Open-World Virtual Reality Headset Tracking

Todd E. Humphreys, Ronnie Xian Thong Kor, Peter A. Iannucci, and James E. Yoder
Radionavigation Laboratory
The University of Texas at Austin

BIOGRAPHIES

Todd Humphreys (BS, MS, Electrical Engineering, Utah State University; PhD, Aerospace Engineering, Cornell University) is an associate professor in the department of Aerospace Engineering and Engineering Mechanics at The University of Texas at Austin, where he directs the Radionavigation Laboratory. He specializes in the application of optimal detection and estimation techniques to secure and robust perception for automated systems and centimeter-accurate location. His awards include The University of Texas Regeants' Outstanding Teaching Award (2012), the National Science Foundation CAREER Award (2015), the Institute of Navigation Thurlow Award (2015), the Qualcomm Innovation Fellowship (2017), and the Presidential Early Career Award for Scientists and Engineers (PECASE, 2019). He is a Fellow of the Institute of Navigation.

Ronnie X.T. Kor is an M.S. student in the Department of Aerospace Engineering and Engineering Mechanics at The University of Texas at Austin, and a member of the Radionavigation Laboratory. His research interests are in estimation and terrestrial position, navigation, and timing security.

Peter A. Iannucci (BS, Electrical Engineering-Computer Science and Physics, MIT; PhD, Networks and Mobile Systems, CSAIL, MIT) is a postdoctoral research fellow in the Radionavigation Laboratory at The University of Texas at Austin, and a member of the UT Wireless Networking and Communications Group (WNCG). His current research interests include collaborative navigation, multi-spectral mapping, and re-purposing broadband Internet satellites for radionavigation.

James Yoder is a graduate student in the department of Aerospace Engineering and Engineering Mechanics at the University of Texas at Austin, and a member of the Radionavigation Laboratory. He received a B.S. in Electrical and Computer Engineering from the University of Texas at Austin in 2019. His research interests currently include GNSS signal processing, estimation, and sensor fusion techniques for centimeter-accurate navigation.

ABSTRACT

A novel outdoor Virtual Reality (VR) concept called Open-World Virtual Reality (OWVR) is presented that combines precise GNSS positioning and a smartphone-grade inertial sensor to provide globally-referenced centimeter-and-degree-accurate tracking of the VR headset. Unlike existing augmented and virtual reality systems, which perform camera-based inside-out headset tracking relative to a local reference frame (e.g., an ad-hoc frame fixed to a living room), OWVR's globally-referenced tracking enables a novel VR experience in which the user's outdoor exploration is robust to extremes in lighting conditions and local visual texture. This paper introduces the OWVR concept and presents a prototype OWVR system with two candidate sensor fusion architectures, one loosely and one tightly coupled. Comparative performance is evaluated in terms of tracking accuracy and availability of an integer-aperture-test-validated fixed tracking solution. For scenarios with degraded GNSS availability, which will be typical for outdoor VR, the tightly-coupled architecture is shown to offer a critical tracking robustness advantage.

INTRODUCTION

When used outdoors, existing virtual reality (VR) headsets perform unreliably. Lighthouse-based headset tracking cannot be used in sunlight, and camera-based inside-out headset tracking requires well-lit conditions with a rich local visual texture so that its simultaneous-headset-tracking-and-local-area-mapping algorithms can extract enough visual features to reliably infer the headset's relative position and orientation (pose).

Nonetheless, the outdoors beckon, offering unlimited space for VR in which every virtual step is a real step. Whereas simulating self-motion in a confined indoor space can result in visually-induced motion sickness [1], [2], outdoor VR could offer a mixed reality experience in which a user's self-motion is governed entirely by natural physical movement.

Preprint of the 2020 ION GNSS+ Conference Online, Sept. 21-25, 2020

Fig. 1. The Open-World Virtual Reality headset. Measurements from two GNSS antennas and a low-cost inertial measurement unit (IMU) mounted on the plate above the user's head are combined for centimeter-and-degree-accurate tracking of the VR headset in a global coordinate frame.

Further immersion is possible when a high-resolution 3D reconstruction of the local environment is available and when headset tracking is precisely referenced to the local structural environment. When these conditions hold, users can transform the real world into a personalized landscape in which they explore and even touch virtually-represented components of the real world. In effect, users can overlay a virtual "skin" of their choice on the physical world.

Existing VR systems have made modest progress toward such a tactile mixed reality. For example, when combined with real-time 3D mapping, current inside-out headsets are capable of rendering a chair in one's living room as a similar-looking chair in the VR experience, with correct pose relative to the VR headset. But the limitations of real-time visual 3D reconstruction cause such real-world objects to be rendered with unappealing coarseness in the virtual landscape. Moreover, inside-out tracking is highly sensitive to visual conditions: it cannot be used in low light or in a visually-homogeneous space such as the grassy area of a large park. Nor does the current approach support persistence: a virtual object cannot be placed permanently at a given location in the physical world.

This paper combines two techniques to unlock the great potential of tactile mixed reality with full object persistence: (1) centimeter-accurate fixed-integer carrier-phase differential GNSS (CDGNSS) anchored VR headset pose determination, and (2) prior high-resolution globally-referenced 3D mapping. Anchoring the VR headset pose to CDGNSS enables continued operation in variable outdoor visual conditions and permits virtual objects to be persistently located in the real world despite radical changes in the visual scene (e.g., snowfall, construction). Moreover, building a globally-referenced 3D map of the local real-world environment beforehand, as opposed to building it on-the-fly with limited computation, permits a beautifully high-resolution landscape to be constructed and enhanced or extended according to users' aesthetic preferences.

As early as 1997, differential GNSS was explored for augmented reality (AR) headset pose estimation [3], [4], but registration with the real world was poor because the GNSS antenna was not mounted on the headset and only pseudorange-differential GNSS was used. Later work fused position and orientation sensors with a visual-based tracker to improve tracking accuracy [5]. Shepard et al. incorporated CDGNSS and inertial measurement unit (IMU) measurements into visual simultaneous localization and mapping for precise and globally-registered AR [6], but this system never achieved real-time operation, and its handheld screen proved an awkward interface.

Compared to AR, in which the real world remains a central and visible backdrop to all digital information, VR fully eclipses the external scene, enabling greater experiential control. The most fully developed prior work in outdoor VR is Microsoft's Dreamwalker, which fuses inside-out tracking with filtered low-accuracy GNSS positions and RGB depth frames [7]. But besides having only 5-m accurate headset pose, which spoils virtual object persistence, Dreamwalker does not situate the user within a facsimile of the local environment, and so does not support the tactile mixed reality

envisioned in this paper.

An extensive body of prior work has explored coupling inertial sensors with CDGNSS processing to improve the reliability of integer fixing. Such fixing is required to obtain the best CDGNSS performance: centimeter accuracy and instantaneous convergence. But almost all such prior work has focused on coupling with tactical-grade [8]–[11] or industrial-grade [12], [13] IMUs, whereas the current paper couples CDGNSS with a low-cost (under 3 USD) smartphone-grade IMU for mass-market acceptance. Bochkati et al. [14] and Henkel et al. [15], [16] also couple CDGNSS with smartphone-grade inertial sensors. However, when operating with IMU coupling under dynamic scenarios, [14] only achieves float solutions, and [15] fails to demonstrate any benefits of IMU coupling: tests were conducted under open-sky conditions in which IMU-free CDGNSS integer fixing is already highly reliable. One interval of degraded GNSS conditions (near a group of trees) is explored in [16], but in this case the coupled CDGNSS and IMU system fails to maintain an accurate trajectory and must be aided by a vision system. By contrast, the present paper demonstrates significant benefits of IMU coupling under severely degraded GNSS conditions.

A companion paper [17] extends this paper's work to include vision sensing for robust pose estimation under more severely degraded GNSS conditions. The coupled IMU-CDGNSS system reported in the current paper can be thought of as a sub-system of the pose estimator in [17].

This paper makes three primary contributions. First, it introduces the Open-World Virtual Reality (OWVR) concept, which aspires to the full potential of outdoor VR. By fusing CDGNSS and IMU data for precise real-time user headset pose estimation, and by situating the user within a detailed globally-referenced 3D map previously constructed from high-resolution photos of a large outdoor space, OWVR enables a novel and compelling mixed reality experience.

Second, this paper presents a working prototype OWVR system that fuses GNSS observables from dual spatially-separated antennas rigidly mounted on a VR headset (Fig. 1) with high-rate sensor data from a smartphone-grade IMU to produce a stream of centimeter-and-degree-accurate pose estimates that are in turn fed to a VR rendering engine driving the stereo headset display.

Third, this paper demonstrates two IMU-CDGNSS fusion strategies—one loosely coupled and one tightly coupled—and evaluates these for robustness in the face of degraded GNSS availability. So far as the authors are aware, this paper's tightly coupled sensor fusion strategy is the first to demonstrate robustified CDGNSS integer fixing via coupling with a smartphone-grade inertial sensor.

