A Robust Real-Time Face Tracking using Head Pose Estimation for a Markerless AR System

Márcio C. F. Macedo (UFBA),
Antônio L. Apolinário Jr. (UFBA),
Antonio C. S. Souza (IFBA)

Mestrado Multiinstitucional em Ciência da Computação - UFBA/UEFS
LABRAGAMES - IFBA

Summary

- Introduction;
- Robust Real-Time Face Tracking using Head Pose Estimation:
  - KinectFusion;
  - Real-Time Head Pose Estimation from Consumer Depth Cameras using Random Regression Forests;
  - Our Approach;
- Results and Discussion:
  - Fast Translation;
  - Fast Rotation;
- Conclusions and Future Work;
- Acknowledgments;
- References

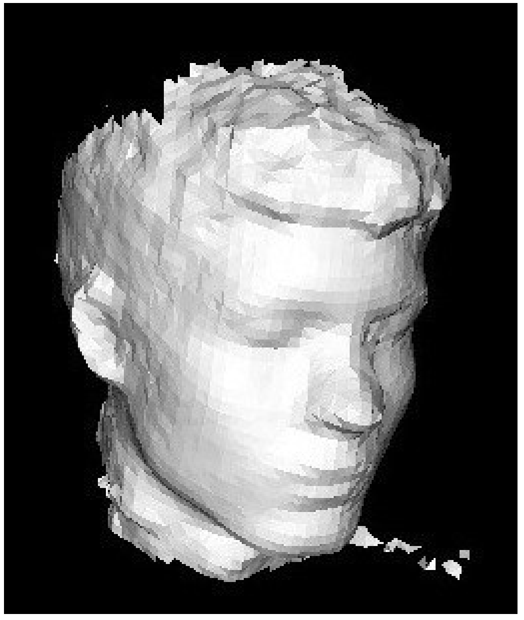
Introduction

- In some AR systems, the user turns his head in front of a camera and the head is augmented with a virtual object;
- We need an algorithm able to track the person's head with enough accuracy and real-time;

http://campar.in.tum.de/Chair/ProjectDepthPerceptionMedicalAR

Introduction

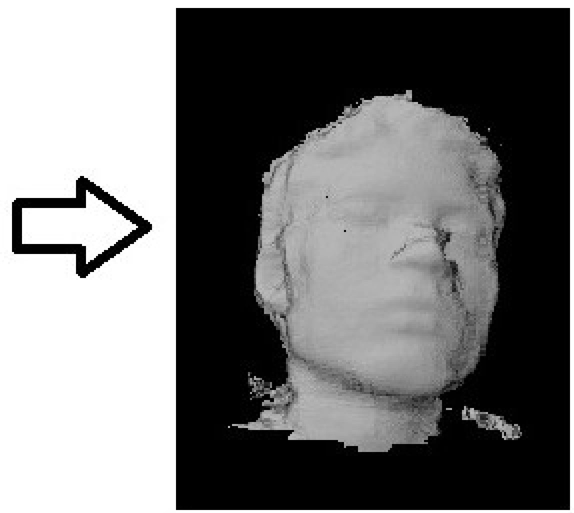
- One way to do it is building a reference 3D model of the user's head and tracking it through all frames.



Introduction

- Our approach consists of some steps:
  - First: A reference 3D model is built with the KinectFusion;

