A 40×40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

Luke R. Everson, Sachin S. Sapatnekar, Chris H. Kim University of Minnesota, Minneapolis, MN

© 2019 IEEE **International Solid-State Circuits Conference**



- Vertex Details - Edge Details Applications Conclusion

© 2019 IEEE **International Solid-State Circuits Conference**

Outline

- Background: Path Planning Algorithms Time-Based A* ASIC
 - Time-Based Primer
 - Top Level Design
- 65nm Test Chip Results





• Formal definition:

© 2019 IEEE **International Solid-State Circuits Conference**



- Graph: set of "objects" and their "connections"
 - $-G = (V, E), V = \{v_1, v_2, ..., v_n\}, E = \{e_1, e_2, ..., e_m\}$
 - V: set of vertices (nodes), E: set of edges (links, arcs)
 - Directed graph: $e_k = (v_i, v_i)$
 - Undirected graph: $e_k = \{v_i, v_i\}$
 - Weighted graph: w: $E \rightarrow R$, w(e_k) is the "weight" of e_k.

Network Analysis



[Wikipedia.org]

© 2019 IEEE International Solid-State Circuits Conference

Why Graphs? **Routing/Path Planning**

[yuchsia.github.io]

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control



Self-Driving Cars



[techrepublic.com]

4 of 62

Graph Algorithms- Breadth First Search • Explores search area equally in all directions Adds new neighbors to queue and visits in order added Simple to implement in software







Graph Algorithms- Breadth First Search Explore entire graph from Source (S) • **Visited** = {**S**} • Queue = {A,D}

© 2019 IEEE International Solid-State Circuits Conference







Graph Algorithms- Breadth First Search Explore entire graph from Source (S) • Visited = {S,A} • **Queue = {D,B}**

© 2019 IEEE International Solid-State Circuits Conference







Graph Algorithms- Breadth First Search Explore entire graph from Source (S) • Visited = {S,A,D} • **Queue = {B,E}**

© 2019 IEEE International Solid-State Circuits Conference





Graph Algorithms- Breadth First Search Explore entire graph from Source (S) • Visited = {S,A,D,B} • **Queue = {E,C}**

© 2019 IEEE International Solid-State Circuits Conference







Graph Algorithms- Breadth First Search Explore entire graph from Source (S) • Visited = {S,A,D,B,E} • **Queue = {C,T}**

© 2019 IEEE **International Solid-State Circuits Conference** 2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control







10 of 62

Graph Algorithms- Breadth First Search Explore entire graph from Source (S) • Visited = {S,A,D,B,E,C} • Queue = $\{T\}$

© 2019 IEEE International Solid-State Circuits Conference





Graph Algorithms- Breadth First Search Explore entire graph from Source (S) • Visited = {S,A,D,B,E,C,T} • **Queue** = {}

© 2019 IEEE **International Solid-State Circuits Conference**







• Similar to BFS when edges have weights neighbor first Keeps track of

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- Dijkstra's Dijkstra's BFS Uniform cost search

Search "cheapest" distance FROM start



Find shortest path from Source (S) to Target (T) • Visited= $\{S_n\}$ • Cost=0• **PriorityQueue=** $\{A_{s_1}, D_{s_1}\}$

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- Dijkstra's

Store Parent Node for traceback

Find shortest path from Source (S) to Target (T) • Visited= $\{S_0, A_{S_1}\}$ • Cost=1 • **PriorityQueue=** $\{D_{S1}, B_{A2}\}$

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- Dijkstra's

Find shortest path from Source (S) to Target (T) • Visited= $\{S_0, A_{S_1}, D_{S_1}\}$ • Cost=1 • **PriorityQueue=**{ B_{A2}, E_{D4} }

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- Dijkstra's

• Cost=2

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- Dijkstra's Find shortest path from • **PriorityQueue=**{ C_{B3} , E_{D4} }

Source (S) to Target (T) • Visited= $\{S_0, A_{S_1}, D_{S_1}, B_{A_2}\}$

Find shortest path from Source (S) to Target (T) • Visited= $\{S_0, A_{S1}, D_{S1}, B_{A2}, B_{A2}, B_{A2}, B_{A2}, B_{A3}, B_{A$ C_{B3} • Cost=3• **PriorityQueue=**{ T_{C4}, E_{D4} }

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- Dijkstra's

Find shortest path from Source (S) to Target (T) • Visited= $\{S_0, A_{S1}, D_{S1}, B_{A2}, B_{A2}, B_{A2}, B_{A2}, B_{A3}, B_{A$ **C**_{B3}, **T**_{C4}} • Cost=3• **PriorityQueue=** $\{E_{D_4}\}$