THE OPEN-WORLD VIRTUAL REALITY CONCEPT

The OWVR concept has three primary ingredients: correspondence, customizability, and persistence.

Correspondence: OWVR seeks to mix the real and virtual worlds by rendering all permanent structural components of the real world as tactilely indistinguishable structures in the virtual world. The 3D-reconstructed scenes displayed by the headset are arranged in precise correspondence with the surrounding real physical environment. This enables the user to reach out and touch the world she sees: a nearby lamp-post appears in the virtual world, positioned just so as match its physical counterpart.

Customizability: OWVR allows users to adapt and extend the 3D virtual world according to their aesthetic preferences: the real lamp-post's drab gray metal can be rendered virtually as colorfully decorated in graffiti.

Persistence: OWVR enables full persistence of virtually-represented objects. The virtual lamp-post and a virtual treasure chest hidden near its base remain permanently at the same location with respect to a well-defined global coordinate system, even if the real lamp-post is torn down and the whole area is paved over and converted into a parking lot.

3D Mapping

The OWVR production pipeline begins with creation of a detailed 3D map of a large outdoor space. Visual 3-dimensional (3D) reconstruction of buildings, historic monuments, and landscapes has become a major theme in the computer vision, photogrammetry, and remote sensing literature in recent years [18]–[27].

The foreseeable, if ambitious, endpoint of large-scale visual reconstruction is a 3D map of the world—the digital realization of the 1-to-1 map of an empire imagined many years ago by Argentine novelist Jorge Louis Borges [28]. For maximum utility, the map should be high-resolution and globally-referenced to a well-defined coordinate system.

Fig. 2. Real (left) and virtual (right) views taken from a similar perspective. Most features, such as the parking lines and the retaining wall to the right, are common to both environments. But the table in the foreground of the virtual scene is purely virtual, having no counterpart in the real world. Similarly, some real-world elements, such as the netted drone cage to the left and distant buildings, were not realistically captured by the 3D mapping process.

For a compelling OWVR experience, millimeter-scale resolution and centimeter-level absolute accuracy are required. Such resolution is several orders of magnitude better than that offered by current global reconstructions such as Google Earth 3D, and such accuracy would improve on the absolute accuracy of Google Earth, which is limited to approximately 2 meters with respect to the WGS-84 reference frame [29], by factor of at least 20. Demonstrating that such mapping can be achieved over large areas via crowd-sourcing with inexpensive equipment will be the subject of later studies related to OWVR.

The OWVR concept envisions a metric textured point-cloud map that is fragmentary but global in scope, with each feature point in each map fragment having millimeter-accurate coordinates relative to nearby feature points and centimeter-accurate coordinates relative to the latest realization of the International Terrestrial Reference Frame (ITRF) (e.g., ITRF2014 [30]). The goal of centimeter-level absolute accuracy is inspired by the state of the art in high-end mobile mapping, whose 5-cm global accuracy has proven enormously useful for 3D mapping of road surfaces, urban structure, and vegetation [31], [32].

Micro aerial vehicles (MAVs) are a key enabler of large-scale high-resolution mapping. Although MAVs are currently used only in small numbers for 3D capture [26], [27], future fully-automated MAV swarms and computational pipelines will support construction of breathtakingly detailed 3D reconstructions of monuments, buildings, campuses, and perhaps eventually cities.

Customizing the Map

OWVR presents a 3D reconstruction of the real world created from a set of high-resolution images to the user through a VR headset in precise correspondence with the user's actual motion through the same outdoor space. Depending on the user's needs and aesthetic choices, the map may be augmented with purely virtual elements, or altered in other ways, but it remains fundamentally structured in agreement with the static physical structure of the local space. By adapting the map's coloring and texture, OWVR users can drape a virtual "skin" of their choice on the physical world. Entirely virtual elements—those having no corresponding element in the real work—can of course also be added to the rendered 3D reconstruction.

Fig. 2 illustrates the OWVR concept with a side-by-side comparison of images from the real and virtual worlds.

EXISTING VR HEADSET TRACKING

There exist three techniques in current use for VR headset tracking. The most precise, used for example by the HTC Vive Pro and the Valve Index, depends on external base stations, also called lighthouses, that emit laser beacons tracked by sensors on the headset. Sample-to-sample RMS tracking repeatability for this technique are impressive: under 0.1 mm and 0.01 degrees for position and orientation, respectively [33]. However, the system can experience large tracking errors when tracking is lost due to lighthouse signal blockage.

The second technique, called outside-in tracking, makes use of reflective or infrared-emitting markers attached to the headset. A set of external cameras mounted around the area of operation track the markers to determine the headset pose. Sub-millimeter accurate and highly reliable, this technique is the standard for scientific research in visual simulation of self-motion [33].

In the third technique, referred to as inside-out tracking, cameras mounted on the headset produce high-rate images from which visual features are extracted. These features become the basis for a simultaneously localization and mapping (SLAM) algorithm that provides pose estimation with respect to the prior map. See [34], [35] and references therein for a further description of this technique.

In recent years, inside-out tracking has seen widespread adoption because it is sufficiently accurate for an immersive VR experience yet is more convenient that lighthouse or outside-in tracking, as it requires no external hardware. Windows Mixed Reality headsets and the latest headsets from Oculus have all adopted inside-out tracking.

Unfortunately, none of the existing VR headset tracking techniques function reliably outdoors and so cannot support the vision of open-world virtual reality. Strong sunlight interferes with lighthouse-based and outside-in tracking, and lack of visual texture or poor lighting conditions, including shadows, challenge inside-out tracking. Moreover, all existing headset tracking techniques confine users to a limited area of operation. Even inside-out tracking, the most recent versions of which boast an extended usable space, still limit users to warehouse-scale exploration.

In principle, the size of the area of operation for inside-out tracking is only limited by the headset's capacity for map storage, which, for a wirelessly-connected headset, becomes virtually limitless, since the 3D point cloud of salient features can be stored in the cloud, subdivided, and retrieved on demand. Such a cloud-based SLAM architecture is the basis of Google's recently-released Live View pedestrian navigation system. But inside-out tracking will also perform poorly in areas that lack distinctive features, such as a large grassy field, or in which such features are arranged in repeating patterns.

Another failing of existing headset tracking techniques with respect to OWVR is their lack of registration to to a global coordinate system, which prevents the OWVR desideratum of fully persistent virtual objects despite radical changes in the visual environment.

OWVR HEADSET TRACKING

One of this paper's primary contributions is demonstration of robust outdoor VR headset pose tracking despite severe GNSS signal blockage and multipath.

The OWVR headset tracking system is illustrated schematically in Fig. 3. The fundamental sensors are an IMU and dual single-frequency GNSS antennas. A detailed description of the measurement and dynamics models of the pose estimator, which is implemented as an unscented Kalman filter (UKF), may be found in the companion paper [17]. This paper's UKF is identical to that of [17] but without any vision-related states or measurements. The following section offers a high-level overview of the components shown in Fig. 3; the reader is referred to [17] for a more thorough discussion.

Headset

The headset houses primary and secondary GNSS antennas rigidly coupled to an IMU. GNSS signals received through the antennas are filtered, amplified, downmixed to an intermediate frequency (IF), and synchronously sampled with the specific force and angular rate measurements from the IMU. To avoid the additional weight and bulk of multi-frequency antennas, the headset's GNSS antennas are single-frequency (L1) only.

