XPS: EXPL: Scalable distributed GPU computing for extremely high-dimensional optimization

Alberto Cano
Department of Computer Science
Virginia Commonwealth University
acano@vcu.edu

I. Project Summary

High-dimensional optimization is a challenging research problem and a computationally intensive task which requires innovative approaches from parallel and distributed computing perspectives. Extremely high-dimensional optimization involves very large problem dimensionalities with millions of variables. However, proposals to date limited the problem dimensionality to a few million variables due to the constraints in memory and computational resources in traditional single GPU computing. This transformative research project advances within the field of efficient, scalable and distributed GPU computing to increase the dimensionality to up to 100 million variables. The problem decomposition into multiple devices allows for the efficient use of the memory and computational resources, while avoiding synchronization delays or heavy data transfer overheads. It involves heterogeneous compute units to combine both the CPU and GPU resources distributed in a networked cluster, and presents a methodology to automatically balance the workload among devices with different capabilities. Developed models will be evaluated in two applications: the IEEE large scale global optimization functions, and increasing the performance of the video motion capture problem.

II. Intellectual Merit

The main contribution of this proposal is to develop a distributed computing model for extremely high-dimensional optimization which divides the computationally intensive problem into multiple heterogeneous CPUs and GPUs. It addresses the challenge of increasing the problem dimensionality 100 times larger than current approaches using single GPU computing. The proposed methodologies efficiently allocate data and millions of threads to collaborative resolve the optimization problem. This research also translates into an application in a real-world problem for speeding up the video motion capture analysis and scale its performance to real-time processing. The PI has demonstrated a strong background and trajectory in the area as well as high performance in the quality and number of publications per funding dollar, which increases his merit to conduct transformative and innovative research.

III. Broader Impacts

The results of this project will be published in a number of top quality publications in journals and international conferences. It involves the training of graduate students in the High-Performance Distributed Systems course, to connect teaching and research experiences conducted in the High-Performance Data Mining Laboratory at the Virginia Commonwealth University. The university offers a unique Dual Ph.D. program in Computer Science with the University of Cordoba in Spain. This research will comprise international collaborative research and the involvement of minority Hispanic students, currently underrepresented in science and engineering. Moreover, the student population at VCU is predominantly African American, of whom a more than half are female. This project would improve the participation of underrepresented minorities and our recruitment efforts.

References

[1] S. Shan and G. Wang, "Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions," *Structural and Multidisciplinary Optimization*, vol. 41, no. 2, pp. 219–241, 2010.

[2] A. Holzinger, M. Dehmer, and I. Jurisica, "Knowledge Discovery and interactive Data Mining in Bioinformatics - State-of-the-Art, future challenges and research directions," *BMC Bioinformatics*, vol. 15, no. S-6, 2014.

[3] S. Bahmann and J. Kortus, "EVO Evolutionary algorithm for crystal structure prediction," *Computer Physics Communications*, vol. 184, no. 6, pp. 1618–1625, 2013.

[4] D. Niu, Y. Wang, and D. D. Wu, "Power load forecasting using support vector machine and ant colony optimization," *Expert Systems with Applications*, vol. 37, no. 3, pp. 2531–2539, 2010.

[5] J. Dudley, E. Schadt, M. Sirota, A. Butte, and E. Ashley, "Drug discovery in a multidimensional world: Systems, patterns, and networks," *Journal of Cardiovascular Translational Research*, vol. 3, no. 5, pp. 438–447, 2010.

[6] W. Shi, G. Wahba, R. A. Irizarry, H. C. Bravo, and S. J. Wright, "The partitioned LASSO-patternsearch algorithm with application to gene expression data," *BMC Bioinformatics*, vol. 13, 2012.

[7] K. Lange, J. C. Papp, J. S. Sinsheimer, and E. M. Sobel, "Next-Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data," *Annual Review of Statistics and Its Application*, vol. 1, pp. 279–300, 2014.

[8] Y. Tan and K. Ding, "A Survey on GPU-Based Implementation of Swarm Intelligence Algorithms," *IEEE Transactions on Cybernetics, In Press*, 2015.

[9] S. Mahdavi, M. E. Shiri, and S. Rahnamayan, "Metaheuristics in large-scale global continues optimization: A survey," *Information Sciences*, vol. 295, pp. 407–428, 2015.

[10] A. LaTorre, S. Muelas, and J.-M. Pena, "A comprehensive comparison of large scale global optimizers," *Information Sciences*, vol. 316, pp. 517–549, 2015.

[11] X. Li, K. Tang, M. N. Omidvar, Z. Yang, and K. Qin, "Benchmark Functions for the CEC'2013 Special Session and Competition on Large-Scale Global Optimization," Evolutionary Computation and Machine Learning Group, RMIT University, Australia, Tech. Rep., 2013.

[12] D. Hill, A. I. Rapoport, R. F. Lehming, and R. K. Bell, "Changing U.S. Output of Scientific Articles: 1988-2003," *National Science Foundation, Division of Science Resources Statistics, NSF 07-320*, 2007. [Online]. Available: http://www.nsf.gov/statistics/nsf07320/

[13] J.-M. Fortin and D. J. Currie, "Big science vs. little science: How scientific impact scales with funding," *PLoS ONE*, vol. 8, 2013.

[14] C. Bloch and M. P. Sorensen, "The size of research funding: Trends and implications," *Science and Public Policy*, vol. 42, no. 1, pp. 30–43, 2015.

[15] A. Cano, A. Zafra, and S. Ventura, "Solving classification problems using genetic programming algorithms on GPUs," in *Proceedings of the 5th International Conference on Hybrid Artificial*

Intelligent Systems (HAIS). Lecture Notes in Computer Science, vol. 6077 LNAI, no. PART 2, 2010, pp. 17–26.

[16] A. Cano, A. Zafra, and S. Ventura, "Speeding up the evaluation phase of GP classification algorithms on GPUs," *Soft Computing*, vol. 16, no. 2, pp. 187–202, 2012.

[17] A. Cano, A. Zafra, and S. Ventura, "Weighted Data Gravitation Classification for Standard and Imbalanced Data," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1672–1687, 2013.

[18] A. Cano, A. Zafra, and S. Ventura, "Parallel evaluation of Pittsburgh rule-based classifiers on GPUs," *Neurocomputing*, vol. 126, pp. 45–57, 2014.

[19] A. Cano, S. Ventura, and K. Cios, "Scalable CAIM discretization on multiple GPUs using concurrent kernels," *Journal of Supercomputing*, vol. 69, no. 1, pp. 273–292, 2014.

[20] A. Cano, S. Ventura, and K. Cios, "Multi-Objective Genetic Programming for Feature Extraction and Data Visualization," *Soft Computing*, vol. In press, 2015.

[21] A. Cano, J. M. Luna, and S. Ventura, "High Performance Evaluation of Evolutionary-Mined Association Rules on GPUs," *Journal of Supercomputing*, vol. 66, no. 3, pp. 1438–1461, 2013.

[22] J. Luna, A. Cano, M. Pecheniskiy, and S. Ventura, "Speeding-up Association Rule Mining with Inverted Index Compression," *IEEE Transactions on Cybernetics*, vol. In Press, 2016.

[23] A. Cano, J. Olmo, and S. Ventura, "Parallel Multi-Objective Ant Programming for Classification Using GPUs," *Journal of Parallel and Distributed Computing*, vol. 73, no. 6, pp. 713–728, 2013.

[24] A. Cano, E. Yeguas-Bolivar, R. Munoz-Salinas, R. Medina-Carnicer, and S. Ventura, "Parallelization Strategies for Markerless Human Motion Capture," *Journal of Real-Time Image Processing*, vol. In press, 2015.

[25] A. Cano, A. Zafra, and S. Ventura, "Speeding up multiple instance learning classification rules on GPUs," *Knowledge and Information Systems*, vol. 44, no. 1, pp. 127–145, 2015.

[26] A. Cano and S. Ventura, "GPU-parallel Subtree Interpreter for Genetic Programming," in *Proceedings of the Conference on Genetic and Evolutionary Computation*, 2014, pp. 887–894.

[27] A. Cano, J. Luna, E. Gibaja, and S. Ventura, "LAIM discretization for multi-label data," *Information Sciences*, vol. 330, pp. 370–384, 2016.

[28] R. Regis, "Evolutionary Programming for High-Dimensional Constrained Expensive Black-Box Optimization Using Radial Basis Functions," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 326–347, 2014.

[29] S. Chen, J. Montgomery, and A. Bolufe-Rohler, "Measuring the curse of dimensionality and its effects on particle swarm optimization and differential evolution," *Applied Intelligence*, vol. 42, no. 3, pp. 514–526, 2015.

[30] X. Li, K. Tang, M. N. Omidvar, Z. Yang, and K. Qin, "Benchmark Functions for the CEC'2013 Special Session and Competition on Large-Scale Global Optimization," 2013.

[31] M. Lastra, D. Molina, and J. M. Benitez, "A high performance memetic algorithm for extremely high-dimensional problems," *Information Sciences*, vol. 293, pp. 35–58, 2015.

[32] C. García-Martínez, F. J. Rodríguez-Díaz, and M. Lozano, "Role differentiation and malleable mating for differential evolution: an analysis on large-scale optimisation," *Soft Computing*, vol. 15, no. 11, pp. 2109–2126, 2011.

[33] D. Swagatam and P. N. Suganthan, "Differential Evolution: a survey of the state-of-the-art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, 2011.

[34] J. Liu and K. Tang, "Scaling Up Covariance Matrix Adaptation Evolution Strategy Using Cooperative Coevolution," in *Proc. of the Intelligent Data Engineering and Automated Learning*, vol. 8206 LNCS, 2013, pp. 350–357.

[35] Y. Wang and B. Li, "Two-stage based ensemble optimization for large-scale global optimization," in *Proc. of the IEEE Congress on Evolutionary Computation*, 2010, pp. 1–8.

[36] P. Korošec, K. Tashkova, and J. Šilc, "The Differential Ant-Stigmergy Algorithm for Large-Schale Global Optimization," in *Proc. of the IEEE World Congress on Computational Intelligence*, 2010, pp. 18–23.

[37] R. Cheng and Y. Jin, "A competitive swarm optimizer for large sclae optimization," *IEEE Transactions on Cybernetics*, vol. 45, no. 2, pp. 191–204, 2015.

[38] J. Fan, J. Wang, and M. Han, "Cooperative coevolution for large-sclae optimization based on kernel fuzzy clustering and variable trust region methods," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 4, pp. 829–839, 2014.

[39] A. LaTorre, S. Muelas, and J. Peña, "Large scale global optimization: experimental results with MOS-based hybrid algorithms," in *Proc. of the IEEE Congress on Evolutionary Computation*, 2013, pp. 1–8.

[40] D. Molina, M. Lozano, and F. Herrera, "MA-SW-Chains: memetic algorithm based on local search chains for large scale continuous global optimization," in *Proc. of the IEEE Congress on Evolutionary Computation*, 2010, pp. 1–8.

[41] M. A. Potter and K. A. De Jong, "Cooperative coevolution: an architecture for evolving coadapted subcomponents," *Evolutionary Computation*, vol. 8, no. 1, pp. 1–29, 2000.

[42] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative coevolution with differential grouping for large scale optimization," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 3, pp. 378–393, 2014.

[43] Z. Yang, K. Tang, and X. Yao, "Large scale evolutionary optimization using cooperative coevolution," *Information Sciences*, vol. 178, no. 15, pp. 2985–2999, 2008.

[44] F. V. den Bergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 225–239, 2004.

[45] X. Li and X. Yao, "Cooperatively coevolving particle swarms for large scale optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 2, pp. 210–224, 2012.

[46] A. Cano, C. García-Martínez, and S. Ventura, "Extremely High-dimensional Optimization with MapReduce: Scaling Functions and Algorithm," *Information Sciences*, pp. 1–31, 2016.

[47] J. Dean and S. Ghemawat, "MapReduce: A flexible data processing tool," *Communications of the ACM*, vol. 53, no. 1, pp. 72–77, 2010.

[48] J. Gomez-Luna, J. M. Gonzalez-Linares, J. I. Benavides, and N. Guil, "Performance models for asynchronous data transfers on consumer graphics processing units," *Journal of Parallel and Distributed Computing*, vol. 72, no. 9, pp. 1117–1126, 2012.

[49] R. K. Thompson and A. H. Wright, "Additively decomposable fitness functions," Dept. of Computer Science, University of Montana, Tech. Rep., 1996.

[50] K. Tang, X. Li, P. Suganthan, Z. Yang, and T. Weise, "Benchmark Functions for the CEC'2010 Special Session and Competition on Large-Scale Global Optimization," Nature Inspired Computation and Applications Laboratory, USTC, China, Tech. Rep., 2009.

[51] J. Dean and S. Ghemawat, "MapReduce: A Flexible Data Processing Tool," *Communications of the ACM*, vol. 53, no. 1, pp. 72–77, 2010.

[52] M. A. Franco and J. Bacardit, "Large-scale experimental evaluation of GPU strategies for evolutionary machine learning," *Information Sciences*, vol. 330, pp. 385–402, 2016.

[53] NVIDIA Corporation, "NVIDIA CUDA Programming and Best Practices Guide," 2016, http://www.nvidia.com/cuda.

[54] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, pp. 90–126, 2006.

[55] B. Rymut and B. Kwolek, "GPU-supported object tracking using adaptive appearance models and particle swarm optimization," in *Proc. of the international conference on Computer vision and graphics: Part II*, ser. ICCVG'10, 2010, pp. 227–234.

[56] T. Krzeszowski, B. Kwolek, and K. Wojciechowski, "GPU-accelerated tracking of the motion of 3D articulated figure," in *Computer vision and graphics*, 2010, pp. 155–162.

[57] L. Mussi, S. Ivekovic, and S. Cagnoni, "Markerless articulated human body tracking from multi-view video with GPU-PSO," in *Evolvable Systems: From Biology to Hardware*, ser. Lecture Notes in Computer Science, G. Tempesti, A. Tyrrell, and J. Miller, Eds., 2010, vol. 6274, pp. 97–108.

[58] B. Rymut, B. Kwolek, and T. Krzeszowski, "GPU-Accelerated Human Motion Tracking Using Particle Filter Combined with PSO," *Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science*, vol. 8192, pp. 426–437, 2013.

[59] Y. Zhou and Y. Tan, "GPU-based parallel particle swarm optimization," in *IEEE Congress on Evolutionary Computation*, 2009, pp. 1493–1500.

[60] R. Ugolotti, Y. S. Nashed, P. Mesejo, pela Ivekovi, L. Mussi, and S. Cagnoni, "Particle swarm optimization and differential evolution for model-based object detection," *Applied Soft Computing*, vol. 13, no. 6, pp. 3092 – 3105, 2013.

[61] S. Ventura, C. Romero, A. Zafra, J. A. Delgado, and C. Hervas, "JCLEC: A Java Framework for Evolutionary Computation," *Soft Computing*, vol. 12, no. 4, pp. 381–392, 2007, http://jclec.sourceforge.net.

[62] A. Cano, J. M. Luna, A. Zafra, and S. Ventura, "A Classification Module for Genetic Programming Algorithms in JCLEC," *Journal of Machine Learning Research*, vol. 16, pp. 491–494, 2015.