Multi-sensor fusion and integration play a major role for robotics and computer vision. The autonomy in navigation and manipulation that robotics pursue as its goal requires ultimately the ability of a robot to recognize and model the environment they are engaged in, often relying on vision. One of the key issues the robotics and vision field is facing today is how to solve the dependability in recognition and modeling against the many real-world variations, such as the variations in illumination, texture, surface reflection, occlusion, form factor, sensor pose, etc., in spite of the presence of fundamental limitations of sensing in, e.g., dynamic range, resolution, measurement error, field of view, etc. Multi-sensor fusion and integration have been regarded as indispensable for solving the issue of dependability addressed above. This chapter presents how multi-sensor fusion and integration can be applied to the dependability in recognition and modeling of environments, in particular, for robotic navigation, manipulation, and interaction with human.

The first two papers address the integration and fusion of two heterogeneous sensors, laser scanners and cameras, to improve the performance of Simultaneous Localization and Map Building (SLAM). The paper entitled “Simultaneous Estimation of Road Region and Ego-Motion with Multiple Road Models,” by Yoshiteru Matsushita and Jun Miura addresses the multi-sensor based simultaneous estimation of road region and ego-motion based on a particle filter. A laser range finder and an omni-directional camera system are integrated to detect and fuse the L-shaped curb and the road boundary lines and roadside regions. For the latter, the intensity gradient and the color gradient images are used, respectively. In addition, particles representing the gradual road type change are incorporated in the particle filter. Autonomous driving of a mobile robot is demonstrated as experimentation. The paper entitled “Visual SLAM in Indoor environments using Autonomous Detection and Registration of Objects” by Yong-Ju Lee and Jae-Bok Song presents how a hybrid grid/vision map can be built by integrating vision detected objects with an IR scanner based grip map. Various 2D visual cues are used to distinguish objects from the background for detection. The authors claim that their approach requires a less number of landmarks than the conventional laser scanner based SLAM.

In their paper entitled “The ‘Fast Clustering-Tracking’ Algorithm in the Bayesian Occupancy Filter Framework,” Kamel Mekhnacha, Yong Mao, David Raulo, and Christian Laugier propose clustering the occupancy and velocity grid into an object
level report, where the occupancy and velocity grid describes the probability distribution of cell occupancy and cell occupancy velocity. Note that a grid of occupancy probabilities and mean velocity estimates representing environments is from the Bayesian occupancy filter presented as a unified framework for sensor integration and fusion by the authors. A fast clustering algorithm is proposed as a means of avoiding the combinatorial complexity in computation. In their paper entitled “Fusion of Double Layered Multiple Laser range Finders for People Detection from a Mobile Robot,” Alexander Carballo, Akihisa Ohya and Shinichi Yuta present a method for simple and accurate detection and tracking of people in an indoor public area based on multi-layered laser range finders. The laser range finders in a double layer configuration provide the 360 degree of surroundings at the human chest and leg levels based on data fusion. The paper has shown that not only the 3D model of people and their positions but also the direction the person is facing at could possibly be obtained.

The paper entitled “Model based Recognition of 3D objects using Intersecting lines” by Hung Q. Truong, Sukhan Lee and Seok-Woo Jang presents the recognition and pose estimation of 3D objects based on perpendicularly intersecting 3D line segments as the cues to match with the model. Probabilities are assigned to all the possible interpretations of the object pose based on the matching scores, such that the probabilities can be updated as more evidences are accumulated in time. The paper entitled “Pedestrian Route Guidance System using Moving Projection based on Personal Feature Extraction” by Takuji Narumi, Yashushi Hada, Hajime Asama, and Kunihiro Tsuji presents a method of the traffic-line guidance for pedestrians by projecting moving images on the site of estimated human gaze by a pan-tile projector. Surveillance cameras are used to estimate the human sight.

In their paper entitled “Behavioral Programming with Hierarchy and Parallelism in the DARPA Urban Challenge and RoboCup,” Jesse G. Hurdus and Dennis W. Hong propose a hierarchical state machine for the efficient construction, organization and selection of behaviors in such a way that a robot can exhibit contextual intelligence. Not only the arbitration of competing behaviors but also the assembly of behaviors into an emergent behavior take place in the proposed hierarchical state machine.

Ruben Smits, Tinne de Laet, Kasper Claes, Herman Bruyninckx, Joris de Schutter present “iTASC: a Tool for Multi-Sensor Integration in Robot manipulation” as a unified framework for integrating instantaneous task specification and geometric uncertainty estimation. iTASC helps specify complex tasks for a general sensor-based robot system based on system constraints. A people tracker based on encoders, a force sensor, cameras, a laser distance sensor and a laser scanner is implemented, where kinematic control and uncertainty estimation equations are derived based on 10 primary constraints, 7 uncertainty coordinates, 6 scalar measurements, and 12 secondary constraints. Mario Prats, Philippe Martinet, Sukhan Lee and Pedro J. Sanz show in their paper entitled “Compliant Physical Interaction based on External Vision-Force Control and tactile-Force Combination” that multi-sensor based compliant physical interaction is feasible based on the task frame formalism. They demonstrate the pull-opening of the handle with a parallel jaw
gripper, as well as the opening of a sliding door with a three-fingered hand, based on the external vision-force coupling and the extracting of a book from a bookshelf by tactile and force integration.

In their paper entitled “Recognizing Human Activities from Accelerometer and Physiological Sensors,” Sung-Ihk Yang and Sung-Bae Cho present the recognition of 9 kinds of human activities: walking, running and exercising, eating, reading, studying, playing, sleeping, based on an armband sensor system integrated with accelerometers and physiological sensors. The latter includes heat flux, galvanic skin response, skin temperature sensors and thermometer. The inability of accelerometer to detect near stationery activities is compensated by other sensors. About 74% accuracy is reported with the fuzzy logic used for decision. In their paper entitled “Enhancement of Images Degraded by Fog using Cost Function based on Human Visual Model,” Dongjun Kim, Changwon Jeon, Bonghyup Kang, and Hanseok Ko present an estimation of an air-light map generated by fog particles in order to enhance the image quality by subtracting the estimated air-light map from the degraded image. The estimation of an air-light map is based on the estimate of luminance distribution variation under the constraints of the sensitivity derived from the human visual model.