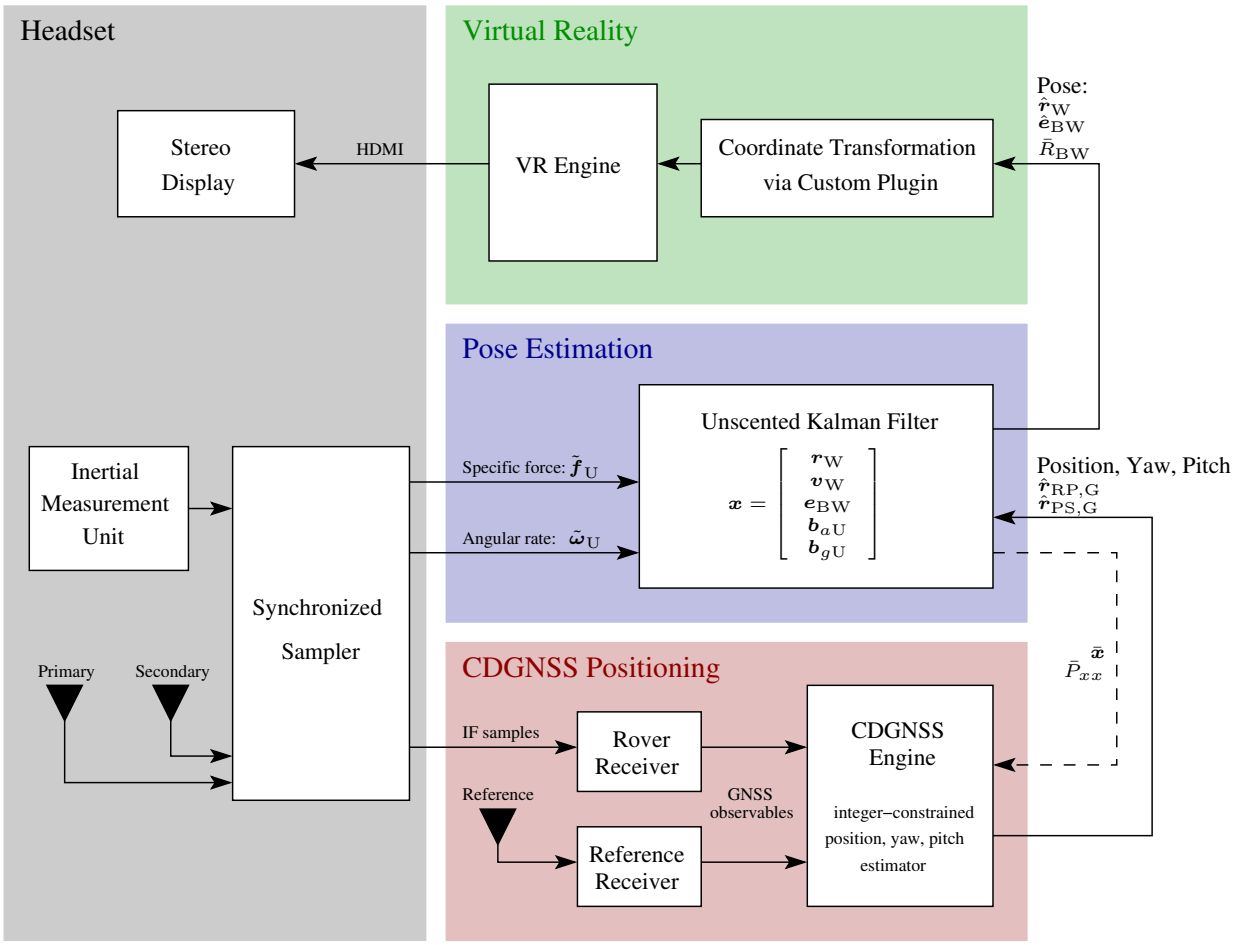

Fig. 3. The OWVR system architecture. The dashed line on the lower right denotes optional data flow from the pose estimation UKF to the CDGNSS engine. When this flow is absent, the GNSS-IMU fusion architecture is loosely coupled; when it is present, the architecture is tightly coupled in the sense that inertial measurements aid the resolution of the CDGNSS integer ambiguities.

CDGNSS Positioning

The CDGNSS positioning block pulls IF samples from the headset and processes these via an advanced software-defined GNSS receiver called GRID developed at The University of Texas [36], [37]. The receiver and front end are capable of multi-frequency GNSS tracking, but only L1-centered GNSS signals are tracked because the headset antennas are single-frequency. GRID feeds observables to a specialized research-oriented CDGNSS engine that has been designed for operation in urban environments. A detailed description of the CDGNSS integer fixing logic may be found in [37]. At each measurement epoch, the CDGNSS engine, an integer-constrained position, yaw, and pitch estimator, produces vectors $\hat{\boldsymbol{r}}_{\mathrm{RP,G}}$ and $\hat{\boldsymbol{r}}_{\mathrm{PS,G}}$. The former is a cm-accurate estimate of the vector from a reference antenna to the headset's primary antenna, expressed in a global Earth-fixed reference frame G. The latter is a constrained-length estimate of the vector from the headset's primary to secondary antenna, also expressed in G. This latter vector's two degrees of freedom are the headset's globally-referenced pitch and yaw angles, each accurate to a few degrees. Covariance matrices (not shown) for the estimated antenna vectors are also produced at each measurement epoch.

Pose Estimation

Pose and twist estimation is performed by a UKF. The UKF takes vectors $\hat{\boldsymbol{r}}_{\mathrm{RP,G}}$ and $\hat{\boldsymbol{r}}_{\mathrm{PS,G}}$, together with their covariance matrices, from the CDGNSS engine and fuses these with IMU measurements to estimate headset position, velocity, and orientation. It also estimates biases for the IMU accelerometer and gyroscope, for a total of 15 scalar states. A 24-state variant of the UKF is run offline to calibrate extrinsic (fixed attitude with respect to headset) and intrinsic (scale factors and static biases) parameters of the IMU [17].

Virtual Reality

Headset pose estimates are fed to the VR block as $\hat{\boldsymbol{r}}_{\mathrm{W}}$, the IMU position in the local ("World") frame, and $\hat{\boldsymbol{e}}_{\mathrm{BW}}$, the Euler angle errors relative to the reference World-to-Body attitude matrix $\bar{R}_{\mathrm{BW}}$. Inside the VR block, a plugin performs the necessary coordinate transformations so that the VR engine's rendered scene corresponds to the input pose in an absolute sense. The VR engine pushes images to the headset's stereo display at a rate between 60 and 90 frames per second. The system's end-to-end latency, from head movement to its effect on the displayed pixels, is under 30 ms.

Loosely- and Tightly-Coupled Architectures

The dashed line on the lower right of Fig. 3 denotes optional data flow from the pose estimator UKF to the CDGNSS engine. Absent this flow, IMU and GNSS data fusion is performed in a standard *loosely-coupled* manner in which the CDGNSS engine operates without the benefit of inertial aiding.

When the indicated data flow is enabled, then at each measurement update the UKF's *a priori* state estimate $\bar{\boldsymbol{x}}$ and its covariance $\bar{P}_{xx}$ are passed to the CDGNSS engine. Transformed versions of $\bar{\boldsymbol{x}}$ and $\bar{P}_{xx}$ replace the internal state and covariance of the CDGNSS engine's *a priori* estimates. In this way, the pose estimator UKF provides a prior constraint on the CDGNSS solutions, increasing the success probability of the integer search process. This propagation and state replacement strategy is similar to the "position seeding" described in [38].

Because the IMU measurements aid the CDGNSS engine's integer resolution, the system can be said to operate in a *tightly-coupled* mode when the optional feedback flow is enabled. But this is not a traditional tightly-coupled CDGNSS-inertial system like the one presented in [9], where real-valued states, IMU biases, and integer states are jointly optimized. Instead, it is an example of what might be called federated tight coupling. This federated architecture, in which discrete estimators pass data back and forth, is suboptimal in terms of estimation performance and consistency compared to traditional tight coupling. Seeding the CDGNSS engine with the UKF's *a priori* state estimate at each measurement update risks introducing unmodeled correlations between the UKF's state errors and errors in the measurements it receives from the CDGNSS engine. However, this "loopy inference" is mitigated by the fact that double-difference carrier phase measurements are millimeter-precise compared to the decimeter-level precision of the UKF-provided prior position. Such precision asymmetry implies that, when conditioned on correct integer ambiguities, the CDGNSS engine's solution errors are dominated by carrier phase multipath, which is uncorrelated with the UKF's state errors. In practice, the precision asymmetry is magnified by inflating the *a priori* covariance $\bar{P}_{xx}$ sent from the UKF to the CDGNSS engine by a scalar factor.

As noted in [38], the integer-conditioning-induced decorrelation breaks down as the number of double-difference carrier phase measurements drops below 3, at which point the 3-dimensional position state becomes truly unobservable via GNSS measurements alone. When this occurs, the returned measurements contain information solely from the prior along any unobservable directions, leading to estimator inconsistency. To avoid this situation, the UKF does not ingest CDGNSS engine outputs made with fewer than 3 double-difference measurements. This restriction has little effect in practice, as situations with only one or two double-difference measurements are typically transient in nature, and short gaps can usually be bridged by inertial positioning.

Of course, if ambiguities are incorrectly resolved, substantial correlation can build up between the CDGNSS engine errors and the UKF state errors. Hence, extracting good performance from this paper's federated tight coupling scheme demands careful aperture testing to validate candidate CDGNSS integer fixes [39]. The integer fixing logic employed in this work, which is described in [37], adopts a layered approach of signal selection, pseudorange-based innovations testing, and a controlled-failed-fixing-rate integer acceptance test.

The federated estimation architecture, although suboptimal, has the virtue of being simple to implement and diagnose and can draw on an existing well-tested CDGNSS engine [38]. Moreover, it enables switching back-and-forth between loose and tight coupling, which allows for convenient examination of the benefits of tight coupling. Finally, it manages to deliver impressive results compared to loose coupling, as will be shown, and so serves as a valuable stepping stone toward future work with traditional tight coupling.