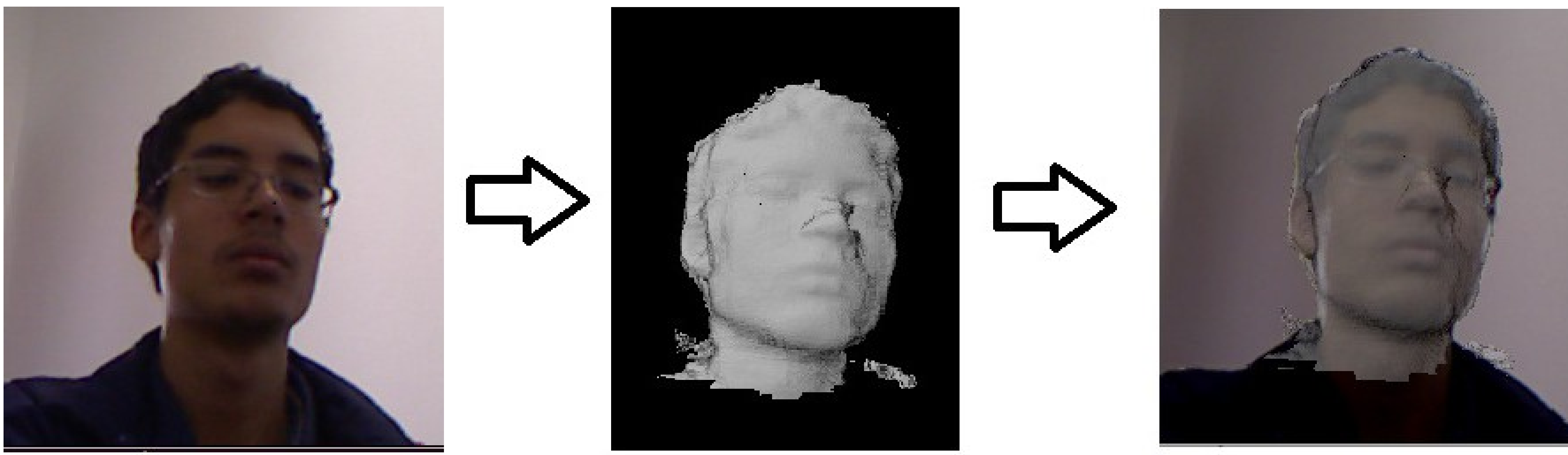
Live Stream



Reference 3D Model

Introduction

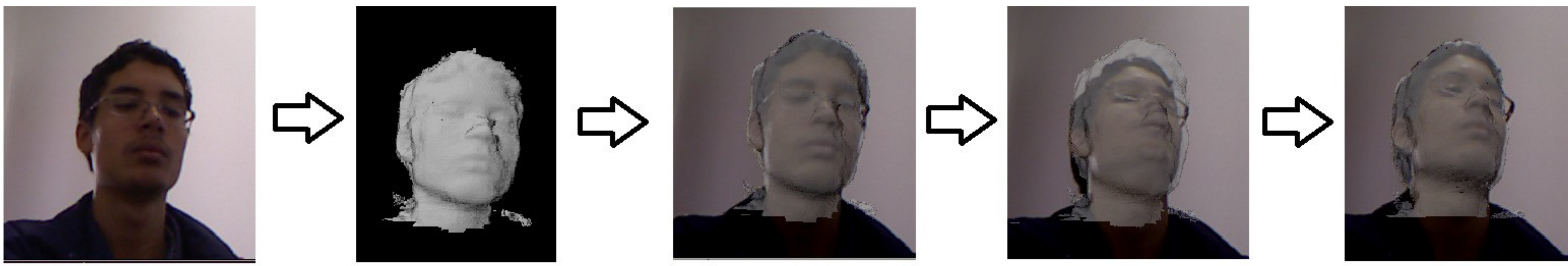
- Second: The Kinect raw data is aligned to the reference 3D model, predicting the current camera pose;



Live Stream   Reference 3D Model   AR with the user's head

Introduction

- Third: A head pose estimator is used to give an initial guess to the tracking algorithm when it fails.