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- Dijkstra's

destination • Cost(n) = F(n) + H(n)Actual Cost

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A*

- A* guides the search towards the

 - F(n) = actual distance from source
 - H(n) = heuristic that predicts the distance to target
 - H(n) = 0 for Dijkstra's
- Provides optimal path if H(n) <=

© 2019 IEEE International Solid-State Circuits Conference

• Let H(n) = ManhattanDistance • $H(n) = abs(x_i-x_T)+abs(y_i-y_T)$

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

21 of 62

Distance • Visited= $\{S_n\}$ • Cost/F(n)=0

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A* Find shortest path from Source (S) to Target (T)

- Let H(n) = Manhattan
- **PriorityQueue=** $\{A_{S_1}\}$

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

to Target (T) Distance • Visited= $\{S_n\}$ • Cost/F(n)=0

© 2019 IEEE International Solid-State Circuits Conference

Graph Algorithms- A* Find path from Source (S) • Let H(n) = Manhattan • **PriorityQueue=** $\{A_{s_1}^5\}$ C(A) = F(A) + H(A) = 1 + 4

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

Distance • Visited= $\{S_n\}$ • Cost/F(n)=0

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A* Find shortest path from Source (S) to Target (T) • Let H(n) = Manhattan

• **PriorityQueue=** $\{A_{S1}^{5}, D_{S1}^{5}\}$

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

 Find shortest path from Source (S) to Target (T) Let H(n) = Manhattan Distance • Visited= $\{S_0, A_{S_1}^5\}$ • Cost/F(n)=1• **PriorityQueue=**{ $B_{\Delta 2}^{5}$, D_{S1}^{5} }

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A*

 Find shortest path from Source (S) to Target (T) • Let H(n) = Manhattan Distance • Visited= $\{S_0, A_{S_1}, D_{S_1}, D_{S_1}, D_{S_1}, D_{S_1}\}$ • Cost/F(n)=1• **PriorityQueue=**{ B_{A2}^{5} , E_{D4}^{6} }

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A*

 Find shortest path from Source (S) to Target (T) • Let H(n) = Manhattan Distance • Visited= $\{S_0, A_{S_1}^5, D_{S_1}^5, B_{A_2}^5\}$ • Cost/F(n)=2• **PriorityQueue=**{ C_{B3}^{5} , E_{D4}^{6} }

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A*

 Find shortest path from Source (S) to Target (T) • Let H(n) = Manhattan Distance • Visited= $\{S_0, A_{S_1}^5, D_{S_1}^5, B_{A_2}^5, B_{A$ C_{B3}^{5} • Cost/F(n)=3• **PriorityQueue=**{ T_{C4}^4 , E_{D4}^6 }

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A*

 Find shortest path from Source (S) to Target (T) • Let H(n) = Manhattan Distance • Visited= $\{S_0, A_{S_1}, D_{S_1}, B_{A_2}, B_{$ C_{B3}^{5}, T_{C4}^{4} • Cost/F(n)=3• **PriorityQueue=** $\{E_{D_4}^6\}$

© 2019 IEEE **International Solid-State Circuits Conference**

Graph Algorithms- A*

- 4-neighbor grid • Directed edges

 - on/off control
 - edge weight
 - runtime

Graph Structure- This Work

Each direction binary 4 bit digital control Programmable at

Outline • Background: Path Planning Algorithms Time-Based A* ASIC **– Time-Based Primer** - Top Level Design - Vertex Details - Edge Details 65nm Test Chip Results Applications Conclusion

© 2019 IEEE **International Solid-State Circuits Conference**

© 2019 IEEE **International Solid-State Circuits Conference**

Digital Computing

Advantages of digital arithmetic: – Binary representation – Less "buy-in" required – Existing IP for rapid SoC development – No calibration

Time-based Computing

Advantages of time-based circuits: Compact area Low power consumption

High precision tunability

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

33 of 62

40×40 A* ASIC

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

• Time \propto Distance Each vertex stores first input **Directly readout** shortest path Multiple start points

40×40 A* ASIC- Vertex

40×40 A* ASIC- Vertex

40×40 A* ASIC - Edge

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

Outline Background: Path Planning Algorithms Time-Based A* ASIC 65nm Test Chip Results Applications • Conclusion

