SceneNet: 3D Reconstruction of Videos Taken by the Crowd on GPU

Chen Sagiv
SagivTech Ltd.

GTC 2015 – San Jose

SagivTech Snapshot

- Established in 2009 and headquartered in Israel
- Core domain expertise: GPU Computing and Computer Vision
- What we do:
  - Technology
  - Solutions
  - Projects
  - EU Research
  - Training
- GPU expertise:
  - Hard core optimizations
  - Efficient streaming for single or multiple GPU systems
  - Mobile GPUs

Mobile Crowdsourcing Video Scene Reconstruction

- If you've been to a concert recently, you've probably seen how many people take videos of the event with mobile phone cameras

- Each user has only one video – taken from one angle and location and of only moderate quality

The Idea behind SceneNet

Leverage the **power of** multiple mobile phone cameras

to create a **high-quality 3D** video experience that is

sharable via social networks

SceneNet as a FET SME collaborative project

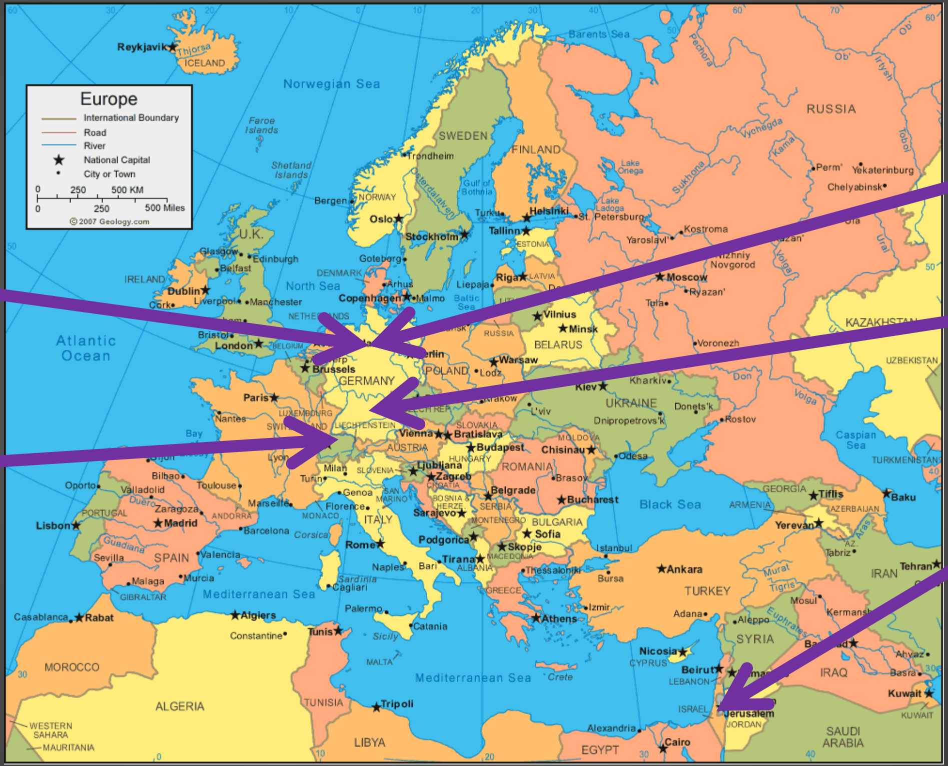
Uni Bremen

SCiLS

ERS

EPFL

SagivTech



SceneNet Advisory Board

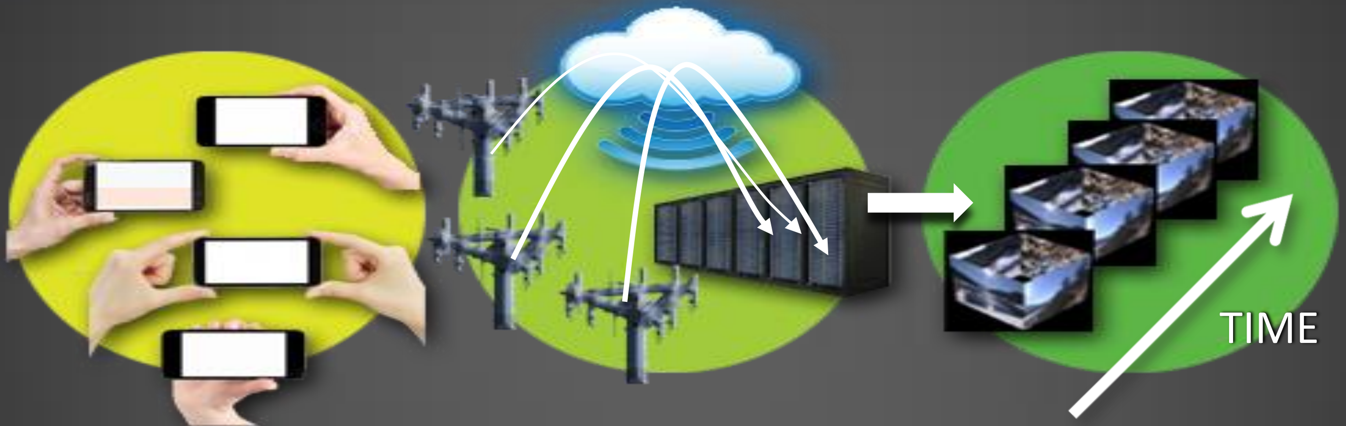
Prof. Dr. Franziska Boehm
Uni. Munster

Mr. Malte Blumenthal
Eventim

Prof. Dr. Hans-Georg Stark
Hochschule Aschaffenburg

Mr. Izhar Ashdot
Musician

Creation of the 3D Video Sequence



The scene is photographed by several people using their cell phone camera

The video data is transmitted via the cellular network to a High Performance Computing server.

