Configuring sensors for the broadest view

Firdaus Janoo

Coverage and feature saliency can be employed to determine optimal orientation of a camera surveillance system.

Pan-tilt-zoom (PTZ) cameras, widely used in surveillance applications, provide extensive coverage at low cost. However, camera location and orientation are critical to system behavior and efficient resource utilization. Multiple coverage for sensitive areas, fault tolerance, and system reconfiguration to camera drop-out are also important. Any system should contend with these issues in robust and predictable ways.

Sensor placement and orientation are usually addressed ad hoc or considered, in terms of computational geometry, as instances of the art gallery problem (AGP), which is known to be NP-hard in two and three dimensions. Yet solutions at present assume unrealistic camera capabilities such as unlimited depth-of-field, resolution, depth-of-focus, and ultra-fast reconfiguration speed. Furthermore, they do not allow specification of additional coverage requirements, such as feature importance and multiple-sensor coverage.

We have developed a method to determine optimal camera orientation that takes into account realistic parameters while ensuring quality of coverage. This allows computer-vision algorithms to best use captured video data for analysis and reconstruction. Additional coverage is specified using a ‘saliency map’. Better than existing techniques, our method can contend with real-world complexities such as occlusions and undulations or irregularities in the ground surface or area under surveillance. The algorithm is designed so that system configuration can be determined on-line in response to changing requirements and camera failure.

We built a model of our environment using 3D modeling software (Maya), as shown in Figure 1, with optimal locations of the cameras determined using a method developed by Murray et al.2

The saliency map is a 2D grid of values that specify the relative coverage of corresponding polygonal cells on the ground surface under surveillance. Saliency values vary from 0.0 to 1.0, where 0.0 implies no coverage of the cell is needed (e.g., the interior of a building) and 1.0 indicates that maximum coverage is required (e.g., a bicycle rack or building doorway). Saliency can be user-specified or, as shown in Figure 2, created automatically by recording human activity in the environment.3

To quantify the quality of camera coverage, we developed a non-linear metric by taking into account sensor resolution and distance to the ground. Determining the set of orientations that best satisfied given requirements was posed as a minimization of cost function that incorporated the parameterized geometry of the camera view, the coverage quality metric, saliency map, and environmental occlusion. The rationale behind the cost-function formulation is that it seeks to maximize coverage of the cells weighted by saliency map values. We performed optimization using the downhill simplex optimization technique with simulated annealing to avoid local minima.

Our project covered a portion of a large urban campus, approximately 500m² with multiple buildings, sidewalks, trees,

Continued on next page
and other occluding objects. The resulting sensor configurations (Figure 3) satisfied real-world security and surveillance concerns. We used various saliency maps, including one extracted from activity captured by cameras mounted on buildings in the campus area. Currently, the optimization takes about 15 minutes to run.

We are also investigating real-time reconfiguration of the sensors, and so are currently evaluating other optimization techniques, and considering trade-offs in terms of speed. Of allied interest is the prospect of incremental updates to solutions as a result of changes in the environment, such as new occlusions. This project is part of a larger effort to develop high-level reasoning surveillance strategies. These include behavioral template matching, path clustering, and recognition of anomalous activity patterns, paths, and individual actions.

This material is based upon work supported by the National Science Foundation under Grant 0428249. We also acknowledge the support of AudoDesk for the use of Maya.

Author Information

Firdaus Janoos
Computer Science and Engineering
Ohio State University
Columbus, OH
http://www.janoos.net
http://www.cse.ohio-state.edu

Firdaus Janoos is a graduate student in Computer Science and Engineering at the Ohio State University. His research interests include computer vision, image analysis, medical image analysis, and computer graphics.

References