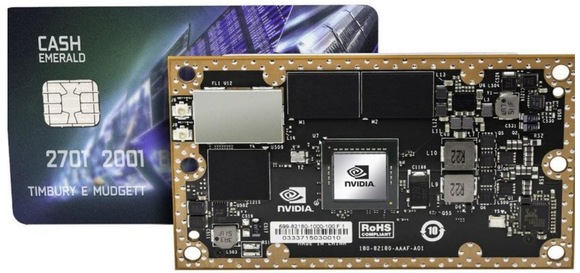
Embedded Real-time Stixel Computation

Daniel Hernández Juárez, *Antonio Espinosa, Juan Carlos Moure*

David Vázquez, and Antonio M. López



Computer Architecture & Operating Systems Department (CAOS)

Autonomous University of Barcelona

caos

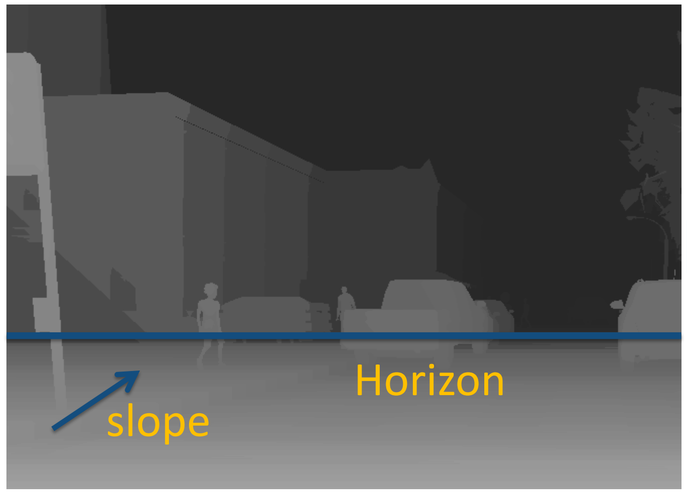
UAB
Universitat Autònoma de Barcelona

CVC
Centre de Visió per Computador

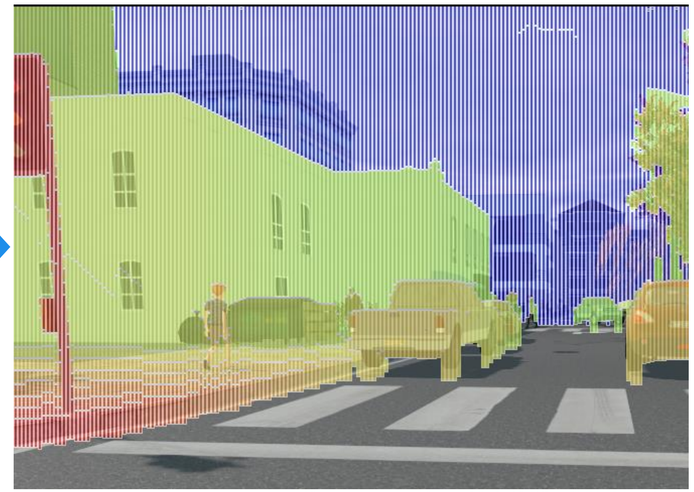
WACV 2017

Stixel World: Compact representation of the world

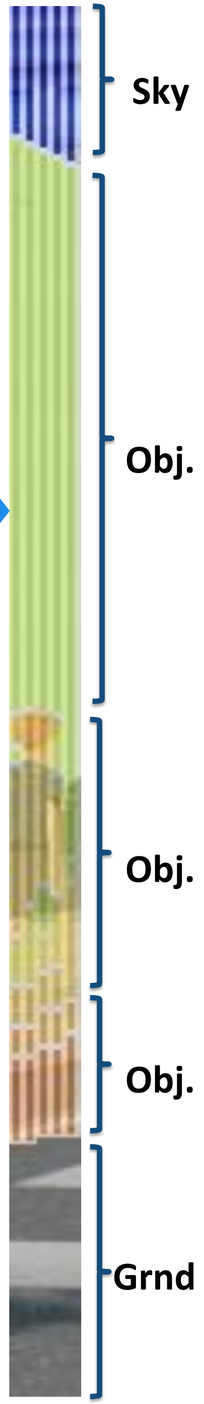
Stereo Images



Stereo + Horizon Line + Road Slope



Stixels



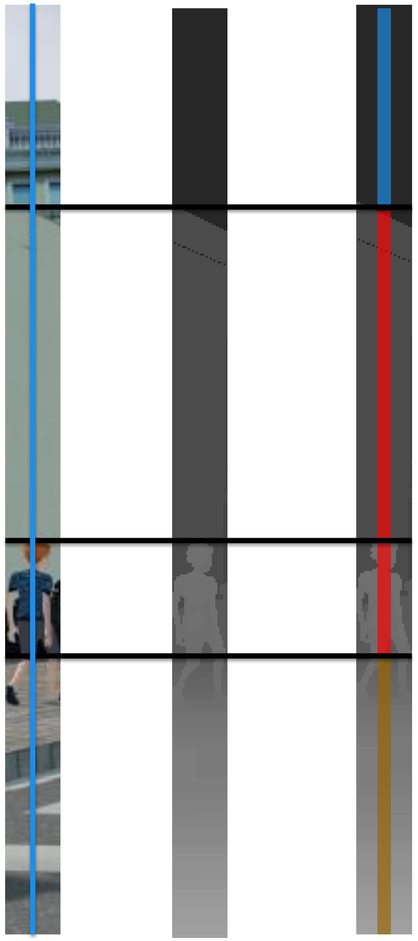
