

# GPU computing to data visualization in data mining

Liang Xu

Department of Electrical and Computer Engineering  
Virginia Commonwealth University  
Richmond, Virginia  
Email: xul4@vcu.edu

**Abstract**—This paper survey is the second assignment for CS691 High performance distributed system in VCU. The topic of this survey is applications of GPU computing to data visualization in data mining. A growing number of data scientists are using GPUs for big data analytics to make better, real-time business decisions. More recently, data visualization has caught attention by data mining field. How to transform data from high dimension to 2D or 3D which can be recognized by human visualizations becomes very popular question.

In this paper, we survey the-state-of-the-art algorithms and applications in GPU computing to data visualization in data mining and discuss the future research opportunities.

**Index Terms**—GPU, Data Visualization, Data Mining

## I. INTRODUCTION

**D**ATA visualization is an essential component of data analysis. It involves the creation and study of the visual representation of data. A primary goal of data visualization is to communicate information clearly and efficiently via statistical graphics. Data visualization is both an art and a science. It is viewed as a branch of descriptive statistics by some, but also as a grounded theory development tool by others. [1]

General-purpose computing on graphics processing units is the use of a graphics processing unit (GPU), which typically handles computation only for computer graphics, to perform computation in applications traditionally handled by the central processing unit (CPU). [2], [3]

## II. BACKGROUND

Data visualization is human and machine collaboration to mass data where extract the most informative features. Visualization provides a bridge from human understanding to the information from a mass data set. This information includes relations between features, structure of the data and revealing the underling relationship between features.[4] GPU is initially designed for game application, however, GPUs demonstrate impressive computing power and high levels of parallelism and are now being used for may applications which not limited by the graph computation. The GPU provides the computing power especially for the algorithm can be parallelize. The GPU assisted algorithms are playing an important role in data vitalization, which include simple matrix manipulation[5], [6], [7], FFT[8], [9], and wavelet transform[10], [11] and so on.[12]

A. history

B. latest development

$$\sqrt{2} \sin x, \sqrt{2} \sin x \quad (1)$$

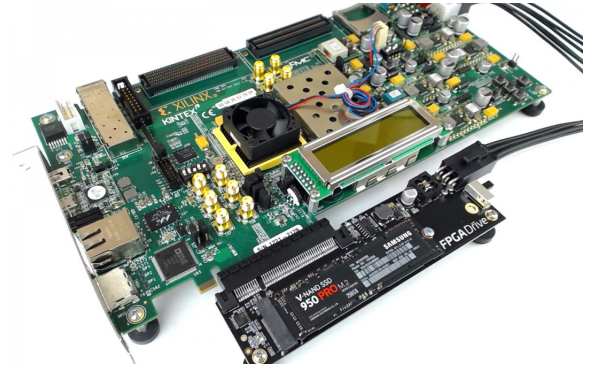


Fig. 1. FPGA Development Board

TABLE I  
TEST TABLE

laign	xu	thsi	is	a
table	table	is	goodf	make
a	table	from	this q	web
is	essay	and	fast	!!!

## III. METHODOLOGY

## IV. GPU CONTRIBUTION FOR DATA VISUALIZATION

## V. CONCLUSION

The conclusion goes here.

## REFERENCES

- [1] S. Ramírez-Gallego, S. García, H. Mouriño-Talín, D. Martínez-Rego, V. Bolón-Canedo, A. Alonso-Betanzos, J. M. Benítez, and F. Herrera, "Data discretization: taxonomy and big data challenge," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 6, no. 1, pp. 5–21, 2016.
- [2] J. Fung, F. Tang, and S. Mann, "Mediated reality using computer graphics hardware for computer vision," in *Wearable Computers, 2002.(ISWC 2002). Proceedings. Sixth International Symposium on*. IEEE, 2002, pp. 83–89.
- [3] NVIDIA, "Cuda c programming guide," 2013.
- [4] I. Icke and A. Rosenberg, "Multi-objective genetic programming for visual analytics," in *European Conference on Genetic Programming*. Springer, 2011, pp. 322–334.
- [5] E. S. Larsen and D. McAllister, "Fast matrix multiplies using graphics hardware," in *Proceedings of the 2001 ACM/IEEE conference on Supercomputing*. ACM, 2001, pp. 55–55.
- [6] J. D. Hall, N. A. Carr, and J. C. Hart, "Cache and bandwidth aware matrix multiplication on the gpu," 2003.
- [7] J. Bolz, I. Farmer, E. Grinspun, and P. Schröder, "Sparse matrix solvers on the gpu: conjugate gradients and multigrid," in *ACM Transactions on Graphics (TOG)*, vol. 22, no. 3. ACM, 2003, pp. 917–924.
- [8] K. Moreland and E. Angel, "The fft on a gpu," in *Proceedings of the ACM SIGGRAPH/EUROGRAPHICS conference on Graphics hardware*. Eurographics Association, 2003, pp. 112–119.
- [9] T. J. B. von Rymon-Lipinski and N. H. E. Kieve, "Fourier volume rendering on the gpu using a split-stream-fft," in *Vision, Modeling, and Visualization*, 2004, p. 395.
- [10] M. Hopf and T. Ertl, "Hardware based wavelet transformations," in *Vision, Modeling, and Visualization*, vol. 99, 1999, pp. 317–328.
- [11] C. Tenllado, J. Setoain, M. Prieto, L. Piñuel, and F. Tirado, "Parallel implementation of the 2d discrete wavelet transform on graphics processing units: filter bank versus lifting," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 3, pp. 299–310, 2008.
- [12] S. Guha, S. Krisnan, and S. Venkatasubramanian, "Data visualization and mining using the gpu," *Tutorial at ACM SIGKDD*, vol. 5, 2005.
- [13] H. Liu, F. Hussain, C. L. Tan, and M. Dash, "Discretization: An enabling technique," *Data mining and knowledge discovery*, vol. 6, no. 4, pp. 393–423, 2002.