A Smart Helmet Framework Based on Visual-Inertial SLAM and Multi-Sensor Fusion to Improve Situational Awareness and Reduce Hazards in Mountaineering

Charles Shi Tan, Concord College, UK*

 https://orcid.org/0009-0005-5901-8613

ABSTRACT

Sensitivity to surrounding circumstances is essential for the safety of mountain scrambling. In this paper, the authors present a smart helmet prototype equipped with visual SLAM (simultaneous localization and mapping) and barometer multi-sensor fusion (MSF), IMU (inertial measurement unit), omnidirectional camera, and global navigation satellite system (GNSS). They equipped the helmet framework with SLAM to produce 3D semi-dense pointcloud environment maps, which are then discretized into grids. Then, the novel danger metrics they proposed were calculated for each grid based on surface normal analysis. The A* algorithm was applied to generate safe and reliable paths based on minimizing the danger score. This proposed helmet system demonstrated robust performance in mapping mountain environments and planning safe, efficient traversal paths for climbers navigating treacherous mountain landscapes.

KEYWORDS

Head, Mountaineering, Multi-Sensor Fusion (MSF), Pointclouds, Route Planning, Simultaneous Localization and Mapping (SLAM)

INTRODUCTION

Mountain scrambling, often called alpine scrambling, is a recreational sport that entails ascending mountain peaks and ridges while sometimes employing one's hands to scale rock walls and navigate challenging terrain (Whymper, 1871). While less technical than rock climbing, scrambling can still be quite dangerous, given the exposure to heights and objective hazards. Some of the potential dangers of mountain scrambling include falling from narrow ledges and cliffs; rockfall when climbing steep rock faces, with loose rocks potentially dislodged onto climbers below; rapid weather changes like rain, wind, and lightning (which can lead to slippery conditions, loss of visibility and hypothermia); getting lost due to lack of distinctive landmarks in mountainous terrain; and exposure to altitude sickness, sunburn, dehydration, and cold weather injuries due to the high altitude environment.

*Corresponding Author

The underlying mountaineering severity can lead to tragic consequences. Analysis of The International Alpine Trauma Registry (IATR) data reveals that out of 306 recorded mountain accidents resulting in multisystem trauma, the majority were due to falls onto solid ground (51.4%), followed by falls in snowfields (10.8%), falls into crevasses (8.1%), and being struck by stones (5.4%). In 24.3% of cases, the cause was unknown (Rauch et al., 2019). These examples remind us of the dangers climbers who venture onto precipitous terrain face. The field would benefit from continued efforts to improve safety protocols and leverage technological innovations to mitigate the risks inherent to these physically and mentally demanding sports.

In this regard, a novel helmet prototype has been invented that presents a viable solution by alerting individuals through a loudspeaker about potential hazards and suggesting the safest route while minimizing travel time. The equipped SLAM technology offers a revolutionary approach to managing the objective risks inherent in mountaineering through the collaborative integration of multiple sensors. At the same time, the Multi-Sensor Fusion (MSF) has great promise to address the limitations of SLAM systems alone for enhanced precision and reliability. This helmet showcases the ability to use real-time data from the surrounding environment to show the best route to the destination.

RELATED WORK

Throughout human history, helmets have played a crucial role in safeguarding the lives of individuals facing dangerous environments and challenges. From ancient civilizations to the present day, the evolution of helmet technology has been driven by a relentless pursuit of safety and protection. In the early days, helmets were forged from basic materials such as leather and bronze. Warriors and adventurers wore them to guard against head injuries in combat and dangerous terrain. As the centuries passed, helmet design and materials advanced, incorporating iron, steel, and carbon fiber. These improvements significantly enhanced protective capabilities, but the helmet is still primarily focused on protecting against physical impacts.

The rise in amateur mountaineers has increased the risk of fatal accidents, necessitating advanced safety equipment. In the 21st century, mountaineers face numerous environmental and navigational hazards, requiring protection beyond physical impacts. Helmets equipped with computer vision and guidance capabilities are essential, aiding in decision-making under extreme conditions and ensuring safer, less strenuous routes, especially in altitude sickness and hypothermia cases.

A principal method used in our helmet framework is Simultaneous Localization and Mapping (Durrant-Whyte & Bailey, 2006; Engel et al., 2018), which is widely used to enable a robot or autonomous vehicle to construct a map of an unfamiliar environment while simultaneously recognizing its position in that environment. The recent advancement in 3D map reconstruction (Grisetti et al., 2010) and SLAM (Ebadi et al., 2022b) not enable robots to precisely positioning and make autonomous decisions in a scalable approach (Kohlbrecher et al., 2011) that can be applied to extreme environments.

This technique facilitates the helmet system in navigating through intricate and unexplored areas by continuously updating its spatial awareness. By utilizing Simultaneous Localization and Mapping for generating detailed 3D terrain mapping and recognizing cliffs, edges, and overhangs, scramblers could more effectively evaluate potential risks of falling or rockfall and choose paths that are less exposed.

Moreover, another primary function of SLAM systems is to map surroundings while locating oneself, which makes this technology highly suitable for mitigating disorientation. Since the lack of noticeable reference points within rugged mountainous settings can frequently result in hikers becoming disoriented, the utilization of precise, up-to-date positioning technology within a constantly evolving diagram of the mountains could significantly enhance understanding of the current situation and lower the incidence of unintentional deviations from the intended path. A few Simultaneous Localization and Mapping (SLAM) algorithms have integrated supplementary

Table 1. Comparative analysis of smart helmets

Causes of Injuries	Strengths	Problems
Skully AR-1	- Augmented reality HUD - Rearview camera - GPS navigation	-Reliability issues - Software bugs - High cost
DAQRI Smart Helmet	- Integrated visor display - Sensors for AR applications - Safety features	- High price - Limited battery life - Customized software development required
CrossHelmet X1	- Rearview camera - Voice control - Companion app support	- High price - Comfort and fit concerns - Helmet weight
Forcite Alpine Smart Helmet	- Built-in cameras - Communication capabilities	- Limited battery life in extreme cold conditions - Durability concerns
Livall BH60SE	- Integrated LED lights - Turn signal indicators - Bluetooth connectivity	- Limited device compatibility - Potential challenges with firmware updates

Note: The smart helmets listed in this table were all launched from 2016 to 2019. The features and limitations of various smart helmet models are outlined, highlighting the common issues of high cost and lack of intelligent decision-making. Our helmet framework is equipped with visual SLAM technology, ensuring a robust and economical solution scalable for mass production.

data like weather patterns in their cartography protocols (Hong et al., 2020). Combining meteorological knowledge with topographical maps presents an opportunity to identify locations of shelter or alternate routes with less susceptibility to unfavorable weather patterns. Such data would significantly mitigate risks associated with precipitation, wind, and lightning strikes while undertaking demanding expeditions.