Stereo of 1280 x 960 = 1,228,800 pixels => **Too much data to process**

Medium-level representation with only relevant information

Fixed width stixels, **variable** number of stixels per column

Stixel = Stick + Pixel

Stixel World: Idea & Computational Complexity



Sky: Very far, all pixels = disparity near 0

Object: All pixels = Constant disparity

Ground: Disparity close to expected model (given ground estimation)

Image Resolution	
640 x 480	147 M ops.
1280 x 480	294 M ops.
1280 x 960	1179 M ops.

Total Computation Work
Θ (Width × Height ^{2})

Computed independently for each column

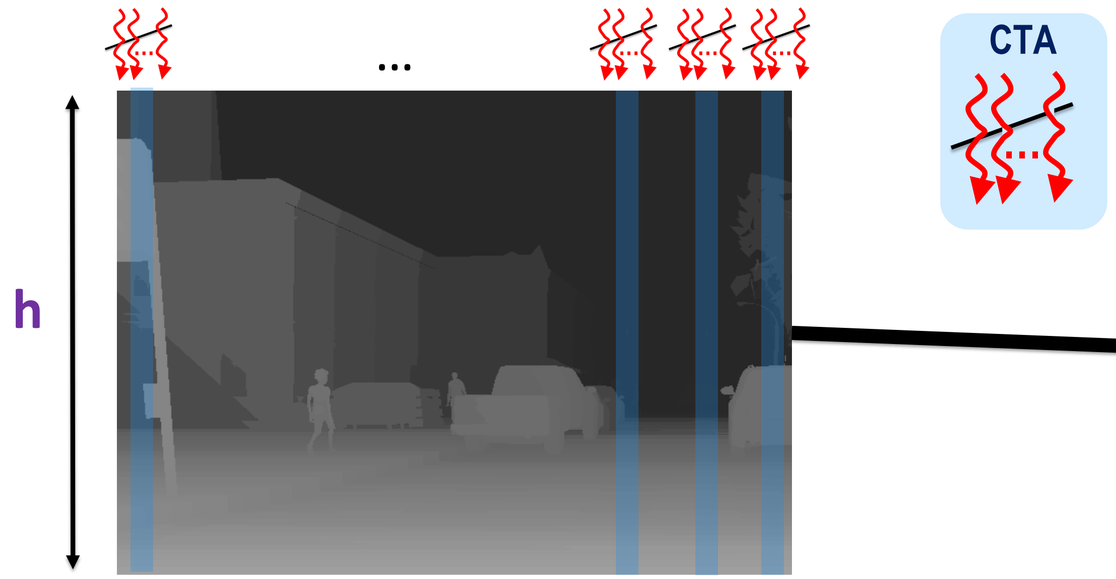
Enforces constraints: no sky below horizon, no neighbors objects at same distance…

