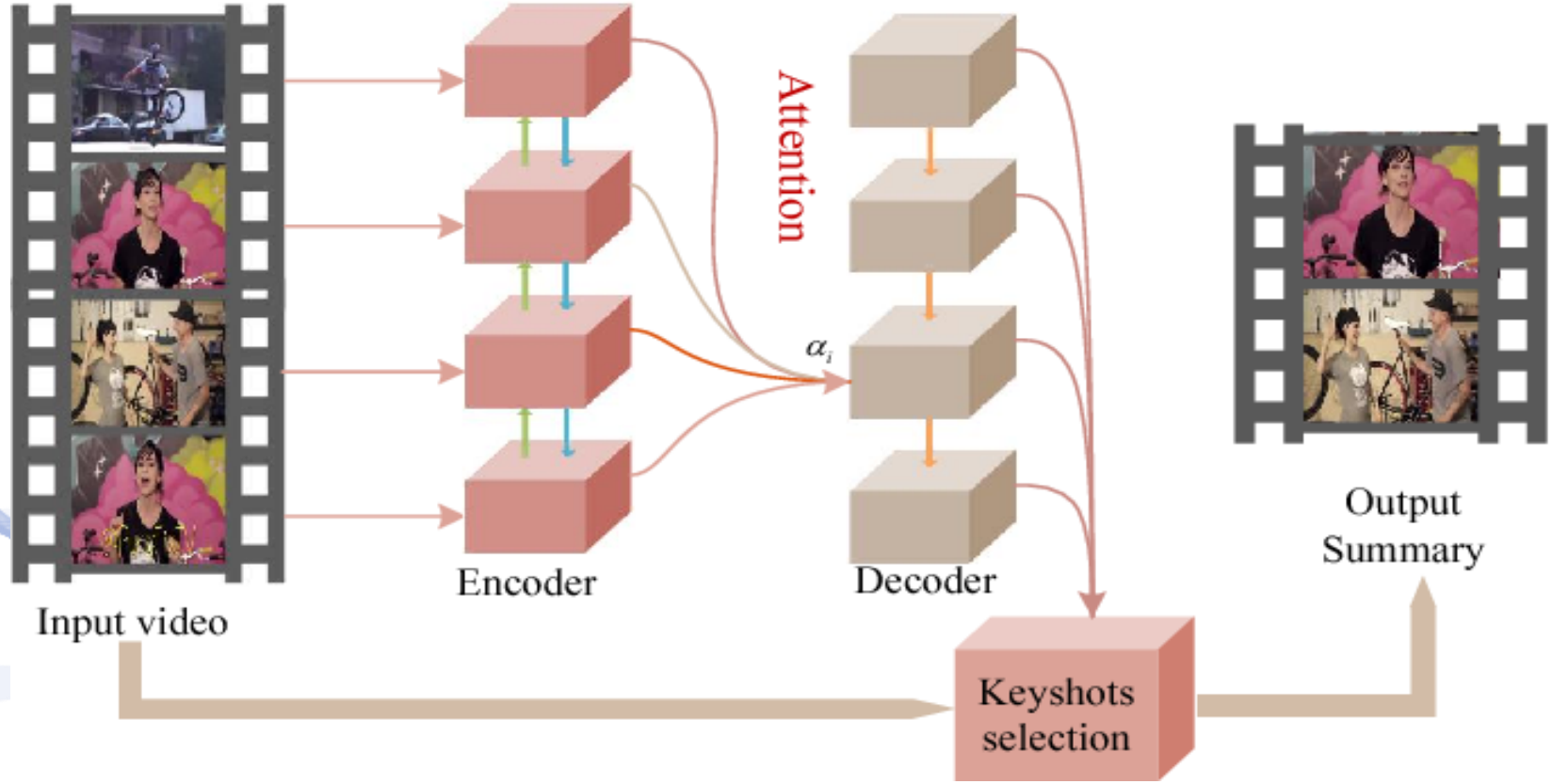


Image from Semantic Scholar

Video Summarization Techniques

- **Clustering-based Video summarization.**

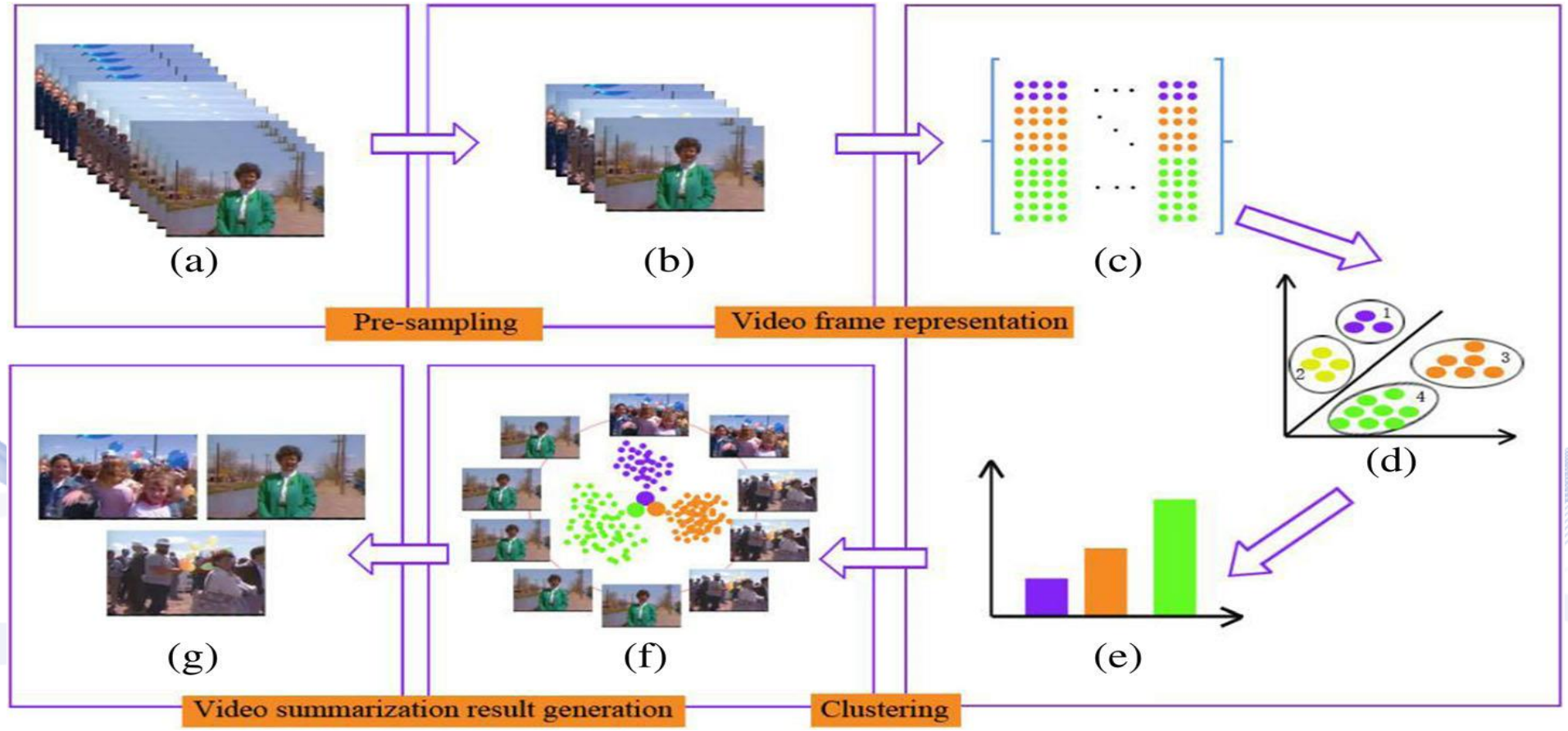


Image from Research Institute for Future Media Computing

Video Summarization Techniques

- **Selection of shot / shot boundaries-based video summarization.**

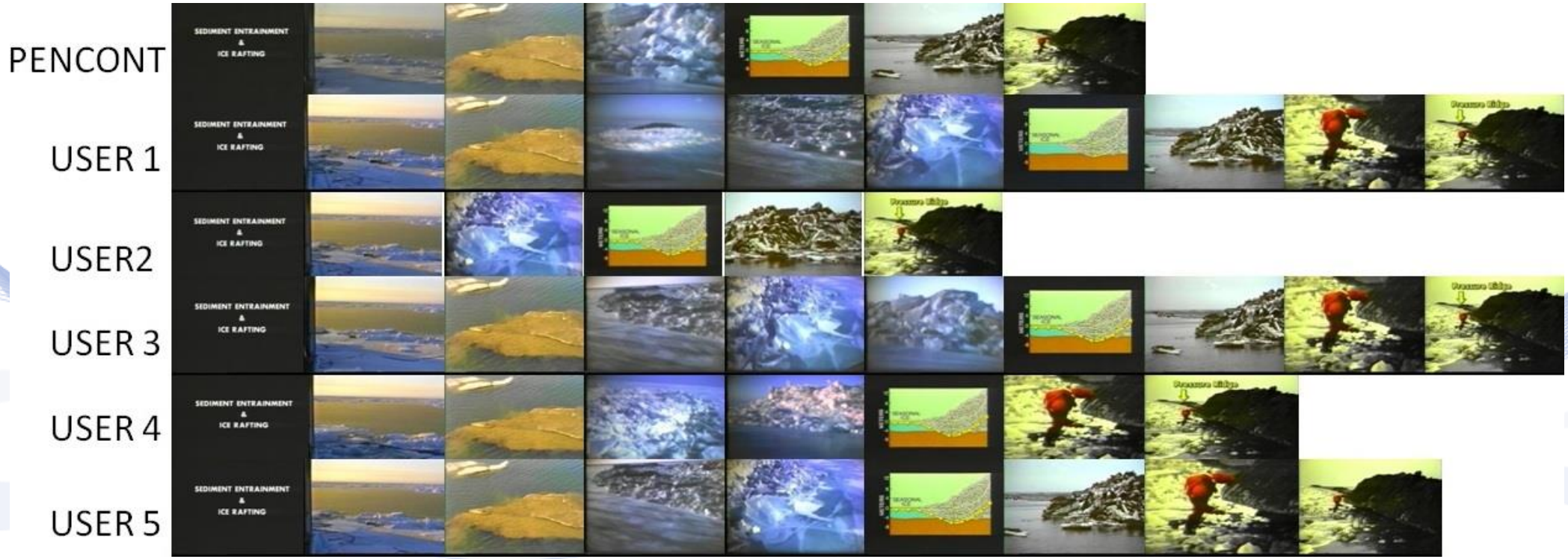


Image from poseidon.csd.auth.gr

Video Summarization Techniques

- **Trajectory-based Video Summarization.**

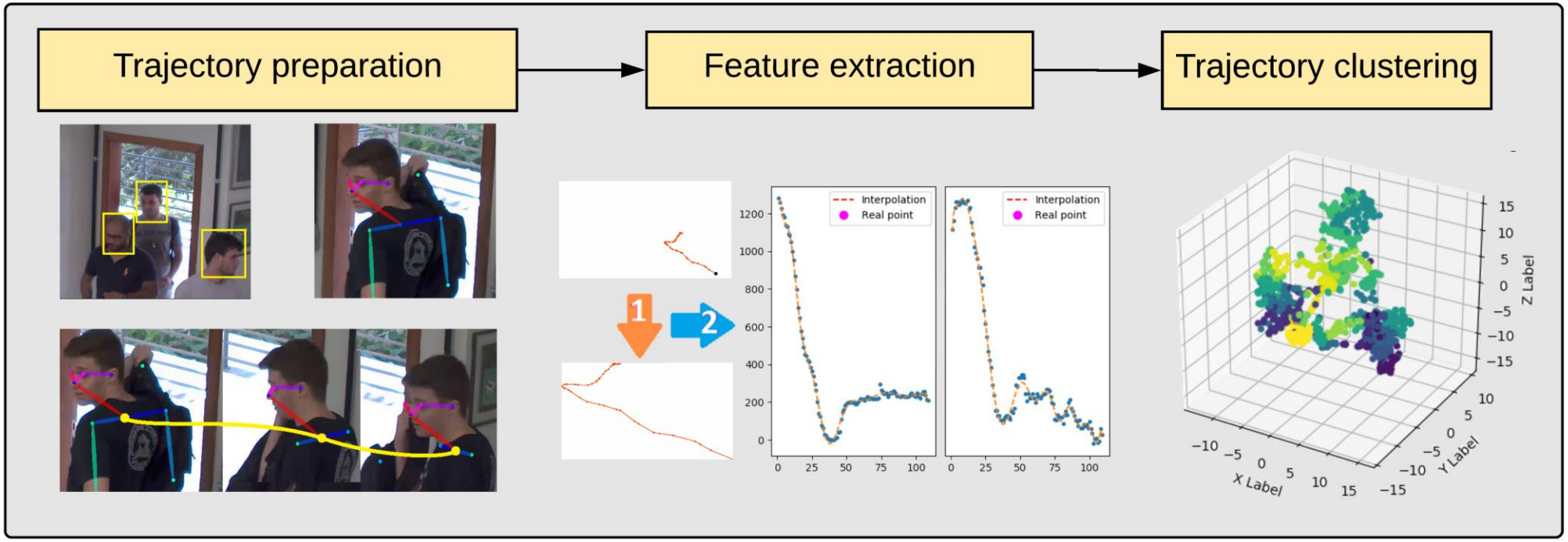


Image from Universidad Católica San Pablo

Video Summarization Techniques