Though SLAM's capabilities show immense promise in enhancing mountain scrambling safety, relying solely on SLAM for localization and mapping during mountain scrambling has limitations, as the algorithm can accumulate errors over time, leading to imprecision and posing uncertainty. Therefore, in this case, we have incorporated sensors such as GNSS, IMU, and barometer to perform the MSF algorithm to ultimately obtain a more accurate, reliable, and stable position and orientation estimation. This is where MSF (Luo et al., 2002) can complement SLAM to create more robust and resilient systems for practical use in unpredictable real-world alpine environments.
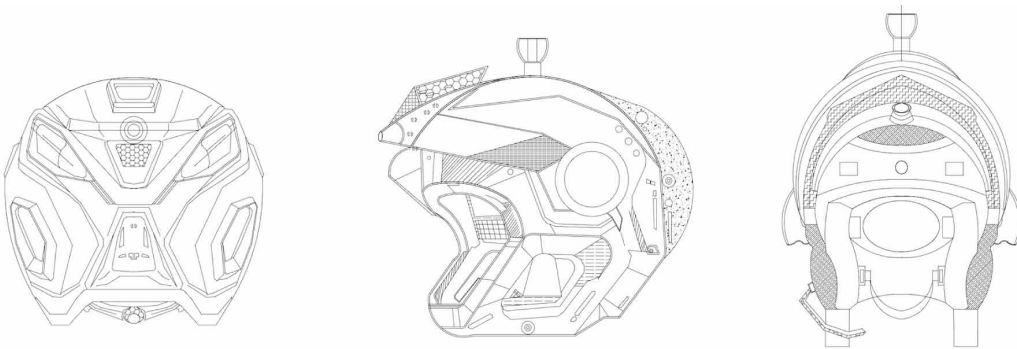
In MSF, data from multiple heterogeneous sensors are combined using intelligent algorithms to leverage the strengths and mitigate the weaknesses of each sensory modality. For instance, visual SLAM may be fused with inertial measurement units (IMUs) containing accelerometers and gyroscopes. While visual SLAM is prone to drift, IMU inputs can provide short-term motion constraints to correct and bind these errors. However, IMUs alone are inaccurate over longer durations due to integration drift. Fusing the two modalities creates a system more reliable than either sensor in isolation. Additionally, other complementary sensors can be integrated, such as LIDAR for long-range mapping, GNSS (Global Navigation Satellite System) (Kaplan & Hegarty, 2017; Zaliva & Franchetti, 2014) coupling SLAM with this array of proprioceptive and exteroceptive sensors through MSF provides multi-faceted observations of the environment to overcome the weaknesses of any individual sensor channel. This leads to greater accuracy and resilience for localization and mapping in challenging mountainous terrain.

Furthermore, MSF also enables redundancy and cross-validation between sensor streams. When one modality is compromised, such as loss of visual tracking, others may be able to provide continuity and observe environmental constraints until the degraded signal recovers. This fail-soft characteristic is essential for robustness in safety-critical mountain scrambling.
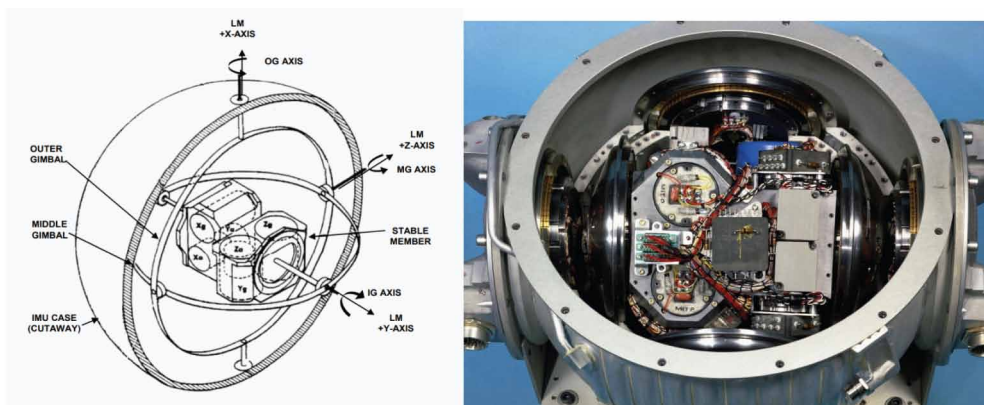
METHODOLOGY

The core functionality of the helmet (Figure 1) is provided by 360-degree cameras (Scaramuzza, 2014) around the rim to provide a detailed 360-degree visual understanding of the surroundings, eliminating blindspots. The video feeds are analyzed using algorithms such as image segmentation (Haralick & Shapiro, 1985) to classify terrain, identify hazards like cliffs and overhangs. The vision system builds a geospatially-registered point cloud and mesh model of the environment updated continuously with wearer movement. This allows assessing slope, proximity to drop-offs, safe paths, and other navigational insights.

Figure 1. The prototype design of the helmet

Note. Figure 1 presents a novel helmet prototype equipped with a multi-faceted sensor suite and intelligent software.

Figure 2. Apollo IMU, where inertial reference integrating gyros (IRIGs, Xg, Yg, Zg) sense attitude changes, and pulse integrating pendulous accelerometers (PIPAs, Xa, Ya, Za) sense velocity changes

Note. Evaluation ratings were made during the 2014 fall academic term. Apollo IMU, where Inertial Reference Integrating Gyros (Xg, Yg, Zg) sense attitude changes, and Pulse Integrating Pendulous Accelerometers (Xa, Ya, Za) sense velocity changes. From National Aeronautics and Space Administration, n.d. (https://wehackthemoon.com/tech/inertial-measurement-unit-mechanical-engineering-wizardry)

- **GNSS:** GPS and other GNSS provide an absolute geographical position fix, which helps build an accurate map of the surroundings and locate the climber precisely on it. This enables planning optimal traversal routes.
- **IMU (Ahmad et al., 2013):** The IMU uses accelerometers and gyroscopes to track the climber's movement and orientation. This aids localization and mapping by complementing the cameras with information on movement and perspective. It can also detect falls by recognizing sudden accelerations.
- **Barometer:** The barometer measures atmospheric pressure changes to estimate altitude, complementing the visual mapping. This improves vertical positioning when visual references are inadequate.