Following time synchronization, resolution normalization and spatial registration, the several videos are merged into a 3-D video cube.

SCENE NET

SAGIVTECH

The Event Community

A 3-D video event is created.

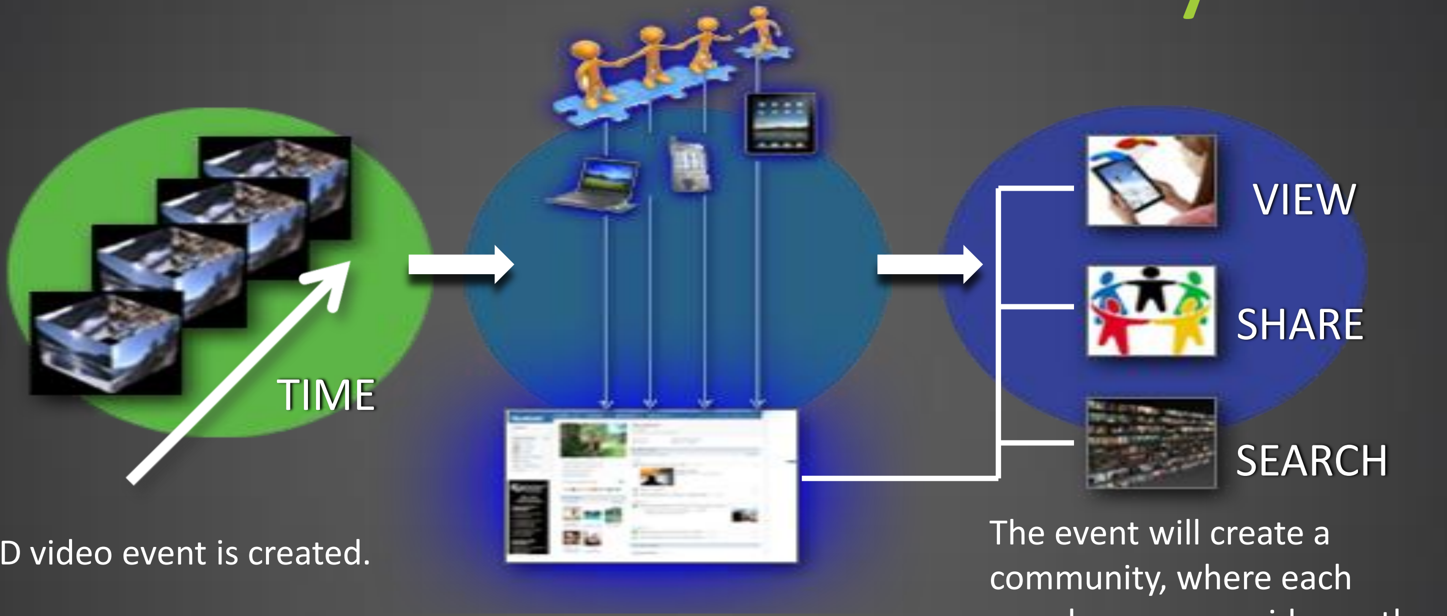
The 3-D video event will be available on the internet as public or private event.

The event will create a community, where each member may provide another piece of the puzzle and view the entire information.

SCENE NET

SAGIVTECH

The Event Community



A 3-D video event is created.

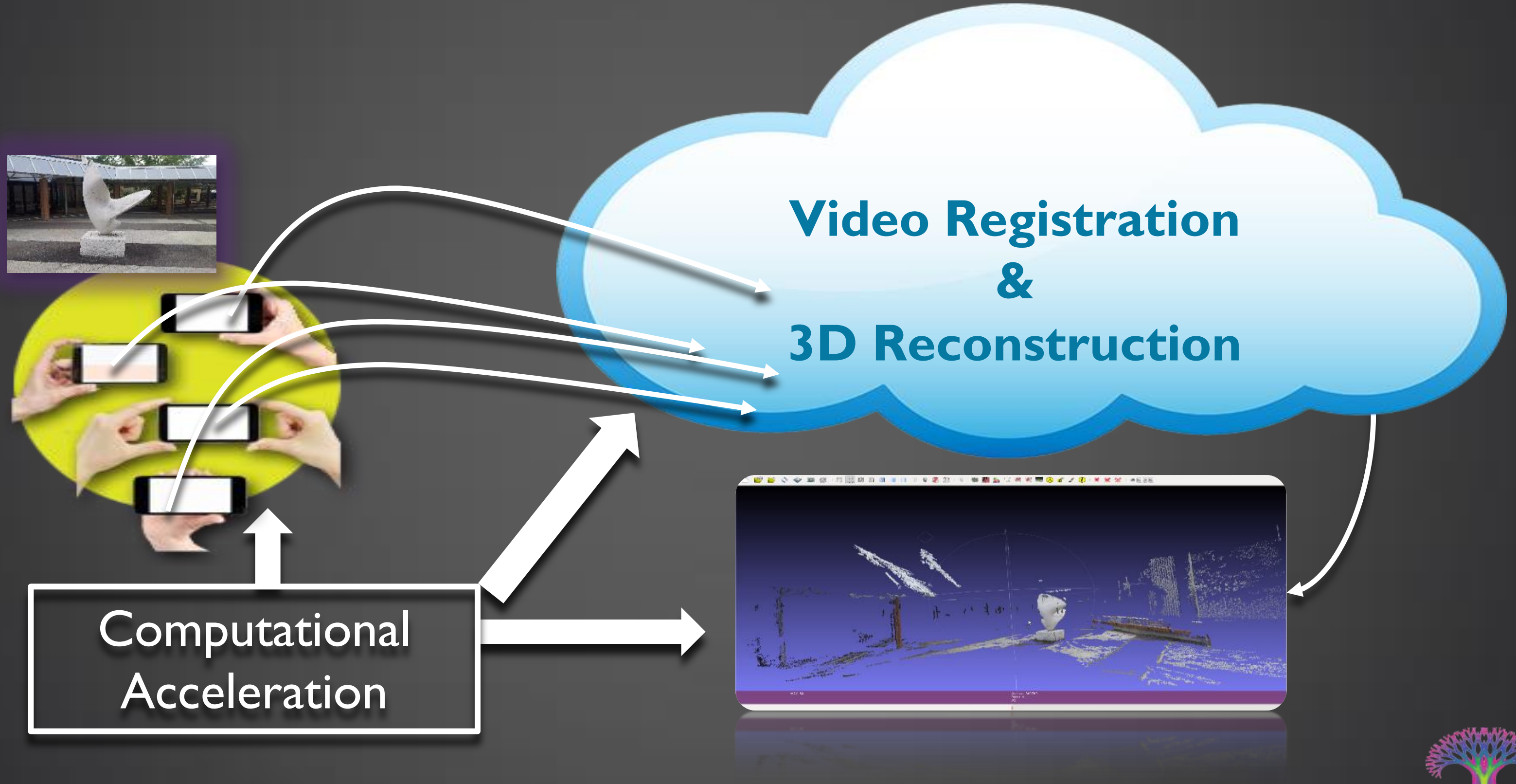
The 3-D video event will be available on the internet as public or private event.

