Contextual processing for pedestrian tracking in GPS-denied environments

Enrico de Marinis[1], Michele Uliana[1], Fabrizio Pucci[1], Fabio Andreucci[1]

[1] DUNE srl, via Tracia 4, Roma

Abstract

This paper introduces the ConTeXtual Processing (CTXP), a novel and powerful concept for pedestrian tracking in GPS-denied environments. Its major advantages are: no need of external, ad-hoc infrastructures, no need of floorpan, low cost/weight/size, no need of calibrations or fingerprinting measurement campaigns, accuracy independent of the walked distance. In addition, CTXP processing is light enough to be hosted in a pocket-size commercial smartphone/tablet. CTXP has been extensively tested by Italian and International Agencies and industries in a widespread ensemble of scenarios (e.g. battleships, vessels, large industrial plants, malls) with experiments durations up to 3 hours and walked distance up to 6 km, always providing end-to-end results compliant with the target requirements.

Keywords

MEMS, Tracking, Processing, Inertial, IMU, Geolocation

1 Introduction

Substantial efforts and resources have been steered in the past decade toward Inertial Navigation Systems (INSs) for human tracking and localization based on Inertial Measurement Units (IMUs) based on Micro Electro-Mechanical Systems (MEMS) technology (Foxlin, 2005), (Leppäkoski et al. 2013). The major attractive is that these devices might provide low-cost, low-power, miniaturized, lightweight and infrastructure-less solutions for the navigation in GPS-denied scenarios. However they suffer significant bias, noise, scale factors, temperature drifts and limited dynamic ranges, resulting into position deviation and magnification of the angular Abbe error. These drawbacks de-facto prevent the use of MEMS IMUs for a long-range localisation. As a consequence, it is not surprising that in the recent years a widespread ensemble of techniques have been proposed to improve the localization capabilities. Most of the techniques rely on the Pedestrian Dead Reckoning (PDR) (Foxlin, 2005), where the walking behaviour is exploited to reset the INS errors by adopting an ECKF (Extended Complementary Kalman Filter). Other approaches achieve better performance by exploiting the presence of ancillary sensors, such compass (Jimenez et al., 2010), or by visual-inertial odometry (Li & Mourikis, 2013). Also pre-existing and independent sources of information are exploited, such as RFID tags (Jiménez et al., 2012) or map-matching techniques (Kaiser, Khidera, Robertson, 2013). The recent trends jointly exploit multiple-sensors readings (e.g. compass, barometer, RFID tags) into UKF (Unscented Kalman Filter) structures (Romanovas et al., 2013).

Scrutinizing the current State of the Art (SoA), the better performing solutions are based on the exploitation of two major sources of information: intrinsic (human-centric) and extrinsic (infrastructure-centric) information. The former is relevant to the operator himself and is basically limited to the information coming from body-mounted sensors (e.g. IMUs, barometer, compass, laser scanners); the latter is the information:

a) extracted from external sources (e.g. radio beacons, UWB devices, 3G/4G base stations, WiFi hotspots) that can be either ad-hoc deployed or can be just sources of opportunity (or both);
b) a priori information (e.g. floorpan, map).

The major advantage of the intrinsic information is that no position-aiding infrastructures are needed (expensive to deploy and maintain), but this comes at the cost of poor performance, as the available body-mounted sensor technology is not yet mature enough to insure an acceptable accuracy in the long term (hours). Moreover, some sensors need a calibration phase before the operations to avoid excessive drifts and errors. Despite the plethora of calibration methods (Ilewicz & Nawrat, 2013), some MEMS-based IMUs and compasses also suffer a long-term obsolescence of the calibration, from a few months to even 1 week. This would imply a re-calibration performed on a regular basis: an unacceptable task from the end-user perspective.

The solely use of extrinsic information has weak points as well, as the infrastructure cost, deployment time and maintenance can hardly meet the CApital Expenditure / Operating Expense (CAPEX/OPEX) budget and the set-up time constraints.