The 360-degree vision achieves comprehensive hazard detection, the IMU tracks climber motion for localization, the GNSS provides precise global positioning to situate the climber on the map, and the barometer aids with altitude estimates when visual cues are limited. The combination enables robust situational awareness in challenging mountain environments.

System Integration

The diverse sensor suite enables the building of a detailed 360-degree model of the environment. The vision algorithms integrate footage from omnidirectional cameras queue depth and inertial data to the appropriate scene region. Then, the hazardous zones, navigation markers, and other real-time overlays are rendered on a compact heads-up display inside the helmet. The algorithm continuously analyzes the 360-degree camera feeds and is designed to be low latency to ensure the augmented environment view responds immediately to user movement and terrain changes.

When the helmet perceives an area of instability like loose rock, thinning ice, or a steep drop-off, it marks the danger zone in the heads-up display. As the climber moves, these warnings remain geospatially fixed, allowing them to be visually tracked. The system can also provide audible alerts when in proximity to hazards. By constantly monitoring the scene, the helmet acts as a vigilant lookout for potentially deadly situations the climber could miss in the high cognitive load of navigation.

The combination of a 3D semi-dense map, motion tracking, and real-time hazard overlay allows navigating mountain faces while continuously assessing upcoming terrain and proximity to cliffs or crevasses.

Rugged Design

The helmet employs a hardened design to handle the harsh conditions of mountaineering. The housing encasing the cameras, sensors, and other fragile electronics is impact-resistant and reinforced to withstand falls, stones, and compressive forces. The optics employ scratch-resistant glass lenses and hydrophobic coatings to maintain visibility in rain and snow. Internal components feature moisture protection, heat sinking, and shock absorption. The helmet maintains reliable operation across a broad thermal range, functioning equally well on frozen summits or under blazing sun. Automated self-tests validate all sensors and indicators to catch failures (Henry & Clarke, 1993). The helmet continues providing vital assistance even after harsh impacts or prolonged exposure by employing resilience engineering.

User Data Collection and Privacy Considerations

The development of our smart helmet is grounded in a conscientious approach to user data collection, ensuring rigorous compliance with privacy standards and explicit user consent (Mertens, 2018). We employ robust anonymization techniques to safeguard climbers' personal information, upholding their privacy rights. Simultaneously, the helmet's real-time hazard analysis algorithms are designed to maximize user safety, meticulously balancing the dual imperatives of risk mitigation and privacy

preservation. Additionally, the Gaussian distribution-based machine learning scheme could be applied to anomaly detection and ensure the security of our data analysis processes (Dwivedi et al., 2021).

We must also address the critical aspect of data collection and sensor reliability. The integration of IoT-based sensors for big data collection, as discussed by Plageras et al. (2018), highlights the necessity of efficient data processing and analysis to enhance the functionality and reliability of the smart helmet. Similarly, the work of Chawra and Gupta (2022) on optimizing wake-up scheduling in 3D-wireless sensor networks, and the efficient data collection scheme for underwater linear sensor networks proposed by Ahmed et al. (2022) underscore the significance of ensuring sensor robustness and data accuracy.
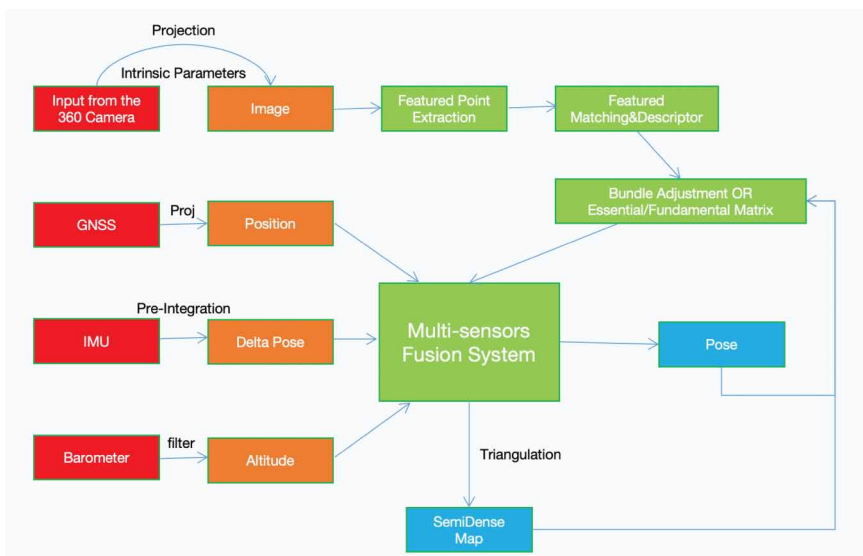
ALGORITHM

To carry out localization, our approach initially involves implementing SLAM to construct a semi-dense map. Subsequently, the map is divided into grids, and we compute a dangerous score for each grid. The subsequent path planning is founded on the computed grid graph utilizing the A* algorithm.

Figure 3 displays the whole SLAM and MSF process. The input to the algorithm is a stream of images from the 360 camera as it moves through the environment. The first step is to perform image projection to undistort those images according to camera intrinsic parameters, and feature extraction and feature matching are applied to get the poses between two keyframes via geometry-based function essential matrix and fundamental matrix (Luong & Faugeras, 1996), followed by the triangulation to get the 3D map points. Furthermore, those expanded 3D map points and the estimated camera pose are fed back into bundle adjustment for further optimization (Triggs et al., 2000). This incremental process repeats as new keyframes are acquired to build a consistent map while estimating the camera trajectory.

The optimized 3D map points and camera poses are combined with additional sensor data from other sensors using MSF techniques. This sensor fusion also allows for estimating the initial camera pose for the next keyframe, which is then used to triangulate and expand the 3D map (SemiDense Map) via features matched to previous keyframes through image projection.

Figure 3. The pipeline of the entire SLAM and MSF process

In summary, image projection, bundle adjustment optimization, multi-sensor fusion, and triangulation are combined in a loop to generate an optimized 3D map along with estimated camera poses at each keyframe.

Image Projection

Map points can be projected into the 360 camera's equirectangular image plane based on the currently estimated camera pose to provide a detailed 360-degree visual understanding of the surroundings. The projected pixel coordinates can then be matched to feature points detected in the equirectangular image.

To establish 2D-3D correspondences between features in the catadioptric omnidirectional camera system, Geyer and Daniilidis (2000) first proposed a four-step projection process to model the geometric relationship between a 3D scene point and its corresponding 2D pixel location in the camera image.