PROTOTYPE SYSTEM

Figs. 1 and 4 show the prototype OWVR system. The base VR headset is an OSVR HDK2. A lightweight carbon fiber antenna plate is cantilevered to the VR headset. Rigidly mounted to the antenna plate are two L1 GNSS dual-feed

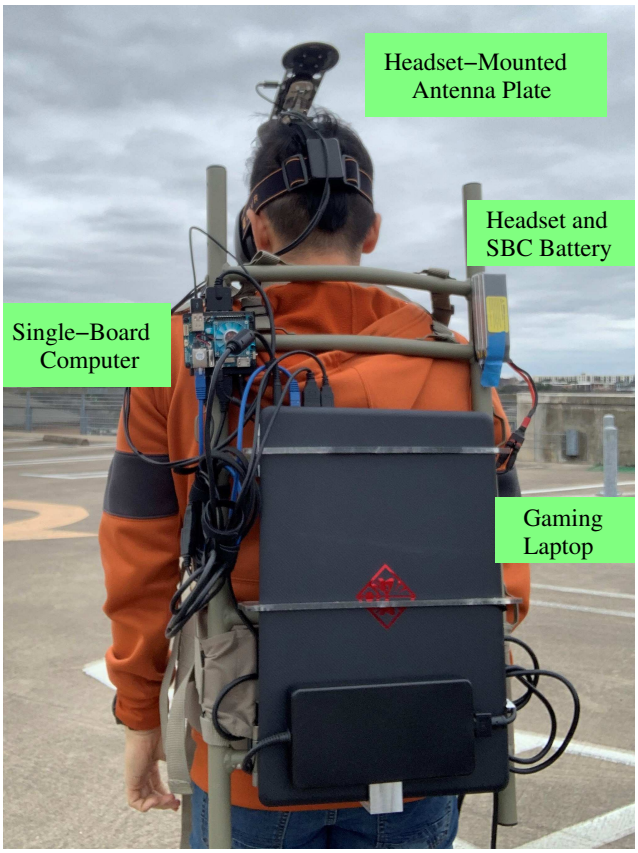

Fig. 4. The prototype OWVR system.

Tallysman TW1421 antennas separated by one GNSS L1 wavelength (19 cm), and a custom-designed GNSS front-end board that incorporates a Bosch BMX055 IMU [40].

GNSS IF samples at 9.6 MHz from each antenna, and IMU specific force and angular rate samples at between 70 and 100 Hz, are synchronously sampled to better than 1 ms, packed together, and sent via a USB interface to a single-board computer (SBC), an Odroid XU4, shown in Fig. 4. The GRID software-defined GNSS receiver, the CDGNSS engine, and the pose estimation UKF have all been optimized to run simultaneously in real time on the SBC's modest applications processor, a Samsung Exynos5422 operating at 2 GHz.

The virtual reality functions shown in Fig. 3 are performed within the gaming laptop, whose powerful graphical processing unit enables display rendering at between 60 and 90 frames per second. A custom OSVR plugin performs the necessary coordinate transformations so that the VR engine's rendered scene corresponds to the input pose in an absolute sense. The VR engine, an instance of Unreal Engine, then renders to the stereo headset display the globally-referenced 3D map of the surrounding area in accordance with the headset pose.

PERFORMANCE CHARACTERIZATION

Pose Estimator Model Validation

Data collected during a VR experience under good GNSS visibility were used for an initial test of the system. In this case, a confident ground truth could be obtained against which experimental offline runs with various artificial impairments could be compared.

A first check is to examine the size of the accelerometer and gyroscope biases estimated by the pose estimator UKF when the system is operating in its loosely-coupled mode (Fig. 5). That these remain small and consistent with the estimator-expected uncertainty over the nearly 5-minute walk with natural VR head movement indicates that the IMU, kinematics, and measurement models assumed within the pose estimator are accurate.

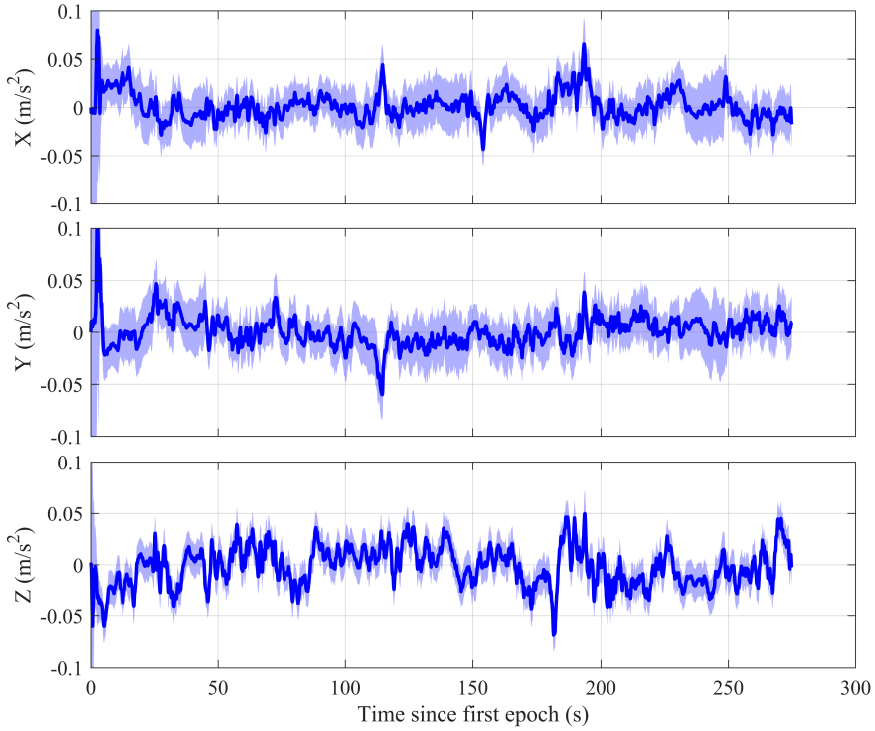

Fig. 5. Estimates of the X, Y, and Z accelerometer biases over a ∼5-minute walk with natural VR head movement under good GNSS visibility. The blue shaded regions indicate the formal (estimator-produced) 1-σ uncertainty envelope of each trace.

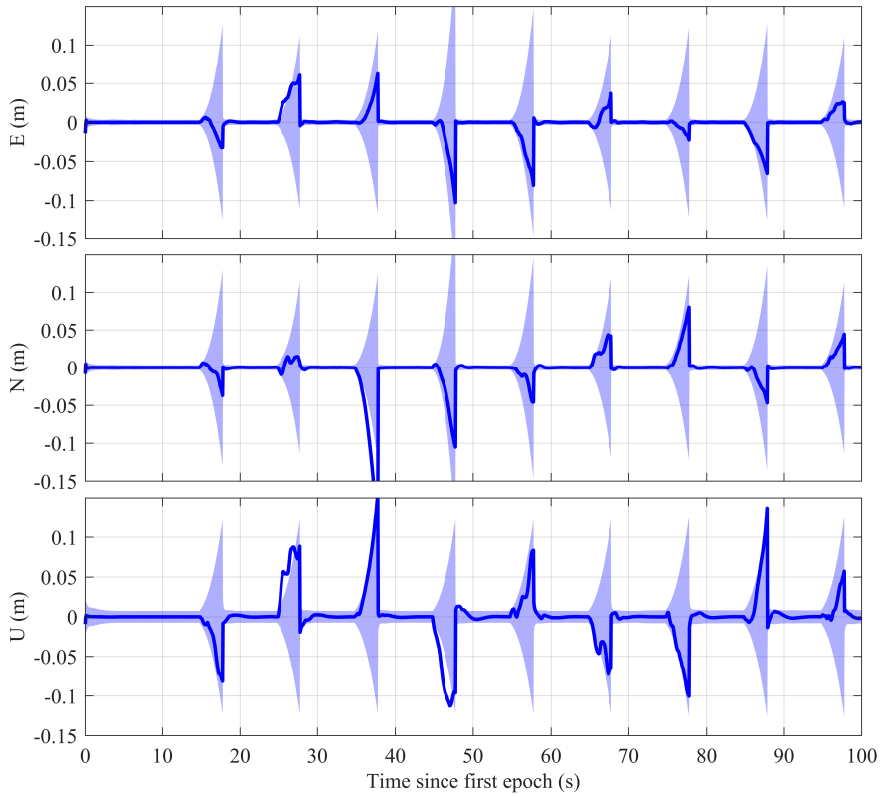

Fig. 6. Pose estimator position errors during a sequence of artificial 3-second full GNSS signal outages under loose coupling. The blue shaded regions indicate the formal (estimator-produced) 1-σ uncertainty envelope of each trace.

