Computer Vision Augmented Geospatial Localization

Ashish Gupta

1 Synonyms

Autonomous navigation, GPS-denied Geo-localization, Simultaneous Localization And Mapping, Visual Odometry, Global Navigation Satellite Systems, Unmanned Aerial Vehicles

2 Definition

Geospatial localization is the estimation of global geographic location using, in part, geospatial analysis. Geospatial analysis uses statistical and other analytic techniques for data that has a geographic or spatial context to it, typically available in Geographic Information Systems (GIS). Geographic location is typically ascertained using Global Navigation Satellite Systems (GNSS) like GPS and GLONASS, which requires simultaneous line-of-sight connection with multiple satellites to estimate location within an error margin of a few metres. These constraints limit the use of GNSS based localization to outdoors with few obstructing structures in close proximity and a tolerance to uncertainty in exact location. In addition to these constaints, in many environments such as indoors, urban canyons, under dense foliage, underwater and underground, there is limited or no GPS access. Besides these naturally occurring constraints, GPS access can be easily blocked by jamming, spoofing and other GPS-denial threats in advesarial environments. Consequently, for Positioning, Navigation and Timing (PNT) applications, GPS must be augmented or supplanted by other sensors and systems. In such cases GPS is used for an approximate localization within a geographic region, which can range from tens to thousands of square metres based on the environment. Alternate techniques are used to acertain

Ashish Gupta
Ohio State University, Columbus OH 43210, United States, e-mail: gupta.637@osu.edu

exact location within this geographic region. Simultaneous Localization And Mapping (SLAM) techniques are popularly used in robotics to estimate the location of a robot in real time based on information acquired from the environment using sensors on board the robot [1]. It is typical to use multiple sensors like Inertial Measurement Units (IMU), single or double video cameras for monocular or stereo vision, Light Detection and Ranging (LIDAR), and Sound Navigation and Raning (SONAR). The choice of sensor suite is based on the environment (aerial, ground, under-water, indoor); the type of mobile platform; and the performance, processing, and cost budget. Vision based sensors are amongst the most popular in SLAM techniques since they are informative and cost-effective. In addition to acquiring sensor information of its vicinity and estimating its position, SLAM also builds a map of the geographic region in real time. There are different types of maps. However, with navigation being the principal objective, topological maps are most relevant. A topological map focusses on the connectivity between important entities in the environment with disregard to their exact location [2]. A metric mapping is used in conjunction with topological maps to compute a topometric map which can be used for exact localization within that geographic region [3]. This technique records sensor information with its registered location in a map database. Subsequently, while moving through that geographic region, sensor information can be used as query to the recorded map database to retrieve geospatial location in real time.

3 Historical Background

Long Range Navigation (LORAN) was a hyperbolic radio navigation system developed prior to the advent of GNSS based PNT. It used low frequency radio waves and covered several thousand miles, but had a poor accuracy of hundreds of metres. It was used for localization of naval vessels while crusing oceans. Such systems are too inaccurate for precision demands of present day PNT dependent systems. To overcome the accuracy issues of GNSS for precision aviation, a Local Area Augmentation System (LAAS) is used for precision aircraft landing in all weather conditions [4]. A VHF signal link from airport transmitters is used by aircraft to correct GPS signal for precise localization. Cellular capable devices can use Assisted GPS (A-GPS) for improved localization using information provided by the cellular network in conjunction with satellite signal for a quicker estimation of location. Localization using cellular tower triangulation is another alternative with a reasonable error of several tens of metres, but it is only feasible outdoors in urban areas. For indoor navigation, the IEEE 802.11 wireless LAN (WLAN) location tracking system is an option [5]. It uses received signal strength indication (RSSI) on a mobile devices and estimates location by comparison with a pre-computed database of RSSI measurements in that indoor environment. It accounts for signal propagation loss and can provide can accuracy of few metres.

These PNT systems are currently operational but are incumbent with high infrastructure cost and have other limitations like availability exclusively in urban environments. Moreover, they have a natural limitation of localization accuracy. In comparison, computer vision based pose estimation techniques used in robotics for SLAM have a comparatively high localization accuracy, but have historically been used for mapping small sized environments. However, success in the DARPA Grand Challenges for autonomous navigation of driverless vehicles over large distances using SLAM based techniques established the viability of computer vision augmented geospatial localization as a viable PNT alternative in GPS-denied or degraded environments [6].