- **Sparse dictionary learning.**

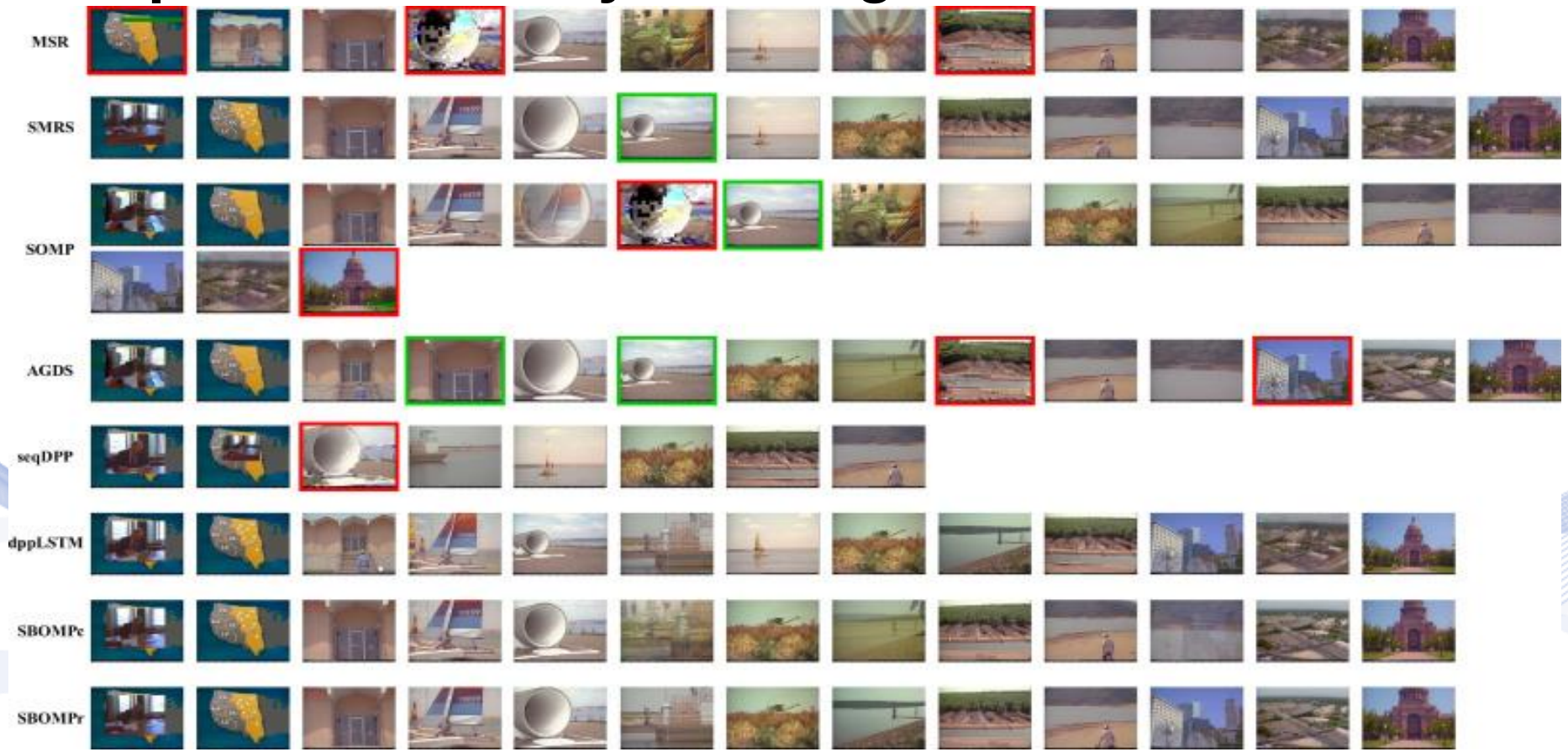


Image from sciencedirect.com

Video Summarization Methods

Within artificial intelligence (AI) and machine learning, there are two basic approaches for Video Summarization:

Supervised learning and Unsupervised learning.

The main difference is one uses labeled data to help predict outcomes, while the other does not. However, there are some nuances between the two approaches, and key areas in which one outperforms the other. [HTT2021] But recently research the **Reinforcement Learning** mechanism also applied to it. [WOR2020]