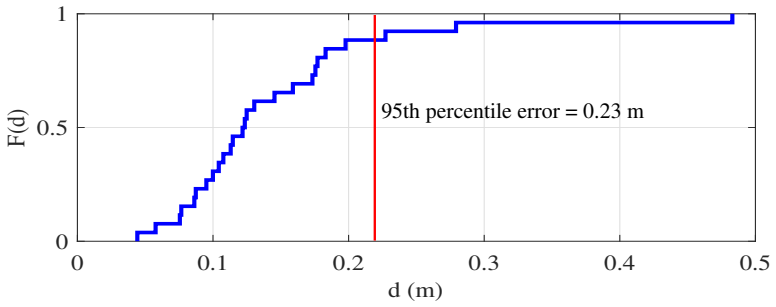

Fig. 7. Cumulative distribution function of the maximum headset position error during the 26 3-second artificial GNSS outages under loose coupling.

Pose Estimator under Intermittent Artificial Full GNSS Denial

In a more challenging test, GNSS measurements entering the pose estimator were artificially suppressed to evaluate the estimator's ability to flywheel through complete GNSS outages in its loosely-coupled mode. In this case, a 3-second outage was simulated every 10 seconds over the ∼5-minute walk, for a total of 26 outage events. Fig. 6, which shows the first 100 seconds of artificial outages, indicates that the east, north, and up headset position errors, shown in dark blue, are consistent with the pose estimator's expected uncertainty, shown in light blue. The cumulative distribution function of the maximum error during all 26 3-second GNSS outages (Fig. 7) reveals that, when properly calibrated, the smartphone-grade IMU supports GNSS-denied positioning over 3-second intervals to better than 23 cm. This remarkable stability owes much to offline calibration of the IMU's extrinsic and intrinsic parameters, and to the CDGNSS measurements' precision, which strengthens estimation of the instantaneous accelerometer and gyroscope biases during the between-gap intervals of GNSS availability.

CDGNSS Positioning under Intermittent Artificial Severe Degradation

GNSS signals are rarely completely denied in the outdoor environment. More often one experiences blockage of a subset of signals. To simulate such conditions while retaining a confident ground truth, the CDGNSS engine was altered to artificially suppress a user-configurable number of incoming double-difference pseudorange and carrier phase measurements. Consider a case in which all but the 3 strongest and highest elevation double-difference signals were eliminated in 10-second intervals with a period of 20 seconds. The following subsections examine this case under loose and tight coupling.

Loose Coupling: Fig. 8 shows position error results under loose coupling with the artificial GNSS degradation described above. The pose estimator experiences positioning errors greater than 2.5 meters during intervals when the CDGNSS engine, operating on only 3 pairs of double-difference observables, loses its integer-constrained fix and falls back to a float solution.

Tight Coupling: Fig. 9 shows position error results for the same data set with 10-second artificial GNSS impairment events on which the loosely-coupled system failed to maintain integer fixing, but now with tight coupling. One can see that vertical errors grow slightly beyond 20 cm in this case when only 3 double-difference measurements are available, but horizontal errors remain below 5 cm and the system never loses its integer fix, clearly demonstrating the benefit of tight coupling.

A Closer Look: It is instructive to closely examine a case in which the loosely-coupled system loses its integer fix to appreciate how the tightly-coupled architecture achieves its advantages.

For this, a brief discussion of the CDGNSS engine's integer fixing algorithm will be necessary. Fig. 10, taken from [37], offers a flow diagram for the algorithm. Double-difference carrier phase and pseudorange observables are initially screened by a combination of quality indicators to exclude signals likely to be outliers due to multipath. A float solution followed by innovations testing is performed on the remaining signals, after which an integer least squares algorithm searches for the integer vector that produces the lowest cost given the prior state estimate and the current epoch's observables.

The candidate integer ambiguities are only accepted if an integer aperture test indicates that they are highly likely to be correct. In particular, the OWVR system's CDGNSS engine applies a difference test with a controlled-failed-fixing

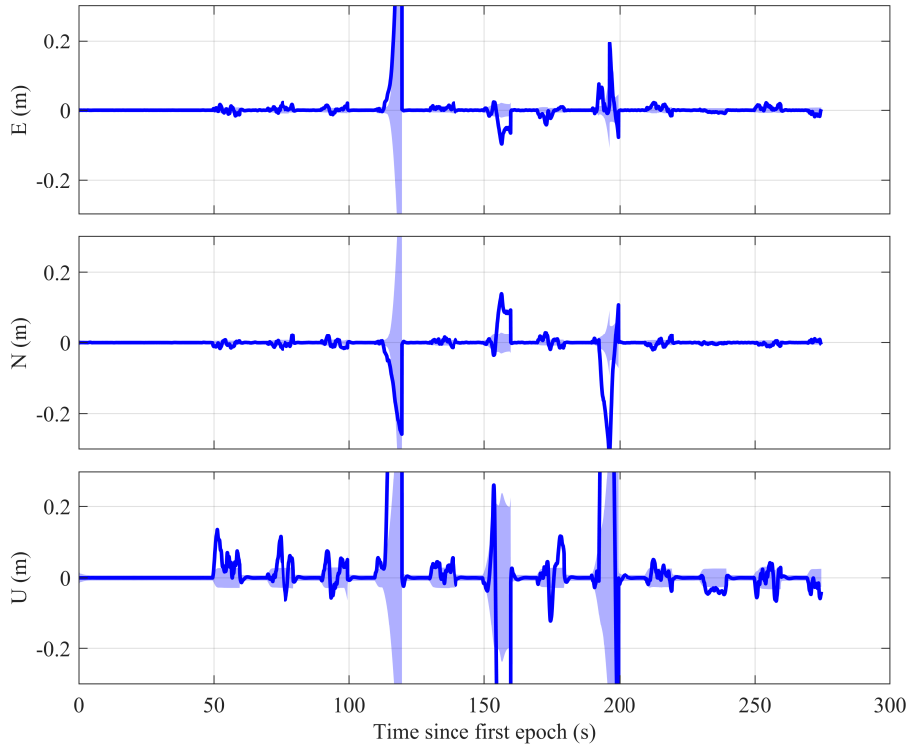

Fig. 8. Pose estimator positioning errors under loose coupling over a ∼5-minute test during which all but 3 double-difference observable pairs are eliminated in 10-second intervals with a period of 20 seconds. The blue shaded regions indicate the formal (estimator-produced) 1-σ uncertainty envelope of each trace.

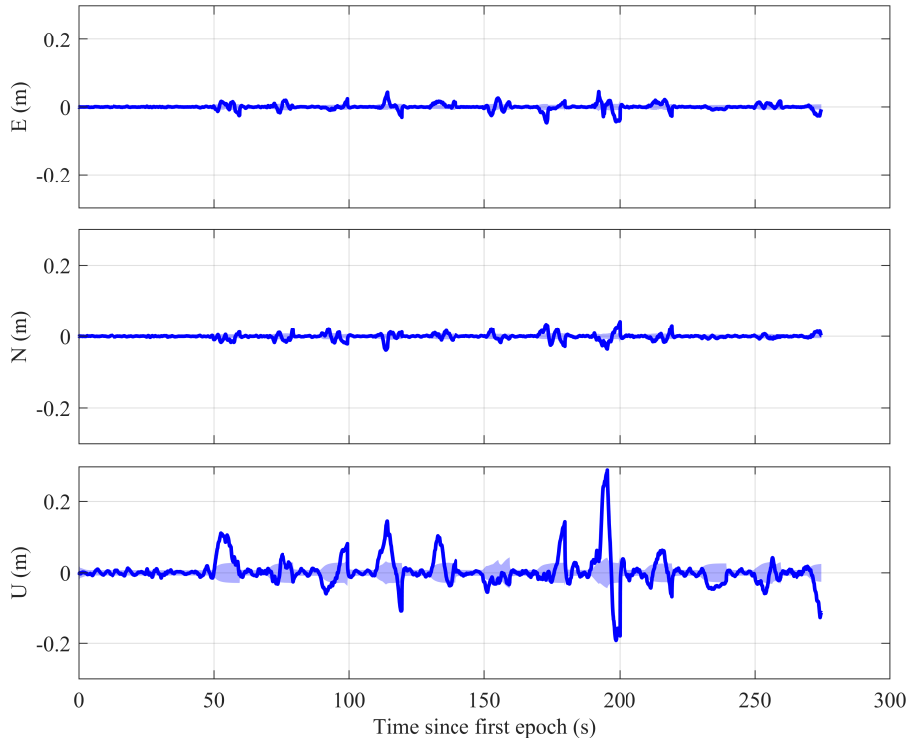

Fig. 9. As Fig. 9 but under tight coupling.

