FAST-FORWARD VIDEO BASED ON SEMANTIC EXTRACTION

Washington L. S. Ramos, Michel M. Silva, Mario F. M. Campos, Erickson R. Nascimento

{washington.ramos, michelms, mario, erickson}@dcc.ufmg.br

Universidade Federal de Minas Gerais (UFMG)

I. Introduction

In this work, we propose a novel methodology capable of transforming raw egocentric videos into watchable fast-forward videos by considering both the pleasantness and relevance of frames to the viewer. Our approach analyses the semantic information extracted from the frames and segments the video by selecting the set of images which maximizes the semantic term, the required speed-up as well as the desired smoothness in the transition between the frames.

II. Methodology

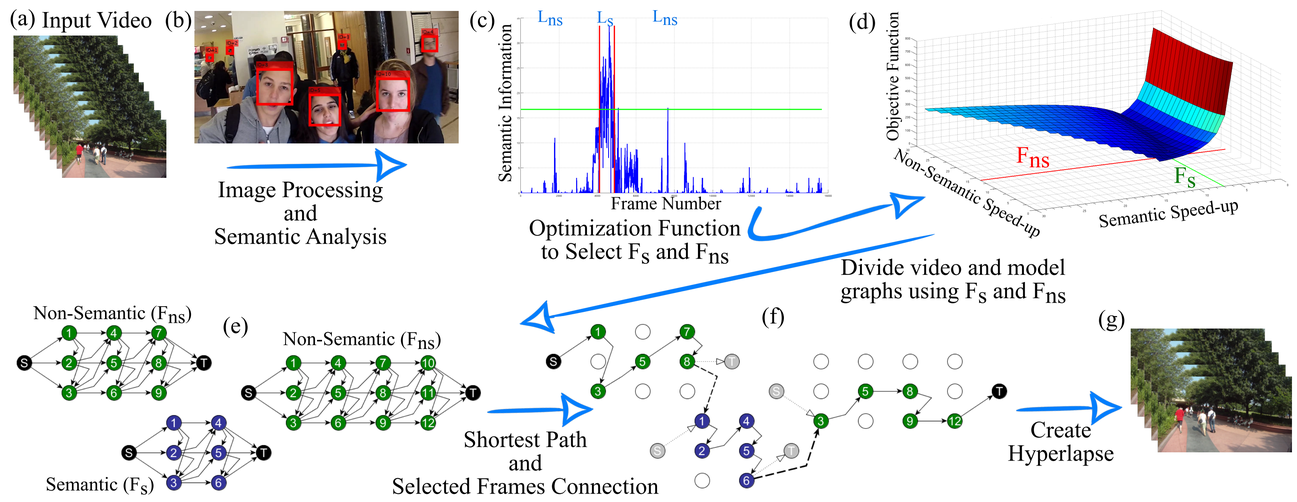


Figure 1: Overview of our fast-forward methodology.

III. Formulation

$$D\left(\mathcal{F}_{ns}, \mathcal{F}_s\right) = \left(\left|\frac{L_s + L_{ns}}{F_d} - \frac{L_s}{\mathcal{F}_s} - \frac{L_{ns}}{\mathcal{F}_{ns}}\right|\right)$$

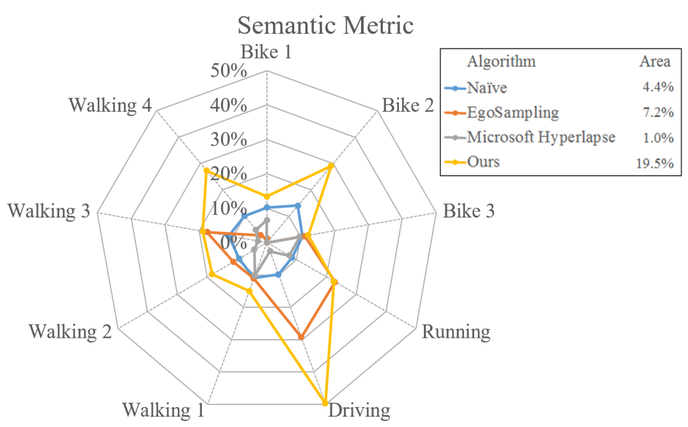
$$\arg\min_{\mathcal{F}_s, \mathcal{F}_{ns}} \left(D\left(\mathcal{F}_{ns}, \mathcal{F}_s\right) + \lambda_1 \cdot \left|\mathcal{F}_{ns} - \mathcal{F}_s\right| + \lambda_2 \cdot \left|\mathcal{F}_s\right|\right)$$

IV. Graph Modeling

$$\mathcal{W}_{i,j} = \left(\alpha \cdot \mathcal{B}_{i,j} + \beta \cdot \mathcal{V}_{i,j} + \gamma \cdot \mathcal{A}_{i,j} + \eta \cdot \mathcal{S}_{i,j}\right) \cdot \left\lceil \frac{(j-i)}{\mathcal{F}} \right\rceil$$

$$\mathcal{S}_{i,j} = \frac{1}{S_i + S_j + \epsilon}, \qquad S_x = \sum_{k \in f_x} C(k) \cdot G_\sigma(k) \cdot A(k)$$

V. Results



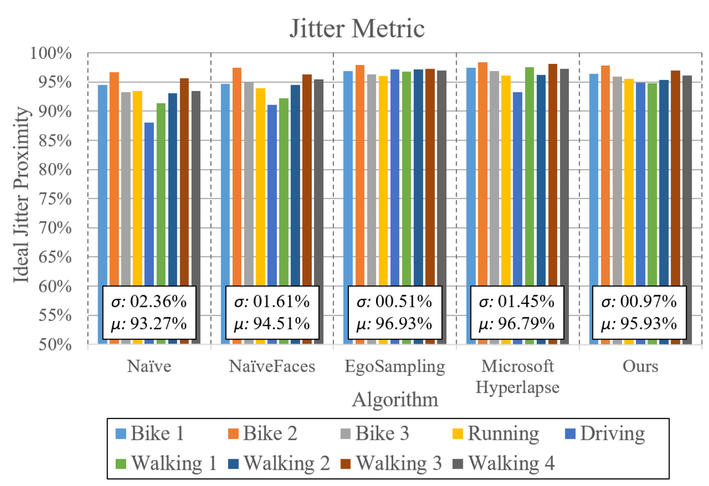


Figure 2: Semantic information in the final video, the values are related to NaïveFaces result, once it selects as much information as possible (left); Jitter analysis using the mean FOE deviation over the sequences (right).

Method	Mean (%)	S.D. (%)
ES	89.71	14.36
MH	**98.77**	**00.47**
Ours	96.49	02.71

Table 1: Speed-up metric analysis of the experiments. The values show how close the algorithms are from the ideal.

VI. Conclusion

We presented a novel method for producing fast-forward videos from egocentric videos focusing on its semantic content. We compared our adaptive selection strategy against prime works in the literature. Results showed our superiority as far as the semantic information is concerned.

References

[1] Yair Poleg, Tavi Halperin, Chetan Arora, and Shmuel Peleg. Egosampling: Fast-forward and stereo for egocentric videos. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pages 4768–4776, June 2015.

[2] Neel Joshi, Wolf Kienzle, Mike Toelle, Matt Uyttendaele, and Michael F. Cohen. Real-time hyperlapse creation via optimal frame selection. *ACM Trans. Graph.*, 34(4):63:1–63:9, July 2015.

Acknowledgements

UFMG · PPGCC Programa de Pós-Graduação em Ciência da Computação DCC-UFMG · VERLab Computer Vision and Robotics · CAPES · CNPq · FAPEMIG · ITV · BR Petrobras