Video Summarization Methods

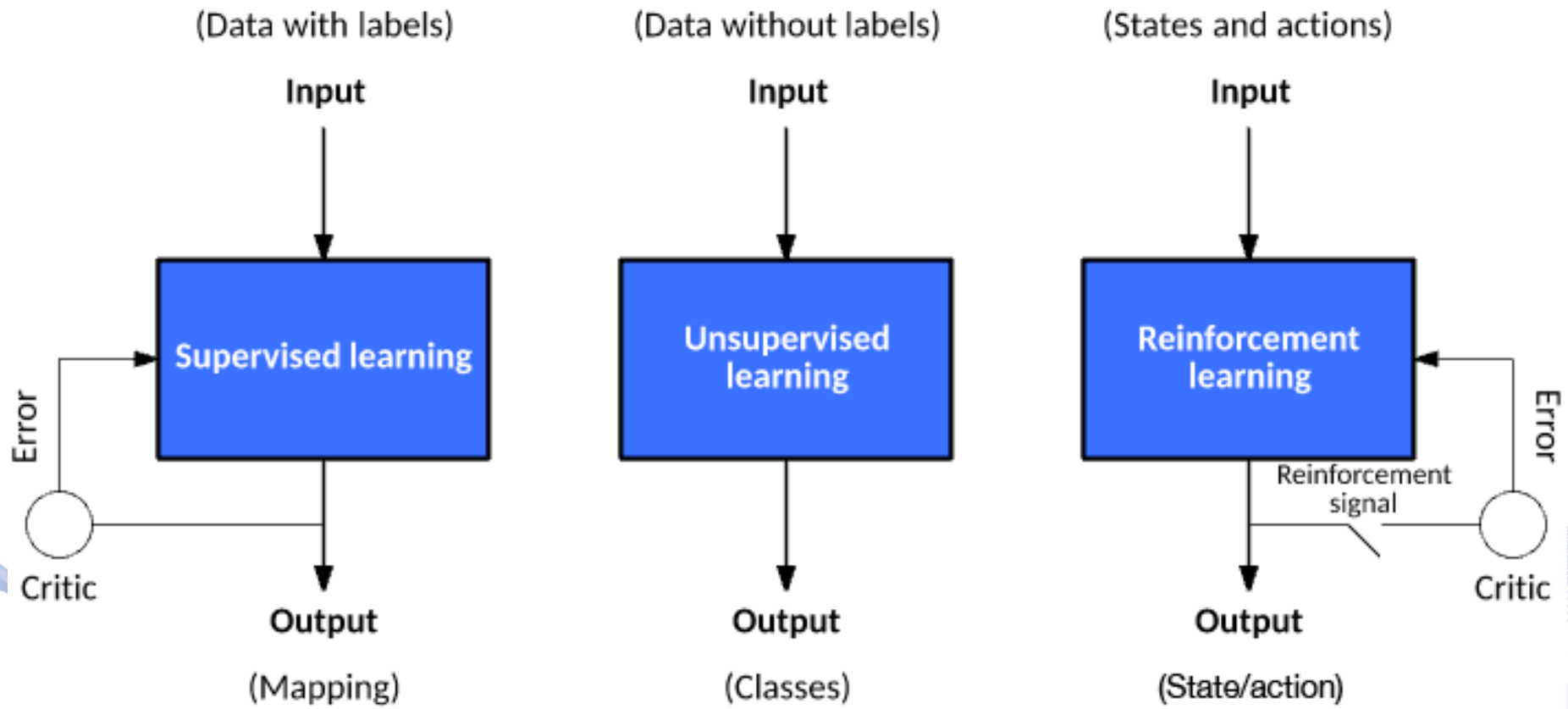


Image from IBM Developer

Video Summarization Methods

- **Supervised learning for Video Summarization**

Classification

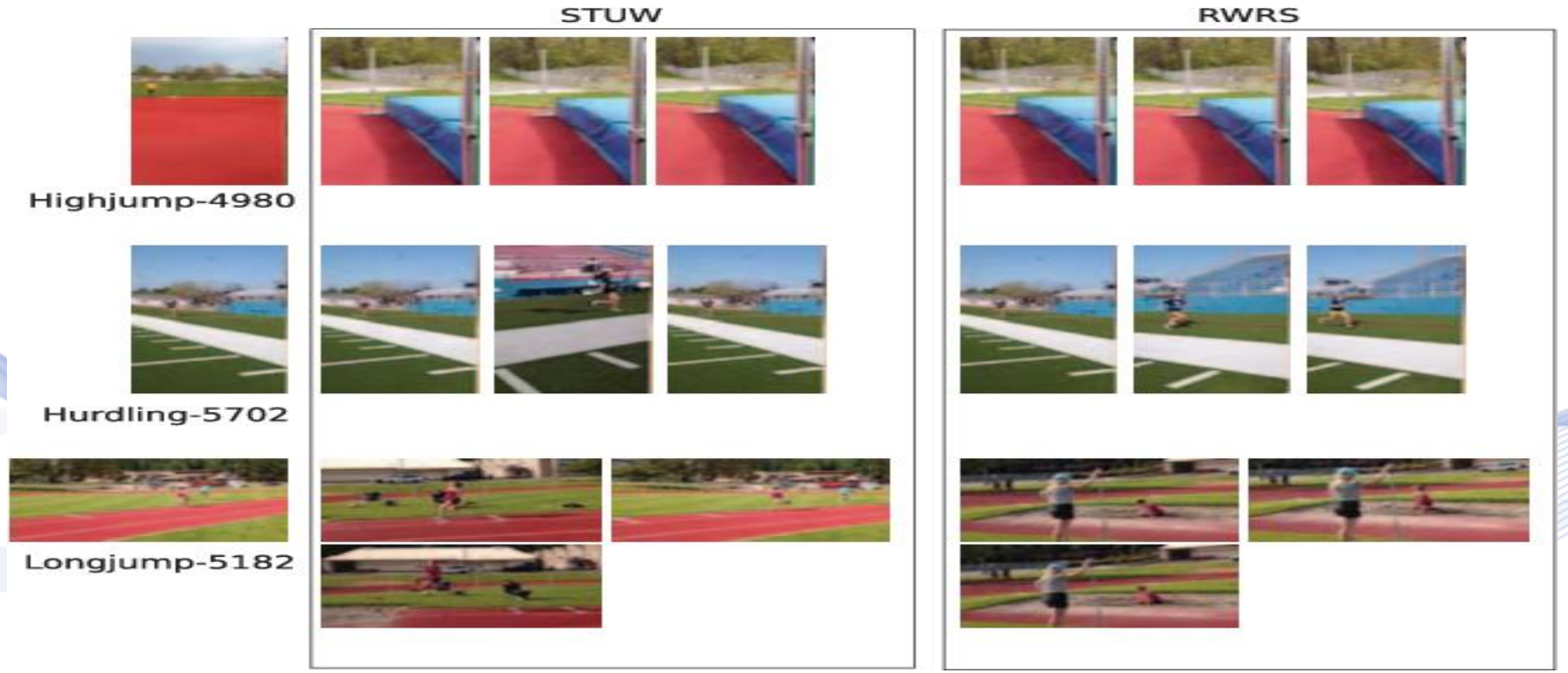


Image from Semantic Scholar

Video Summarization Methods

- **Unsupervised learning for Video Summarization**
- **Clustering**

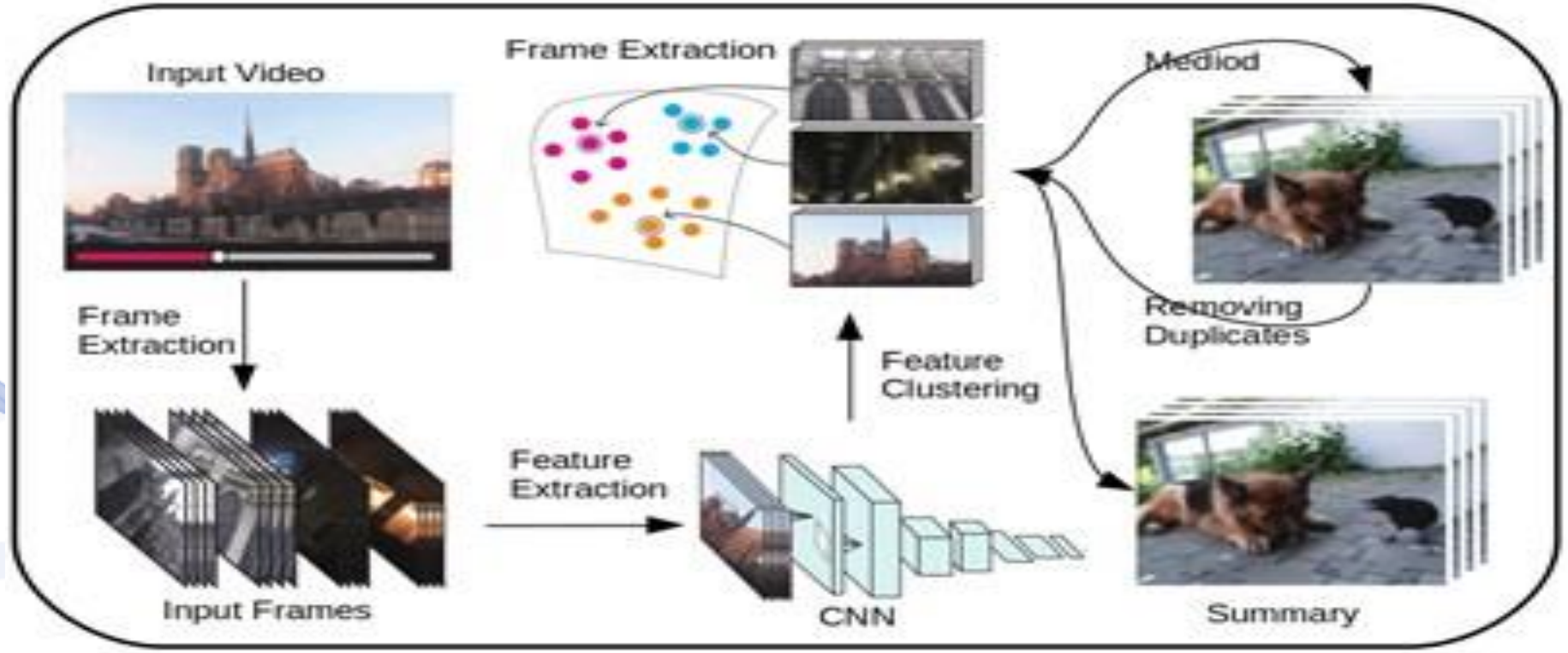


Image from papers.ssrn.com

Video Summarization Methods

- **Reinforcement Learning Methods**

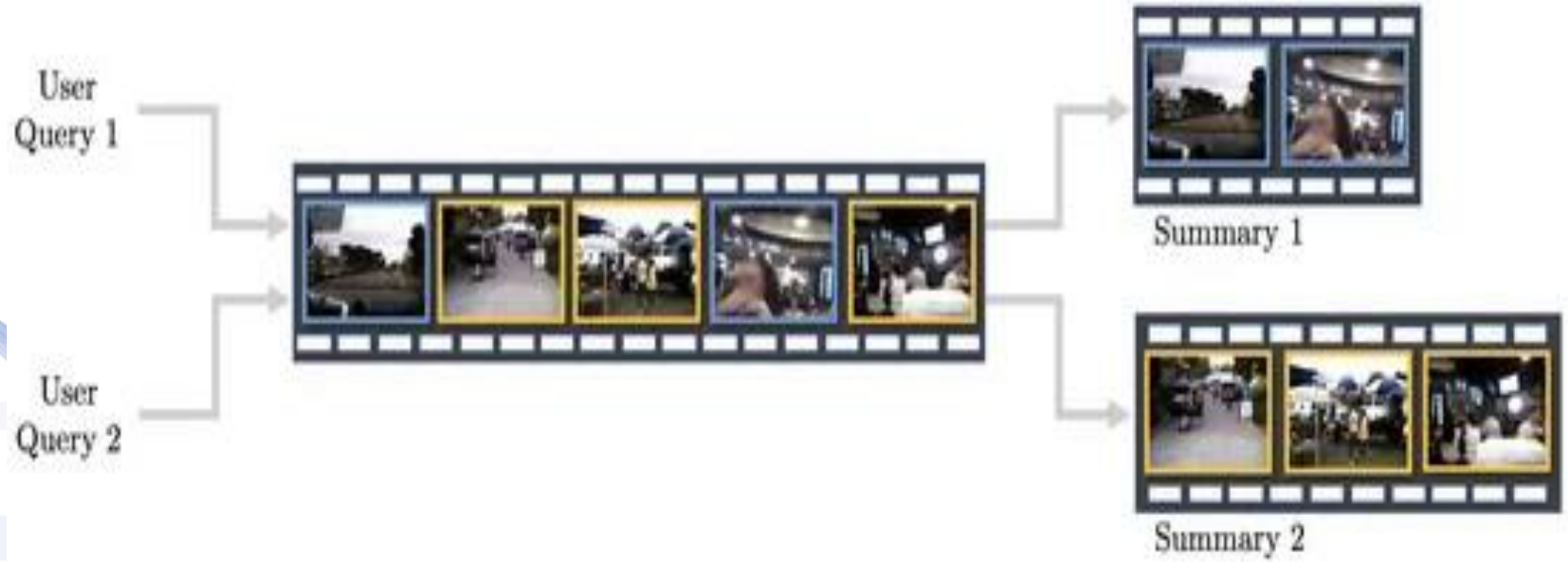


Image from mdpi.com

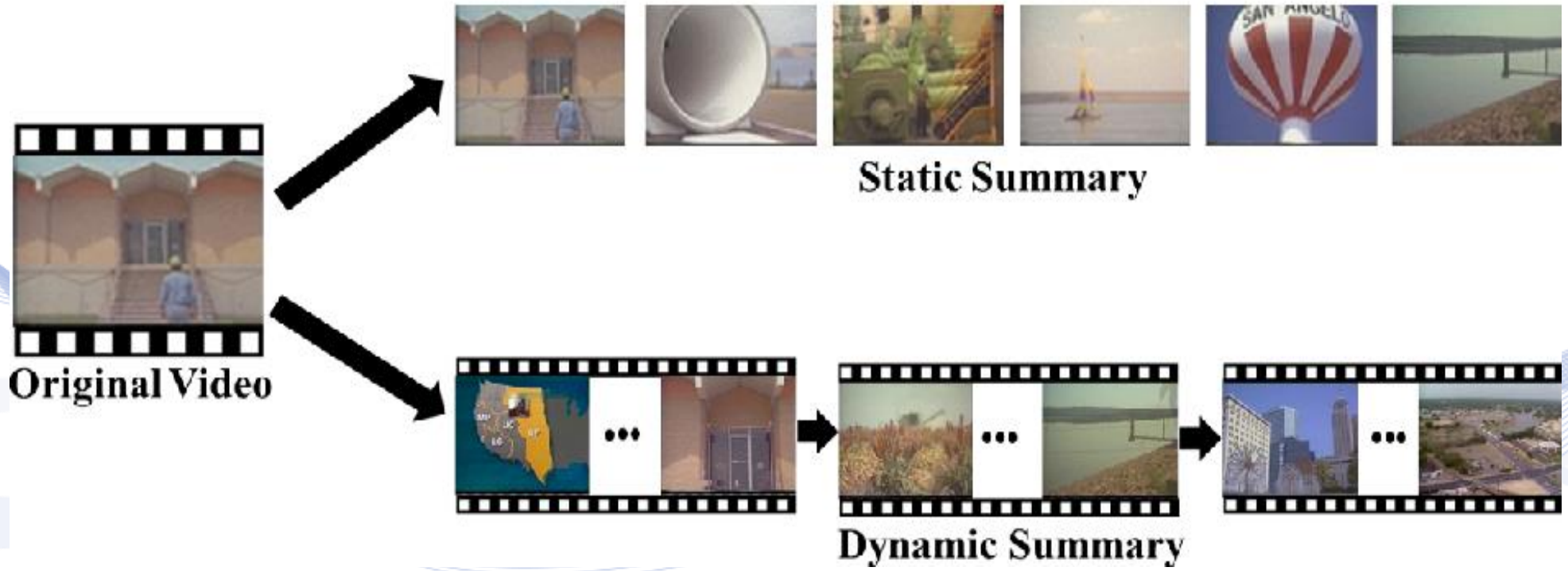