Consider a scene point $P = (x, y, z)$ in the mirror's reference frame. For convenience, we assume the axis of symmetry of the mirror aligns with the camera's optical axis, and the x and y axes of the camera and mirror are also aligned. Hence, the only difference between the camera and mirror reference frames is the translation along z.
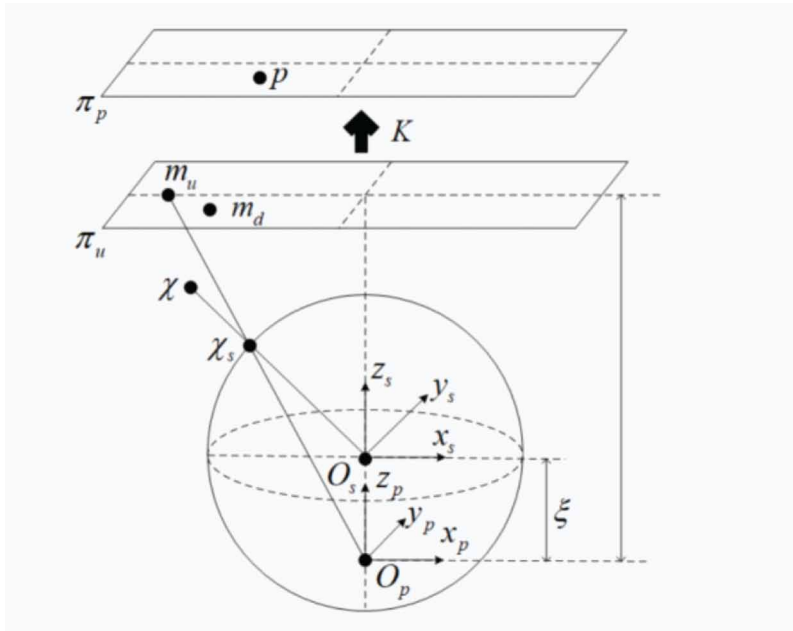
Projecting the scene point onto the unit sphere:

$$P_b = \frac{P}{\|P\|} = \left(x_s, y_s, z_s\right)$$

The point coordinates are then changed to a new reference frame centered in $C = \left(0, 0, -\varepsilon\right)$:

$$P_s = \left(x_s, y_s, z_s + \varepsilon\right)$$

Figure 4. The image projection for the omnidirectional camera

The P_s is then projected onto the normalized image plane in C_s:

$$\bar{m} = \left(x_m, y_m, 1\right) = \left(\frac{x_s}{z_s + \varepsilon}, \frac{y_s}{z_s + \varepsilon}, 1\right) = g^{-1}\left(P_s\right)$$

The point m is then mapped to the camera image point $\bar{p} = \left(u, v, 1\right)$ through the intrinsic-parameter matrix K.

Therefore:

$$\bar{p} = K\bar{m}$$

where K is:

$$K = \begin{bmatrix} \alpha_u & \alpha_u \cot\left(\theta\right) & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

However, Geyer and Daniilidis' fisheye projection model varies from camera to camera and depends on the field of view of the lens, and approximation of fisheye lenses by catadioptric cameras has limited accuracy.

To tackle that problem, Scaramuzza (2014) proposed a unified model that uses the Taylor polynomial to overcome the lack of knowledge of a parametric model for fisheye cameras, whose coefficients and degrees are found through the calibration process. Accordingly, the relation between the normalized image point $\bar{m} = \left(x_m, y_m, 1\right)$ and the unit vector P_s in the fisheye (mirror) reference frame can be written as:

$$P_s = g(m)\infty \begin{bmatrix} x_m \\ y_m \\ \alpha_0 + \alpha_2\rho^2 + \ldots + \alpha_N\rho^N \end{bmatrix}$$

where $\rho = \sqrt{x_m^2 + y_m^2}$, Scaramuzza (2014) has also emphasized that polynomials of the third or fourth order can accurately represent all catadioptric and several types of fisheye cameras currently available in the market. The ability of this model to be used effectively with a diverse range of commercial cameras is the fundamental reason for its great future potential.

Feature Point Detection (FAST)

Feature point detection extracts distinctive keypoints for tracking across frames. A commonly used detector is Features from Accelerated Segment Test (FAST). FAST (Mair et al., 2010) examines local pixel neighborhoods to rapidly identify high-contrast corners. This efficiency suits real-time SLAM, especially in complex mountain environments. Real-time performance is critical but challenging due to factors, such as variable lighting and terrain. FAST's speed enables fast keypoint extraction despite these difficulties. By only considering local pixel circles rather than expensive smoothing.

ORB Descriptor

While FAST efficiently detects keypoint locations, we also need descriptors to characterize the local visual appearance at each keypoint. Some Robust descriptors can be used to facilitate matching keypoints between frames by representing the distinct image patch around each point.

A highly effective descriptor that complements FAST detection is ORB (Oriented FAST and Rotated BRIEF), which builds upon FAST to extract oriented FAST corners, followed by describing these keypoints using a compact binary string via the BRIEF (Binary Robust Independent Elementary Features) method (Calonder et al., 2010).

BRIEF compares pixel intensities in a smoothed image patch at predefined locations relative to the keypoint. The comparisons produce a bit string that summarizes the salient gradients. ORB modifies BRIEF to make it rotation-invariant for matching under viewpoint changes. Together, the oriented FAST detection and rotated BRIEF description in ORB provide a feature that is fast to compute, invariant to rotation, and robustly matched across frames. Moreover, those characteristics make ORB ideal for real-time SLAM in dynamic environments with continuously changing camera orientation.

Feature Matching

Descriptor matching between keyframes is essential for tracking visual features across different viewpoints in mountain visual SLAM. As the camera explores mountainous terrain, the perspective changes dramatically, causing the appearance of features to vary significantly across frames. The ORB descriptor vectors are compared using Hamming distance (Norouzi et al., 2012) to identify matches between keyframes.

We can establish a keypoint match between frames if the descriptors have a small Hamming distance, indicating they likely correspond to the same physical 3D point viewed from different mountain terrain perspectives. We can track the same features across multiple keyframes taken at distinct camera poses and lighting conditions by finding descriptor matches. This would allow incrementally estimating the camera motion and constructing a map of the mountain environment by tracking matches over sequential views.

Initially, we have two keyframes I_1 and I_2 that depict overlapping views of a scene. We can then establish 2D-to-2D correspondences between points in the two frames by leveraging epipolar geometry constraints to estimate the fundamental matrix F and initial relative pose (Zhang et al., 1995). For example, pixel P_1 in I_1 matches to P_2 in I_2.