The better performing systems employ intrinsic information augmented by extrinsic information (e.g. joint use of IMU body-mounted sensors and Ultra Wide Band (UWB) transmitters deployed in the environment). On one hand the accuracy boost can be significant, on the other hand the drawbacks and limitations due to the CAPEX/OPEX and deployment time of the infrastructure(s) can easily jeopardize the real adoption of such systems. In addition, not all the areas/buildings are allowed to provide a clearance to deploy a location-aiding infrastructures (for safety or privacy constraints).

The authors, starting from the results of the RESCUE. PATH-SAFE and EXPLORERS projects (co-funded by Italian research initiatives), beside the intrinsic and extrinsic domains, introduce a novel information domain: the contextual domain that has shown (more than 120 on-field experiments) unprecedented performances in terms of accuracy, cost, weight.

2 The Contextual Processing

2.1 The Contextual domain

The contextual domain is the information coming from the behavioural features extracted from the walked path. For instance, when a pedestrian, after some time, crosses again a point previously crossed, this generates a contextual information. Indeed, a typical PDR processing, because of mismatches and drifts, would estimate these points (actually matching in the ground truth) as separate points in a 3D space. The same happens by analysing the spatial pattern of specific elements of the walked path; for instance, when some stairs are waked upward and then, after some time, are walked downward, the two typical “stairs-shaped” patterns are estimated, because of the drifts and scale factors, as dislocated in space by a PDR processing.

In other words, the contextual information is related to points or elements of

the path crossed again after some time: each time a contextual point is detected, the DUNE proprietary ConTeXtual Processing (CTXP) software estimates the (many) mismatches and drift factors, providing a backward compensation of the estimated path and, at the same time, a superior compensation of the forthcoming track, still to be walked.
On a statistical base, the presence of contextual points has been verified in practically all the scenarios and situations analysed so far. In some scenarios the presence of contextual points is very high (e.g. ships and vessels); in other scenarios is pretty good (e.g. malls). Only in some very specific cases are hardly to happen (e.g. inspector visiting a nuclear plant), but in this special case the cooperating inspector purposely crosses again a few previously visited points, so to generate the contextual information.

2.1 The DUNE system

The whole system is made of simple elements: a foot-mounted MEMS-based IMU sensor unit (20 g, 30x30x25 mm) also hosting the PDR processing; the uncompensated, drift-affected, estimated path is transmitted via Bluetooth (BT) link to a commercial smartphone/tablet, where the DUNE proprietary Android App provides a first drift compensation based on barometer, compass and GPS data (if available and if reliable) as well as the logging of the detected context points (that can be automatic or manual, depending on the application). The uncompensated track and detected contextual points are then fed into the CTXP software, which provides a final drift removal.

2.1.1 System architecture

In Figure 1, the general overview of the system, along with the processing blocks, is presented. The architecture can be conceptually split into three major blocks: the Local Processing Block (LPB) and the Remote Post-processing Block (RPB) and the Contextual Processing (CTXP). The core of the former is an ECKF and its processing cycle is performed at the same sample rate of the IMU signals (typically from 100 Hz to 1 kHz); the RPB is based on a iterative drift compensation driven by the information coming from the ancillary sensors (e.g. compass, GPS, Altimeter), if available and reliable. The CTXP block is then in charge of providing the estimation of the drift factors coming from the contextual information.
The terms “local” and “remote” indicate the suggested, but not mandatory, physical location of the processing blocks: the data rate of raw IMU data (input of the LPB) rages from 30 to 150 kbps/user; whereas the output of the LPB (that is the input of the RPB) is approx 0.1 kbps/user. This suggests that, in case of a radio link between LPB and RPB, is by far better to perform the LPB on the same device hosting the IMU, thus requiring a very limited bandwidth to transmit the necessary data to the device where the operations of the RPB are performed (e.g. a smartphone).
The Local Processing Block: it performs the basic (almost “standard”) processing to get the (drift-affected) walked path from the solely inertial data.
The sampled signals (temperature-compensated accelerometers and gyroscopes) from the foot-mounted IMU unit are first fed into a step-detection algorithm (Benzerrouk, 2014); the detected stance, swing and stride phases of the foot, along with the sampled signals, are the input of a 15 states ECKF (Foxlin, 2005) that performs the PDR and thus provides a first, rough estimation of the track.

