COM-US

Authors: Craig Carthel, Stefano Coraluppi

Email: stefano.coraluppi@compunetix.com

Platform: Windows

Prerequisites: None

COM-US: SUMMARY

Our approach includes distinct detection and tracking stages. The detection stage includes image pre-processing (convolution-based low-pass filtering), image segmentation (with a computationally-efficient iterative histogram-processing approach), and object extraction (baricenters of all large-enough contiguous object regions). The automatic tracking employs the multiple-hypothesis tracking paradigm, and includes data association (with small hypothesis-tree depth), sequential track extraction (with feedback to data association processing to favor confirmed tracks), and track post-processing (to identify parent links). Tracking is based on a nearly constant position object motion model. We have not yet introduced feature-aided tracking for these datasets. Further refinement to track post-processing will be required for improved performance.

COM-US: PREPROCESSING

Each 2D or 3D frame is smoothed by image convolution with a unity-valued disk (2D) or ellipsoid (3D) followed by normalization. This removes spurious spikes from the image data.

COM-US: SEGMENTATION

The image segmentation process is motivated by the heavy computational burden associated with processing of the fairly large 2D or 3D images. Thus, we start by forming an intensity histogram for each image. Next, we identify the best-fit N-point approximation to the histogram. This is defined as a set of N intensity values, chosen such that the average displacement from an intensity value to the closest of the N points is minimized. The pixels that are mapped to the highest and lowest of the N values are retained and discarded, respectively. The remaining pixels are then approximated again with an N-point approximation. The procedure continues until the remaining histogram contains N or fewer values. The middle value defines the detection threshold. All pixel values that match or exceed the detection threshold define the object region; all other pixel values define the background region. The object region leads immediately to object detections, as each set of contiguous pixels defines an object. The barycenter of each contiguous object region is an object detected. Small objects are discarded.

COM-US: TRACKING

The automatic tracker takes as input the sequence of detection sets that result from the detection processing described above. The automatic tracker is based on a *multiple-hypothesis tracking* (MHT) paradigm that we have enhanced over the years, principally for defense surveillance applications; see [1] and references therein. Our approach relies on (1) sequential data-association and (2) track-extraction processing, with some feedback from the track-extraction module to the data-association module as described below. In the data-association module, all detections are accounted for in multiple-hypothesis processing that partitions the set of all detections. Each subset includes at most one detection per time point. Association decisions are based on identifying the *maximum a posteriori* (MAP) global hypothesis with a fixed delay; all competing global hypotheses that conflict with the MAP solution far enough in the past are discarded, while hypothesis diversity in the recent past is maintained. Track-oriented MHT does not require explicit enumeration of global hypotheses. Hypothesis management logic limits the hypothesis space by disallowing sufficiently unlikely associations and considering either a missed detection or an object death hypothesis in the absence of a measurement update. Hypothesis selection relies on local (track) hypothesis scores that in turn, depend on birth and death statistics, detection statistics, and object motion and localization statistics that are used in recursive Kalman filtering. Spurious false alarms are removed in the subsequent track-extraction stage. The processing entails a sliding-window *M-of-N* test that promotes tentative tracks to nearly-confirmed and, ultimately, to confirmed track status. This impacts track scoring in the data-association stage, thus introducing a feedback mechanism that enhances overall tracking performance by favoring nearly confirmed and confirmed tracks over competing ones. At most *K* missed detections are allowed before track termination.

COM-US: POST-PROCESSING

Although tracker processing does not rely on the detailed shape characteristics of object detections, it is important for these to be available at the tracker output for **SEG** performance evaluation. Thus, when a detection is available, the object state estimate is replaced by the latest detection. When a detection is not available (i.e., track coast), the object state estimate defines the single pixel that is labeled in the output file. While the single-pixel objects represent a node mismatch with respect to **TRA** processing, it is necessary to ensure proper individual and parentage link identification in our current processing scheme. Indeed, the final processing stage considers each object track in turn. For each track, we identify the closest potential sibling at the time of object birth. If the potential sibling is close enough, a parent link to the track at the previous time is introduced, if the track exists. Additionally, a parent link between the track at the previous time and the identified sibling replaces the individual link. A more effective track post-processing scheme that avoids the need for single-pixel coast-track objects could be considered as a future improvement.

REFERENCES

1. Coraluppi S, Carthel C. Modified scoring in multiple-hypothesis tracking, *ISIF Journal of Advances in Information Fusion* **7**, 153-164 (2012).