Stixel = {v_{Base}, v_{Top}}, **all possible** v_{Base} and v_{Top} **combinations tested** => complexity is **O(wh^2)**

Combinatorial explosion (of possible configurations): Dynamic programming to evaluate efficiently

GPU Implementation: Proposal

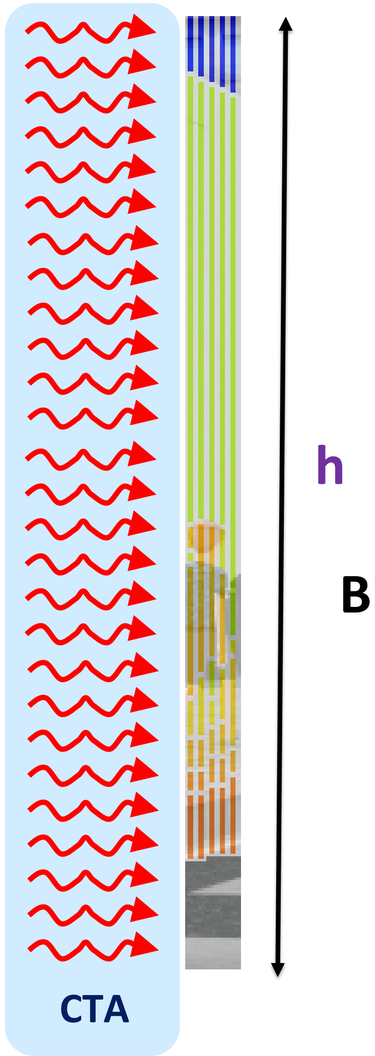
First level parallelism (big granularity):



Stereo Disparity

Independent / Task level: Typical CPU parallelization
Each image column is processed in parallel by a CTA
CTA = Cooperating Threads Array

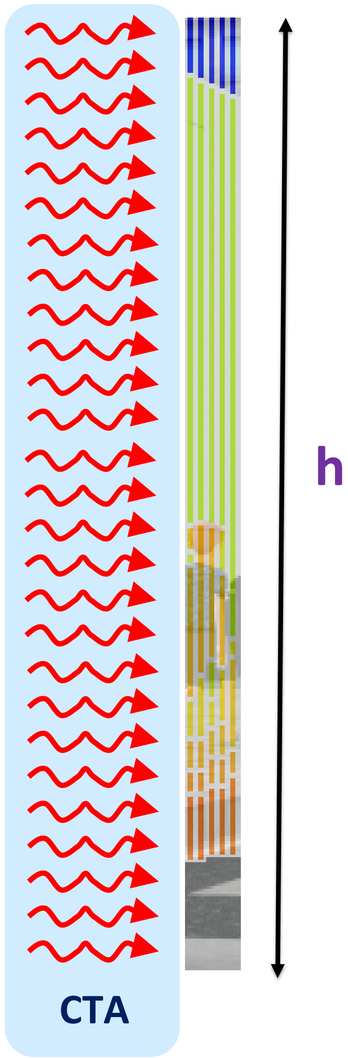
Second level parallelism (fine-grained):