Using these feature matching, we can compute the fundamental matrix F that encodes the epipolar constraint:

$$P_2^T F P_1 = 0$$

Decomposing F gives us the essential matrix E and an initial estimate of the relative camera motion between the two poses P_1 and P_2.

By Single Value Decomposition $E = T_x R$, where T is the translation and R is the rotation from frame I_1 to I_2. From this, we get our initial pair of camera matrices P_1 and P_2.

Next, we triangulate the feature matches to reconstruct 3D points X_j in the scene. For a matched pair of pixels P_1 and P_2:

$$p_1 = P_1 X_j$$
$$p_2 = P_2 X_j$$

By back-projecting rays from the feature bearings in both frames, their intersection gives the 3D coordinate X_j.

Repeating this process for all matches allows the construction of a map of 3D points corresponding to tracked visual features.

We now have a set of 3D-to-2D correspondences between reconstructed points X_j and their pixel coordinates in the keyframes.

As the camera explores new viewpoints I_n, we can utilize the Perspective-n-Point (Lu, 2018) algorithm to estimate an initial pose P_n using established 3D-to-2D matches with points visible in frame I_n.

Finally, bundle adjustment globally refines all camera poses $\{P_1, P_2, P_3, ..., P_n\}$ by minimizing the total reprojection error between the 3D points and corresponding observed pixel coordinates across all keyframes. This jointly optimizes the pose estimates to obtain a globally consistent reconstruction.

Back-End Fusion With GPS and IMU Pre-Integration

Every keyframe state at time t is defined as $x_t = \left[p_t, q_t, v_t, ba_t, bw_t\right]$ compromising position p, orientation q as a quaternion, velocity v, and IMU biases ba for the accelerometer and bw for the gyroscope.

IMU measurements between times *t* and *t + 1* are pre-integrated into a relative motion increment *z: t → t + 1* (Förster et al., 2017):

$$z : t \rightarrow t + 1 = \left[\Delta p_t, \Delta q_t, \Delta v_t\right]$$

where Δ denotes the change in each state variable between *t* and *t* + 1 predicted by integrating the IMU. This pre-integration accounts for the IMU biases by including them in the state as:

$$\Delta p_t = \Delta p_t \left(ba_t, bw_t\right)$$

Figure 5. Triangulate to obtain map point depth

When a GNSS measurement y_{t+1} arrives, it observes an absolute position that can constrain pose drift. Furthermore, the v_{t+1} is the pose stemmed from our visual SLAM bundle adjustment; we then incorporate this by minimizing the cost:

$$C\left(p_t, p_{t+1}\right) = \left\|p_{t+1} - v_{t+1}\right\|_{\sum x} + \left\|p_{t+1} - y_{t+1}\right\|_{\sum y} + \left\|p_{t+1} - p_t - \Delta P_t\left(ba_t, bw_t\right)\right\|_{\sum z}$$

Here $\left\|\cdot\right\|_{\sum}$ denotes the Mahalanobis distance (De Maesschalck et al., 2000) with covariance Σx for SLAM, Σy for GNSS (Global Navigation Satellite Systems) and Σz for the pre-integrated IMU increment. This optimizes over x_t and x_{t+1}, adjusting the state estimate to best satisfy the constraints from both GNSS and IMU pre-integration. Critically, the pre-integrated IMU measurements are not recomputed, as the optimization only touches the biases ba_t, bw_t, avoiding unnecessary recalculations.

This way, GNSS global position measurements are fused with locally estimated pose increments through an optimization leveraging pre-integration. This provides globally accurate and locally precise state estimates.

In circumstances where GNSS suffers from intermittent errors caused by temporary obstruction of the satellites by mountains, groves, etc. and fails to achieve precise positional data, there is a possibility of experiencing drift in the vertical position estimation. To tackle this issue, adding a barometer can be utilized to measure atmospheric pressure. By incorporating pressure measurements, we can obtain a constraint for the altitude h from the z direction, thus enhancing the accuracy of pose measurements.

The barometer measures atmospheric pressure, providing an independent altitude estimate $f(x)$ which convert the pressure signal to height value and may be varied according to a different model:

$$h_{t+1} = f\left(p_{t+1}\right)$$

We can get the pose state by:

$$x_{t+1} = \left[p_{t+1}, q_{t+1}, v_{t+1}, ba_{t+1}, bw_{t+1}\right]$$

We then fuse the barometer into the optimization, and minimize the cost:

$$C\left(p_t, p_{t+1}\right) = \left\|p_{t+1} - v_{t+1}\right\|_{\sum x} + \left\|p_{t+1} - y_{t+1}\right\|_{\sum y}$$
$$+ \left\|p_{t+1} - p_t - \Delta P_t\left(ba_t, bw_t\right)\right\|_{\sum z} + \left\|p_{t+1} - h_{t+1}\right\|_{\sum h}$$

In summary, fusing barometer altitude GNSS, especially for vertical positioning, when GNSS references are impaired due to lack of internet connectivity. The barometer helps maintain 3D accuracy when external data is limited.

Normal Vector Calculation From the Semi-Dense Map

To extract normal vectors from a semi-dense Point Cloud map, we first divide the 3D space into grids at a resolution that captures local surface elements. With the grid produced from the segmentation, the next step is analyzing each neighborhood's point distribution by the covariance matrix.

The core analysis (Sanchez et al., 2020) phase involves Principal Component Analysis (PCA) on each region. For each local point set $\left\{ P_1, P_2, P_3, \ldots, P_n \right\}$, the 3x3 covariance matrix C_v is computed between the x, y, z coordinates.

C_v encapsulates the correlation structure of the surface points. Eigen decomposition of C_v yields eigenvalues $\left\{ \lambda_1, \lambda_2, \lambda_3 \right\}$ and eigenvectors $\left\{ v_1, v_2, v_3 \right\}$. The eigenvector v_i paired with the minimum eigenvalue λ_i corresponds to the axis of least variance, as the normal vector estimate.

