A prior-based approach to 3D face reconstruction using depth images

Donny Tytgat, Sammy Lievens, Erwin Six
Alcatel-Lucent Bell Labs

Setting the scene

3D face reconstruction …

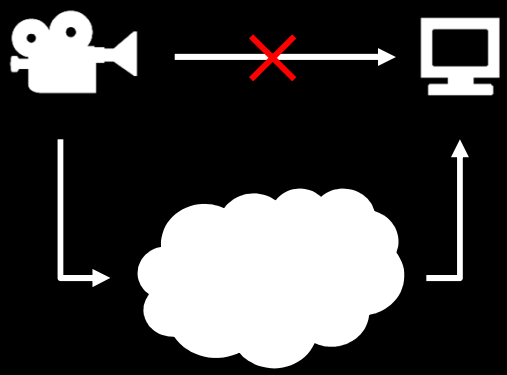
Usefull for decoupling the capturing from the display phase

- Extract intrinsic features of a face
- Reconstruct by imposing a new context
  - Different viewpoint(s), other light conditions, etc.

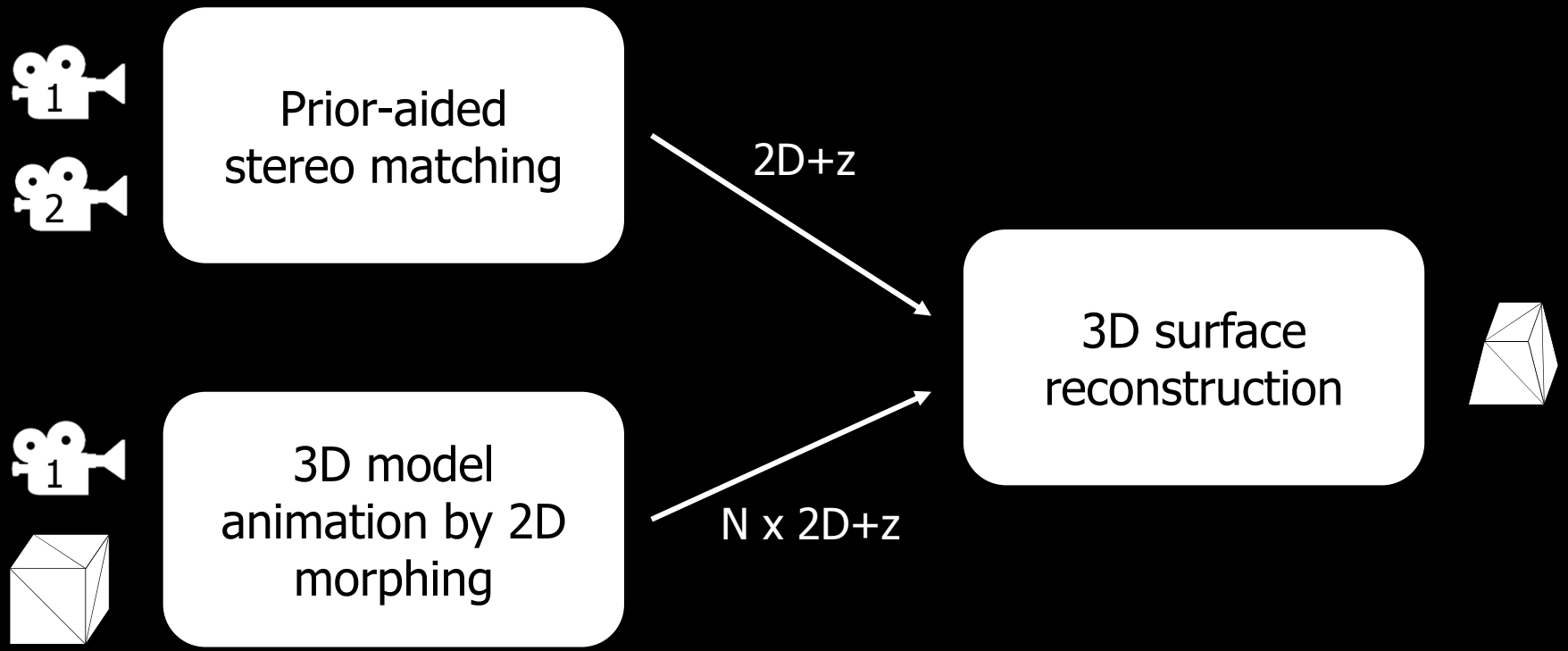
… for video communication

Live: real-time, low latency execution

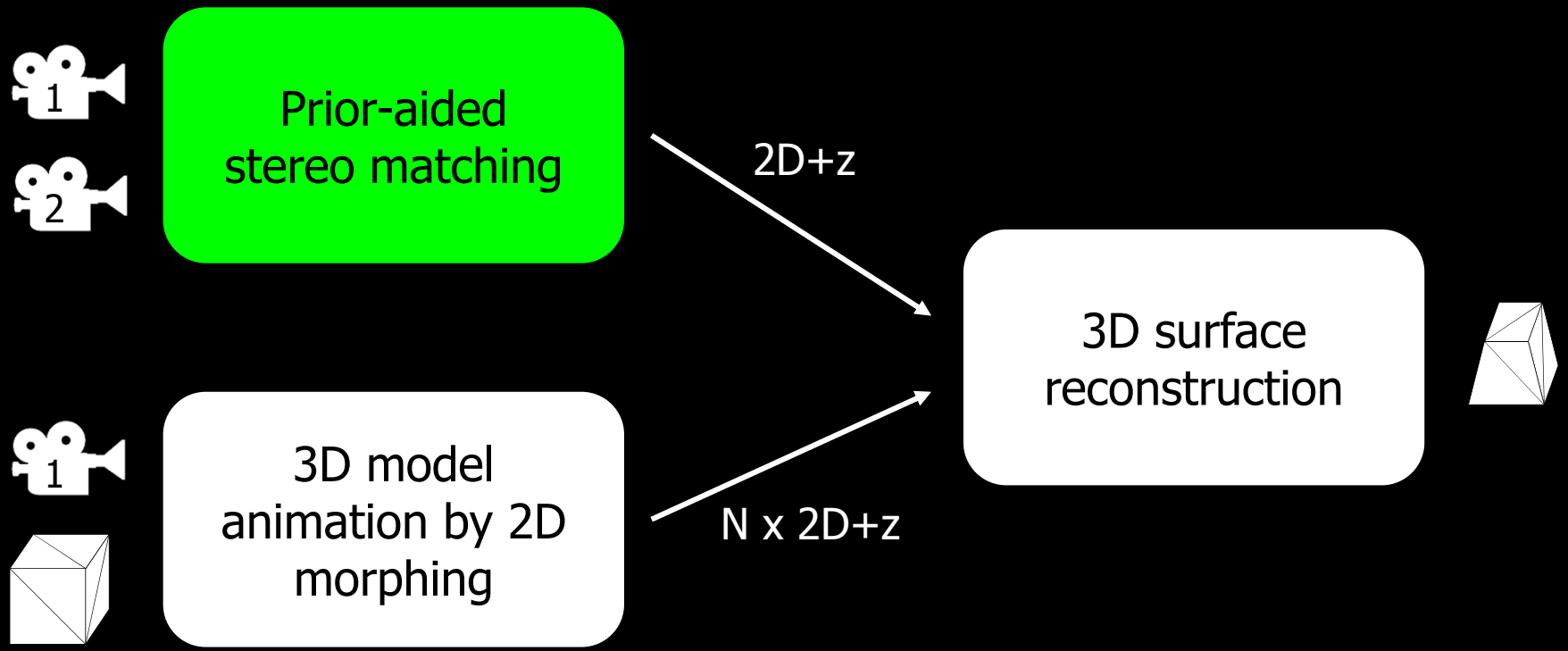
Photo-realistic: quality that approaches reality



System overview



System overview



Prior-aided stereo matching

- Based on [Zhang et al]
  - Block matching with dynamic windows (locally adaptive polygon approximation)
  - Winner-takes-all using SAD

- Accelerated SAD calculation using **integral images**
- Ingested **prior** information: sparse correspondences
  - Active Appearance Models (AAM)
  - Enables a reduction of the search space
- Per-pixel **Gaussian mixture model** (GMM) iso. winner-takes-all
  - Include temporal- and spatial influences