The event will create a community, where each member may provide another piece of the puzzle and view the entire information.

SAGIVTECH

The Combined Model: Mobile & Cloud Computing

GPU Computing in SceneNet



Motivation for Computational acceleration in SceneNet - on the device

- The mobile device performs several tasks for SceneNet alongside its usual tasks (calls, graphics etc.)
- Pre processing algorithms run at < 1 frame per second on the CPU of the mobile device
- Need significant speedup !
- Hence, GPU to the rescue !

Convolution Acceleration on Tegra K1

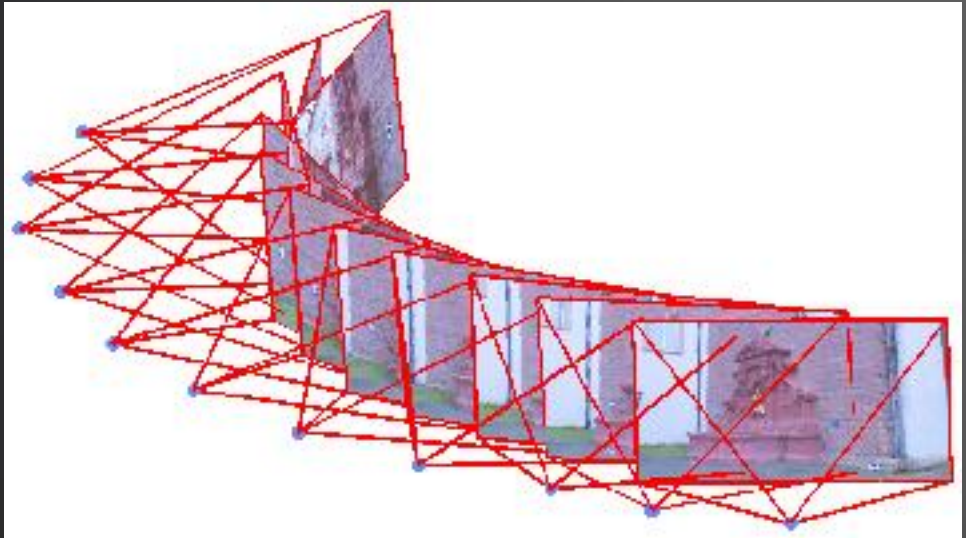
Image Size	CPU Gold	1 CPU Thread	4 CPU Threads	GPU	Speedup
1024 x 1024	142	18	10.2	4	X2.6
2048 x 2048	740	100	50	9	X5.5

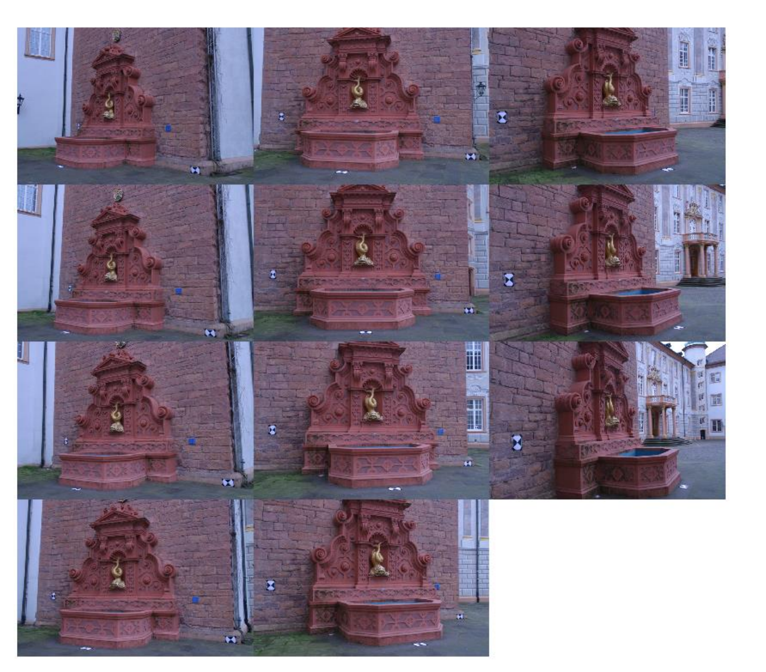
Motivation for Computational acceleration in SceneNet - on the server

