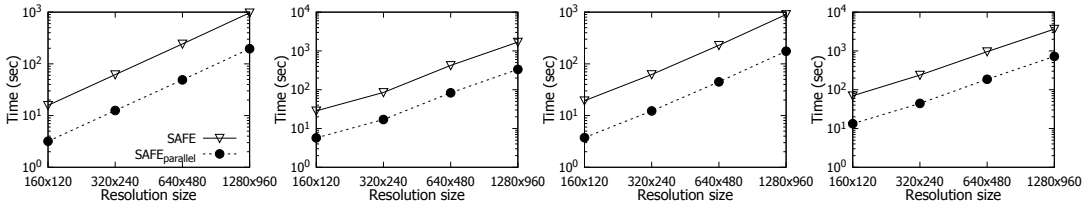
A Parallel Approach for SAFE

Tsz Nam Chan[1], Pak Lon Ip[2,3], Leong Hou U[2,3], Byron Choi[1] and Jianliang Xu[1]

[1]Hong Kong Baptist University
[2]University of Macau
[3]State Key Laboratory of Internet of Things for Smart City

Domain experts (e.g., geoscientists [6], criminologists [4], and transportation experts [5]) may only have limited computational resources. To generate KDV, these users [6, 5, 4] normally adopt QGIS [2] and ArcGIS [1] software packages on a workstation. In order to further improve the user experience, we ask a question: How can we fully utilize a workstation to significantly improve the efficiency for our SAFE method? In this supplementary file, we provide an affirmative answer to this question.

In Algorithm 3 of our work [3], observe that different pixels $\mathbf{q}$ do not need to access the same computational resources for calculating the kernel density values $\mathcal{F}_P^{(b)}(\mathbf{q})$ (i.e., those variables in lines 5 to 10 in Algorithm 3 do not depend on the variables for other pixels). Therefore, we develop the method, SAFE$_{\text{parallel}}$, which assigns each thread to handle different pixels in parallel. Figure 1 shows the experimental results of our SAFE and SAFE$_{\text{parallel}}$ methods with $L = 80$ and different resolution sizes, using the Intel i7 3.19GHz CPU with 6 cores. Observe that our SAFE$_{\text{parallel}}$ method is fully parallel, which achieves five times speedup compared with SAFE. The implementation of SAFE$_{\text{parallel}}$ is available at the Github repository `https://anonymous.4open.science/r/77b4ac95-23a7-4f44-9847-371c22b2580e/`. By adopting the similar concept, we can also extend this parallel approach for supporting SAFE$_{\text{all}}$ and SAFE$_{\text{exp}}$ with on-the-fly bandwidth values.



(a) Chicago (b) Seattle (c) New York (d) San Francisco

Figure 1: Response time of the methods SAFE and SAFE$_{\text{parallel}}$ with $L = 80$, varying the resolution sizes (from 160×120 to 1280×960).

References

[1] ArcGIS. `http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-kernel-density-works.htm`, 2021.

[2] QGIS. https://docs.qgis.org/2.18/en/docs/user_manual/plugins/plugins_heatmap.html, 2021.

[3] T. N. Chan, P. L. Ip, L. H. U, B. Choi, and J. Xu. SAFE: A share-and-aggregate bandwidth exploration framework for kernel density visualization. *Proc. VLDB Endow.*, 2022 (Submitted).

[4] T. Hart and P. Zandbergen. Kernel density estimation and hotspot mapping: examining the influence of interpolation method, grid cell size, and bandwidth on crime forecasting. *Policing: An International Journal of Police Strategies and Management*, 37:305–323, 2014.

[5] K. Xie, K. Ozbay, A. Kurkcu, and H. Yang. Analysis of traffic crashes involving pedestrians using big data: Investigation of contributing factors and identification of hotspots. *Risk Analysis*, 37(8):1459–1476, 2017.

[6] L. Xu, S. Zhao, S. S. Chen, C. Yu, and B. Lei. Analysis of arable land distribution around human settlements in the riparian area of Lake Tanganyika in Africa. *Applied Geography*, 125:102344, 2020.