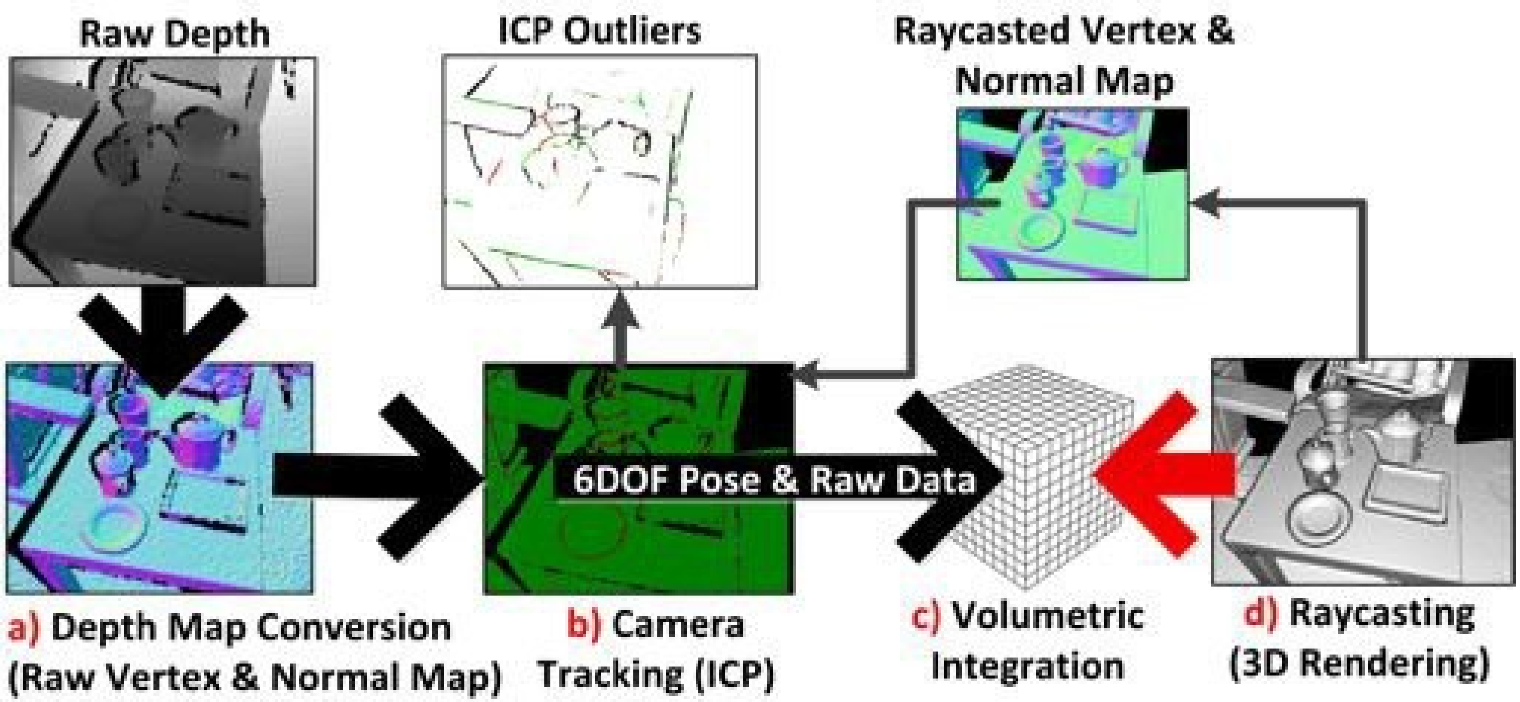


Live Stream → Reference 3D Model → AR with the user's head → Fast Motion and ICP Failure → ICP Recovery with Head Pose Estimation

Robust Real-Time Face Tracking using Head Pose Estimation

- Our method is inspired by two recent works:
  - KinectFusion: used to reconstruct the user's head (Izadi et al. 2011);
  - Real-Time Head Pose Estimation using Random Regression Forests: used to estimate the head pose (Fanelli et al. 2011);

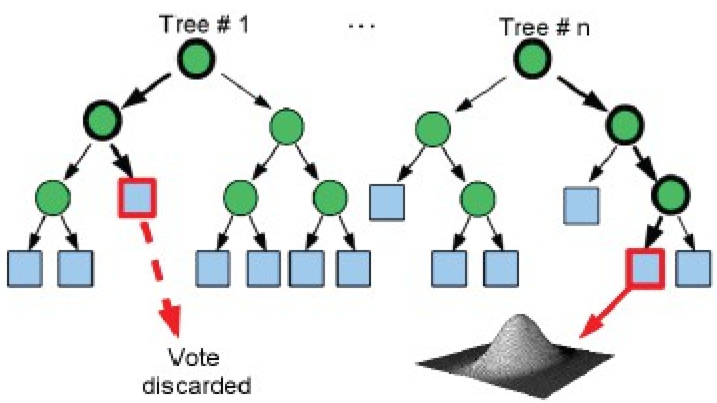
KinectFusion

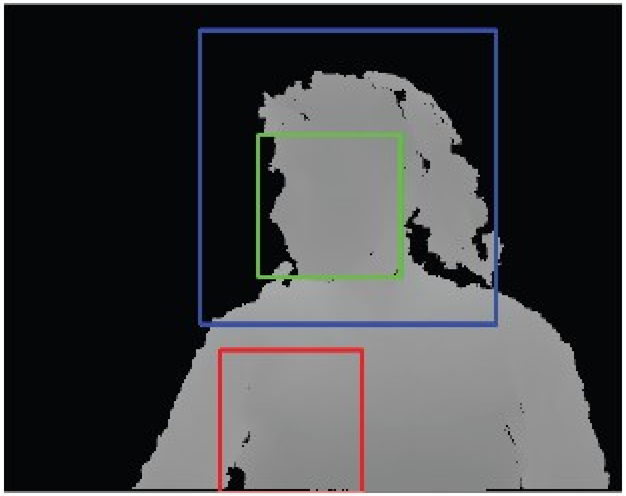


Izadi et al. (2011) “KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera”

Real-Time Head Pose Estimation from Consumer Depth Cameras using Random Regression Forests

- Fanelli et al (2011) trained random forests to estimate head pose from low-quality depth images.
- To train the trees, each depth map was annotated with labels indicating head center and Euler rotation angles.
- This probabilistic approach achieves high-quality and runs in real-time using only CPU.





Fanelli et al. (2011) “Real Time Head Pose Estimation from Consumer Depth Cameras”

Our Approach

- For each new depth frame:
  - We segment the user's head by applying a Z-axis threshold of 1.3m + Viola-Jones face detector;
  - We apply the ICP algorithm to compute the current camera pose;
  - When the ICP fails, the head pose estimation is used to give a new initial guess to the ICP.