- The server related tasks are mainly video synchronization and registration and 3D reconstruction
- Currently, the compute power does not allow for near real time implementations
- **Solution: GPU acceleration is an enabler for the server based operations**
- This relies on SagivTech Background IP Infrastructure where we accommodate it to the needs of SceneNet

SAGIVTECH

Spatial Calibration





SAGIVTECH

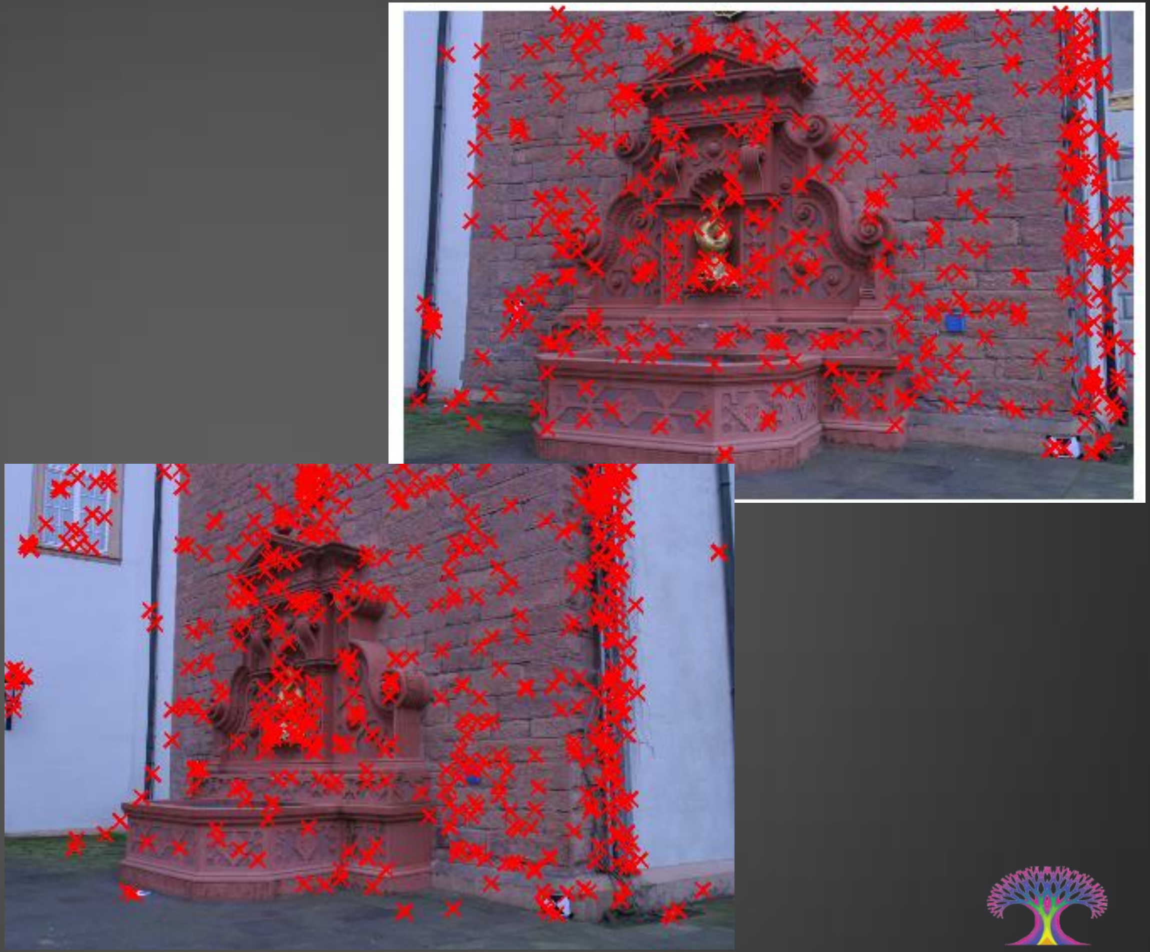
Feature detection + Matching

↓

Fundamental matrix estimation

↓

Global registration



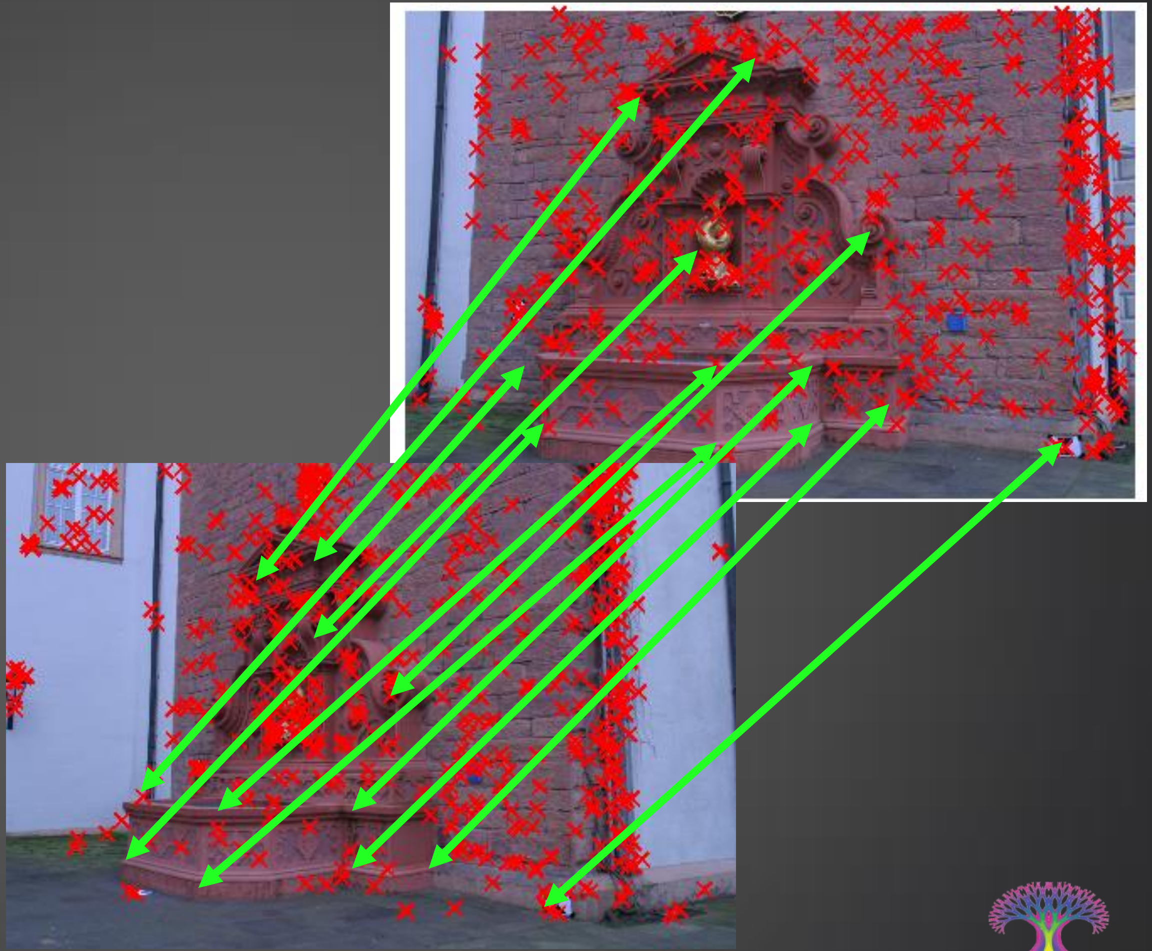
Universität Bremen*

SAGIVTECH

Feature detection + Matching

Fundamental matrix estimation

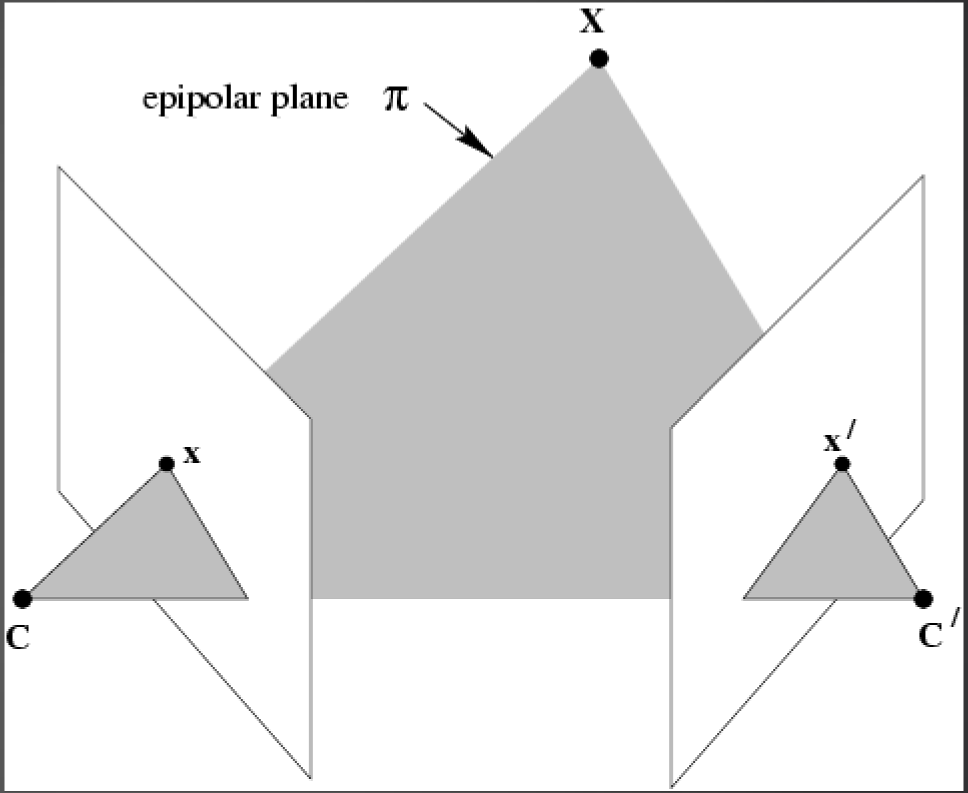
Global registration



Feature detection + Matching