Video Summarization Models

Video summarization can be represented into **two models**:

A static video summary (storyboard) and a dynamic video summarization (skimming).

In one hand, static video summary represents a video sequence in a static imagery form (one or more selected representative frames from the original video, or a synthesized image generated from the selected keyframes). [TRU2007]

Static video summarization and Dynamic video summarization



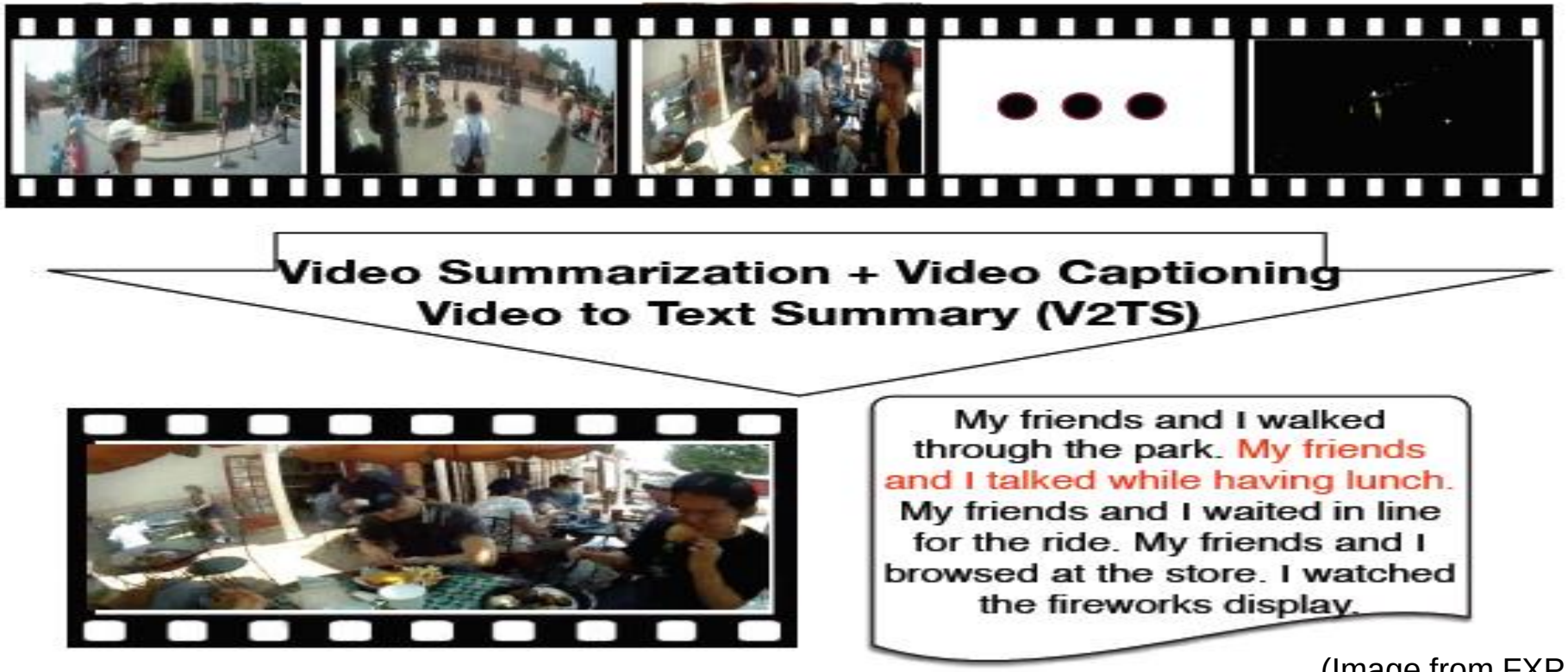
(Image from Semantic Scholar)

Static Video Summarization

There are two kinds of Static video summarization:

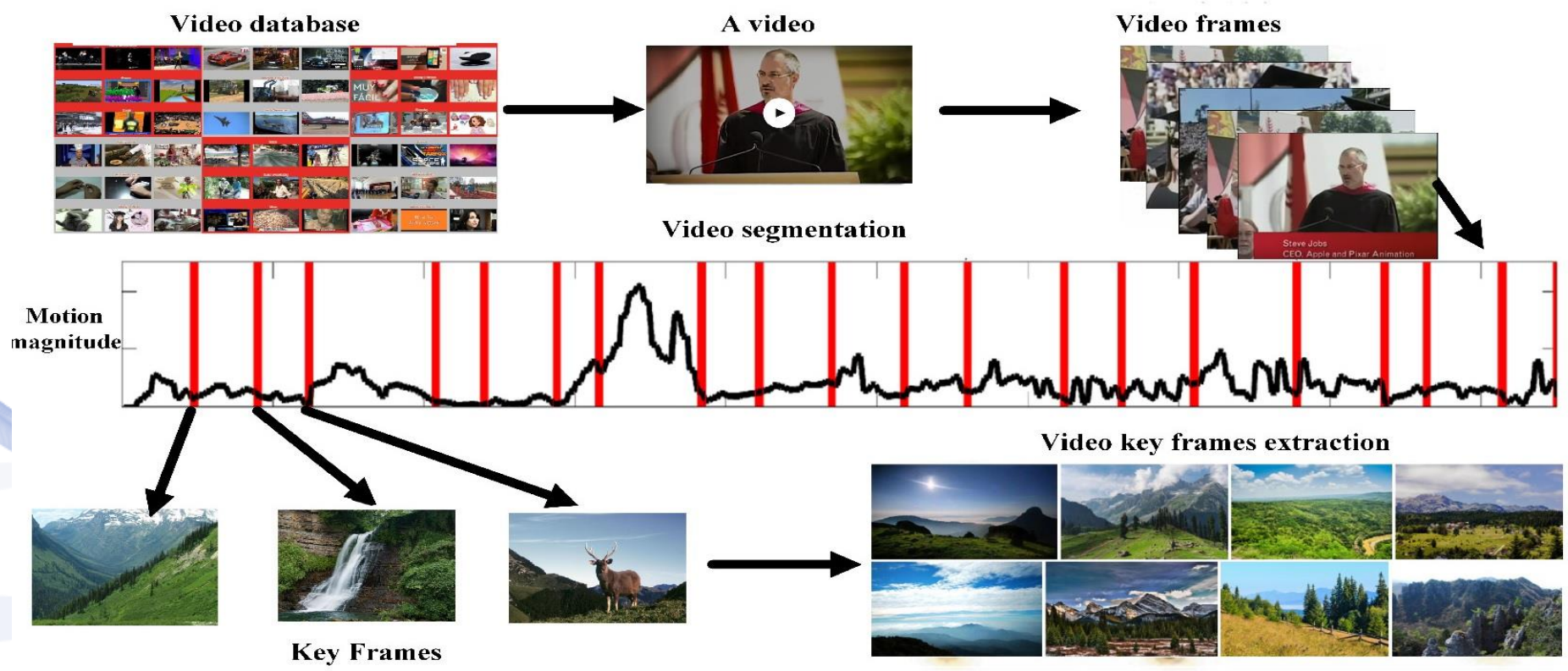
- **Video Captioning:** generating a textual description for a given video content.
- **Key framing extraction:** Key images are taken from the video in order to create the summary[TRU2007]

Static Video Summarization- Video Captioning



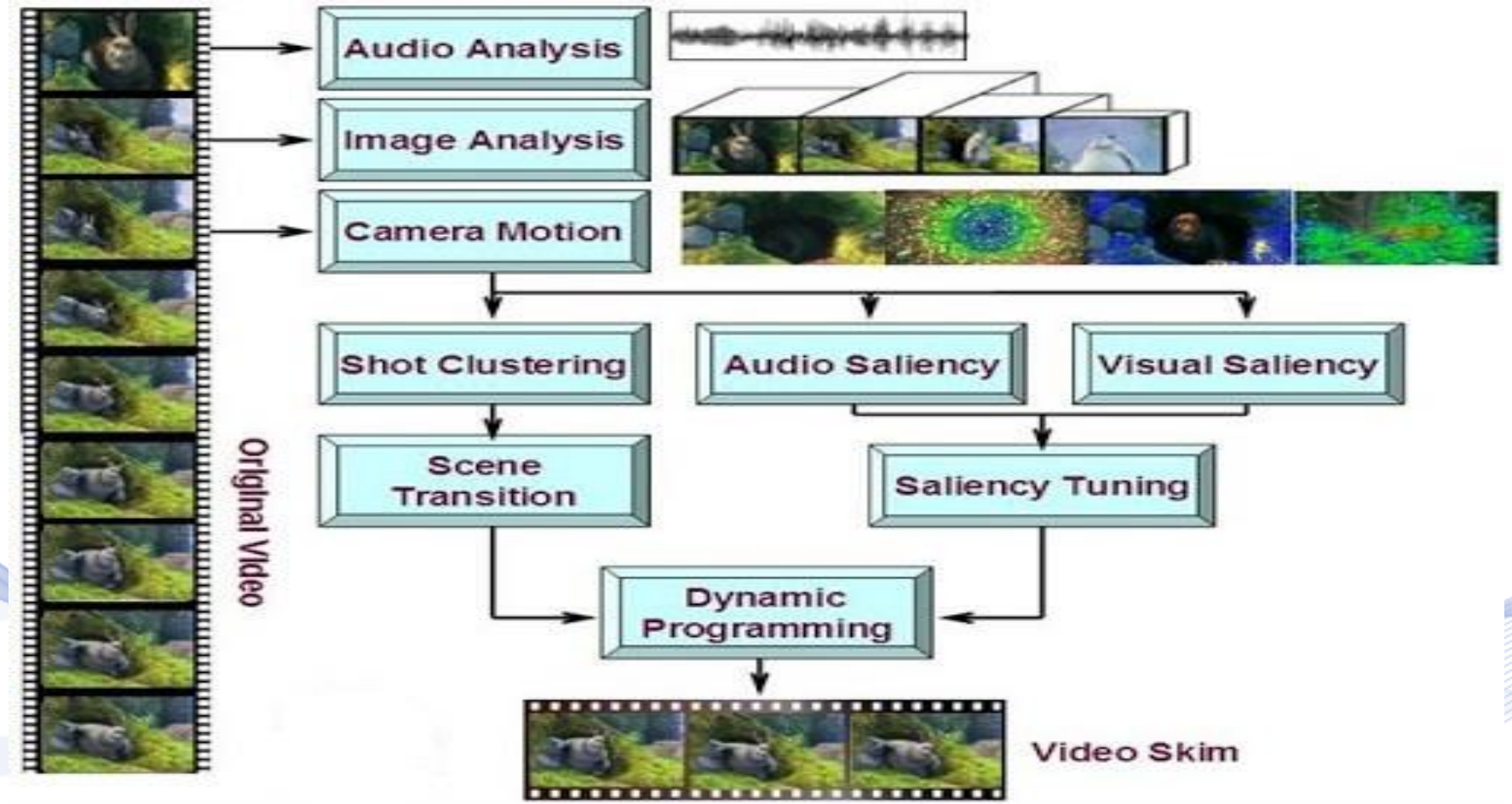
(Image from FXPAL)

Static Video Summarization- Key framing extraction



(Image from Willey Online Library)

Dynamic Video Summarization



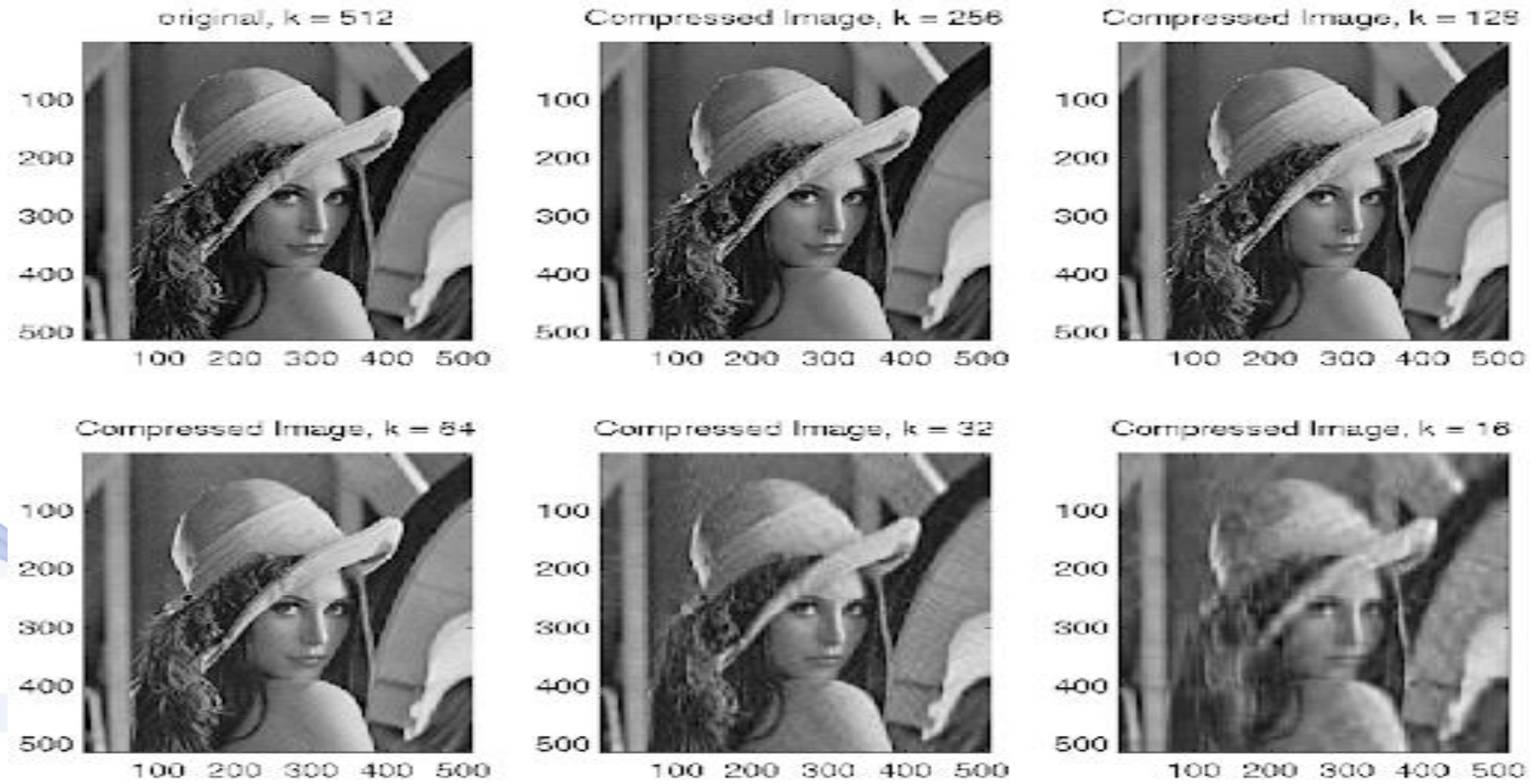
(Image from www.ifp.illonois.edu)

Dynamic Video Summarization

Dynamic video summarization is a combination of audiovisual components extracted from the original videos.