Block of threads computes collaboratively

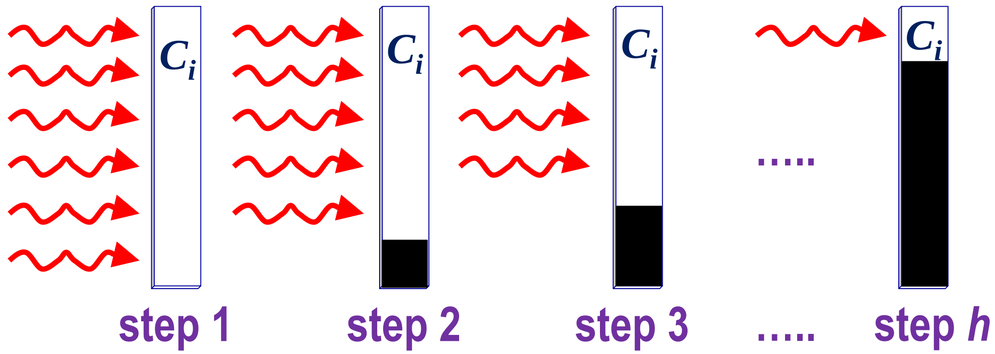
GPU Implementation: Second Level Parallelism

Second level parallelism (fine-grained):



Computational Analysis	
Thread Parallelism	$h \times w$
Compute Work per thread	h
Total Global Data Reads	$h^2 \times w$
Total Global Data Stores	$h \times w$