Prior-aided stereo matching

a. base

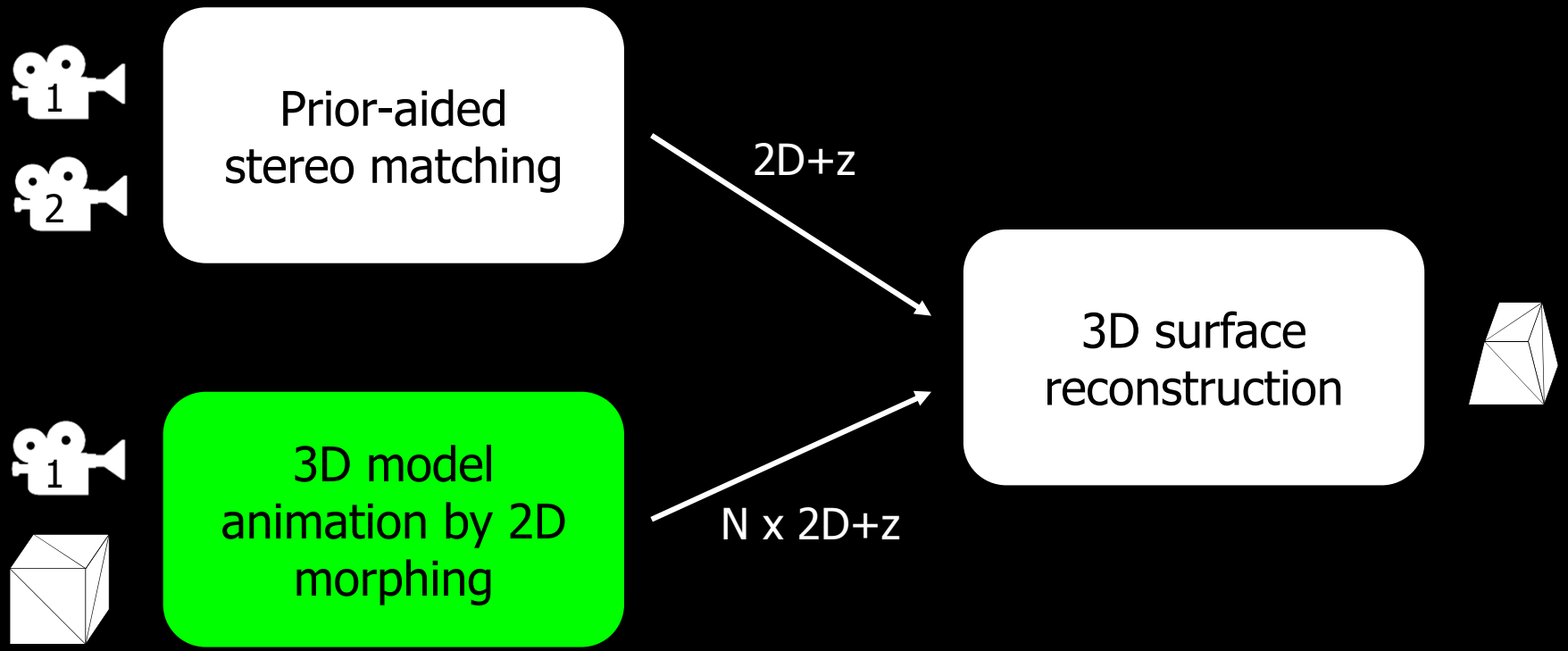
b. adding restricted search

c. adding temporal influence

d. adding neighbour influence *

* 22 FPS @ Intel core i7 2820QM

System overview



3D model animation by 2D morphing

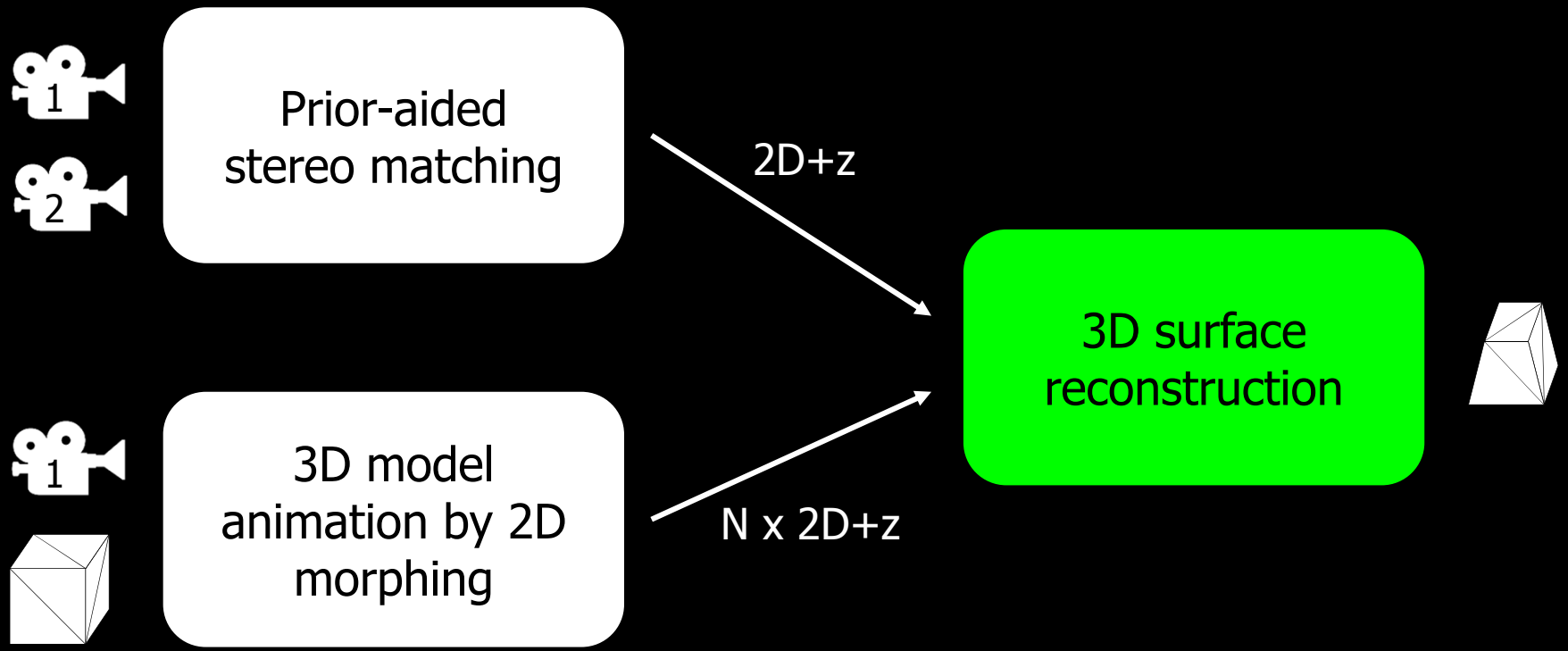
- Input
  - Static 3D mesh model
    - Currently a personal scan; can be replaced by fitted face-space model
  - 2D sparse feature set from live feed
- Output
  - N x 2D+z images of coarsely animated mesh

- Method
  - Register 3D model to match the viewpoint of the 2D feed
  - Transfer sparse feature locations, estimate 3D positions
  - Create N 2D+z projections by warping the static mesh to the new feature locations