Our Approach

Algorithm 1 Use of the head pose estimation

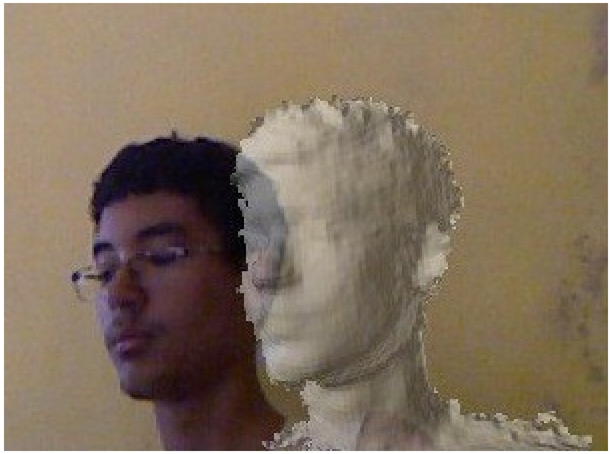
estimate head pose of D_{prev}.
$R_{prev} \leftarrow$ extract rotation matrix estimated from D_{prev}.
$Hc_{prev} \leftarrow$ extract global head center from D_{prev}.
estimate head pose of the D_{curr}.
$R_{curr} \leftarrow$ extract rotation matrix estimated from D_{curr}.
$Hc_{curr} \leftarrow$ extract global head center from D_{curr}.
$R_{inc} \leftarrow R_{curr} * R_{prev}^{-1}$.
$\Delta t \leftarrow Hc_{prev} - Hc_{curr}$.
$t \leftarrow t + \Delta t$.
raycast the implicit surface to generate a new view.
rotate the raycasted view around Hc_{curr} with R_{inc}.

D_{prev} – Previous depth frame
D_{curr} – Current depth frame

Results and Discussion

- We tested our algorithm with real data captured with a Kinect sensor using a grid with volume size of 50cmx50cmx130cm.
- We can analyze the qualitative performance for two cases: fast translation and rotation of the user's face.

Results and Discussion – Fast Translation



A

B

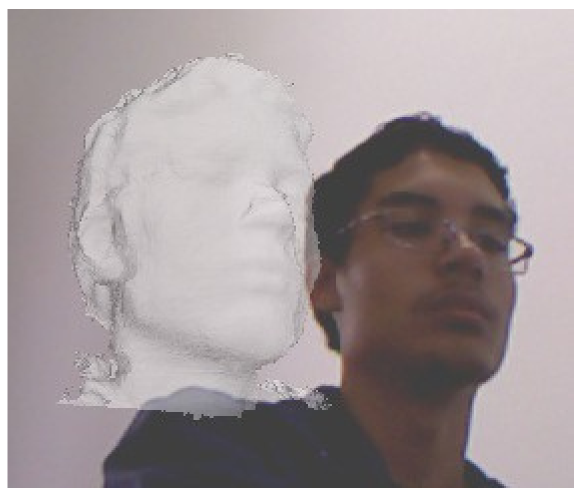
- A) The user translated his face fast. A small number of points were at the same image coordinates and the ICP failed;
- B) By applying our approach we solved this problem.

Results and Discussion – Fast Rotation

- The algorithm slighty improved the tracking performance;

Results and Discussion

- There are cases that the tracking algorithm fails even if the head pose estimation provides the initial guess.
- In this case, the user needs to reposition his face.



Results and Discussion

- The accuracy of the head pose estimation is the same as the Fanelli's approach:
  - Angle error: about 8º in each axis;
  - Head center error: 10mm

Conclusions and Future Work

- We presented an approach for robust real-time face tracking using head pose estimation for a markerless AR system;
- We used the KinectFusion and extended its tracking algorithm using a head pose estimation algorithm;
- We showed that this approach can handle more face pose changes than the original KinectFusion's tracking.

Conclusions and Future Work

- Encouraged by the work of Meister et al. (2012), for future work we plan to analyse the accuracy of the system to check if this method can be used for medical applications.
- Further improvements:
  - Deformable registration algorithm;
  - Improve the ICP matching algorithm;

Acknowledgments

- We are grateful to Gabriele Fanelli and the PCL project for providing the open-source implementations.
- This research is financially supported by FAPESB and CAPES.

Thanks...

?

Márcio C. F. Macedo
(marciocfmacedo@gmail.com)
Antônio L. Apolinário
(apolinario@dcc.ufba.br)
Antonio C. S. Souza
(antoniocarlos@ifba.edu.br)