Instructor Tracking

Radu Dondera

09.28.2006

Tracking for Distance Learning

- Tracking is the continuous computation of the position (2D or 3D) of an object of interest

- Distance Learning project needs
  - track the center of the head of the instructor
  - update image on back screen, based on 3D position
  - create the illusion of a window into a virtual extension of physical classroom

- Tracking components
  - camera calibration
  - appearance detection
  - head position computation

Related Work

- Various approaches to various human tracking problems
  - temporal differencing for controlling an active camera [1]
  - optical flow + head model to determine head pose [6]
  - skin color detection for face tracking [7]
  - stereo + plan-view model for tracking in multi-person sequences [2]
  - feature & template based tracking … [3]
  - 3D model … [4]
  - background subtraction … [5]

- Devising an approach
  - consider problem setting specifics: indoor environment, slow motion of subject
  - consider needed result: 3D position only

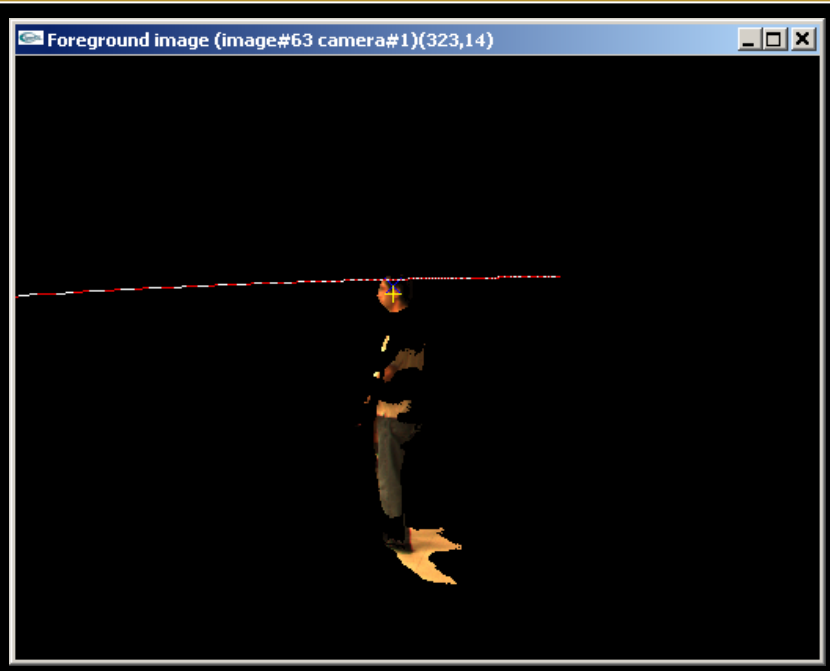
Our Method

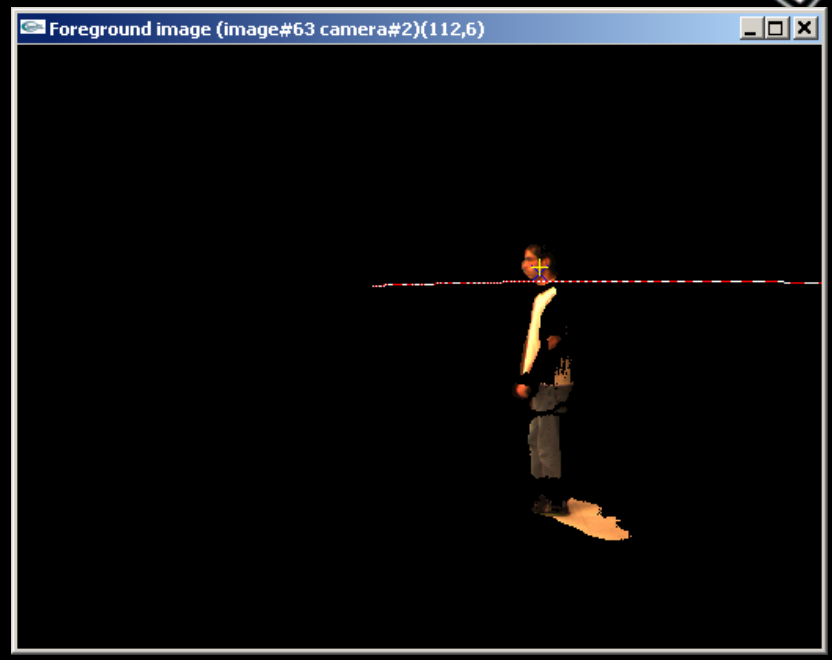
- Background subtraction in normalized color space
  - (r,g) color space to minimize the effects of illumination changes
  - non-adaptive background model because physical setting is static
  - set high thresholds for the silhouette to be eroded
- Approximation of head center
  - compute bounding box of foreground
  - compute head center as average position of foreground points in “top region” of bounding box
- Computation of 3D position of head center
  - have 2 calibrated cameras, perform above steps on both images
  - triangulate the 2 head centers to get 3D position

Problems Addressed

- Instructor appearance and posture
  - the instructor may wear any type of clothing
  - the instructor may rotate arbitrarily (face away from the cameras)
  - the instructor may make many types of gestures
- Illumination effects
  - effects of shadows, multiple local reflections are minimized
  - the majority of illumination effects appear near bottom of foreground bounding box
- No need for stereo hardware

Results





- Image resolution: 512x384
- CPU: 3.4 GHz
- Frame rate: 32 Hz
- Avg imprecision: 0.028 m; max imprecision: 0.719 m

Conclusions and Future Work

- Computing the 3D position of the instructor's head is a significantly simpler problem than head pose estimation
  - a simple scheme that tracks the head centers in images from two cameras and triangulates the 3D point is a promising solution

- Future work directions
  - Extend algorithm to handle multiple persons
    - students may appear in the classroom images
  - Improve the robustness of head center detection
    - the instructor may occlude his/her head with hand
  - Devise a better method to evaluate results of tracking
    - simulation
  - Use coherence to improve algorithm speed

Bibliography

- [1] Arseneau, S. and Cooperstock, J.R. (1999). Presenter Tracking in a Classroom Environment. IEEE Industrial Electronics Conference, Session on Cooperative Environments, Vol. 1, pp. 145-148.
- [2] Harville, M. (2005). Stereo person tracking with short and long term plan-view appearance models of shape and color. Proceedings IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 522-527.
- [3] Yang, R. and Zhang, Z. (2002). Model-based Head Pose Tracking With Stereovision. Proceedings Fifth IEEE International Conference on Automatic Face and Gesture Recognition.
- [4] Gavrila, D.M. and Davis, L.S. (1996). 3-D model-based tracking of humans in action: a multi-view approach. Proceedings IEEE Computer Vision and Pattern Recognition.
- [5] Stauffer, C. and Grimson, W.E.L. (1999). Adaptive Background Mixture Models for Real-Time Tracking. Proceedings IEEE Computer Vision and Pattern Recognition, pp. 246-252.
- [6] Black, M. J. and Yacoob, Y. (1995). Tracking and recognizing rigid and non-rigid facial motions using local parametric model of image motion. Proceedings of International Conference on Computer Vision, pp. 374-381.
- [7] Yang, J., Stiefelhagen, R., Meier, U., and Waibel, A. (1998). Real-time face and facial feature tracking and applications. Proceedings of Audio Video Speech Processing.