Static Video Summaries

Ígor Chagas Marques

Script

- Why?
- What?
- How?
- Metrics
- Results
- Discussion

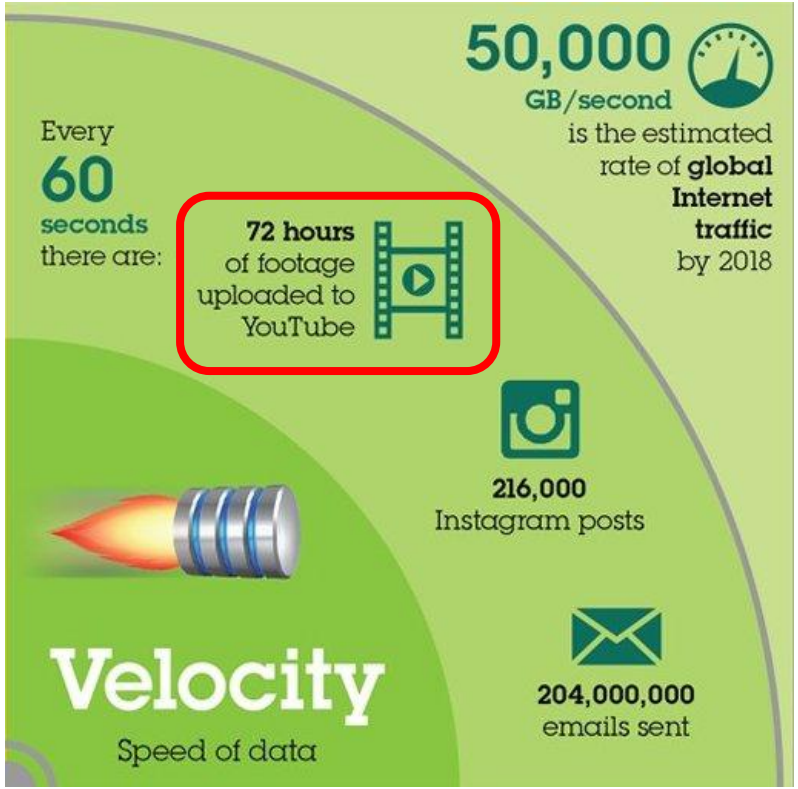
Why?

Video as a Very Popular Media Format

-

- Movies
- Documentaries
- Tutorials
- Music Videoclips
- Gameplays
- Memes/Humour
- Vlog
- TV Series