3D model animation by 2D morphing

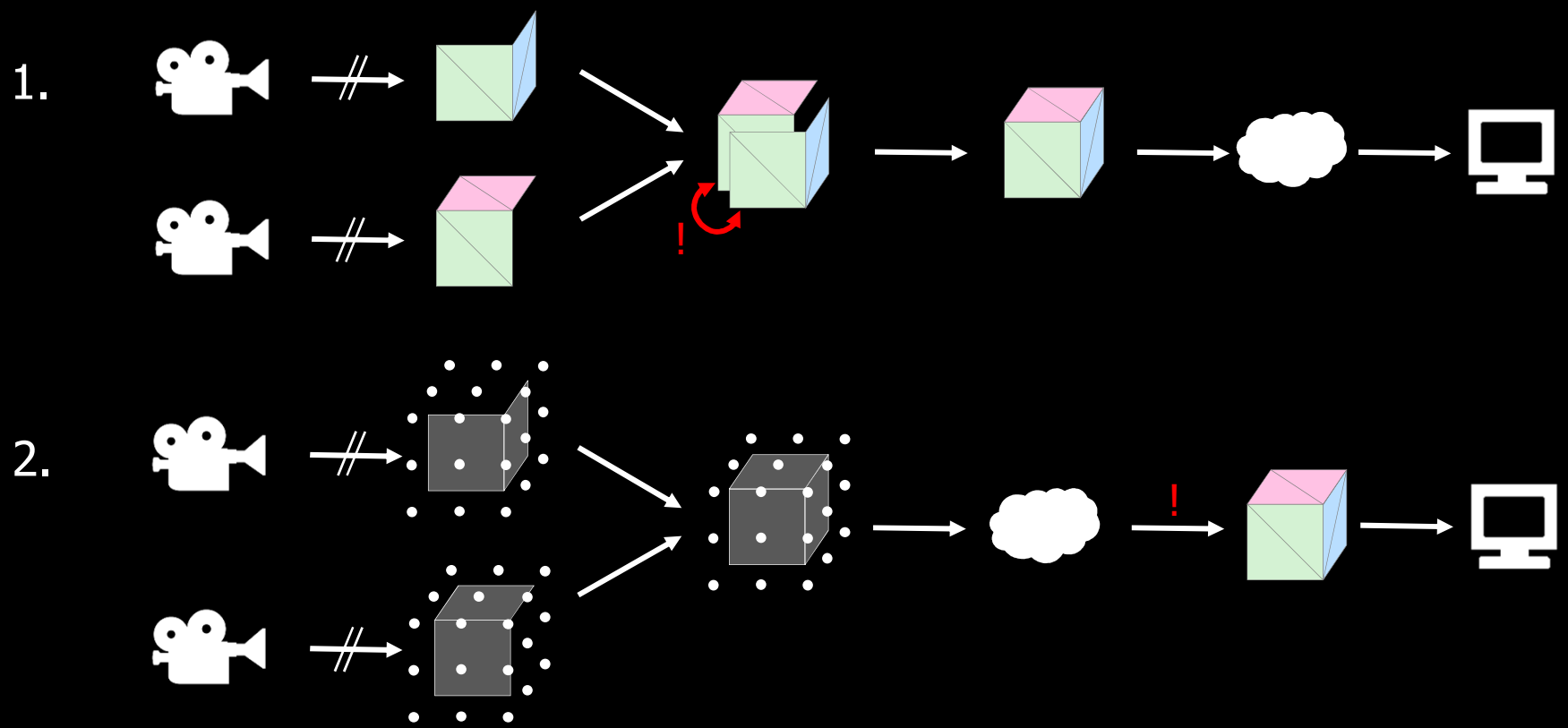
3D model animation by 2D morphing

System overview



3D surface reconstruction

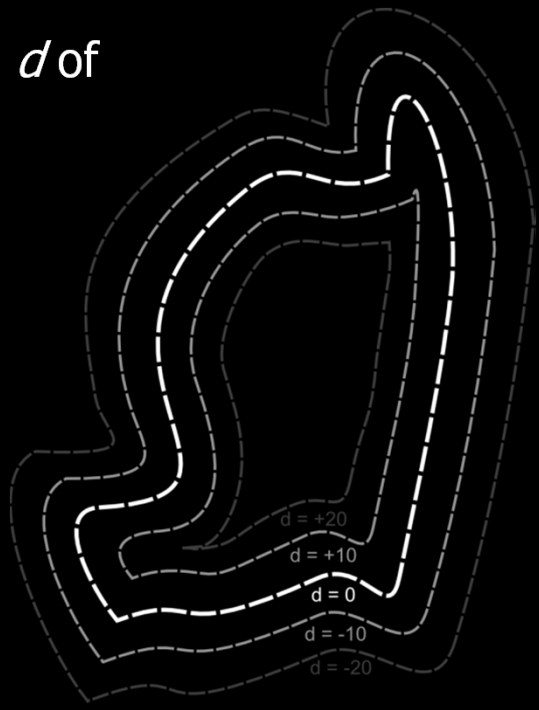
Aggregating multiple depth images into a consistent surface model



3D surface reconstruction

Aggregating multiple depth images into a consistent surface model

- Implicit surface
  - Continuous function that specifies the signed distance *d* of each point in space to the surface
  - Surface can be constructed by building the iso-surface over this function (*d = 0*)



- Why implicit surface?
  - Correspondence requirement is less strict
  - Surface scalability

3D surface reconstruction

Making it work… fast

- Adaptive sampling
- Local distance metric estimation
  - Project 3D point to the depth images, select distance with best score
- Global distance metric propagation by implicit surface filter