© 2019 IEEE **International Solid-State Circuits Conference**

Test Setup Custom Test PCB

Edge Delay Linearity

Edge Code

STEP

DNI

[T_{STEP}]

Z

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

Architecture	This Work	FPGA [4]	μProcessor	CPU [5]			
Product	ASIC	Xilinx Virtex	ARM Cortex-M0	Intel Xeon E5630	NVIC		
Technology	65nm	20nm	40nm	32nm			
Voltage	1.2V		1.1V	0.7-1.35V			
Peak Power	26.4mW	24.22W	127µW	20W/core			
Throughput [MTEPS]	559	731	5.34*10 ⁻⁴	0.83			
Energy per Node	0.328pJ*	33nJ	89.1nJ	24.1µJ			
Normalized Energy	1 x	10 ⁵	2.7x10 ⁵	1.19x10⁶			
^{55%} from SRAM Program (does not include cache access energy) Energy/Node=Unit Delay*Unit Power							

[4]S. Zhou, et al, "High-Throughput and Energy-Efficient Graph Processing on FPGA," IEEE International Symp. on Field-Programmable Custom Computing Machines, pp. 103-110, 2016. [5] Y. Zhou and J. Zeng, "Massively Parallel A* Search on a GPU," Conference on Artificial Intelligence *(AAAI)*, pp. 1248-1254, 2015.

Comparison Table

MTEPS = Million Traversed Edges Per Second

Outline Background: Path Planning Algorithms Time-Based A* ASIC 65nm Test Chip Results Applications Conclusion

© 2019 IEEE International Solid-State Circuits Conference

• Applications: [Wikipedia]

© 2019 IEEE **International Solid-State Circuits Conference**

Voronoi Diagrams

۲

•

۲

- Segmentation of plane such that distance to seeds is maximized

 - kNN classification
 - Biological structures (bone and cells)
 - Computational fluid dynamics meshes
 - Autonomous vehicles

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

۲

Collision Avoidance

Collision Avoidance

Collision Avoidance

© 2019 IEEE International Solid-State Circuits Conference

Path Planning - Dijkstra's

5				
0	1	2	3	
1	2	3	4	
2	3	4	5	
3	4	5	6	
4	5	6	7	
5	6	7	8	

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

All diagonal paths are equal

© 2019 IEEE International Solid-State Circuits Conference

Path Planning - Dijkstra's

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

Paths above blockage dominate paths below

49 of 62

.8

© 2019 IEEE International Solid-State Circuits Conference

ath Planning - 1					
	TVV S 2		VTVV 5.4	VTV 6.8	
	0		2	3	
			6.8		
			3		
	11	9.4	8		9.
	6	5	4		
	12	10.6			10.
	7	6			
	13.2	11.6	12.2	11.4	10.
	8	7	8	9	
	14.2	12.4	12	11.4	1
	9	8	9	10	

Path Planning Finding shortest path w/o gradient 0.9V**0.9V** 0.7

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

Finding shortest path w/ gradient 0.7

							L_ L_ L_ L_ L_	
		LLLLL			LLLL			LLI
5			La La La La La	La La La La La	La La La La La 1	La La La La La I	- La L <mark>a La</mark> La	L. L. 1_
-								L- L- 1
		LLLLL						LLI
10				to to to to to	to to to to to 1	La La La La La t	- La La La La	t_ t_ 1_
		to to to to to to	La La La La La	to to to to to	to to to to to	La La La La La L	- La La La La	t. t. 1
							_ L_ L_ L_ L_	
				La La La La La 1 1 1 1 1		La La La La La L		
15								
IJ	1 -1 -1 -1 -1 -1		La La La La La	to to to to to	La La La La La 1	La La La La La L	- t- t- t- t-	L L 1
		- La La La La La	La La La La La		to to to to to !	La La La La La L		L- L- 1
								L_ L_ 1
20						L. L. L. L. L. L.		
20								
		LLLLL						LLI
	للے لے لے لے ا			L. L. L. L. L.	L. L. L. L. L. !	L. L. L. L. L. I		L L 1
		to to to to to	Le Le Le Le Le	Le Le Le Le Le	to to to to to	La La La La La L	- to to to to	L- L- 1
25								
				La La La La La 1 1 1 1 1	1 + 1 + 1 + 1	La La La La La L	- L- L- L- L- 1 1 1 1	
								LLI
	1							L L 1
30			to to to to to	L. L. L. L. L.	to to to to to	Le Le Le Le Le I		L. L. I_
~~								L L !
25		LLLLL			4444			L L L
55		44444		4444	4 4 4 4 4			++1
		FFFFFF	FFFFF	FFFFF	FFFFF	rrrrr		FFI
	1 -1 -1 -1 -1 -1 -1	rrrrr						F F 1
10	1				rrrrr.			F F 1
40								
	5	10	15	20	25	30	35	40