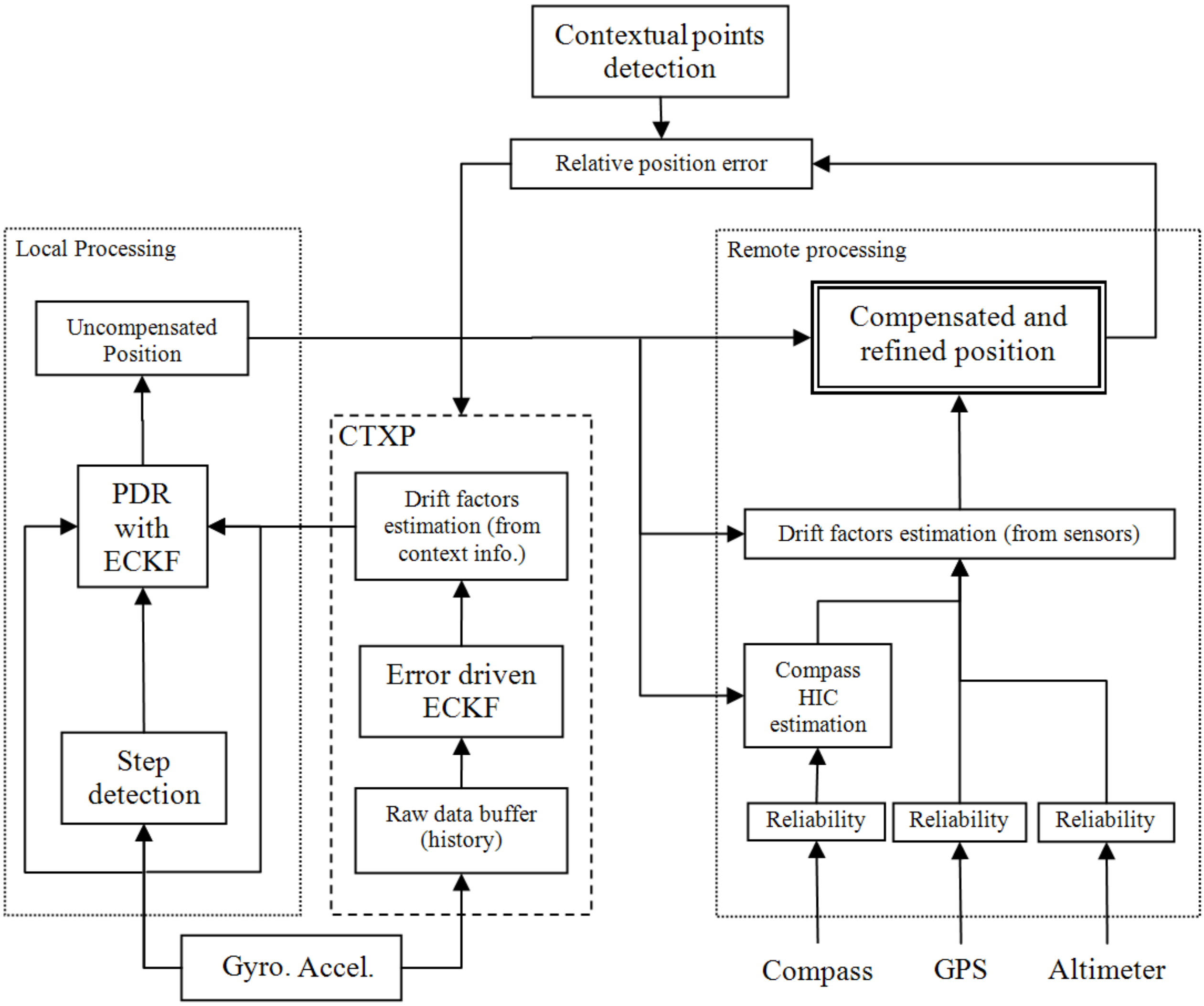


Figure 1. overview of the complete system, with the relevant processing blocks.