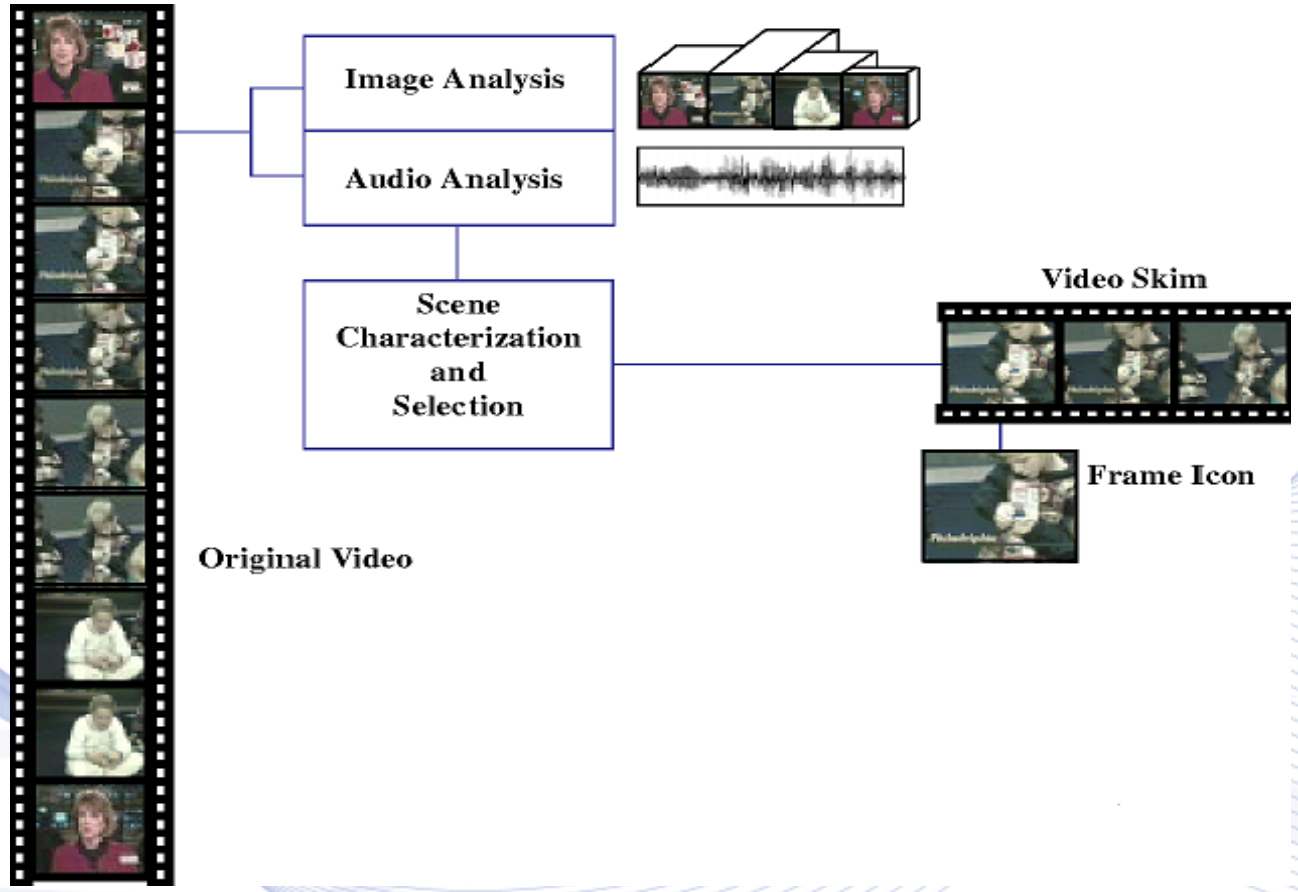
The basic idea of the skim video, which is a short video, consists of informative scenes from the original video presented to the user so that he can receive a summary of the video story but in video format. [BAL2019]

Singular Value Decomposition (SVD)



(Image from http://fourier.eng.hmc.edu)

Dynamic Video Summarization - Video Skim



Skim video for drastic reduction in viewing time without loss in content. The most significant frames from a select scene are chosen for browsing. A single frame is selected from the skim for iconic representation.

(Image from Semantic Scholar)

Video Summarization with Deep Neural Networks

Neural Networks are networks used in Machine Learning that work similar to the human nervous system. It is designed to function like the human brain where many things are connected in various ways.

They are two kinds of Neural Networks:

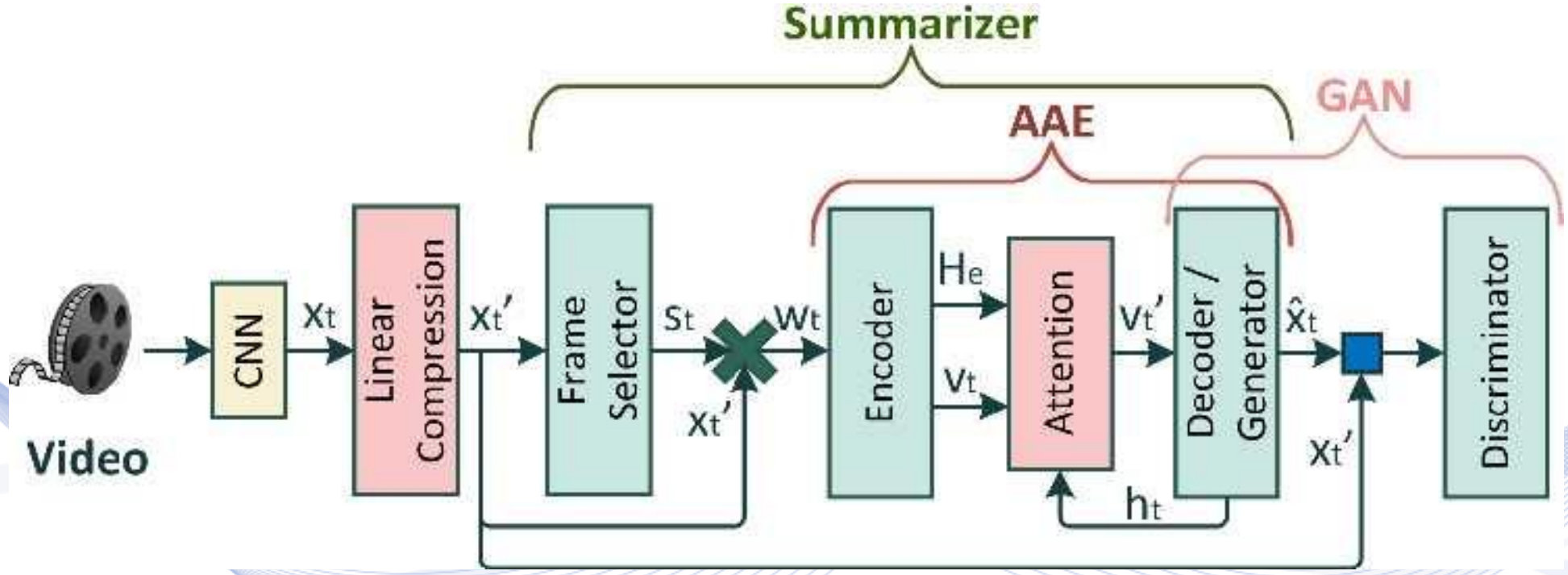
- Shallow Neural Networks
- Deep Neural Networks

Deep Architectures for Video Summarization

Video summarization is a long-standing problem.