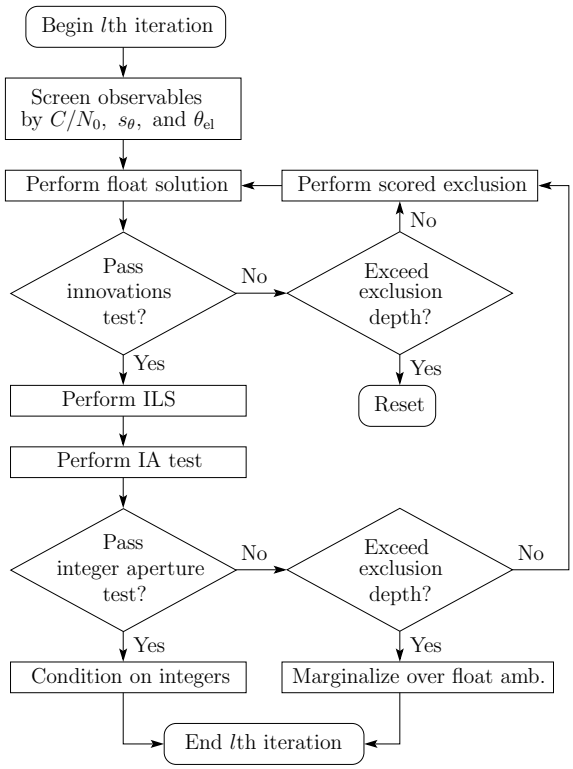

Fig. 10. Flow diagram for the CDGNSS engine's signal exclusion and integer fixing logic. From [37].

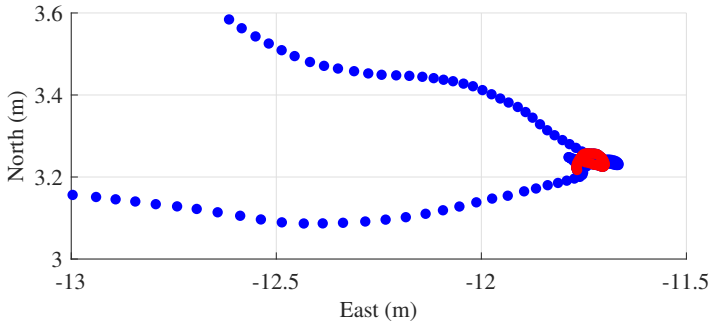

Fig. 11. Horizontal displacement for the "hairpin turn" case study. Blue dots show 20-Hz true headset locations. The red trace covers a headset maneuver that caused a failure of loosely-coupled integer fixing. The 2-second interval of the red trace is analyzed in Fig. 12.

rate [39]. If the test is successful, then the engine conditions its real-valued state elements on the assumed-correct integers and outputs the precise solution. Importantly, integer states are completely discarded from one epoch to the next, whether after conditioning or marginalization. This avoids the need to detect and correct cycle slipping, which occurs often in the urban conditions under which OWVR will be expected to operate. Further details on the CDGNSS integer fixing logic are found in [37].

What exactly causes the aperture test to fail for the loosely-coupled system under artificial impairment? Consider a case study of a hairpin turn. The true headset trajectory is shown by the blue dots in Fig. 11. Severe artificial signal blockage is induced during the full trajectory, leaving only 3 double-difference measurements out of 12 originally available. The loosely coupled system maintained a correct integer fix despite the partial outage over the first few seconds. But during a headset maneuver that took place over the interval indicated by the red trace in Fig. 11, which corresponds to the time interval of Fig. 12, the loosely-coupled system's aperture test statistic collapsed, a fixed solution was abandoned, and the pose estimator's position errors grew to several meters as the CDGNSS engine struggled to provide even a float solution. Meanwhile, the tightly-coupled system suffered only a slight loss of accuracy over the same interval, and had no failed integer fixes.

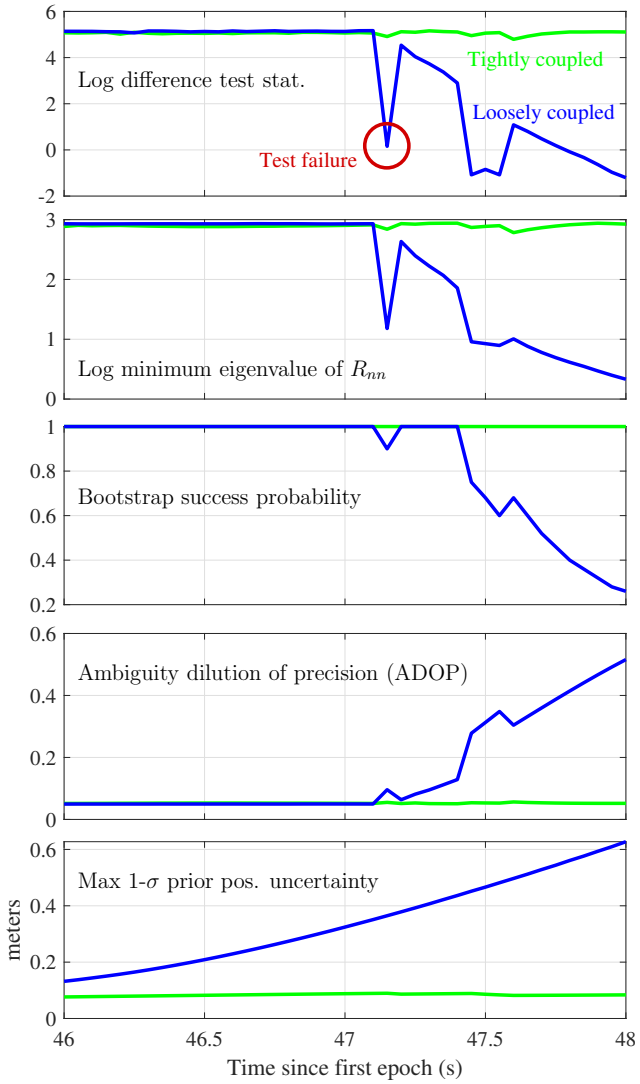

Fig. 12. Various data-based and model-based indicators over the interval of the red trace in Fig. 11. Green traces: tightly coupled; blue traces: loosely coupled. Top panel: The natural logarithm of the CDGNSS engine's position-solution difference test statistic (higher is better). Second panel: The natural logarithm of the minimum eigenvalue of the square-root information matrix R_{nn} that gets passed to the integer least squares solver (higher indicates greater minimum information) [41]. Third panel: The bootstrap integer fixing technique success probability. Fourth panel: The ambiguity dilution of precision (lower ADOP indicates a stronger solution). Bottom panel: The maximum 1-σ prior position uncertainty along any direction before the CDGNSS measurement update.

What precipitated the loosely-coupled system's sudden collapse, and why did the tightly-coupled system avoid such a collapse even when processing identical data?

The collapse was triggered when one of the three double-difference signal pairs was replaced by another with unfavorable geometry, which suddenly weakened the model. One can appreciate this by examining the various indicators shown in Fig. 12. The top panel shows the natural logarithm of the CDGNSS engine's difference test statistic for the tightly-coupled (green) and loosely-coupled (blue) cases. The difference test failed in the loosely-coupled mode at the point indicated, after which the test statistic briefly rebounded. But because the CDGNSS engine refused to condition its real-valued states on the resolved integers for the measurement epoch when the test failed, it arrived at the next epoch with a weaker prior, which, under the extreme artificial GNSS degradation, precipitated further degradation and eventual collapse of the test statistic.

Why was the tightly-coupled solution spared a similar fate? First consider the second, third, and fourth panels from the top in Fig. 12. These are so-called model-based indicators, which characterize the strength of a CDGNSS solution from the point of view of the prior and the measurement geometry, but not the actual measurement values. The three

shown here all indicate a degradation for the loosely-coupled system at the same time when its test statistic collapsed. Thus, irrespective of any multipath errors or other noise in the double-difference measurements, the model on which the CDGNSS engine was operating was suddenly weaker at the point when the test failed.

The tightly-coupled system suffered from the same weakening of signal geometry, but experienced a negligible change in these indicators at the point of the loosely-coupled system's test failure. This is because the tightly-coupled system's prior position uncertainty was much tighter than the loosely-coupled system's, as can be seen in the bottom panel of Fig. 12. Thus, the difference in outcomes can be explained succinctly as follows: The tightly-coupled system's tighter position prior provided it a bulwark against a collapse of measurement constraints.

Real-World Degraded GNSS Scenario

To further demonstrate the tightly-coupled system's robustness, it was subjected to real-world rather than artificial signal impairments. An open-world VR user walked along the east wall of the UT Aerospace Engineering building, under the building's overhanging eaves, as shown in Fig. 13. The user spent a total of about one minute subject to highly obstructed reception of most GNSS signals and strong multipath from the building wall. During the worst moments, only three (briefly two) double-difference signals remained available.

A ground truth trajectory is not available for this scenario, but one can nonetheless evaluate system performance on the metrics shown in Figs. 14 and 15. As can be appreciated in Fig. 14, the loosely-coupled architecture failed to maintain an integer fix in the degraded conditions along the wall. A gap in the data indicates a loss of fix, and the large discontinuity in the top panel's vertical position trace indicates that when an integer fix was regained, it was incorrect.

By contrast, the tightly-coupled architecture maintained solution availability throughout the test (Fig. 15). The integrity check failure indicator—a combination of the IA test, an innovations test, and other sanity checks—never asserted. Moreover, the continuity of the headset's vertical position suggests that a correct integer fix was maintained throughout the test despite the long interval with few signals. This remarkable result is consistent with the results of the foregoing artificial degradation tests. It highlights the robustness of the tightly-coupled architecture and its clear superiority over the loosely-coupled variant.

Note that a similar test under the eaves of the UT Aerospace Engineering building was conducted in this paper's companion paper [17]. In that test the tightly-coupled CDGNSS-inertial (sans vision) system did experience a loss of integer fix. The conditions in that test were more challenging than in the test reported here, for two reasons. First, for the system reported in [17] the IMU is not rigidly connected to the antenna plate. Instead, an assembly composed of an SBC and the GNSS front end is suspended by rubber dampers from the vehicle's chassis. This mechanical isolation is necessary to attenuate motor vibration sensed by the IMU, but it has the ill effect of slightly shifting the IMU-to-Body attitude as the vehicle maneuvers, preventing IMU bias estimation as accurate as for the OWVR system, whose IMU is rigidly mounted to the antenna plate.

Second, the GNSS conditions for the final test in [17] were less favorable than for the test examined here: there were fewer satellites available under open-sky conditions, leaving fewer than 3 double-difference signals available during intervals along the building's east wall under its eaves. Fewer than three GNSS double-difference signals are insufficient to constrain position. As a result, the pose estimator's position solution diverged as IMU drift accumulated along the unobservable direction. As noted in [17], vision-based sensing offers essential additional robustness under such nearly-complete GNSS-denied conditions.

CONCLUSIONS

This paper made three contributions. First, it introduced the concept of open-world virtual reality (OWVR). Second, it presented a working prototype OWVR system. Third, it offered the first demonstration of which the authors are aware of tight inertial-CDGNSS coupling using a smartphone-grade IMU in which the IMU was shown to aid CDGNSS integer fixing. In tests with severe artificial and real-world GNSS degradation, the tightly-coupled OWVR headset tracking system maintained a continuous CDGNSS integer fix even when the number of participating double-difference signals fell to 3. Such robust headset tracking is essential for reliable open-world virtual reality operation under outdoor and urban conditions.

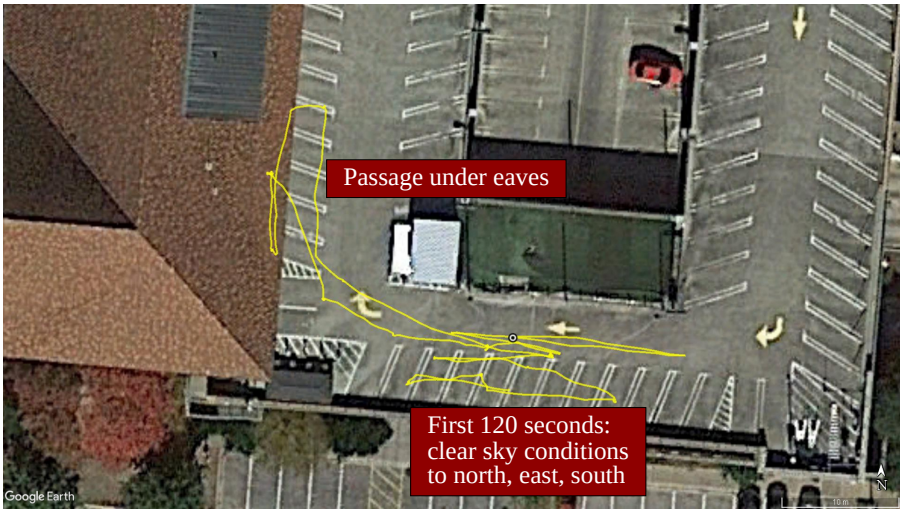

Fig. 13. Trajectory of the VR headset over a ∼360-second walkabout that included both conditions of good GNSS visibility and of severe signal blockage while the user walked along the outside wall of a building under overhanging eaves.

ACKNOWLEDGMENTS

This work has been supported by the National Science Foundation under Grant No. 1454474 (CAREER) and by the UT Wireless Networking and Communications Group (WNCG). The authors wish to thank Jonathan Baurer for developing the custom Unreal Engine plugin and Saura Elghonimi for mechanical design of the OWVR headset and backpack.

REFERENCES

[1] Hettinger, L. J. and Riccio, G. E., "Visually Induced Motion Sickness in Virtual Environments," *Presence: Teleoper. Virtual Environ.*, Vol. 1, No. 3, Jan. 1992, pp. 306–310.

[2] DeAngelis, G. C. and Angelaki, D. E., "Visual–vestibular integration for self-motion perception," *The neural bases of multisensory processes*, CRC Press/Taylor & Francis, 2012.

[3] Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A., "A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment," *Personal Technologies*, Vol. 1, No. 4, Dec 1997, pp. 208–217.

[4] Azuma, R., Hoff, B., Neely, H., and Sarfaty, R., "A motion-stabilized outdoor augmented reality system," *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, IEEE, 1999, pp. 252–259.

[5] Karlekar, J., Zhou, S. Z., Nakayama, Y., Lu, W., Loh, Z. C., and Hii, D., "Model-based localization and drift-free user tracking for outdoor augmented reality," *2010 IEEE International Conference on Multimedia and Expo*, July 2010, pp. 1178–1183.

[6] Shepard, D., *Fusion of Carrier-Phase Differential GPS, Bundle-Adjustment-Based Visual SLAM, and Inertial Navigation for Precisely and Globally-Registered Augmented Reality*, Master's thesis, The University of Texas at Austin, May 2013.

[7] Yang, J. J., Holz, C., Ofek, E., and Wilson, A., "DreamWalker: Substituting Real-World Walking Experiences with a Virtual Reality," *User Interface Software and Technology (UIST) 2019*, ACM, October 2019.

[8] Petovello, M., Cannon, M., and Lachapelle, G., "Benefits of Using a Tactical-Grade IMU for High-Accuracy Positioning," *Navigation, Journal of the Institute of Navigation*, Vol. 51, No. 1, 2004, pp. 1–12.

[9] Scherzinger, B. M., "Precise robust positioning with inertially aided RTK," *Navigation*, Vol. 53, No. 2, 2006, pp. 73–83.

[10] Zhang, H. T., *Performance comparison on kinematic GPS integrated with different tactical-grade IMUs*, Master's thesis, The University of Calgary, Jan. 2006.

[11] Kennedy, S., Hamilton, J., and Martell, H., "Architecture and system performance of SPAN—NovAtel's GPS/INS solution," *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, IEEE, 2006, p. 266.

[12] Godha, S., *Performance Evaluation of Low Cost MEMS-Based IMU Integrated With GPS for Land Vehicle Navigation Application*, Master's thesis, The University of Calgary, 2006.

[13] Li, T., Zhang, H., Gao, Z., Chen, Q., and Niu, X., "High-Accuracy Positioning in Urban Environments Using Single-Frequency Multi-GNSS RTK/MEMS-IMU Integration," *Remote Sensing*, Vol. 10, No. 2, 2018, pp. 205.

[14] Bochkati, M., Sharma, H., Lichtenberger, C. A., and Pany, T., "Demonstration of Fused RTK (Fixed)+ Inertial Positioning Using Android Smartphone Sensors Only," *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, IEEE, 2020, pp. 1140–1154.

[15] Henkel, P. and Hentati, H., "Reliable RTK positioning with tight coupling of 6 low-cost sensors," *Proceedings of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2016)*, 2016, pp. 733–741.

[16] Henkel, P., Sperl, A., Mittmann, U., Bensch, R., Färber, P., and Günther, C., "Precise Positioning of Robots with Fusion of GNSS, INS, Odometry, Barometer, Local Positioning System and Visual Localization," *Proc. of the 31st Intern. Technical Meeting of The Satellite Division of the Institute of Navigation*, 2018, pp. 3078–3087.

[17] Yoder, J. E., Iannucci, P. A., Narula, L., and Humphreys, T. E., "Multi-Antenna Vision-and-Inertial-Aided CDGNSS for Micro Aerial Vehicle Pose Estimation," *Proceedings of the ION GNSS+ Meeting*, Online, 2020.

[18] Everaerts, J. et al., "The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 37, 2008, pp. 1187–1192.