Calculation of Each Grids' Dangerous Score

After converting the mountainous terrain into grids and obtaining the normal vectors for each grid, the subsequent stage involves examining the data related to the normal vectors to evaluate the danger level. Various indicators of danger can be derived from the normal vectors, including multiple metrics:

- **Incline angle:** The angle between the normal vector and the global vertical axis. Steeper slopes generally denote a greater risk of falling or sliding. The dot product between the normal and vertical yields the cosine of the incline angle.
- **Roughness:** Measured by the variance of normal vectors within a grid. High variance implies an irregular, bumpy surface that can be treacherous.
- **Edge proximity:** Distances to nearby grids with substantially differing normals may mark a cliff or overhang. Smaller distances flag a potential edge hazard.
- **Isotropy (Kamash & Robson, 1978; Rivlin & Ericksen, 1997):** How close the normals are to an ideal plane normal. High isotropy indicates loose, unstable terrain like gravel or sand.
- **Obstacle density:** The number of distinct objects distinguished by separate normals. More obstacles mean more hazardous navigating and an increased chance of tripping.

To integrate these factors into an overall danger score D for each grid v, we propose a weighted function:

$$D_v = w_1 f\left(A_v \right) + w_2 g\left(R_v, R_{v-neighbors} \right) + w_3 h\left(E_v, C_v \right) + w_4 k\left(I_v, V_v \right) + w_5 l\left(O_v \right)$$

where:

$$f\left(A_v \right) = 1 - e^{-\beta A v}$$

Models with increasing risk with slope angle A_v exponentially, where β is a scaling factor:

$$g\left(R_v, R_{v-neighbors} \right) = R_v \frac{1 + R_{vn}}{N}$$

Incorporates the roughness of neighboring grids R_{vn}:

$$h\left(E_v, C_v \right) = \frac{1}{E_v} + \tanh\left(C_v \right)$$

Edge proximity E_v weighted by cliff score C_v.

The cliff score C_v is computed as follows.

For each grid v, cast rays outward in directions (such as front or sides), and for each ray, find intersecting grids w via raycasting.

Compute dot product of v's normal N_v and ray direction D.

If $N_v \cdot D < threshold\ \tau$

$$C_v = \sum \left[\alpha \Delta h \left(1 - N_v \cdot D \right) \right]$$

where Δh = height difference between v and w, α is a scaling factor:

$$k\left(I_v, V_v\right) = \frac{I_v}{V_v}$$

We use the ratio of isotropy to grid surface area as an indicator of loose terrain:

$$l\left(O_v\right) = \log\left(1 + O_v\right)$$

Logarithmic scale obstacle density:

A_v = Incline angle

R_v = Roughness

R_{vn} = Neighbor roughness

E_v = Edge proximity

C_v = Cliff score

I_v = Isotropy score

V_v = Grid surface area

O_v = Obstacle density

This danger metric integrates additional grid information like volume and cliff scores with non-linear transformations of the metrics to better model danger response and the exponential and logarithmic functions to help account for sensitivity and saturation effects.

The metrics can be normalized before combining based on expected value ranges. The weighting coefficients w_i are optimized per environment using techniques, such as grid search and cross-validation against hazard data (Yan et al., 2022).

Once calculated for all grids, the danger scores are classified into qualitative levels like low, medium, and high. The environment map can be visualized with risk-coded colors to support planning safe navigation routes and activities. As the normal vectors get updated dynamically from new

SLAM mapping data, the danger scores can be recalculated accordingly, providing an up-to-date representation of the terrain hazards.

The Path Planning

Once the SemiDense Cloud map has been generated and the danger scores of each voxel have been determined, the environment may be represented as a weighted graph $G(V, E, D)$ to identify the optimal path. The objective of this path-finding algorithm is not only to ensure the safety of mountaineers and minimize the distance traveled:

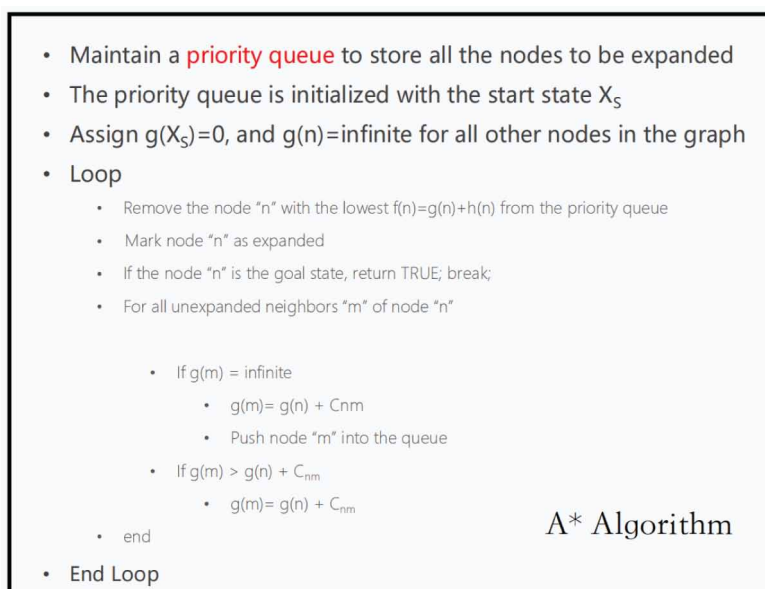$V = \{v_1, v_2, v_3, \ldots, v_n\}$ is the set of grids

$E = \{(v_i, v_j)\}$ is the set of edges between grid pairs

$D(e) = $ danger score of $e \in E$

Dijkstra's algorithm can then find the shortest paths on this graph. We adapt this to use A* search to focus on safety, which optimizes for the lowest total danger along the path (Hart et al., 1968). It uses a heuristic $h(v)$ that estimates the remaining distance to the goal based on Euclidean distance. This helps guide the search towards the goal efficiently. At each step, A* expands the grid v with the lowest $f(v) = g(v) + h(v)$, where $g(v)$ sums the danger scores along the path from the start. By minimizing $f(v)$, A* converges on the optimal path P*, minimizing total danger to reach the goal.

The initial path P* can be refined by smoothing waypoints and tuning the trade-off between danger and path length. As the environment changes, danger scores are updated dynamically, and P* is adjusted incrementally to remain optimal. Replanning only occurs when necessary to limit computations. During execution, the current danger is monitored along P*, and the path is aborted and replanned if thresholds are exceeded. This adaptation allows for responding to dynamic hazards.