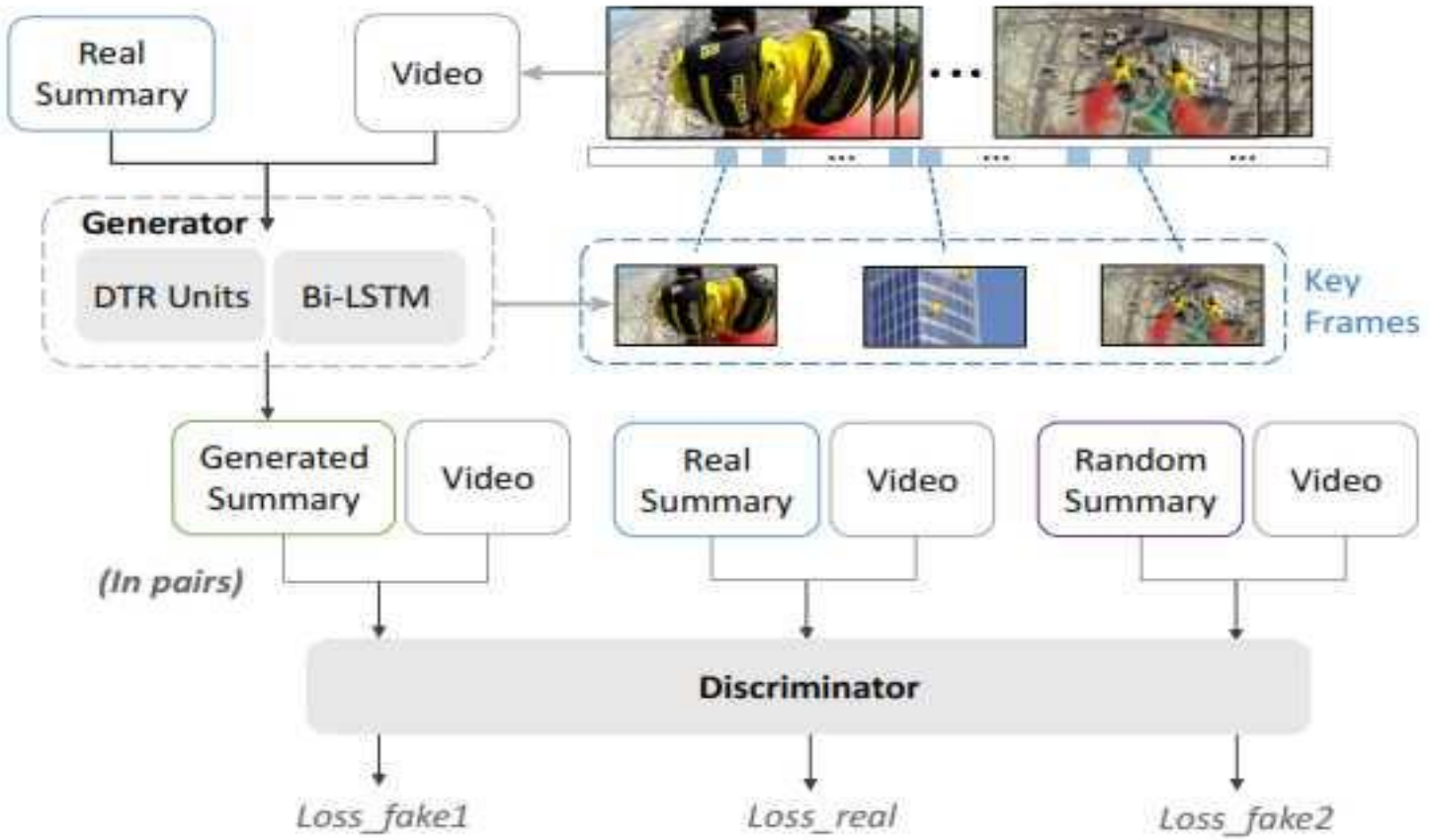
Generative Adversarial Networks (GANs) have been used for image-understanding problems and frame prediction / generation.

The SUM-GAN-AAE model architecture



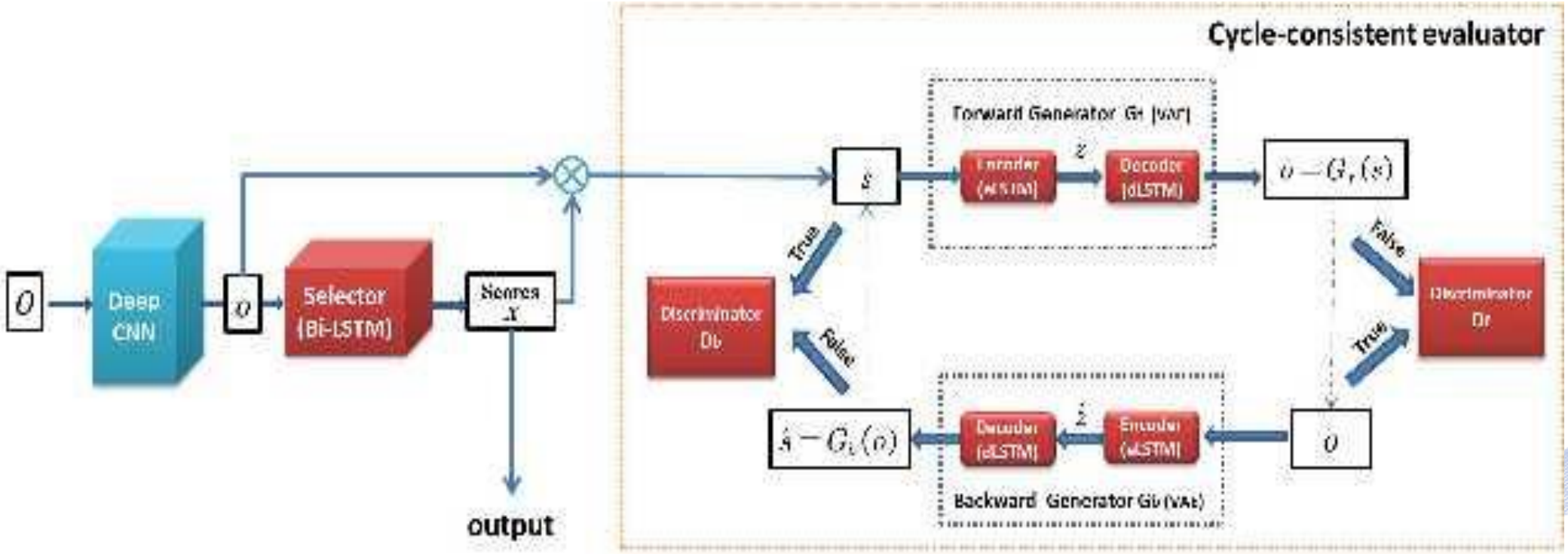
The architecture of SUM GAN-AAE (Image from SlideShare)

Dilated Temporal Relational Adversarial Network model



DTR-GAN (Image from ResearchGate)

Cycle-SUM architecture.



Cycle-SUM architecture. (Image by GroundAI)

Video Summarization Applications

- Movie trailer (film industry)
- Advert creation (Advertisement)
- Football highlights (Recreation means)
- Anomaly detection from video surveillance(security)
- Remove redundancy
- Reduce computational time, storage requirements
- Data visualization, Labeling
- Search, Retrieval, Recommendation [WOR2020]