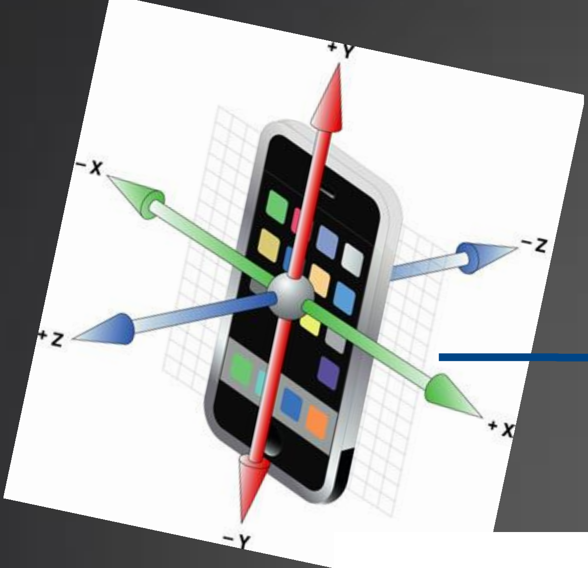
Fundamental matrix estimation

Global registration

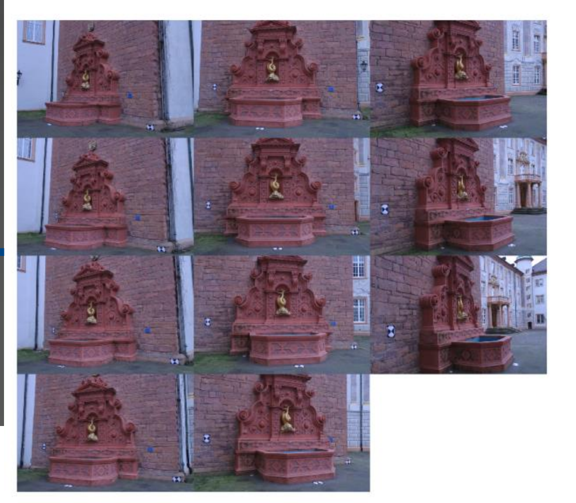


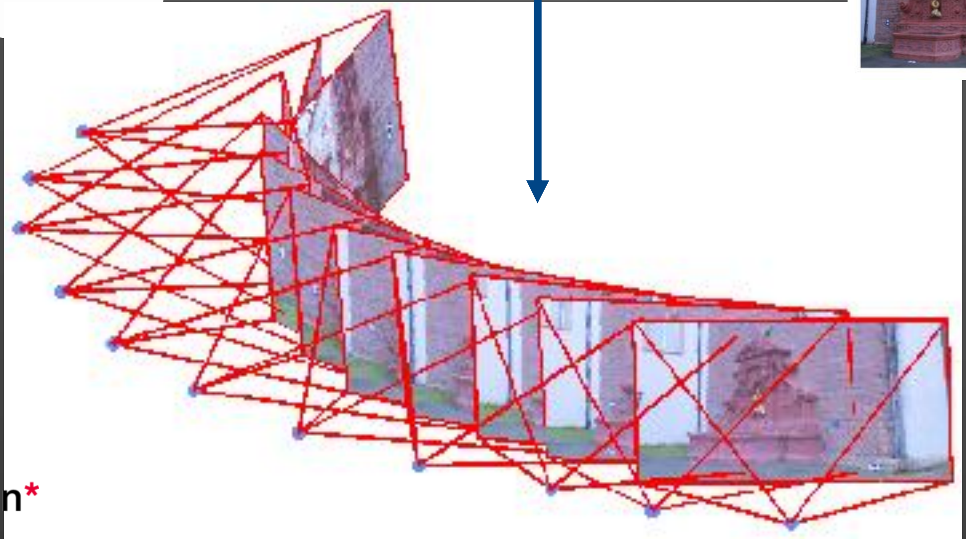
$$\mathbf{x}'^T F \mathbf{x} = 0$$

$$F \in \mathbb{R}^{3\times 3}$$



Spatial Calibration





Universität Bremen*

SAGIVTECH

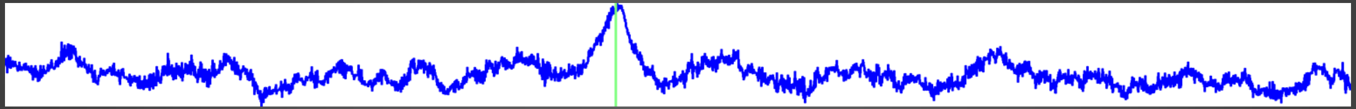
Time & Audio Synchronization

3D model reconstruction

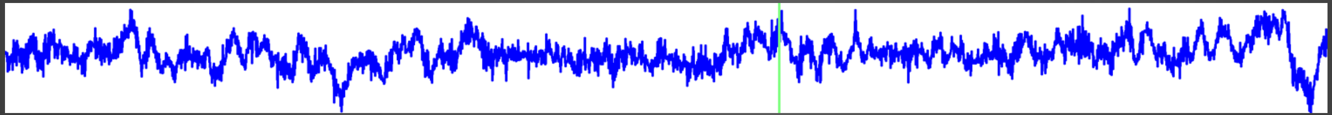
- Precise : epipolar matching is both fast and accurate
  - Dense multi-scale description of the images using binary descriptors

3D model reconstruction

- Precise : epipolar matching is both fast and accurate
  - Empirical probability density check to discard false positives at occlusion points