Figure 6. The A* iteration

- Maintain a priority queue to store all the nodes to be expanded
- The priority queue is initialized with the start state X_S
- Assign $g(X_S)=0$, and $g(n)=$infinite for all other nodes in the graph
- Loop
    - Remove the node "n" with the lowest $f(n)=g(n)+h(n)$ from the priority queue
    - Mark node "n" as expanded
    - If the node "n" is the goal state, return TRUE; break;
    - For all unexpanded neighbors "m" of node "n"
        - If $g(m) = $ infinite
            - $g(m)= g(n) + C_{nm}$
            - Push node "m" into the queue
        - If $g(m) > g(n) + C_{nm}$
            - $g(m)= g(n) + C_{nm}$
    - end

    A* Algorithm
- End Loop

In summary, A* search optimized for minimum danger is used to plan optimal safe paths. The path is refined and adjusted incrementally as the environment and danger scores change. Danger is monitored during execution, and replanning occurs if thresholds are exceeded.

EXPERIMENT RESULT

The helmet system was tested in simulations and real-world environments, with experiments demonstrating robust capabilities in obstacle detection, 3D mapping, localization, and route planning for mountain navigation. Quantitative results showed improved localization accuracy by fusing SLAM, GNSS, and barometric data, while the system also consistently identified optimal safe traversal routes through analyzed terrain hazards.

We started the experiment in a real-world environment. Figure 7 is the moving object detection under this keyframe, and the semi-dense point cloud is then reconstructed from each keyframe, showcasing the helmet's ability to create a detailed 3D representation of the mountain terrain, providing critical situational awareness. The following experiment highlighted the improvements in localization accuracy when fusing SLAM, GNSS, and barometric data. These enhancements significantly reduce position and orientation errors, vital for precise navigation in mountainous terrain. The route planning experiment demonstrated the helmet's ability to calculate safe traversal paths by analyzing terrain hazards. It utilizes the generated grid graph to consistently find optimal minimum-risk routes, balancing safety and efficiency. Collectively, these experiments validate the smart helmet's capacity to enhance situational awareness through real-time sensor data fusion.

The following are details for each experiment respectively.

Moving Object Detection

An essential capability demonstrated was precise moving object detection and image projection to distortion input from the 360-degree camera input. As shown in Figure 7, the system identified and tracked multiple people during mountaineering. By projecting image features onto the frame using the unified projection model, robust detection was achieved.

Localization Accuracy

The position errors show the drift as a percentage of the total 100m traversal distance. With SLAM alone, the position drifts 1.2m over 100m of travel, giving a 1.2% error. Adding GNSS reduces this to 0.8% over the same distance. Integrating the barometer further improves accuracy to 0.4% drift.

The orientation errors are measured in absolute degrees of deviation from the ground truth. SLAM has 4.5° drift, improved to 2.1° with GNSS and 1.2° with barometer added, where the GNSS provides periodic position corrections and the barometer constraints drift along the vertical axis. This results in more precise and robust state estimation in the challenging mountain environment.

Figure 7. Moving object detection

Note. The experiment was conducted at Bodafon Mountain in Wales

Table 2. Localization accuracy with different sensor combinations

Condition	Position Error (m)		Orientation Error (deg)	
	Scene 1	Scene 2	Scene 1	Scene 2
Visual SLAM only	1.23	3.71	0.95	1.33
Visual SLAM + GNSS	0.87	2.69	0.74	0.81
Visual SLAM + GNSS + Barometer	0.43	2.10	0.59	0.63

Note: The values presented in this table represent the localization errors measured over a traversal distance of 100m. Scene 1 is located in Bodafon Mountain (Wales), providing a certain set of challenges and environmental conditions, while Scene 2 is situated in Garnedd Ugain Mountain (Wales), offering a different landscape and set of obstacles.

Semi-Dense Map Reconstruction

The reconstructed semi-dense point cloud provided a detailed representation of the mountain environment geometry. As depicted in Figure 8, the incremental bundle adjustment process leveraged matched features across keyframes to estimate camera poses and triangulate world points. This produced a geospatially registered 3D model reflecting the mountain structure.

Route Planning

The semi-dense 3D point cloud generated via SLAM was divided into grids, and we computed danger metrics, including slope, roughness, and proximity to cliffs. Figure 9 illustrates the resulting grid graph with danger color coding. The optimal minimum-risk route was identified using A* search optimized for safety. The system could consistently find safe trajectories, balancing hazard avoidance with traversal efficiency.

These results validate the helmet prototype's ability to enhance situational awareness through real-time sensory data fusion. Combining computer vision, SLAM mapping, and environmental analytics provides a comprehensive platform for informed navigation in hazardous mountain regions.

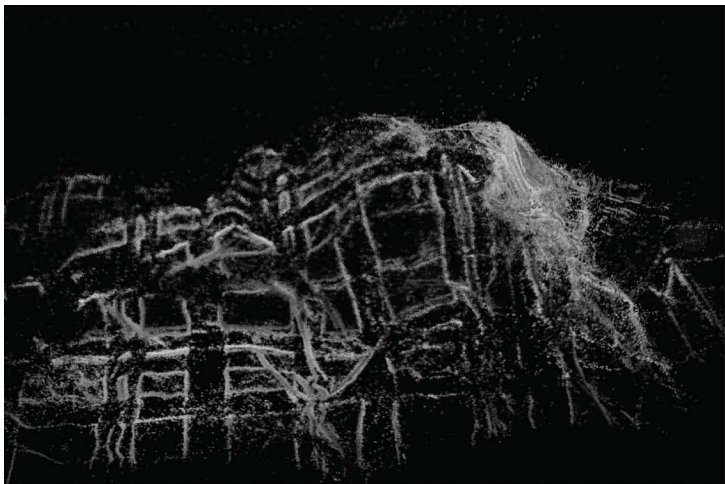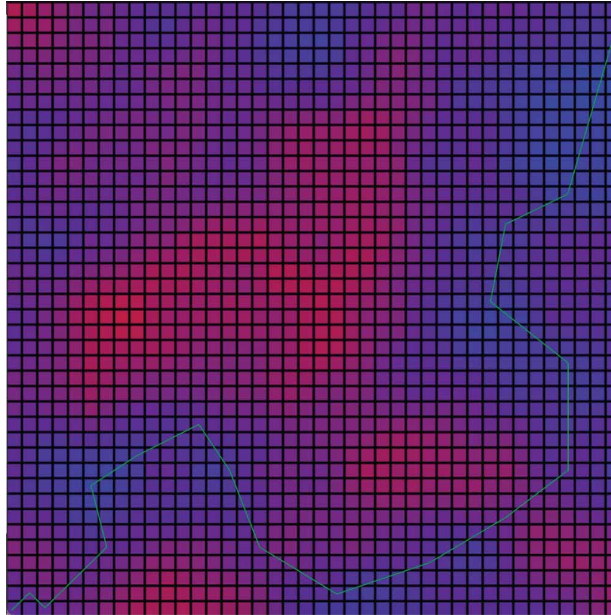Figure 8. Semi-dense grid map reconstruction via SLAM