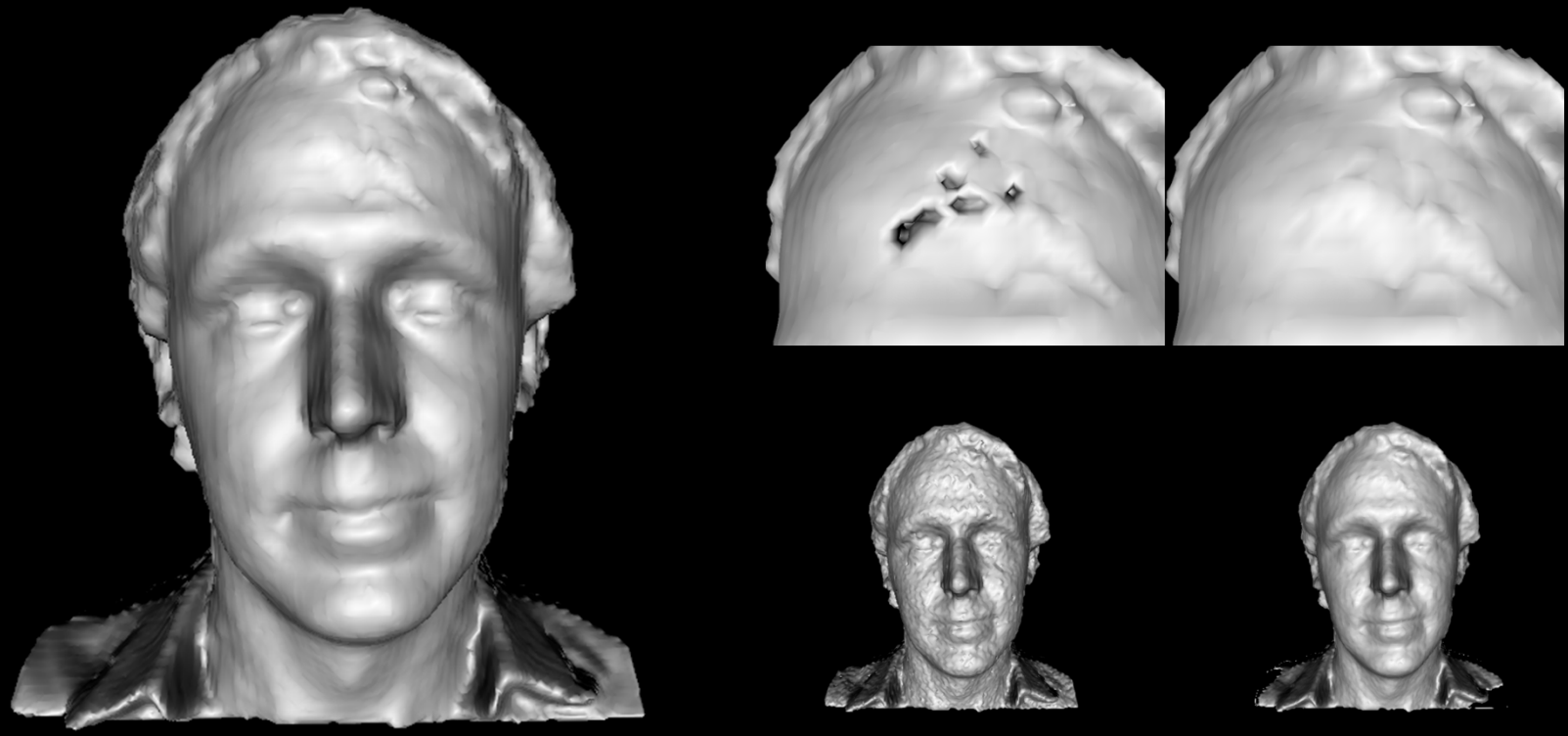
$$D'(p) = R(p)^{\alpha} D(p) + (1 - R(p)^{\alpha}) D_n(p)$$

$$D_n(p) = \frac{\sum_{(x,y) \in N(p)} (R(x) + R(y)) \frac{(D(x) + D(y))}{2}}{\sum_{(x,y) \in N(p)} (R(x) + R(y))}$$

3D surface reconstruction

Generated from 5 depth images, 27k triangles, 9FPS

3D surface reconstruction



Putting it all together…

Conclusions and future work

- Real-time 3D face reconstruction is possible
- Priors are important for
  - improving computational speed,
  - robustness,
  - completeness

- Replace static personal mesh by a fitted model
- Improve rendering quality
- Learn model intrinsics over time in order to reduce live data dependency

Thank you

www.alcatel-lucent.com