Correct match : max peak above other local max



Wrong match : max peak similar to other local max

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

SAGIVTECH

3D model reconstruction

- Robust : works even with a minimal set of inputs
  - two viewpoints already sufficient for dense reconstruction
  - very few erroneous points

3D reconstruction

3D model reconstruction

- Parallelizable : designed with GPU acceleration in mind

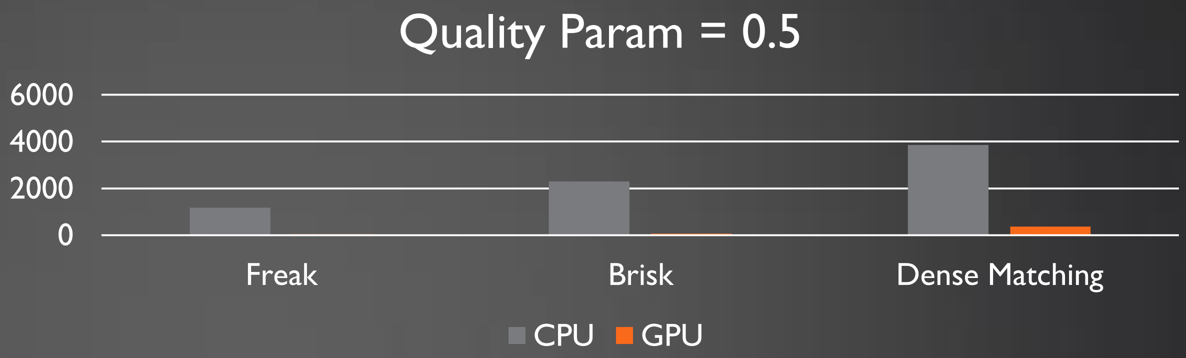
  The algorithm can be parallelized at multiple levels
  - Local operations :
    - Dense description
    - Epipolar matching
    - Triangulation
  - Semi-local operation :
    - Probability density consistency check
  - Multi-view fusing : parallel instead of iterative method