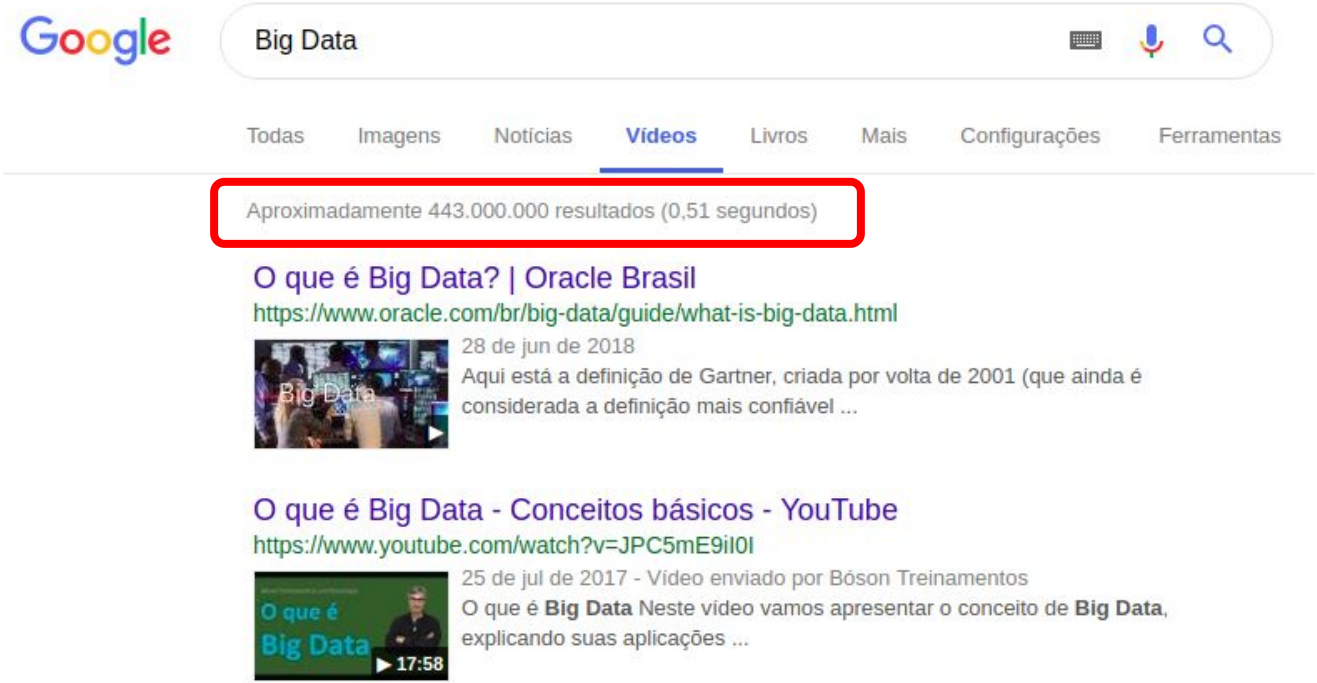
Talking About Numbers



According to IBM, 72 hours of video were uploaded to YouTube every minute[1]. (2018)

[1] https://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data

About Numbers (Yet)



Would you watch 443 MILLION videos to know more about some specific topic?

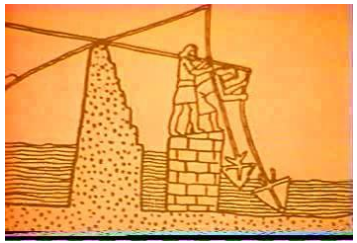
A Tiny Allegory

Scientific papers have an abstract text.
So there is no need to read the whole text to figure out about them.

What?

Static Video Summary





Set of sorted key-frames without sound effects that can represent the video.

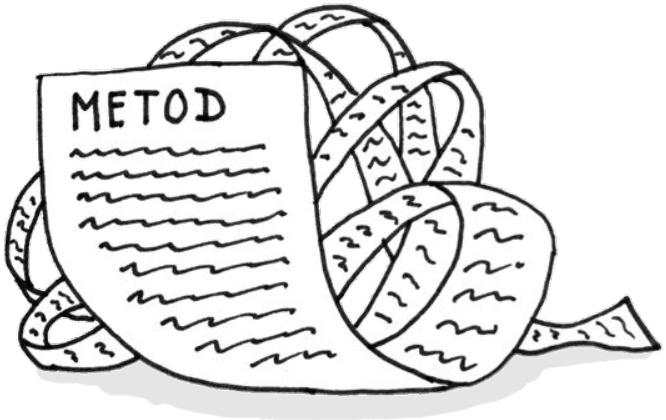
[2]Summary from VSSUM Project.

How?

Thresholding

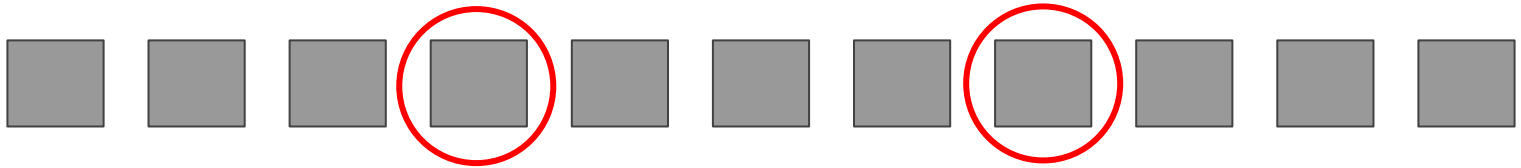
- Based on **Sheena *et al*** Statistical Method

The Method



1. Pre-sampling
2. Feature Extraction
3. Distance Between Frames Calculation
4. Threshold Calculation
5. First Frame Filtering
6. Second Frame Filtering