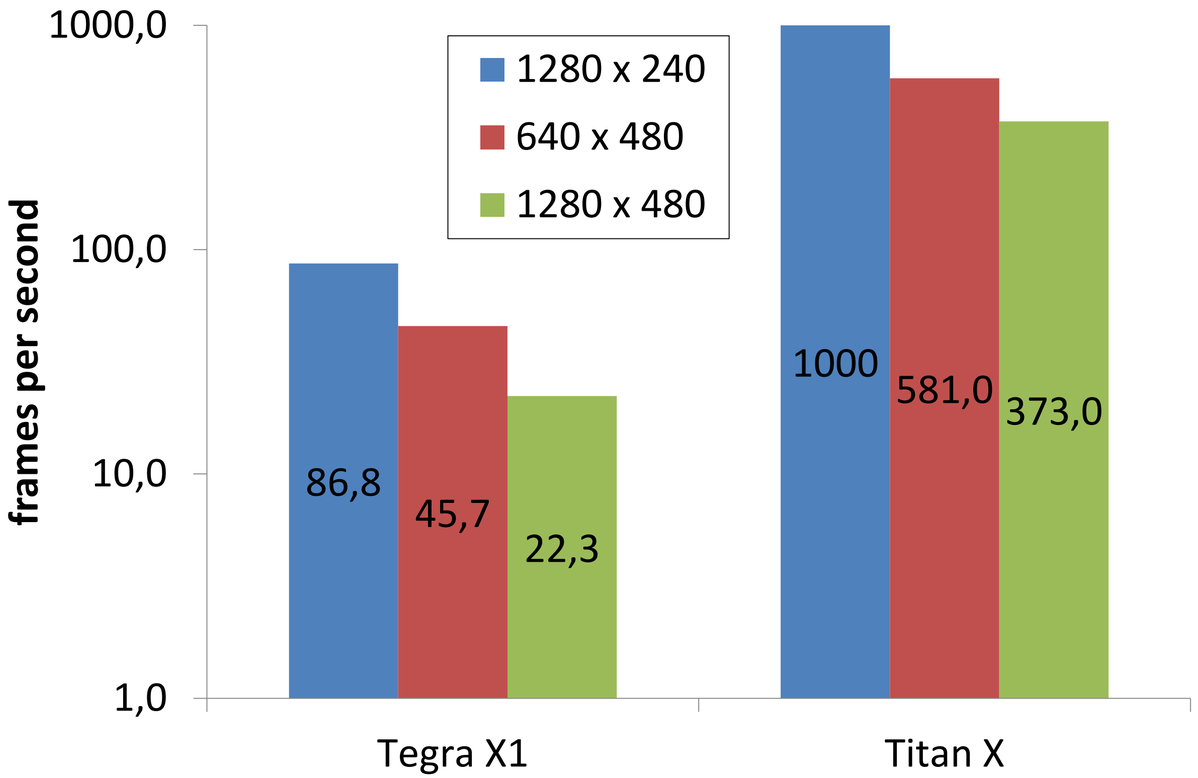
More details



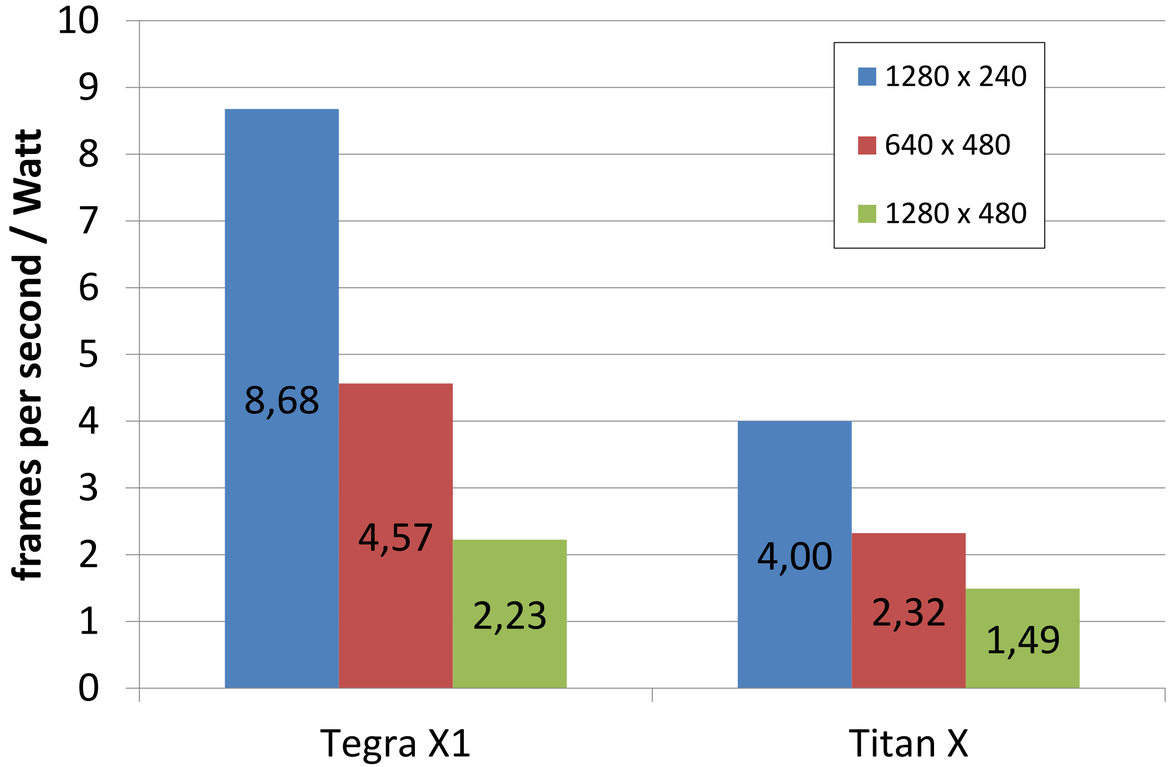
- Extra parallelism level needed for efficient GPU use
- Sequentially perform *h* (image height) steps
- CTA threads collaborate sharing info each step
- Decreasing Parallelism: Each step uses one thread fewer

GPU Implementation: Performance & Energy Efficiency

GPU Performance (Frames/Second, fps)



Energy Efficiency



- Real-time performance for energy efficient GPU: NVIDIA Drive PX
- NVIDIA Drive PX has better energetic efficiency than high-end GPUs

Thank you

Daniel Hernández Juárez, *Antonio Espinosa, Juan Carlos Moure, David Vázquez, and Antonio M. López*

More information:
www.cvc.uab.es/people/dhernandez

Autonomous University of Barcelona

caos

UAB
Universitat Autònoma de Barcelona

CVC
Centre de Visió per Computador