Multicore Evaluation (1/5) Global Map

Multicore Evaluation (2/5) Global Map 15

					·			-
								L 1
-			- L- L- L- L	- L. L. L. L. L.				
5			- L- L- L- L- L					L i
	1		- La La La La L		. La La La La La			L-1
				- L- L- L- L- L- 1 1 1 1 1				
10			LLLL		LLLL	LLLL	LLLL	L i_
10	1	- L. L. L. L. L. L.		- L. L. L. L. L.	. La La La La La		La La La La La	L 1
			- L. L. L. L. L. - 1. 1. 1. 1. 1	- L. L. L. L. L. - 1- 1- 1- 1- 1-	. L. L. L. L. L. L.			
								Li
15		- Le Le Le Le Le Le I	e Le Le Le Le L	- [- [- [- [- [-	. La La La La La	Le Le Le Le Le	La La La La La	L. !_
				- L. L. L. L. L.				
								L i
	1-1-1-1-1-1		- La La La La L	- La La La La La	. La La La La La			L_ 1
20								
								LI
		L to to to to to t	. L. L. L. L. L	- L. L. L. L. L.				L 1
05				- L- L- L- L- L-				
25								
		- Lo toolo to to to t	- t- t- t- t- t	- L. L. L. L. L.	. La La La La La	t. t. t. 📩		
			- L- L- L- L- L	- L- L- L- L- L-				
20								
30						L. L. L.		
			- L. L. L. L. L. - L. L. L. L. L	- L- L- L- L- L- - L- L- L- L- L-		La La La		
35						LLL		
~~								
				- L. L. L. L. L.				
								L 1
40		to to to for to to t		+ L+ f+ L+ L+ L+	. La fa La La La	La fa La La La	to for to to to	L- 1-
	5	10	15	20	25	30	35	40

© 2019 IEEE International Solid-State Circuits Conference

Multicore Evaluation (3/5) Global Map 15

0

© 2019 IEEE International Solid-State Circuits Conference

Multicore Evaluation (4/5) Global Map 15

Multicore Evaluation (5/5) Global Map 15

# = = = = = = = = = = = = = = = = = = =				FF!
15 20	25	30	35	40

Optics Experiment

		FFFF	
<u> </u>	י <mark>ד רררו</mark> ידיררי		

F								
5								
10				· · · · · · · · · · · · · · · · · · ·	·	$ \begin{bmatrix} $		
15	$ \begin{bmatrix} - & - & - & - & - & - & - & - & - & - &$							
20		$ \begin{array}{c} \cdot $						
25								
30						$\begin{array}{c} 1 \\ $		
35						$\begin{array}{c} 1 \\ $		
40	<u></u>	<u>יב ב ב</u> 10	<u></u>	20	<u>2222</u> 25	22 <mark>2</mark> 55 30	<u>35</u>	<mark>בן</mark> 40

Optics Experiment

Outline Background: Path Planning Algorithms Time-Based A* ASIC 65nm Test Chip Results Applications Conclusion

© 2019 IEEE **International Solid-State Circuits Conference**

© 2019 IEEE International Solid-State Circuits Conference

Die Photo and Chip Summary

Applications	A* shorte st path, obstacle avoidance, scientific computation (optics)
Technology	65nm LP CMOS
Architecture	Time-based
# of Vertices	1600
Unit Area	249µm ²
# of Edges	6400
Edge Resolution	4b + Analog Gradient
Voltage	1.2V
Peak Power	26.4mW
Delay per Node	1.79ns @ [V _B =.9V, V _{DD} =1.2V]
Power per Node	183.1μW
Energy per Node	0.238pJ

- A* heuristic implemented with analog bias gradient
- Vertex cells asynchronously evaluate and lockout
- Diverse range of applications
- Scalable to multicore for larger maps
- Orders of magnitude improvement in energy efficiency

© 2019 IEEE **International Solid-State Circuits Conference**

Conclusion

• 40×40 A* ASIC in 65nmLP

This research was supported in part by the National Science Foundation under award number CCF-1763761.

2.5: A 40 × 40 Four-Neighbor Time-Based In-Memory Computing Graph ASIC Chip Featuring Wavefront Expansion and 2D Gradient Control

62 of 62