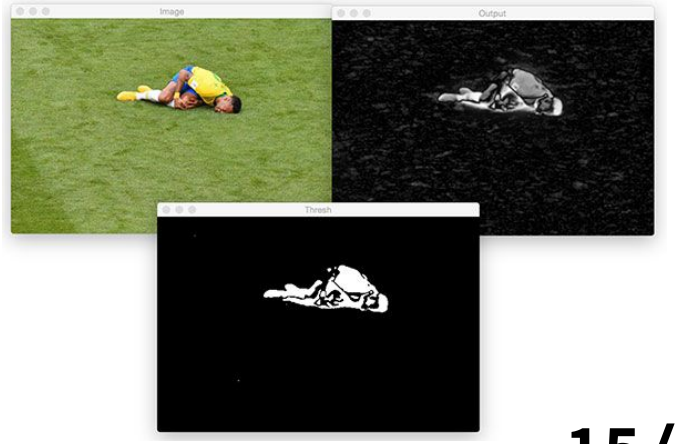
Pre-sampling



Feature Extraction

1) 16 bins hue component from HSV relative histogram

2) Saliency center position (x,y) on screen



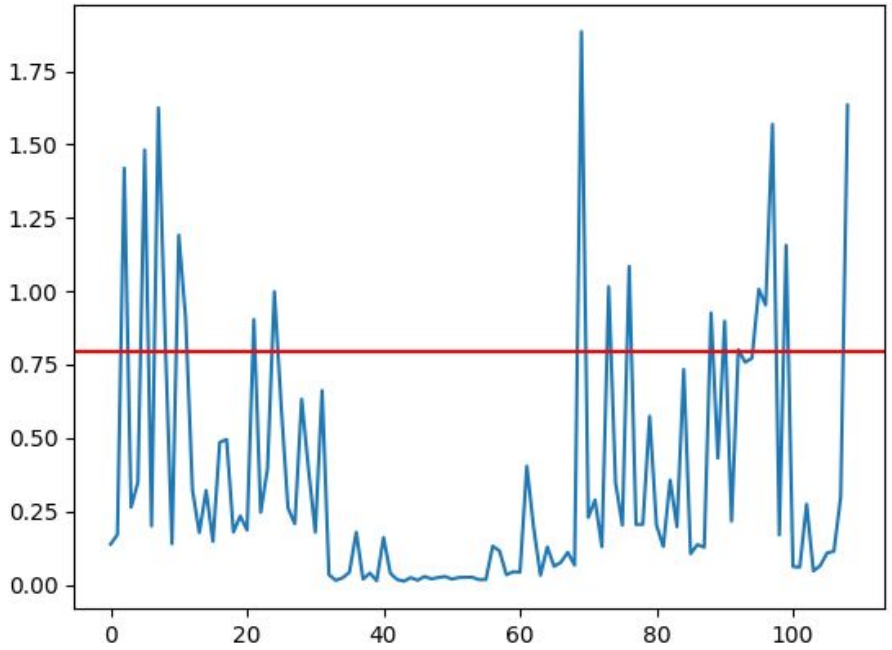
Distance Between Frames Calculation

- Taxi Cab Distance

Threshold Calculation

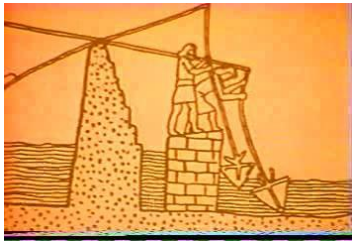
$$t = \mu + \sigma$$

First Frame Filtering



Second Frame Filtering

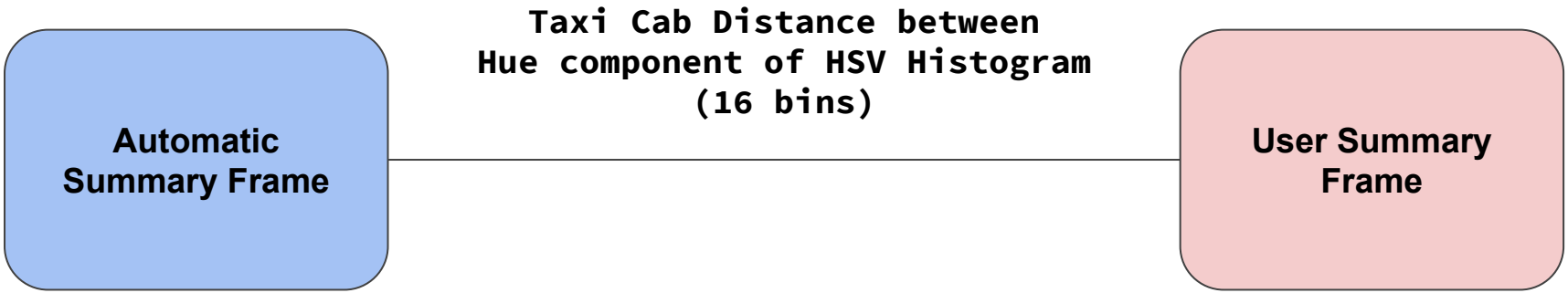




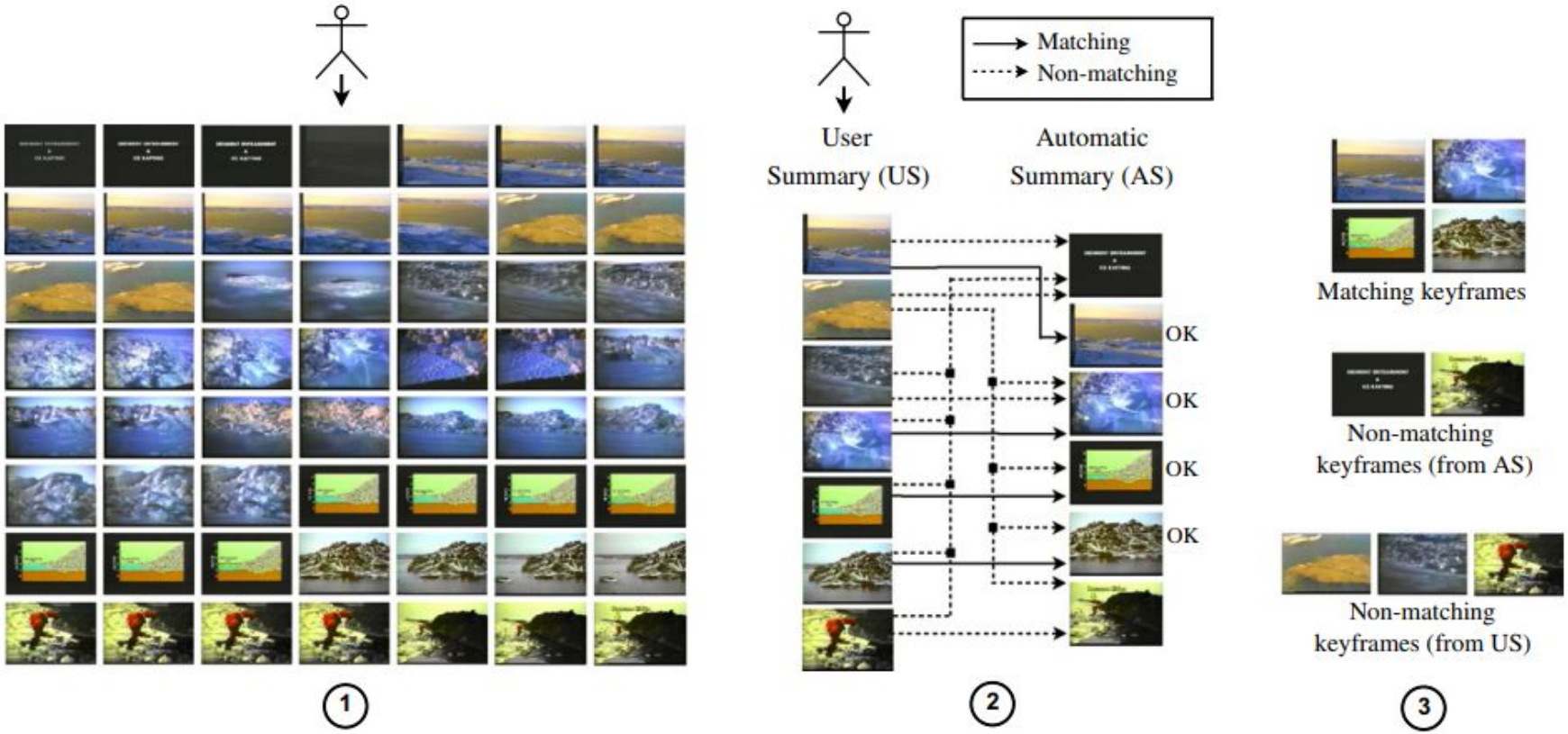
Metrics

An Appropriate Metric

- Ground Truth
- User Summaries
- Comparison
- Quality Metric



If the Distance is smaller than 0.5, we assume it is the “same” frame.


Matching
Non-matching
User
Summary (US)
Automatic
Summary (AS)
OK
OK
OK
OK
Matching keyframes
Non-matching
keyframes (from AS)
Non-matching
keyframes (from US)
1
2
3

The Metrics itself

$$CUS_A = \frac{n_{mAS}}{n_{US}},$$

$$CUS_E = \frac{n_{\bar{m}AS}}{n_{US}},$$

Results

The Open Video Data Base

- MPEG format
- 30 fps
- 352 x 240 pixels
- several genres
- 1 ~ 4 minutes
- 50 videos
- 5 user summary per video

Numbers

Feature Set	CUS_A	CUS_E
Color	0.65	0.60
Saliency	0.52	0.60

TABLE I
RESULT METRICS FOR FEATURE SETS

Discussion

Conclusion and Future Work

The results achieved are interesting. However, unfortunately, they are not so good as Avila *et al* work (VSUMM)[2].
Therefore as future work we propose to check other feature sets and also try to combine Color information and Saliency information in a weighted way.

[2] VSSUM1 achieved CUS_a = 0.85, CUS_e = 0.38.
VSSUM2 achieved CUS_a = 0.70, CUS_e = 0.27.

References

[1] M. Srinivas, M. M M, and R. M. Pai, “An improved algorithm for video summarization a rank based approach,” Procedia Computer Science, vol. 89, pp. 812-819, 12 2016.

[2] P. Kaur, “Analysis of video summarization techniques,” International Journal for Research in Applied Science and Engineering Technology, vol. 6, pp. 1157-1162, 01 2018.

[3] S. Avila, A. Paula Brando Lopes, A. da Luz, and A. Arajo, “Vsumm: A mechanism designed to produce static video summaries and a novel evaluation method,” Pattern Recognition Letters, vol. 32, pp. 56-68, 01 2011.

[4] S. C V and N. Narayanan, “Key-frame extraction by analysis of histograms of video frames using statistical methods,” Procedia Computer Science, vol. 70, pp. 36-40, 12 2015.

References

[5] N. Ejaz, T. Tariq, and S. Baik, “Adaptive key frame extraction for video summarization using an aggregation mechanism,” Journal of Visual Communication and Image Representation, vol. 23, p. 10311040, 10 2012.

[6] Key-Frame Extraction Using Weighted Multi-view Convex Mixture Models and Spectral Clustering, Aug 2014.

[7] B. C. Ivan Laptev, “Recognition of human actions,” 2005, access: 06/10/2019. [Online]. Available: http://www.nada.kth.se/cvap/actions/

[8] F. Cardeal, “Images in the spatial domain,” 2019, access: 06/10/2019. [Online]. Available: http://cardeal.piim-lab.cefetmg.br/Teaching/cvision/ Slides-02.pdf

References

[9] A. Tremau, S. Tominaga, and K. Plataniotis, “Color in image and video processing: Most recent trends and future research directions,” EURASIP J. Image and Video Processing, vol. 2008, 05 2008.

[10] A. Rosebrock, “Opencv saliency detection,” 2018, access: 06/10/2019. [Online]. Available: https://www.pyimagesearch.com/2018/07/16/opencv-saliency-detection/

[11] B. Davida, “Opencv static saliency detection in a nutshell,” 2019, access: 06/10/2019. [Online]. Available: https://towardsdatascience.com/opencv-static-saliency-detection-in-a-nutshell-404d4c58fee4

Fork me on Github

`github.com/daegonny`