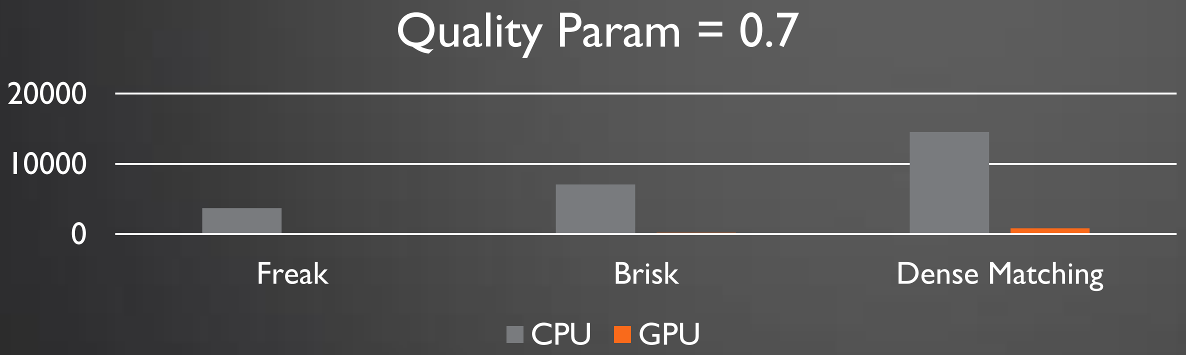
Computational power bounded by a huge number of small independant tasks

3D model reconstruction

- Impressive performance boosts when migrated to a NVIDIA K20Xm discrete card
- Overall performance speedup is from x5.8 to x7.4 depending on the quality parameter
- Percentage of time used by the CPU when running the GPU code, varies from 40% to 51%
- Freak/Brisk on GPU speedup shows about a x35 factor
- Further optimizations could be applied in the future

3D model reconstruction

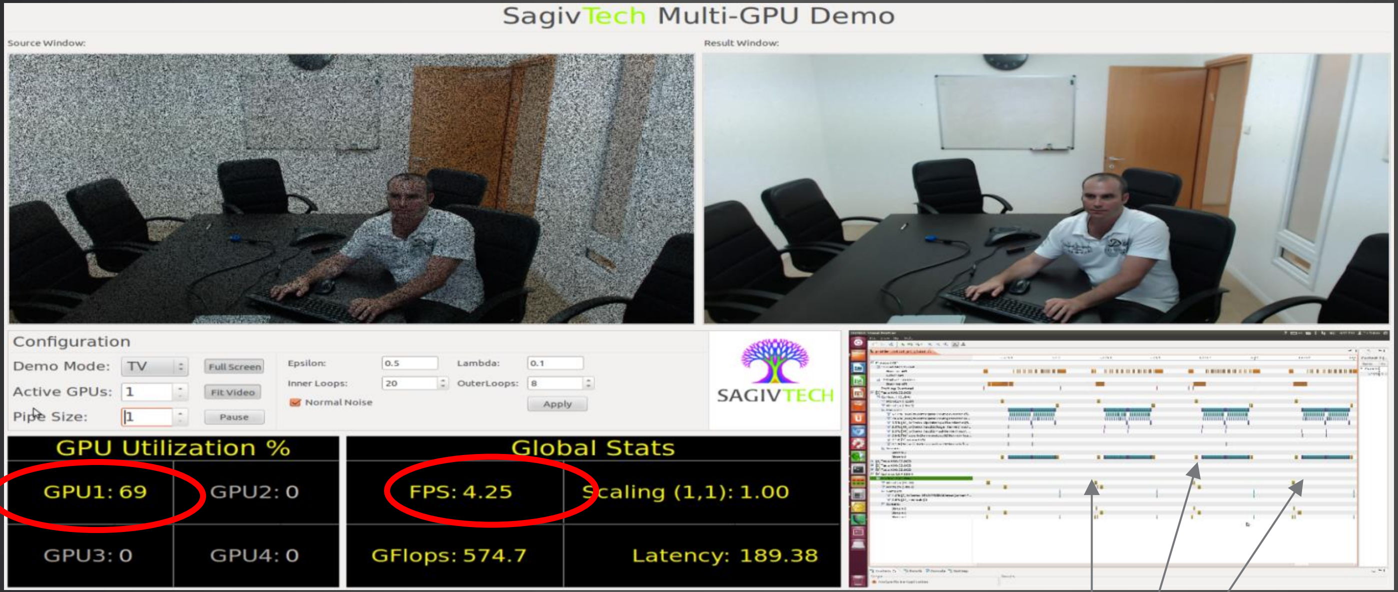




Main Attributes of SagivTech's Streaming Infrastructure

- **Pipelining**: hides memory transfer overhead between CPU and GPU
- **Asynchronous work**: allows job launch on multiple GPUs without waiting for one GPU to finish
- **Peer-to-peer communication**: enables transfer of data between multiple GPUs within the same system