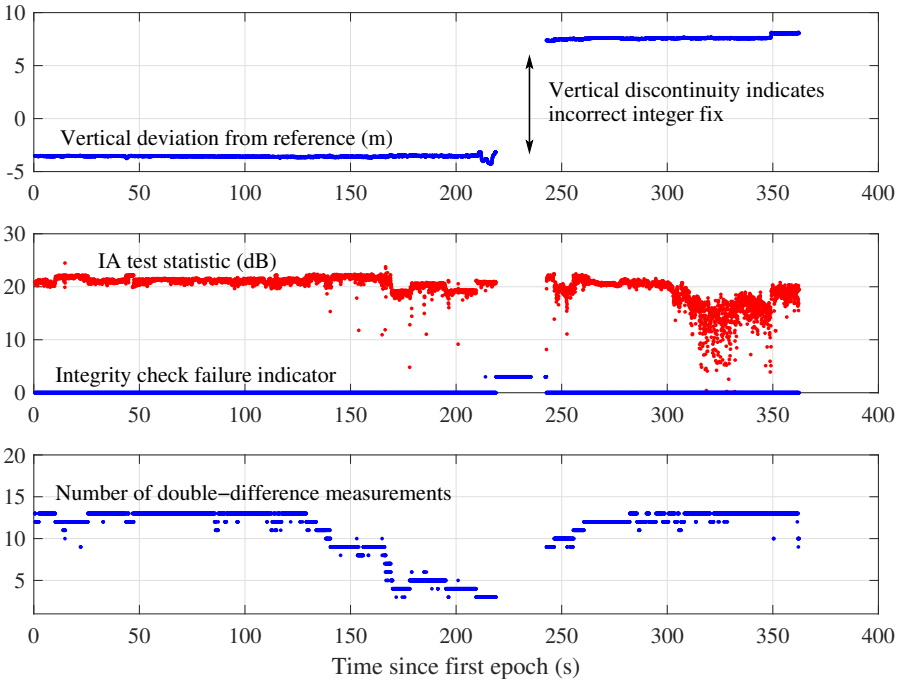

Fig. 14. Real-world performance of the loosely-coupled architecture over the trajectory shown in Fig. 13. Top: Vertical height of the headset's primary GNSS antenna above a reference height as calculated by the CDGNSS engine. Center: Integer aperture (difference) test statistic (red, in dB) and the integrity check failure indicator (blue). Bottom: The number of double-difference signals participating in the CDGNSS position solution.

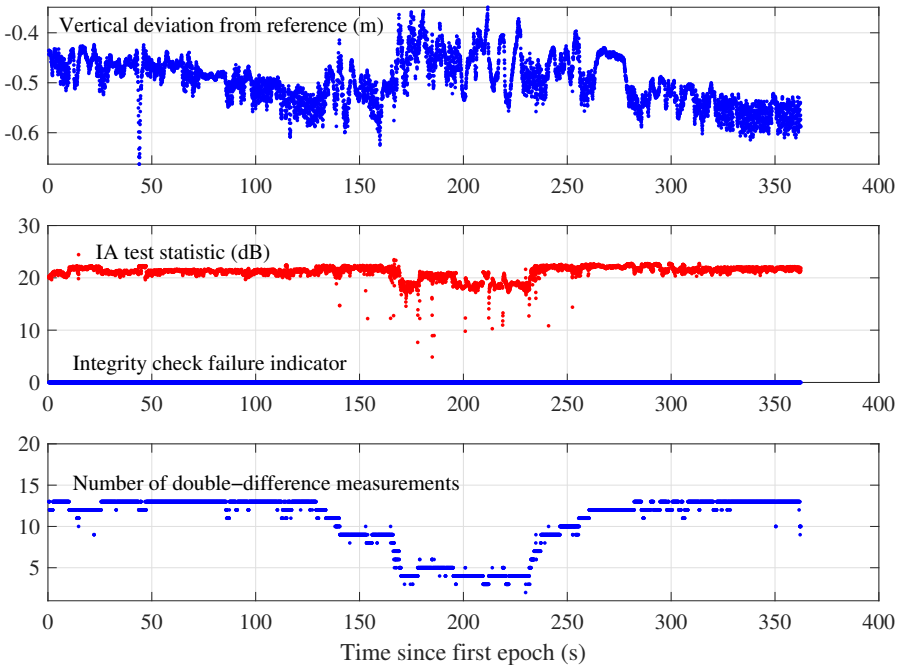

Fig. 15. As Fig. 14 but for the tightly-coupled architecture.

[19] Strecha, C., Pylvänäinen, T., and Fua, P., "Dynamic and scalable large scale image reconstruction," *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, IEEE, 2010, pp. 406–413.
[20] Neitzel, F. and Klonowski, J., "Mobile 3D mapping with a low-cost UAV system," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, Vol. 38, 2011, pp. 1–6.
[21] Crandall, D. J., Owens, A., Snavely, N., and Huttenlocher, D. P., "SfM with MRFs: Discrete-continuous optimization for large-scale structure from motion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 35, No. 12, 2013, pp. 2841–2853.
[22] Irschara, A., Hoppe, C., Bischof, H., and Kluckner, S., "Efficient structure from motion with weak position and orientation priors," *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*, IEEE, 2011, pp. 21–28.
[23] Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S. M., and Szeliski, R., "Building rome in a day," *Communications of the ACM*, Vol. 54, No. 10, 2011, pp. 105–112.
[24] Musialski, P., Wonka, P., Aliaga, D. G., Wimmer, M., van Gool, L., and Purgathofer, W., "A Survey of Urban Reconstruction," *Computer Graphics Forum*, Vol. 32, No. 6, 2013, pp. 146–177.
[25] Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., and Sarazzi, D., "UAV photogrammetry for mapping and 3d modeling–current status and future perspectives," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 38, No. 1, 2011, pp. C22.
[26] Nex, F. and Remondino, F., "UAV for 3D mapping applications: a review," *Applied Geomatics*, Vol. 6, No. 1, 2014, pp. 1–15.
[27] Nolan, M., Larsen, C., and Sturm, M., "Mapping snow-depth from manned-aircraft on landscape scales at centimeter resolution using Structure-from-Motion photogrammetry," *The Cryosphere Discussions*, Vol. 9, 2015, pp. 333–381.
[28] Borges, J. L., "Museum, on Exactitude in Science," *Collected fictions*, 1998, pp. 325–327.
[29] Wirth, J., Bonugli, E., and Freund, M., "Assessment of the Accuracy of Google Earth Imagery for use as a Tool in Accident Reconstruction," Tech. rep., SAE Technical Paper, 2015.
[30] "ITRF2014," http://itrf.ign.fr/ITRF_solutions/2014/.
[31] Optech, "Lynx SG1 Mobile Mapper," 2015, http://www.optech.com/wp-content/uploads/specification_lynx-sg1.pdf.
[32] Haala, N., Peter, M., Kremer, J., and Hunter, G., "Mobile LiDAR mapping for 3D point cloud collection in urban areas—A performance test," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, Vol. 37, 2008, pp. 1119–1127.
[33] Niehorster, D. C., Li, L., and Lappe, M., "The accuracy and precision of position and orientation tracking in the HTC Vive virtual reality system for scientific research," *i-Perception*, Vol. 8, No. 3, 2017.
[34] Klein, G. and Murray, D., "Parallel Tracking and Mapping for Small AR Workspaces," *6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE, 2007, pp. 225–234.
[35] Shepard, D. P. and Humphreys, T. E., "High-Precision Globally-Referenced Position and Attitude via a Fusion of Visual SLAM, Carrier-Phase-Based GPS, and Inertial Measurements," *Proceedings of the IEEE/ION PLANS Meeting*, May 2014.
[36] Lightsey, E. G., Humphreys, T. E., Bhatti, J. A., Joplin, A. J., O'Hanlon, B. W., and Powell, S. P., "Demonstration of a Space Capable Miniature Dual Frequency GNSS Receiver," *Navigation, Journal of the Institute of Navigation*, Vol. 61, No. 1, 2014, pp. 53–64.
[37] Humphreys, T. E., Narula, L., and Murrian, M. J., "Deep urban unaided precise Global Navigation Satellite System vehicle positioning," *IEEE Intelligent Transportation Systems Magazine*, 2020.
[38] Scherzinger, B., "Quasi-tightly Coupled GNSS-INS Integration," *Navigation: Journal of The Institute of Navigation*, Vol. 62, No. 4, 2015, pp. 253–264.
[39] Wang, L. and Verhagen, S., "A new ambiguity acceptance test threshold determination method with controllable failure rate," *Journal of Geodesy*, Vol. 89, No. 4, 2015, pp. 361–375.
[40] Bosch, *Small, versatile 9-axis sensor module, P/N: BMX055*, 2014.
[41] Psiaki, M. and Mohiuddin, S., "Global Positioning System Integer Ambiguity Resolution Using Factorized Least-Squares Techniques," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 2, March-April 2007, pp. 346–356.