4 Scientific Fundamentals

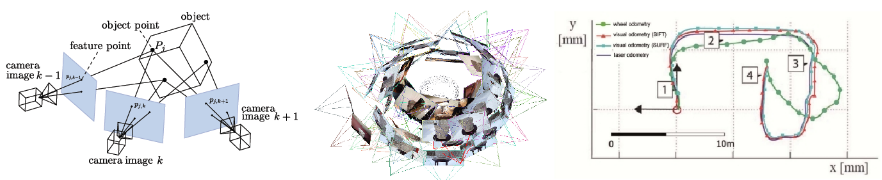

Fig. 1 Estimation of pose of sensors onboard moving vehicle. Trajectory of moving vehicle is estimated based on sensor locations in current and previous frames. Localization accuracy depends on the type of sensor and type of features computed from the data stream.

Visual sensor based localization using computer vision comprises two main parts in its approach: metric and topological localization. Metric location is estimated by computing the coordinates of the location of the sensor on board the vehicle. These could be geographic coordinates of latitude and longitude. The coordinates of the vehicle pose are typically computed by triangulation, using methods like Structure from Motion (SfM) [7] or Visual Odometry (VO) [8]. An illustrative example is shown in Figure 1. Pose of the sensor is estimated based on matching features across different frames in the data stream of a moving vehicle. The sequence of poses is used to estimate a 3D trajectory of this moving vehicle. In topological localization, the position of the sensor on board the vehicle is retrieved from a finite set of possible locations. Topological localization provides a coarse location estimate. Topological maps are typically stored as graph structures, where nodes indicate possible locations and edges are connections between locations. An example is shown in Figure 2 for the city of Washington DC, where the transport network layer acquired from OpenStreetMap for the county area has been abstracted as a graph. The weight of an edge can indicate the similarity or proximity between locations. The size of the finite set of locations is typically kept small so that efficient retrieval in real time applications is a tractable problem. While Metric approaches provide accurate localization results, they tend to fail and drift over time as the vehicle traverses big distances in its geographic region. On the other hand, due to its finite state space, topological approaches provide a robust localization but only rough position estimates. A fusion of the metric and topological approaches achieves accurate metric results while maintaining the robustness of topological matching, which is a technique typically referred to as topometric localization. It uses a fine-grained topological map, where each node has an associated coordinate of its real metric location. Such topological maps can be acquired from sources like GIS databases for outdoor navigation. Finding the node of the current location translates to finding the metric coordinate of the vehicle.

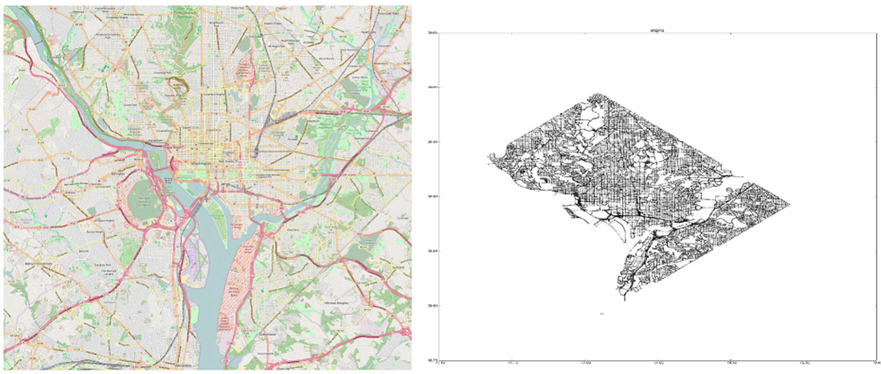

Fig. 2 Map data is abstracted as graph data structure. The transport network layer from OpenStreetMap of Washington DC is abstracted as a graph, where edges typically represent roads and nodes represent intersections and other relevant points in the map.

A generic topometric localization algorithm involves the two stages of map creation and then localization. A vehicle equipped with cameras, IMU and GNSS capable device first traverses the routes to be mapped. GPS and inertial sensors are used to create a graph of this environment. The graph is metric in the sense that the nodes contain the exact location of the vehicle. From the acquired images using an on board camera in monocular or stereo configuration, visual local features are extracted. These features are processed and stored in a database with a reference to the node corresponding to its real location. At runtime, the vehicle drives over the routes included in the a priori map. Video imagery is processed online to obtain features. As the vehicle moves, these visual features are matched with those in the database. Since, there are potentially multiple feature matches from different parts of the mapped region, a method like, Bayes filter is utilized to estimate the probability density function of the position of the vehicle. This facilitates pruning of false positive matches and provides accurate localization and a smooth estimated trajectory of the moving vehicle.