ST MultiGPU Real World Use Case



One GPU

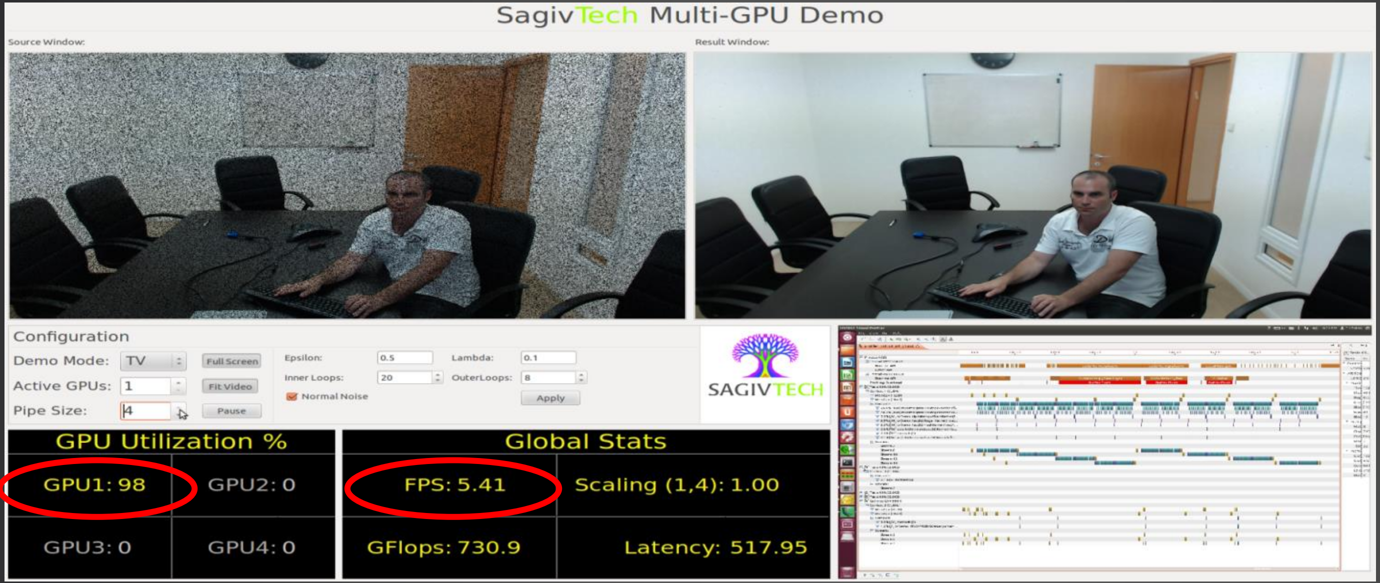
One pipe

Utilization: ~70%

- FPS: 4.25
- Scaling: 1.00
- Note the gaps in the profiler

SAGIVTECH

ST MultiGPU Real World Use Case

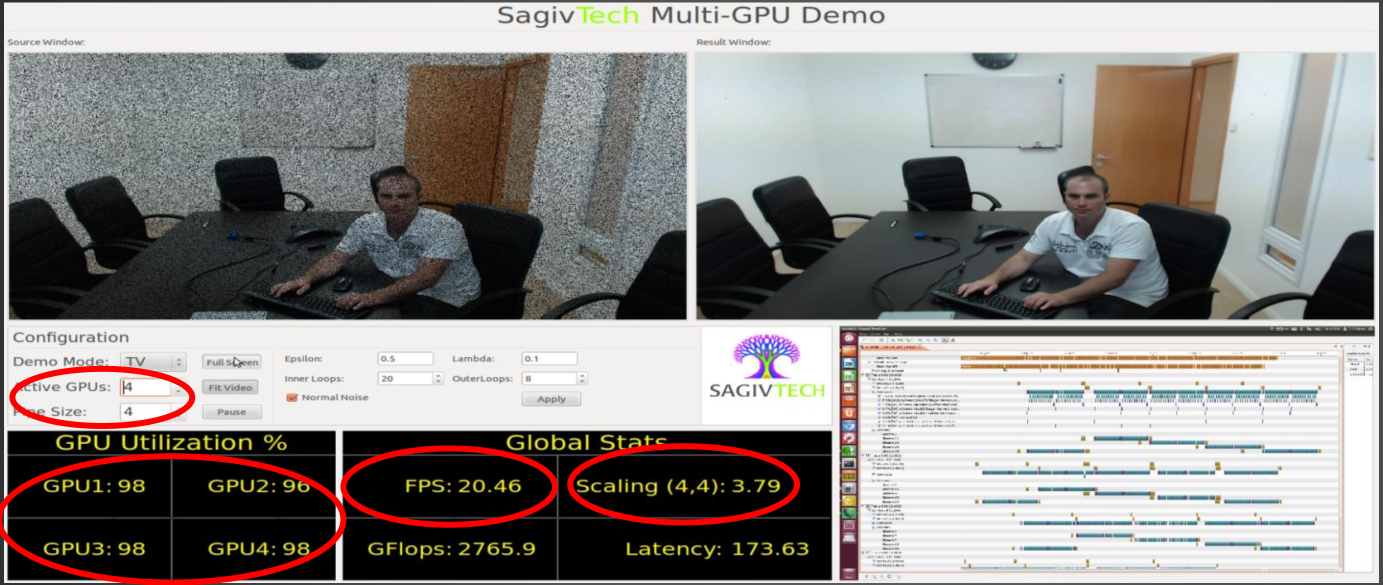


One GPU
4 pipes
Utilization: 95%

- FPS: 5.41
- Scaling: 1.27
- Better utilization using pipes

SAGIVTECH

ST MultiGPU Real World Use Case



Four GPUs
Four pipes
Utilization: 96%+

- FPS: 20.46
- Scaling: 3.79 – Near linear Scaling!
- Note NO gaps in the profiler

SAGIVTECH

Mobile Crowdsourcing Video Scene Reconstruction

This project is partially funded by the European Union under the 7th Research Framework, programme FET-Open SME, Grant agreement no. 309169.

Thank You

For more information please contact

Nizan Sagiv

nizan@sagivtech.com

+972 52 811 3456

SAGIVTECH

SAGIVTECH