The Remote Processing Block: the estimated track from the LPB is the major input to the RPB. The RPB is in charge of providing, if possible, a drift estimation and compensation by processing the information coming from the other body-mounted sensors (GPS, altimeter, compass). The data from these sensors are first fed into a reliability check process, providing a weighting metric about the quality of the sensed information: the lower a metric, the less that information is accounted for the drift estimation process. Employing a proprietary algorithm of DUNE, the compass data are also processed to provide a runtime Hard Iron Calibration (HIC) of the sensor itself. Finally, the multi-sensor fusion of the drift-affected track and the weighted information coming from the ancillary sensors provides a first level of drift compensation by comparing: a) the absolute attitude estimated by the LPB and the one provided by the compass; b) the differential horizontal position estimated by the LPB and the one coming from the GPS; c) the differential vertical position from the LPB and the same information given by the altimeter. At this stage, with no CTXP activated, the error between the estimated path and the ground truth is still linearly growing as 1% of the walked distance (estimate at the 95th centile).
The Contextual Processing Block: from the conceptual point of view, the CTXP is “between” the LPB and the RPB, exploiting the information of both.

Once a context point is detected, the analysis of the track of the RPB (embedding a first drift removal, if possible) gives the estimated error between the locations that should match in the ground truth, but are estimated as dislocated in the 3D space, because of the residual, still uncompensated drifts. This error information is fed into another ECKF that, based on the re-processing of the buffered past IMU data, provides the equivalent Recursive Least Squares (RLS) estimation of the drift factors that would minimise the estimated error between all the detected context points. As these drift factors come from a stronger information, they are employed to refine the already waked path (via backward processing) and supersede the factors previously estimated by the LPB, thus providing a better estimation of the path still to be walked.

2 Experimental results

The CTXP has undergone so far to more than 120 experiments and in all of them (purposely) no information coming from the GPS or the compass has been ever employed. The scenarios range from malls, large urban areas, industrial plants to battleships and vessels (in harbour and in navigation) and underground caves. The experiments have been carried out by different walkers, with or without still phases as long as 1 hour (e.g. emulating a meeting or a dinner) and with no need of calibrations or mandatory starting path shapes, with durations up to 3 hours and walked distance up to 6 km.
Most experiments have been performed in “double blind”, as the path to be waked was not planned by the walker himself (who had no a priori knowledge of the planned path) and the operator in charge of the processing was not aware about the track actually walked.
The next Figure 2 and 3 show the results of a mixed indoor/outdoor experiment, (1h:49mins., 5.300 m), with 5 context points detected. Figure 2 represents the path estimated by a state of the art PDR algorithm, without GPS or compass information; the significant residual drift makes the estimated path useless after just a few hundred meters. Figure 3 represents the same path, but with the adoption of the contextual information; start-end error is 0,5 m for the 5,3 km path.
Figures 4 and 5 show details of the CTXP results of Figure 3. The former is the detail of an indoor/outdoor segment: in the multi-floors mall the GPS is totally obstructed and in the outdoor track is affected by significant errors (multipath). A similar situation is depicted in Figure 5, where 7 floors are walked up and down inside the building and the outdoor part is again affected by a significant multipath, leading to very poor GPS fixes.
As additional example, Figure 6 shows an experiment carried out in a multi-deck vessel (36 mins., 1.420 m, GPS absent, compass data always unreliable). Again, the path without CTXP is affected by a significant drift (start/end error: 34.5 m) that has been fully compensated by the CTXP (start/end error: 0.9 m).
Similar results have been obtained from all the experiments performed so far, from which it has been possible to estimate that the achievable 3D accuracy is largely independent of the walked distance, hardly exceeding 3-5 m or 1 m in the horizontal plane (without and with minimal presence of infrastructures, respectively) and 0.5 m in the vertical plane.