Automatic video summarization applications

- **Surveillance Videos**

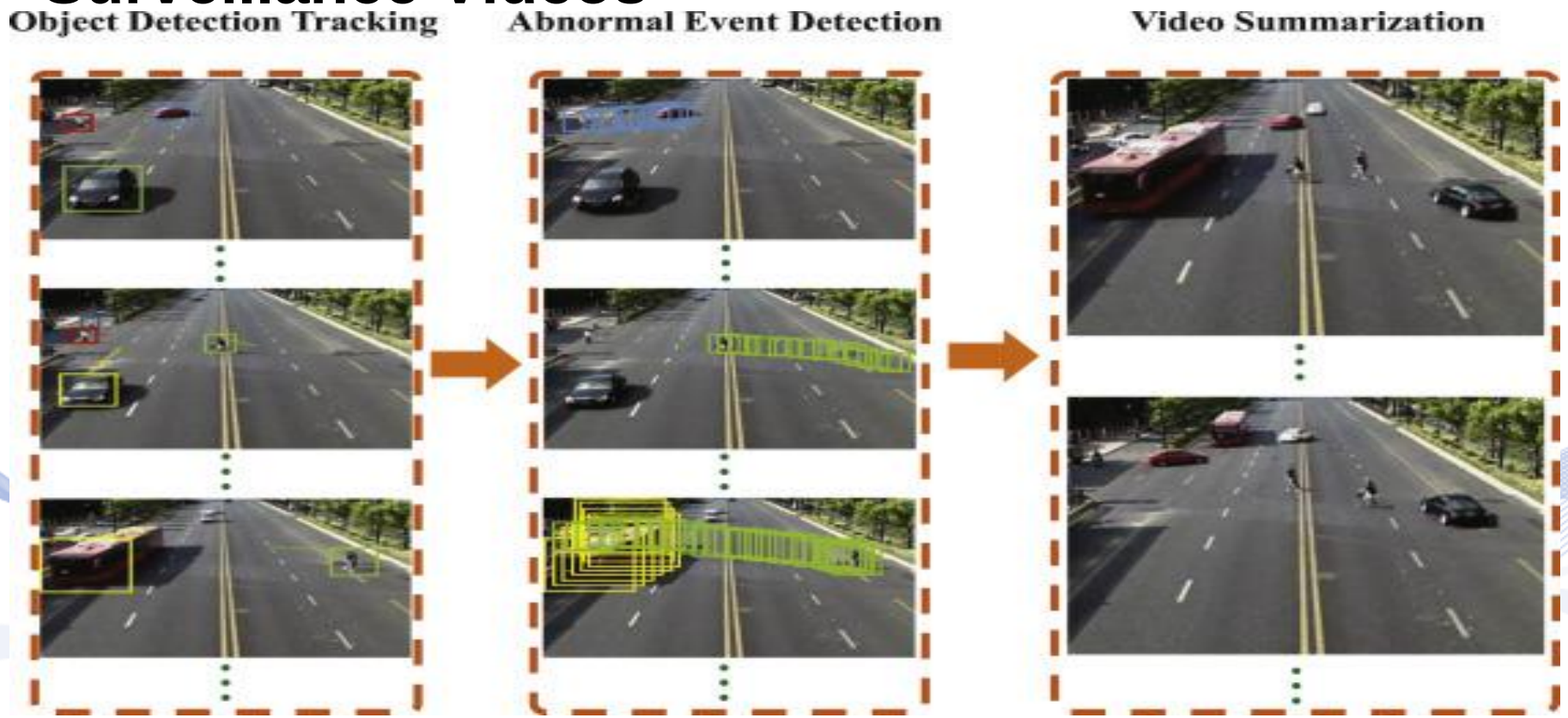


Image from ScienceDirect

Automatic video summarization applications

- **Egocentric Videos**

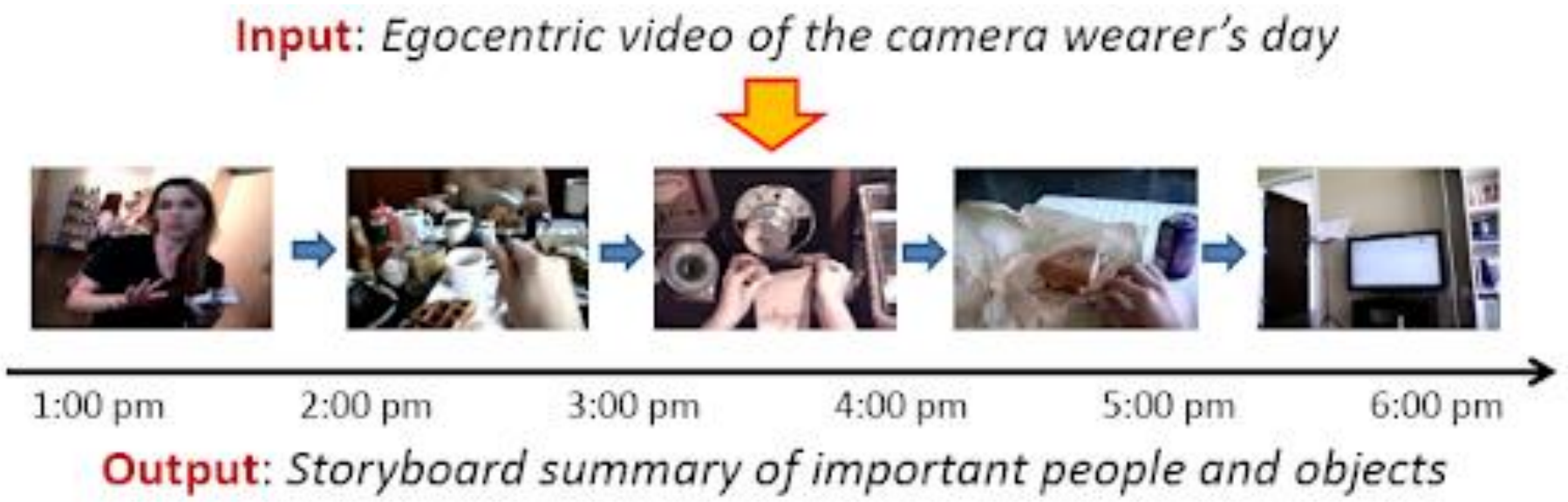


Image from vision.cs.utexas.edu

Automatic video summarization applications

- **Medical Videos**

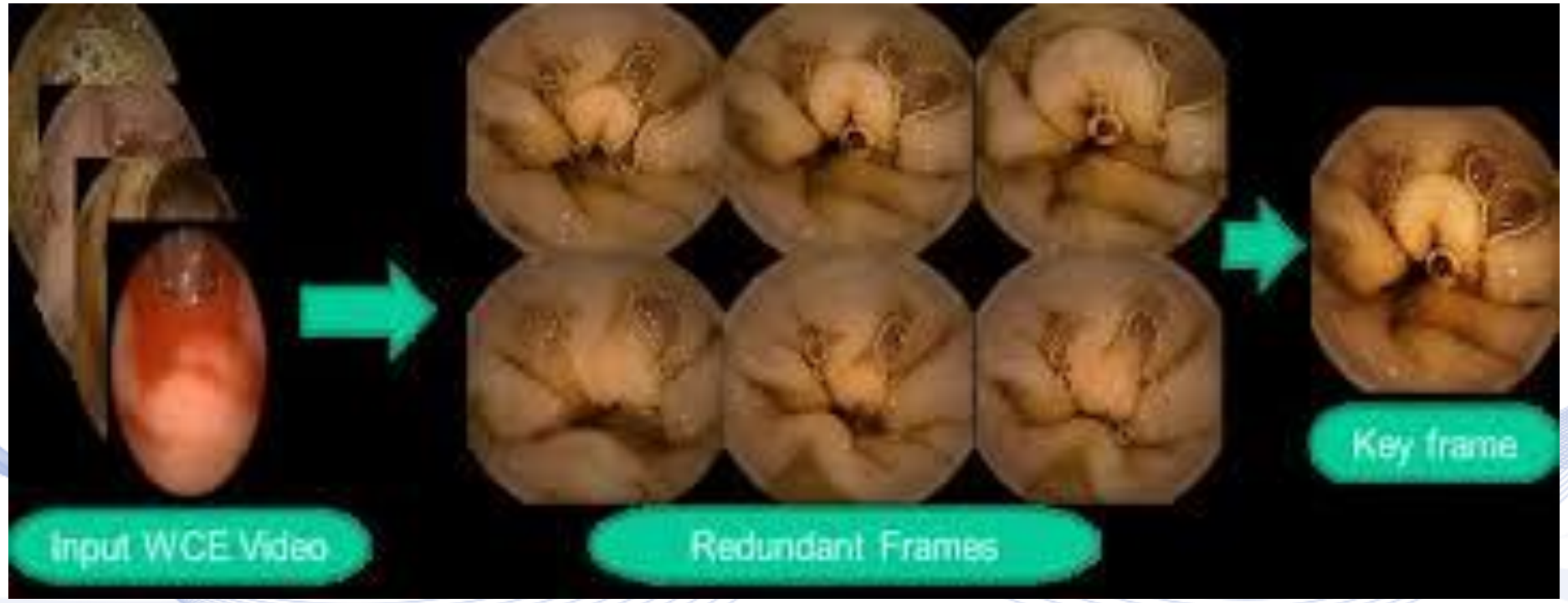


Image from E3S Web of Conferences

Automatic video summarization applications

- **Movies Videos**

Image from ScienceDirect

Automatic video summarization applications

- **Internet Videos**

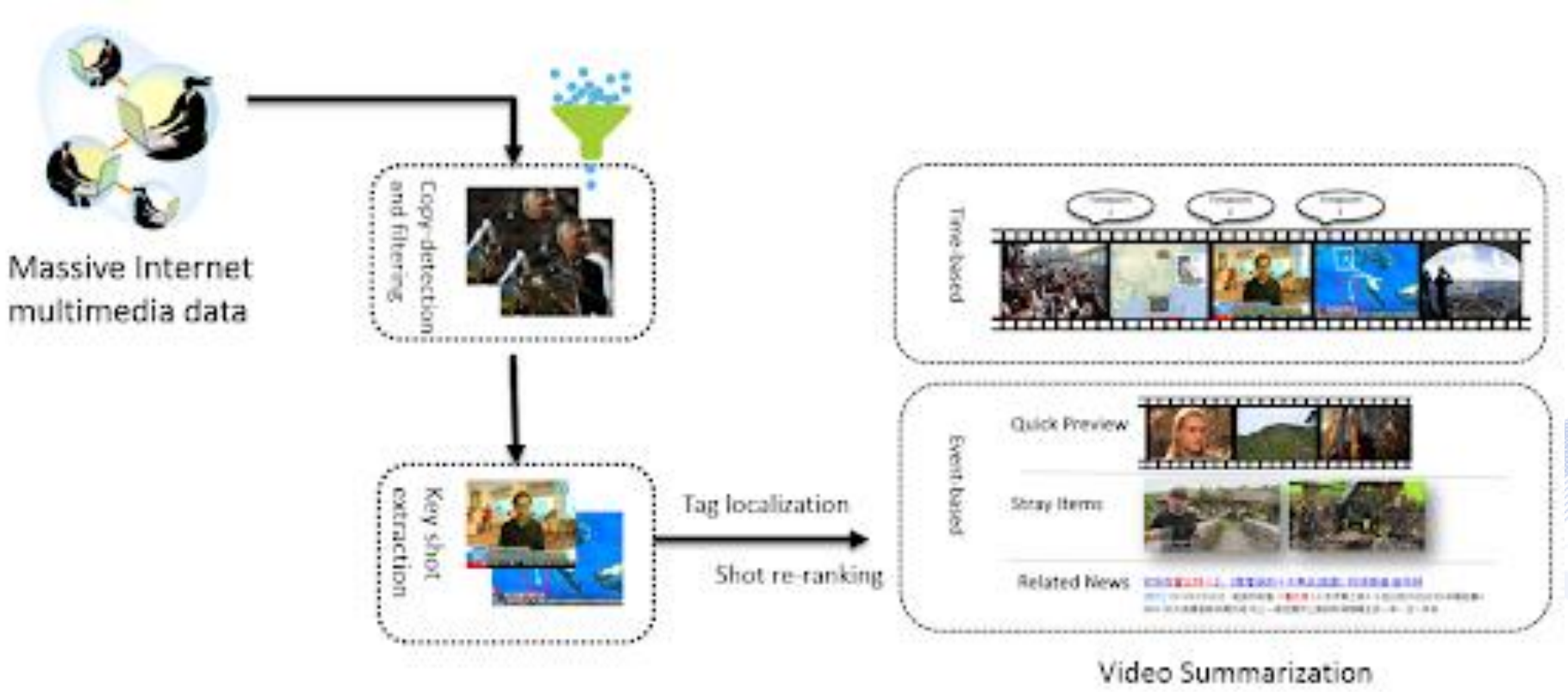


Image from ise.thss.tsinghua.edu.cn

Automatic video summarization applications

- **Drone and Robots**

Cameras are being integrated into recently developed devices like drones and robots equipping them to record the videos at many places which are impossible to reach by humans. Summarizing these videos will ease the interpretation of this new class of videos. [HTT2020]

Bibliography

[PIT2016] I. Pitas (editor), “Graph analysis in social media”, CRC Press, 2016.
[PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.
[PIT2017] I. Pitas, “Digital video processing and analysis” , China Machine Press, 2017 (in Chinese).
[PIT2013] I. Pitas, “Digital Video and Television” , Createspace/Amazon, 2013.
[NIK2000] N. Nikolaidis and I. Pitas, “3D Image Processing Algorithms”, J. Wiley, 2000.
[PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr