ADAPTIVE VIDEO SUBSAMPLING FOR ENERGY-EFFICIENT OBJECT DETECTION

[1]D. Mohan*, [2]S. Katoch*, [2]S. Jayasuriya, [2]P. Turaga, [2]A. Spanias

[1]University of California, Berkeley, [2]SenSIP Center, School of ECEE, Arizona State University

MOTIVATION

- Excessive use of high-resolution videos for object detection in current literature;
- Lack of Energy-efficient object detection solutions in Embedded and Mobile platforms;
- Need for understanding pixel relevance while sensing, to perform computer vision task efficiently.

PROJECT AIM

- Reduce energy expense during image capture while maintaining computer vision task accuracy.
- Develop Adaptive video subsampling technique for energy-efficient object detection.
- Explore and quantify several subsampling strategies.

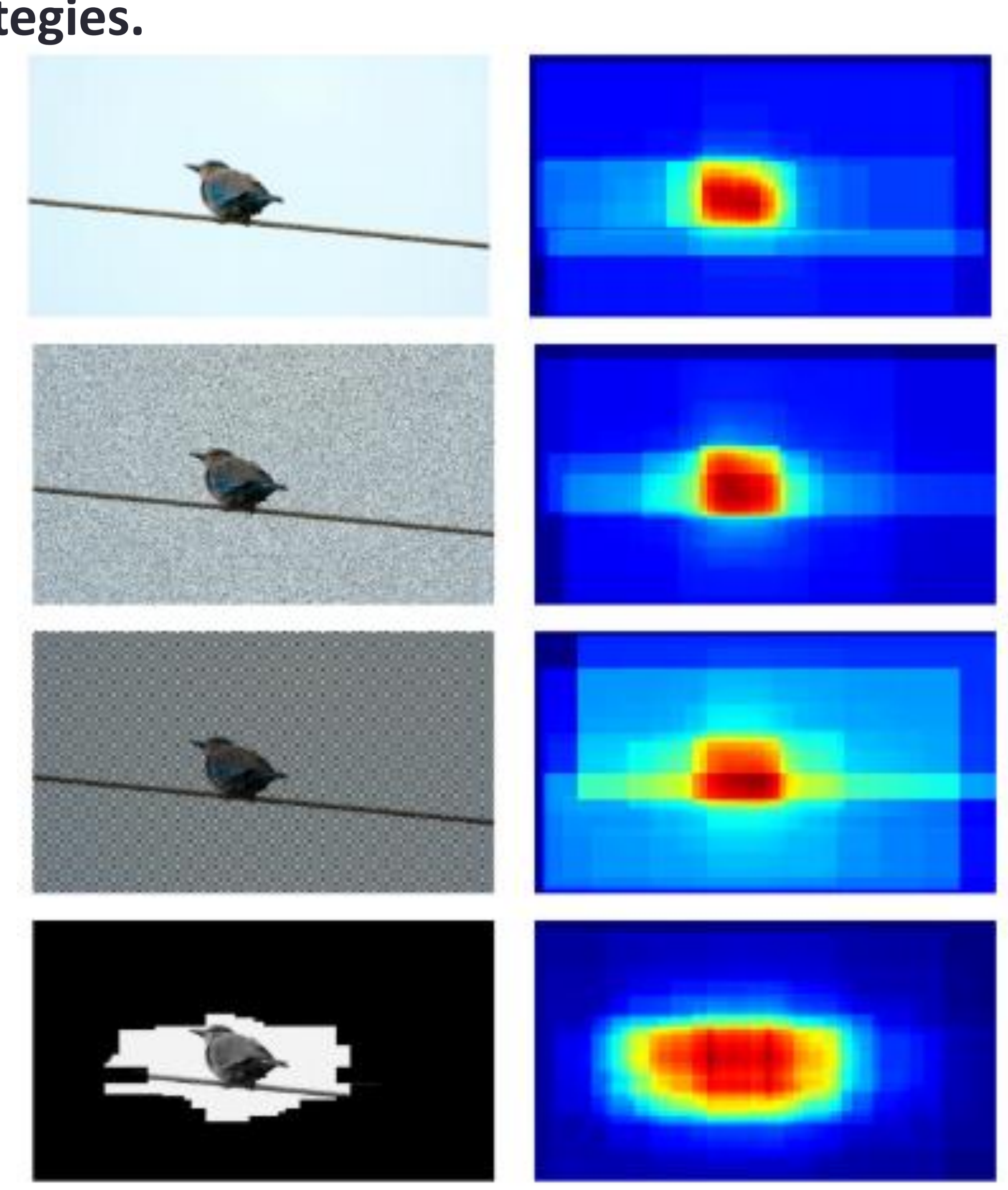

Fig. 1: The first row shows the original image and its resulting objectness map. The next three rows show the same process for three different forms of image subsampling on the original image: Random Pixelation, Checkerboard Mask, and Adaptive Video Sampling.

ADAPTIVE VIDEO SUBSAMPLING ALGORITHM

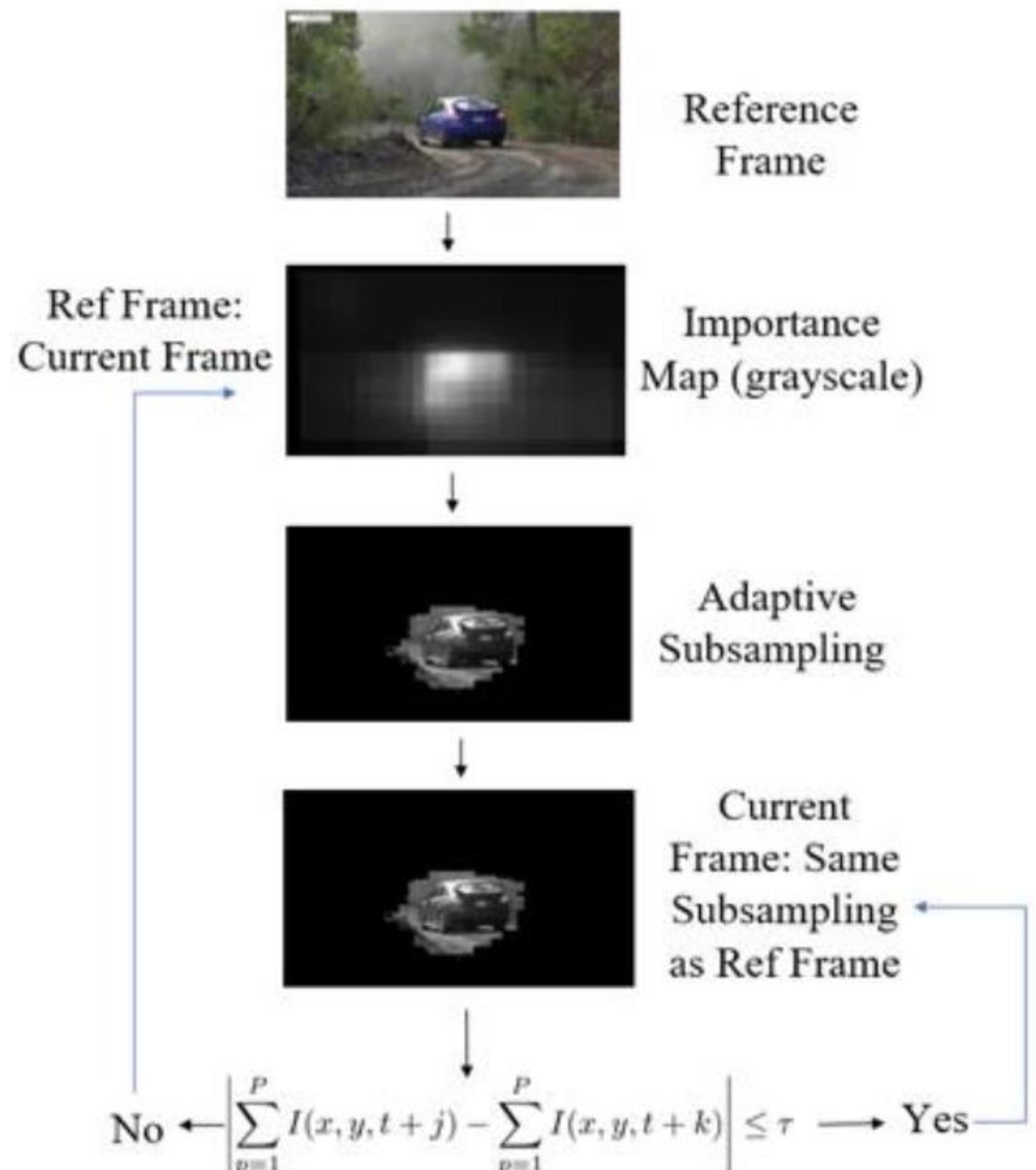

Fig. 2. Diagram illustrating the adaptive video subsampling algorithm

- I(x, y, t) represent a video where (x, y) are the locations of the pixels and t represents the frame index in time.
- A reference frame has its objectness/importance map calculated by considering four image cues: multi-scale saliency, color contrast, edge density and straddleness [1].
- Using Otsu's method, an objectness threshold is used to obtain a binary mask [2].
- Consequently, the binary mask is used to turn-off the pixels in the consecutive frames subject to the shown intensity-based criteria (Fig. 2.) which tells the algorithm to update the reference frame and the importance map. Another experiment with Optical flow magnitude criteria was also conducted.

RESULTS

Experiments are conducted on ILSVRC2015 Image Vid Dataset.

Subsampling Strategies	Fully Sampled	Random Subsampling (α)			Adaptive Subsampling (Otsu's Objectness Threshold + Frame Intensity Threshold)			Adaptive Subsampling (Flow Magnitude Threshold (10^{-3}))			Adaptive Subsampling (Objectness Threshold + Frame Intensity Threshold)
		0.15	0.25	0.35	0.1	0.3	0.5	1.5	5.0	15.0	0.15 + 0.1
mAP	55.5	15.4	5.9	0.9	40.1	37	38	41.8	41.7	28.6	50.1

Table 1. mAP scores for different subsampling strategies.

Subsampling Strategies	Random Subsampling (α)			Adaptive Subsampling (Otsu's Objectness Threshold + Frame Intensity threshold)			Adaptive Subsampling (Flow Magnitude Threshold (10^{-3}))			Adaptive Subsampling (Objectness Threshold + Frame Intensity threshold)
	0.15	0.25	0.35	0.1	0.3	0.5	1.5	5.0	15.0	0.15 + 0.1
Bird	14.16	22.75	30.80	87.43	86.64	86.43	92.23	92.24	92.22	54.04
Watercraft	13.26	22.32	31.62	79.80	79.91	80.09	83.15	83.01	88.30	50.04
Dog	17.71	29.39	40.59	11.83	11.87	11.86	68.76	68.76	68.76	18.44
Car	18.13	30.66	42.86	30.42	30.10	30.18	90.44	90.72	88.94	67.87
Horse	21.21	34.96	48.01	25.82	26.26	26.46	75.65	75.65	75.90	38.85
Train	22.24	29.60	37.41	21.05	21.02	21.07	71.19	71.19	71.19	55.97

Table 2: Energy efficiency in terms of turned off pixel percentage in a video for different subsampling strategies.

- If frame intensity threshold τ is very big, it can lead to conditions where the subsampling strategy neglects the changes due to object motion and if it's too small it will lead to subsampling calculation of every consecutive frame which will result in high computation time.
- Adaptive Subsampling with tuned objectness threshold and frame intensity threshold has the best mAP score of 50.1% and saves upto 67% of energy in terms of turned-off pixels.

CONCLUSION

- The proposed method requires very little resources to implement on an embedded vision platform.
- The method is agnostic to the type of feature detection and could be modified for a particular application domain.
- A promising improvement is to warp the ROI to follow motion of the object(s) using either Kalman filtering or optical flow warping.

REFERENCES

[1] B. Alexe et al., "Measuring the objectness of image windows," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 34, no. 11, pp. 2189–2202, 2012.
[2] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62–66, 1979.
[3] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," 1981.

ACKNOWLEDGEMENTS

This material is based upon work supported in part by the NSF under CNS1659871, 1646542 as well as ASU's SenSIP Center. (* - equal contribution)

Sensor Signal and Information Processing Center
https://sensip.asu.edu