A 32x32 Time-Domain Wavefront Computing Accelerator for Path Planning and Scientific Simulations

Chengshuo Yu[1,2], Yuqi Su[1], Jaeho Lee[1], Kevin Chai[2], Bongjin Kim[1,3]

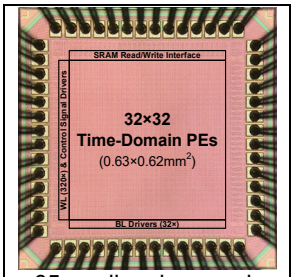
[1]Nanyang Technological University, [2]A*STAR, Singapore
[3]now with University of California, Santa Barbara, CA

Recent hardware accelerators based on FPGA [1-2] and ASIC [3-4] have demonstrated high throughput and energy-efficient processing of path planning for applications, including the manipulation of arm-robots on a high-resolution map [1] and the autonomous navigation of battery-powered micro-robots [4]. The accelerators have focused on demonstrating the modified shortest path planning algorithms, such as extended-stride A* [1] and dual-tree rapid-exploring random tree (RRT) [4], and achieved much higher throughput and energy efficiency (two-to-three orders of magnitude higher) than traditional implementations with CPU and GPU. A two-dimensional processing element (PE) array [5] was proposed as an alternative computing paradigm for searching the shortest path based on the expansion of wavefront in the time-domain. The time-domain delay accumulation was used for finding the shortest paths by accumulating time delays while propagating through the PEs that correspond to the pixels in a two-dimensional searching space. However, it still follows the existing A* algorithm using an extra gradient control circuit with resistor ladders to estimate the remaining distances, resulting in inaccurate results (i.e. sub-optimal paths) with extra hardware and energy overhead.

In this work, we propose a true wavefront computing accelerator that finds the shortest paths using time-domain wavefront propagations only. It does not require estimating distances, and hence, we can find all the shortest paths from a start point to an arbitrary location purely based on the accumulated time delays. Time-domain PEs in a two-dimensional array are connected via the rising-edge signal repeaters (i.e. edge repeaters), and they form either King's or lattice graphs for different purposes. The King's graph with eight interconnects per PE finds the shortest path, while the simpler lattice graph having four interconnects can be used for solving a complex maze problem. We demonstrate Huygen's principle in time-domain by using single or multiple sources (i.e. start points) for circular or planar wavefront propagations.

Fig. 1 shows the concept of wavefront computing accelerator with time-domain PEs connected based on King's graph configuration. Each PE and eight input and output connections correspond to a vertex, and eight interconnects (per vertex) of the King's graph. A PE comprises an edge detector at the center, eight edge repeaters directed towards eight directions (i.e. four cardinal: N, W, S, E, and four diagonal: NW, NE, SE, SW directions), and 10bit SRAM. The edge detector asserts the 'Lock' signal based upon the first rising edge input from the eight-directions. It then triggers the edge repeaters, and the next rising edge signals are generated and distributed towards three (for lattice graph mode) or five (for King's graph mode) neighbor PEs. Note that the delay can be programmed by adjusting supply levels of voltage-controlled delay lines (VCDLs) of each cardinal and diagonal direction edge repeaters. Two SRAM bitcells at the top of a 10bit SRAM column are dedicated to programming each PE. We can program an arbitrary PE as a start (Start=1), an active path (En=1), or an obstacle (En=0). The remaining eight SRAM bitcells are used for storing the triggered output directions for readout while tracing back the shortest path. The graph mode is reconfigured by 'Graph' signal, enabling or disabling diagonal repeaters.

Fig. 2 (top) describes time-domain PE delay components for both cardinal and diagonal directions. Each delay path (i.e. a delay in 'Lock' signal generations from PE to PE) comprises five combinational logic gates, two latches, and a single VCDL. The delay in cardinal (t_C) or diagonal (t_D) direction is programmed by supply voltages (VDD_C for cardinal and VDD_D for diagonal direction) of their VCDLs. Fig. 2 (bottom-left) shows the simulated delays and their ratios while sweeping VDD_D from 0.8 to 1.2V (with fixed VDD_C=1.2V). Here, we choose VDD_D=1V, where t_D is 0.75ns and the delay ratio is 1.5 (i.e. t_C is 0.5ns at VDD_C=1.2V). Fig. 2 (bottom-right) shows the measured wavefront propagation using two maps with obstacles. The gradient color is determined based on each pixel's shortest path after tracing back to a start point. It is a unique property of our work that we can find 32x32 shortest paths for all pixels in the map, while conventional algorithms can only find a single shortest path by designating a goal regardless of the map size. Using the gradient color and the shape of the wavefront, we estimated delays from the start to the farthest goal on the map.