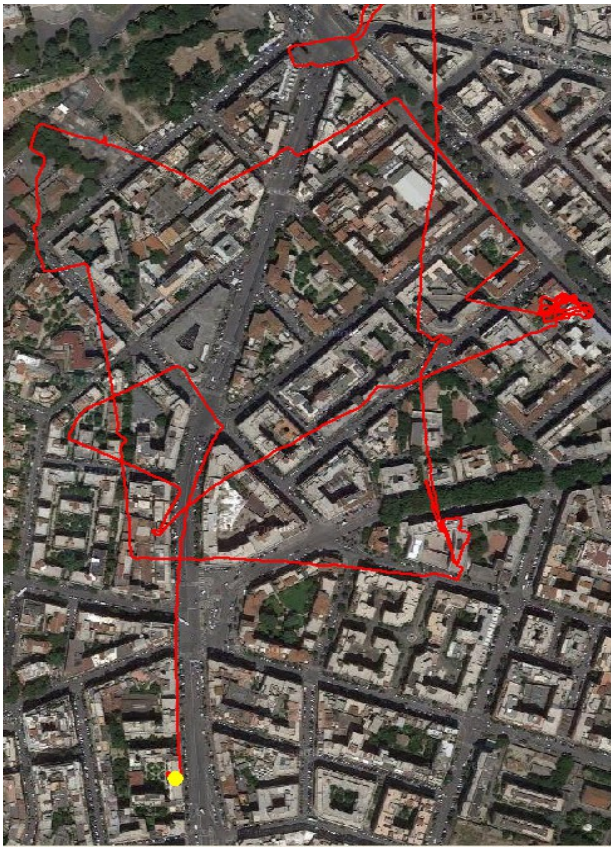


Figure 2. Path estimated by a PDR based on a 15 states ECKF

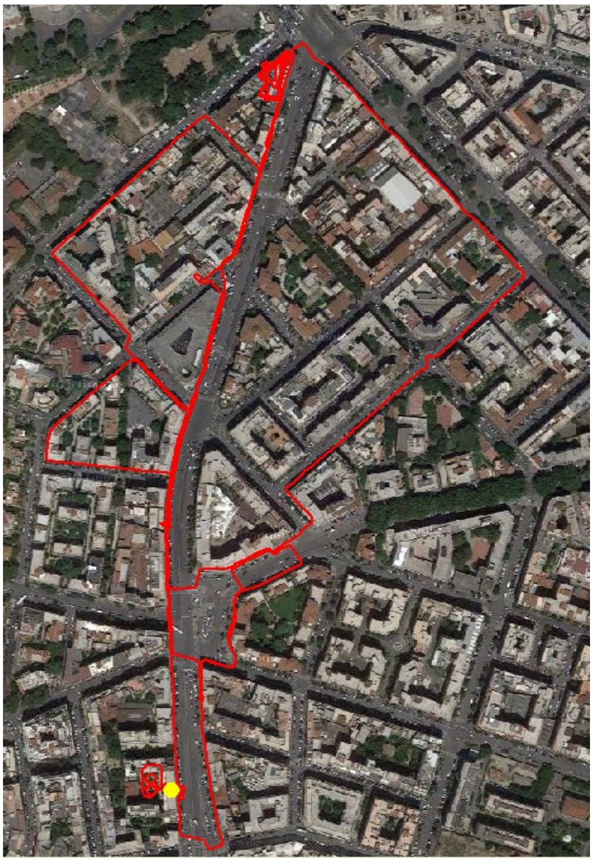


Figure 3. The same path of Figure 2, but processed with the CTXP

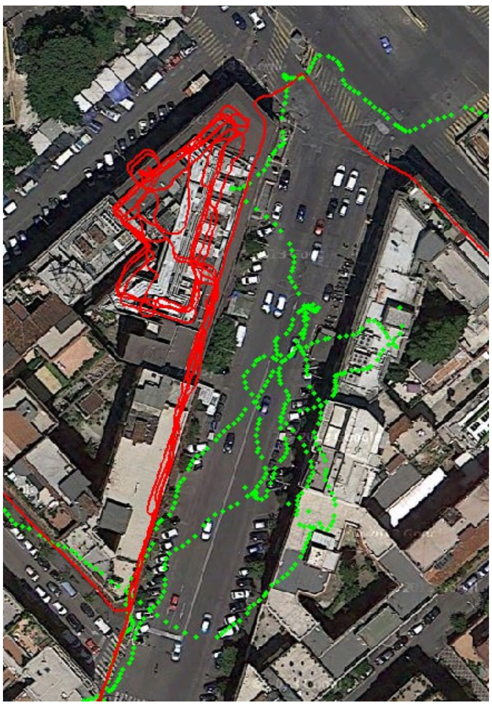