5 Key Applications

Geospatial localization using sensor suite that includes visual sensors in a GNSS denied environment has several applications in numerous scenarios. Since it provides very accurate location information in real time, it can be used for autonomous navigation for self-driving cars, unmanned aerial vehicles (UAV), unmanned ground vehicles (UGV), and unmanned under-water vehicles (UWV). A vision based system provides rich real-time information that allows an autonomous navigation system to tackle a dynamic environement, such as apperance of unexpected objects in the vicinity of the vehicle that were absent during the mapping phase, which is not available in other PNT systems.

A robust GPS alternative is particularly important for military applications since GPS signals can easily be denied to mission critical navigation systems on several assests, especially in contested territory. Relatively cost-effective vision based geo-localization can be alternatively used by guidance systems on weapons platforms like missiles, drones, and UGVs.

Community driven map generation projects like OpenStreetMap are extremely popular [9]. Accurate and information rich vision based localization can simultaneously correct registration errors in these maps and also annotate the maps with geo-referenced objects like buildings, road-signs, vegetation and other geographic entities.

Vision based localization is typically unhampered by its environment, unlike radio signals which suffer issues of multi-path and propagation path losses by absorption. Since, it can be used in most environments, it can also be used ubiquitously with disregards to change in environments like transitioning from outdoors to indoors or driving through tunnels, etc which otherwise typically require a hand-off between different PNT systems operational in their respective environments.

6 Future Directions

Computer vision augmented geospatial localization is a rapidly emerging technique. Future developments include improved sensor fusion where vision, inertial, LIDAR, SONAR, magnetometer, gravimeter sensors will be efficiently combined for ubiquitous navigation while travelling across different environments without degradation in localization accuracy. The quality of information in GIS databases and accuracy of geospatial localization are synergistic where one improves the other and vice versa. Cross-referencing and registration of visual information from different mobile platforms including UAV and UGV can improve GIS databases and provide a ground and aerial map of a geographic region for accurate 3D geospatial localization.

References

1. Lategahn, H., Geiger, A., Kitt, B.: Visual SLAM for autonomous ground vehicles. Proceedings - IEEE International Conference on Robotics and Automation (2011) 1732–1737
2. Paul, R., Newman, P.: FAB-MAP 3D: Topological mapping with spatial and visual appearance. In: 2010 IEEE International Conference on Robotics and Automation. (2010) 2649–2656
3. Badino, H., Huber, D., Kanade, T.: Visual topometric localization. IEEE Intelligent Vehicles Symposium, Proceedings (2011) 794–799
4. Enge, P.: Local area augmentation of gps for the precision approach of aircraft. Proceedings of the IEEE **87**(1) (Jan 1999) 111–132
5. Emery, M., Denko, M.: Ieee 802.11 wlan based real-time location tracking in indoor and outdoor environments. In: Electrical and Computer Engineering, 2007. CCECE 2007. Canadian Conference on. (April 2007) 1062–1065
6. Thrun, S., Montemerlo, M., Dahlkamp, H., Stavens, D., Aron, A., Diebel, J., Fong, P., Gale, J., Halpenny, M., Hoffmann, G., Lau, K., Oakley, C., Palatucci, M., Pratt, V., Stang, P., Strohband, S., Dupont, C., Jendrossek, L.E., Koelen, C., Markey, C., Rummel, C., van Niekerk, J., Jensen, E., Alessandrini, P., Bradski, G., Davies, B., Ettinger, S., Kaehler, A., Nefian, A., Mahoney, P.: Stanley: The robot that won the darpa grand challenge: Research articles. J. Robot. Syst. **23**(9) (September 2006) 661–692
7. Koenderink, J.J., van Doorn, A.J.: Affine structure from motion. J. Opt. Soc. Am. A **8**(2) (Feb 1991) 377–385
8. Alonso, I.P., Llorca, D.F., Gavilan, M., Pardo, S.a., Garcia-Garrido, M.a., Vlacic, L., Sotelo, M.a.: Accurate Global Localization Using Visual Odometry and Digital Maps on Urban Environments. Ieee Transactions on Intelligent Transportation Systems **13**(4) (2012) 1535–1545
9. Floros, G., Van Der Zander, B., Leibe, B.: OpenStreetSLAM: Global vehicle localization using OpenStreetMaps. Proceedings - IEEE International Conference on Robotics and Automation (2013) 1054–1059