65nm die micrograph

Fig. 3 shows the measured circular and plane wavefront propagations using a single (for circular wavefront) and ten parallel start points (for plane wavefront) in a map with no obstacles. It demonstrates the Huygen's principle that describes a wavefront propagation (i.e. every point on a wavefront is the source of spherical wavelets and their tangents form a new wavefront). Note that the wavefront shape (circle or plane) is determined based on the number of sources and their formation. Each time-domain PE works as a source of the spherical wavelet moving forward direction. As a result, we can see either a circular or a plane wavefront, as shown in the measurement results of Fig. 3 (top).

Fig. 4 shows the measured shortest path searching problem using both lattice and King's graph. A 32x32 map with obstacles in Fig. 4 (left) is programmed into a 32x32 time-domain PE array while the chip is reset to 1. After programming the map, we started an asynchronous time-domain wavefront propagation by disabling the reset signal (i.e. reset=0). The measured gradient maps differ between the lattice and the King's graph in their wavefront shapes. The estimated delays based on the measured shortest paths are 36ns for the lattice graph and 30.2ns for King's graph. The measured shortest paths and their delays clearly show the King's graph's advantage in searching the fastest route.

Fig. 5 shows another application of the wavefront computing accelerator and highlights its scalability. Using the lattice graph with only four cardinal directions, we can solve a complex maze having a single start and an exit. To solve a 64x64 maze problem, we first map it into four 32x32 sub-mazes. Each cycle, we solve a sub-maze and find a solution (or multiple solutions) from its start to an exit (or multiple exits). Note that the exits from the previous sub-maze cycles become new starts in the next cycles. We can revisit the same sub-maze we solved previously, as shown in cycle #3, #4, and #7. After seven cycles, we were able to find a unique maze solution by combining the results from the sub-maze operations.

Fig. 6 (top) shows the measured dynamic path searching problem. Thanks to the low-latency in the time-domain wavefront propagation (<50ns), we can repeatedly search the shortest paths with minimal overhead in latency and energy consumption. We can observe the selected paths differ from cycle to cycle due to the updated locations of start points and obstacles. Fig. 6 (bottom-left) demonstrates wavefront propagations when there are double sources at the center of the open space and a source located in double boxes with small openings. Fig. 6 (bottom-right) provides the performance summary and comparison with state-of-the-art path planning accelerators. A PE occupies 384μm^2 and consumes 0.758pJ using 1.2V/1V supplies at the King's graph mode. A 32x32 PE array consumes 776pJ per task, and the estimated latency is <50ns.

References:

[1] A. Kosuge et al., "A 1200x1200 8-Edges/Vertex FPGA-based Motion-Planning Accelerator for Dual-Arm-Robot Manipulation Systems," IEEE Symp. VLSI Circuits, Jun. 2020.

[2] S. Zhou et al., "High-Throughput and Energy-Efficient Graph Processing on FPGA," IEEE International Symposium on Field-Programmable Custom Computing Machines, 2016.

[3] Y. Kim et al., "A 0.55V 1.1mW Artificial-Intelligence Processor with PVT Compensation for Micro Robots," ISSCC, Feb. 2016.

[4] C. Chung et al., "A 1.5μJ/Task Path-Planning Processor for 2D/3D Autonomous Navigation of Micro Robots," ISSCC, Feb. 2020.

[5] L. Everson et al., "A 40x40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control," ISSCC, 2019.

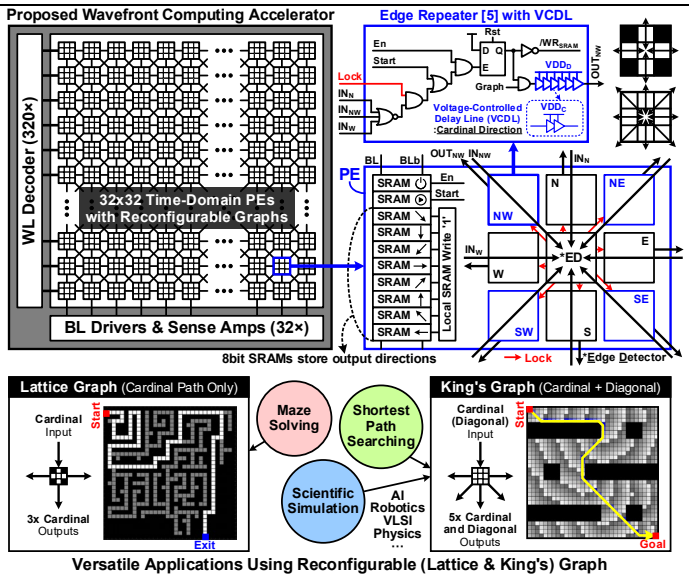


Fig. 1. Concept of the proposed wavefront computing accelerator based on King's graph time-domain processing elements.