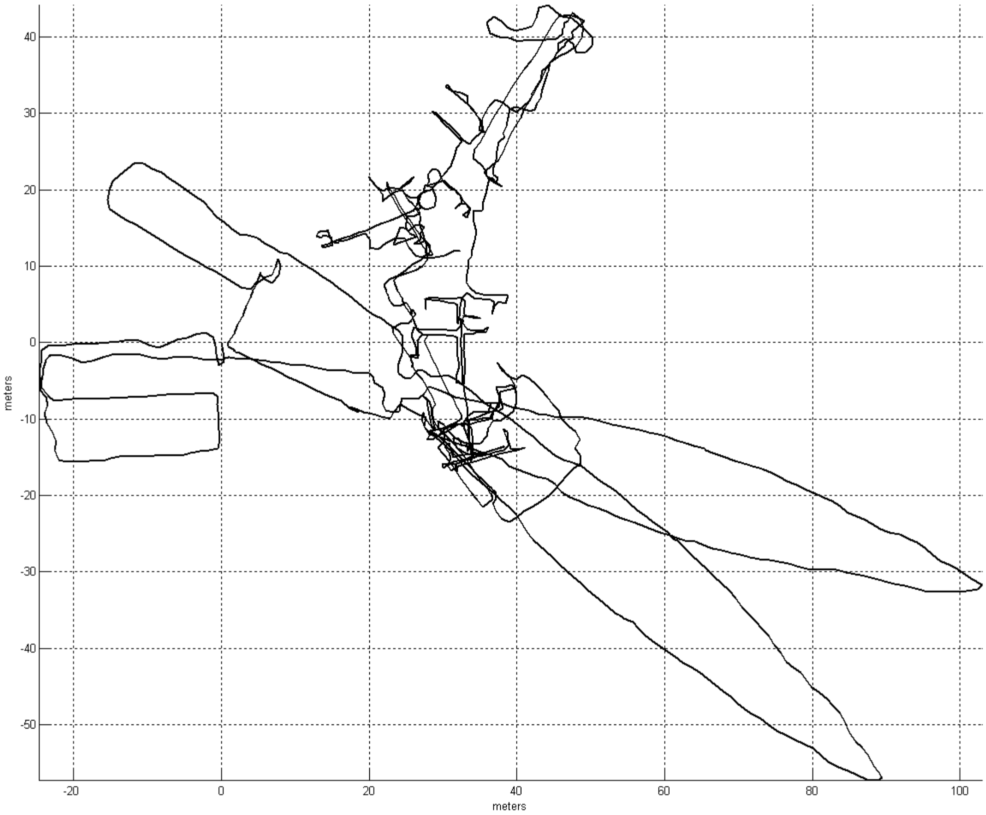


Figure 4. Detail of Figure 3: path walked indoor/outdoor. The green dots represent the GPS fixes.

Figure 5. Detail of Figure 3: path walked indoor/outdoor. The green dots represent the GPS fixes.