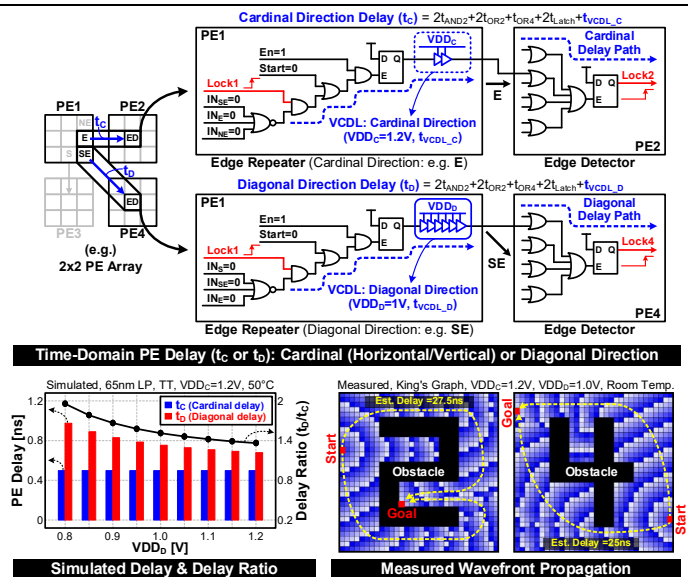


Fig. 2. Proposed circuits (top) and simulated results (bottom-left) of time delays, and wavefront propagation examples (bottom-right).

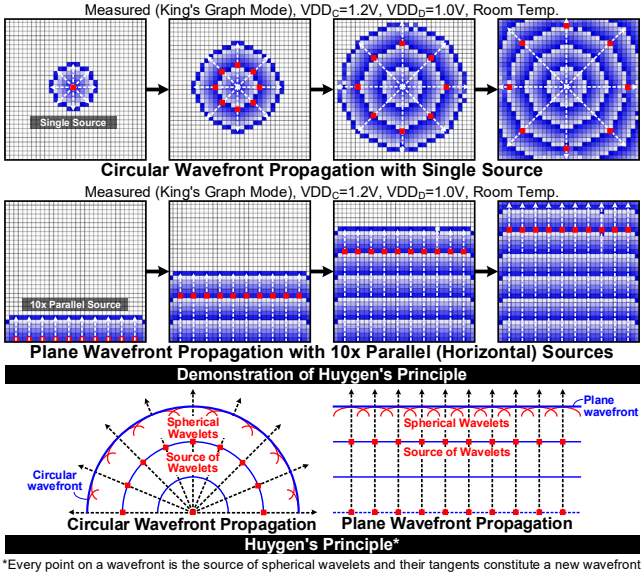


Fig. 3. Demonstrated Huygen's principle. The shape (circle or plane) is determined based on the number of sources and their formation.

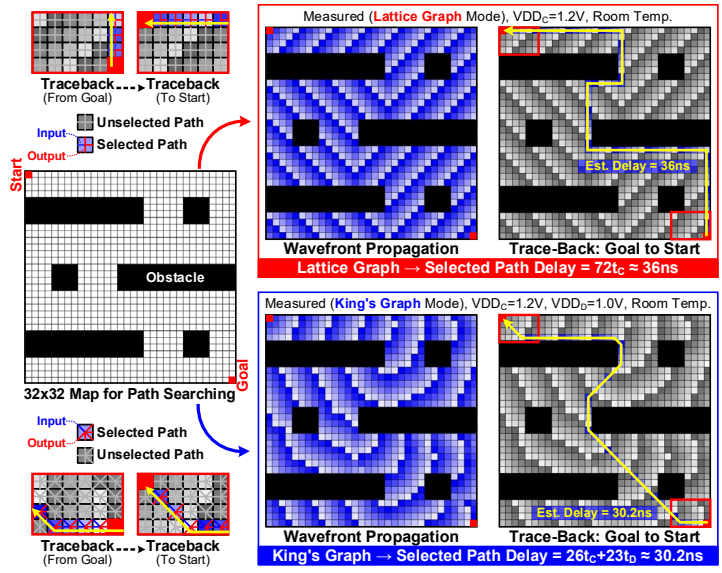


Fig. 4. Measured shortest path searching problem using lattice graph (top) and King's graph (bottom).

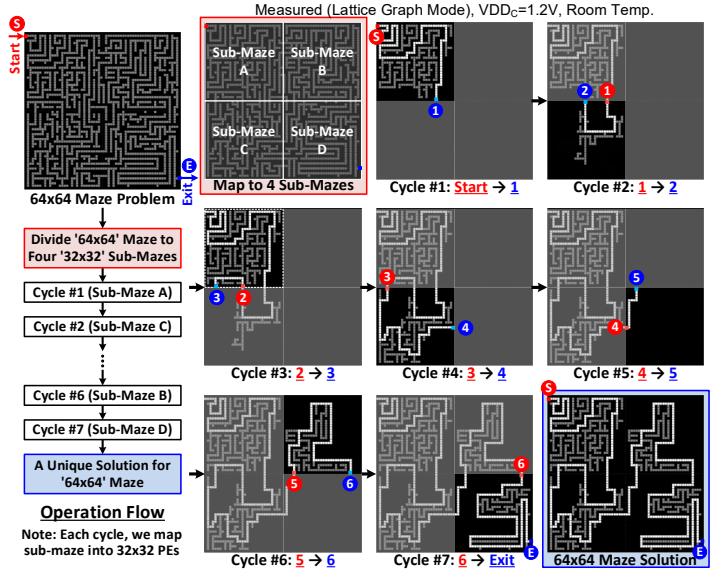
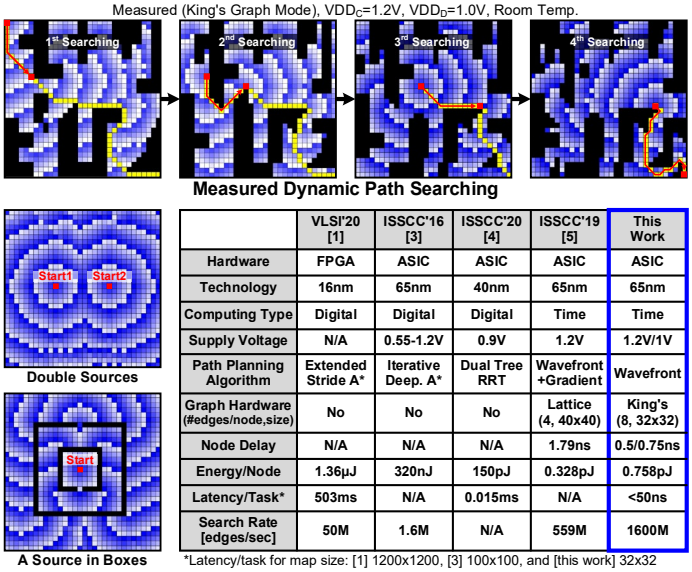


Fig. 5. Measured 64x64 maze solving problem using four 32x32 sub-mazes mapped into 32x32 time-domain PEs.



	VLSI'20 [1]	ISSCC'16 [3]	ISSCC'20 [4]	ISSCC'19 [5]	This Work
Hardware	FPGA	ASIC	ASIC	ASIC	ASIC
Technology	16nm	65nm	40nm	65nm	65nm
Computing Type	Digital	Digital	Digital	Time	Time
Supply Voltage	N/A	0.55-1.2V	0.9V	1.2V	1.2V/1V
Path Planning Algorithm	Extended Stride A*	Iterative Deep. A*	Dual Tree RRT	Wavefront +Gradient	Wavefront
Graph Hardware (#edges/node,size)	No	No	No	Lattice (4, 40x40)	King's (8, 32x32)
Node Delay	N/A	N/A	N/A	1.79ns	0.5/0.75ns
Energy/Node	1.36μJ	320nJ	150pJ	0.328pJ	0.758pJ
Latency/Task*	503ms	N/A	0.015ms	N/A	<50ns
Search Rate [edges/sec]	50M	1.6M	N/A	559M	1600M

*Latency/task for map size: [1] 1200x1200, [3] 100x100, and [this work] 32x32

Fig. 6. Measured dynamic path searching (top), wave propagations (bottom-left), and comparison table (bottom-right).