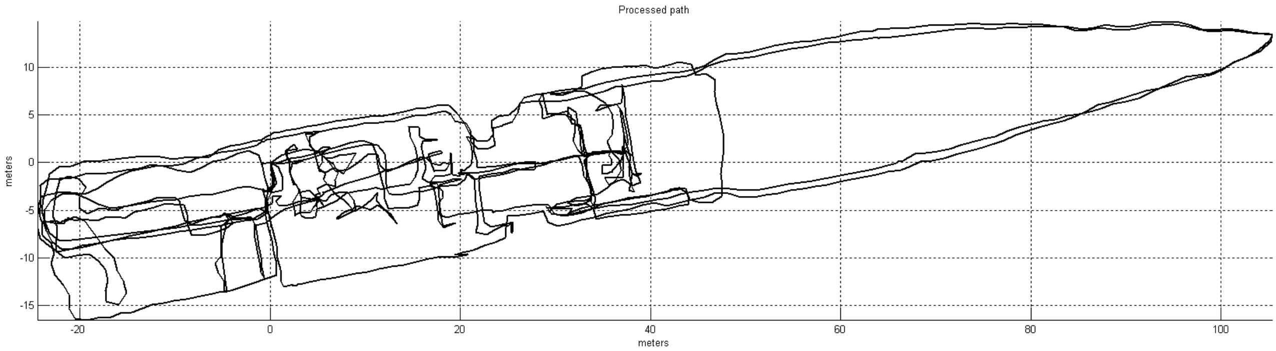


Figure 6. Experiment in a multi-deck vessel; upper plot: path estimated by a PDR based on 15 states ECKF; lower plot: path estimated by CTXP.

The following Figure 7 summarises the indicative accuracy with and without the CTXP block. The figures refer to the 95th centile, 1 hour walk in a dense urban environment (i.e. poor GPS performance and high magnetic pollution). More specifically:

- Without CTXP
  - In the absence of GPS, the processing provides an error below 1% of the walked distance (i.e. unbounded error);
  - When GPS is available, as a function of its reliability, the processing provides an error independent of the walked distance, ranging from 0.5 to 5 m.

- With CTXP
  - without any external infrastructure, (thus requiring manual interaction on the operator's side), the accuracy can be below 5 m, independently of the walked distance.
  - When a minimal infrastructure (e.g. BT tags) is allowed, the CTXP works in a drop-and-forget mode (no interaction with the operator). In this configuration, the accuracy can be below 1 m, independently of the walked distance.

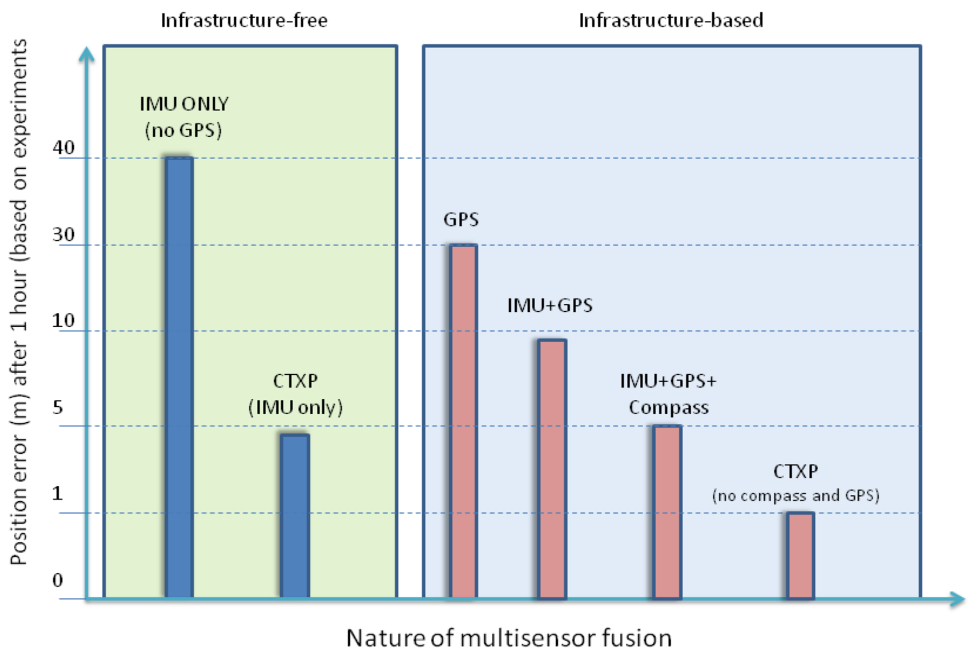


Figure 7. indicative accuracy achieved after 1 hour walk in a dense urban environment (i.e. poor GPS performance and strong magnetic pollution).

The robustness of the CTX has been also tested against the accuracy of the detected context points (i.e. when a context point has an error w.r.t. the unknown ground truth). This performance has been estimated over an ensemble of 20 on-field experiments, each processed 10 times; a set of context points has been given a uniform random error ranging from 2 to 10 m (by adding the error in the processing software), whereas a different set has been geo-located and used as a benchmarking set. The results show that the maximal error on the benchmarking points never exceed the maximal error injected on the context points and that at the 95th centile the mean value of the errors on the benchmarking set approximately equals the mean value of the errors given on the mislocated context points.

2 Conclusions

The CTXP has been designed to cope with the major drawbacks coming from the infrastructure-free and infrastructure-based localisation approaches. It exploits the contextual (behavioural) information domain to achieve a localisation error independent of the walked distance (typical of the infrastructure-based systems) with no actual need of a pre-existing location-aiding infrastructure.

Tested so far in more than 120 on-field experiments (covering a widespread ensemble of different conditions), in all of them the CTXP has achieved results fully compliant with the target requirements, thus demonstrating to provide a real and interesting improvement w.r.t. the state of the art of the localisation systems in GPS-denied environments.

References

Benzerrouk, H., Nebylov, A., Salhi, H., & Closas, P. (2014). MEMS IMU/ZUPT Based Cubature Kalman Filter applied to Pedestrian Navigation System. *In Proceedings of International Electronic Conference on Sensors and Applications*.

Foxlin, E. (2005). Pedestrian tracking with shoe-mounted inertial sensors. *IEEE Comput. Graph. Appl., 25*(6), 38–46. doi:10.1109/MCG.2005.140

Ilewicz W., Nawrat A., (2013). Direct Method of IMU Calibration. In Nawrat A., Simek K., Świerniak A. (Ed.) Springer Berlin Heidelberg. *Advanced Technologies for Intelligent Systems of National Border Security*. (pp 155-171), doi:10.1007/978-3-642-31665-4_13

Jimenez A.R., Seco F., Prieto J.C., Guevara J. (2010), Indoor Pedestrian Navigation using an INS/EKF framework for Yaw Drift Reduction and a Foot-mounted IMU. *WPNC 2010: 7th Workshop on Positioning, Navigation and Communication*, 135-143. doi:10.1109/WPNC.2010.5649300

Jiménez A.R., Seco F., Prieto J.C., Guevara R. J. (2012). Accurate Pedestrian Indoor Navigation by Tightly Coupling Foot-Mounted IMU and RFID Measurements. *IEEE Transactions on Instrumentation and Measurement, 61*(1)*, 178–189*. doi:10.1109/TIM.2011.2159317

Leppäkoski, H., J. Collin, J. Takala (2013). Pedestrian Navigation Based on Inertial Sensors, Indoor Map and WLAN Signals. *Journal of Signal Processing Systems, 71*(3)*, 287-296*. doi:10.1007/s11265-012-0711-5

Li, M., Mourikis A. I., (2013). High-precision, consistent EKF-based visual–inertial odometry. International Journal of Robotics Research, 32(6). doi: 10.1177/0278364913481251

Kaiser S., Khidera M., Robertson P. (2013). A pedestrian navigation system using a map-based angular motion model for indoor and outdoor environments. *Journal of Location Based Services, Special Issue: Indoor Positioning and Navigation. Part III: Navigation Systems, 7*(1), 44-63. doi:10.1080/17489725.2012.698110

Romanovas M., Goridko V., Klingbeil L., Bourouah M., Al-Jawad A., Traechtler M., Manoli Y. (2013). Pedestrian Indoor Localization Using Foot Mounted Inertial Sensors in Combination with a Magnetometer, a Barometer and RFID. In Jukka M. Krisp (Ed.) Springer Science & Business Media. *Lecture Notes in Geoinformation and Cartography: Progress in Location based Services*, (pp. 151-172).