Figure 9. Planned safe traversal route by A* search

CONCLUSION

This innovative smart helmet prototype demonstrates promising capabilities in enhancing mountaineer safety through visual-inertial SLAM and Multi-Sensor Fusion. The main contributions of this work lie in its comprehensive framework, which incorporates algorithms to achieve precise localization, robust mapping, and safe path-finding in extreme terrains, ensuring that mountaineers can navigate safely even in the most challenging conditions. While further refinements are needed, the prototype showcases an essential step toward wearable enhanced reality assistance in hazardous alpine settings. The low cost of visual SLAM technology, compared to its LiDAR counterparts, presents an economic advantage, making it a more feasible option for large-scale production. Visual SLAM components potentially range around 50-500 USD, compared to over 2000 USD for LiDAR-based systems. Enhanced machine learning techniques, such as reinforcement learning, could be vital in optimizing system behaviors for user safety and comfort. The system could adapt and improve its performance over time by continuously learning from user interactions and environmental conditions, leading to a more intuitive and user-friendly experience.

However, our algorithm is dependent on clear visual conditions, and its performance could be compromised in fog, heavy snow, or low-light scenarios. Integrating multiple sensors and advanced processing capabilities demands significant power, potentially resulting in limited battery life. The helmet's capability to collect and process vast amounts of data in real-time could lead to information overload, making it challenging for users to make quick and informed decisions.

The potential applications of this technology extend far beyond mountaineering, promising benefits across various high-risk activities and industries. In search and rescue operations, the helmet's precise localization and mapping capabilities could significantly enhance rescue teams' efficiency and safety when navigating challenging terrains. The construction and mining sectors could see improved worker safety with the helmet's real-time situational awareness and hazard detection. Outdoor adventure enthusiasts, including rock climbers and backcountry skiers, could leverage the helmet for safer navigation in unfamiliar terrains.

REFERENCES

Ahmad, N., Ghazilla, R. A. R., Khairi, N. M., & Kasi, V. (2013). Reviews on various Inertial Measurement Unit (IMU) sensor applications. *International Journal of Signal Processing Systems*, 256–262. 10.12720/ijsps.1.2.256-262

Ahmed, Z., Ayaz, M., Hijji, M., Abbas, M., & Rahim, A. (2022). AUV-based efficient data collection scheme for underwater linear sensor networks. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–19. doi:10.4018/IJSWIS.299858

Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010). BRIEF: Binary robust independent elementary features. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010* (pp. 778–792). Springer. doi:10.1007/978-3-642-15561-1_56

Chawra, V. K., & Gupta, G. P. (2022). Optimization of the Wake-Up scheduling using a hybrid of Memetic and Tabu search algorithms for 3D-Wireless sensor networks. *International Journal of Software Science and Computational Intelligence*, *14*(1), 1–18. doi:10.4018/IJSSCI.300359

De Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. (2000). The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, *50*(1), 1–18. doi:10.1016/S0169-7439(99)00047-7

Durrant-Whyte, H., & Bailey, T. (2006). Simultaneous localization and mapping: Part I. *IEEE Robotics & Automation Magazine*, *13*(2), 99–110. doi:10.1109/MRA.2006.1638022

Dwivedi, R. K., Kumar, R., & Buyya, R. (2021). Gaussian distribution-based machine learning scheme for anomaly detection in healthcare sensor cloud. *International Journal of Cloud Applications and Computing*, *11*(1), 52–72. doi:10.4018/IJCAC.2021010103

Engel, J., Koltun, V., & Cremers, D. (2018). Direct sparse odometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(3), 611–625. doi:10.1109/TPAMI.2017.2658577 PMID:28422651

Förster, C., Carlone, L., Dellaert, F., & Scaramuzza, D. (2017). On-Manifold preintegration for Real-Time Visual—Inertial odometry. *IEEE Transactions on Robotics*, *33*(1), 1–21. doi:10.1109/TRO.2016.2597321

Geyer, C., & Daniilidis, K. (2000). A unifying theory for central panoramic systems and practical implications. In D. Vernon (Ed.), *Computer Vision—ECCV 2000* (pp. 445–461). Springer. doi:10.1007/3-540-45053-X_29

Gong, Z., Liu, P., Wen, F., Ying, R., Ji, X., Miao, R., & Xue, W. (2021). Graph-based adaptive fusion of GNSS and VIO under intermittent GNSS-degraded environment. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–16. doi:10.1109/TIM.2020.3039640

Grisetti, G., Kümmerle, R., Stachniss, C., & Burgard, W. (2010). A tutorial on Graph-Based SLAM. *IEEE Intelligent Transportation Systems Magazine*, *2*(4), 31–43. doi:10.1109/MITS.2010.939925

Haralick, R. M., & Shapiro, L. G. (1985). Image segmentation techniques. *Computer Vision Graphics and Image Processing*, *29*(1), 100–132. doi:10.1016/S0734-189X(85)90153-7

Hart, P. D., Nilsson, N. J., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, *4*(2), 100–107. doi:10.1109/TSSC.1968.300136

Henry, M., & Clarke, D. (1993). The self-validating sensor: Rationale, definitions and examples. *Control Engineering Practice*, *1*(4), 585–610. doi:10.1016/0967-0661(93)91382-7

Kamash, K. M. A., & Robson, J. (1978). The application of isotropy in road surface modeling. *Journal of Sound and Vibration*, *57*(1), 89–100. doi:10.1016/0022-460X(78)90283-3

Kaplan, E. D., & Hegarty, C. (2017). *Understanding GPS/GNSS: Principles and Applications.* Artech House. https://trid.trb.org/view/1472240

Kohlbrecher, S., Von Stryk, O., Meyer, J., & Klingauf, U. (2011). A flexible and scalable SLAM system with full 3D motion estimation. *IEEE International Workshop on Safety, Security, and Rescue Robotics (SSRR)*. doi:10.1109/SSRR.2011.6106777

Lu, X. X. (2018). A review of solutions for Perspective-n-Point problem in camera pose estimation. *Journal of Physics: Conference Series*, *1087*, 052009. doi:10.1088/1742-6596/1087/5/052009

Luo, R. C., Yih, C., & Su, K. (2002). Multisensor fusion and integration: Approaches, applications, and future research directions. *IEEE Sensors Journal*, *2*(2), 107–119. doi:10.1109/JSEN.2002.1000251

Luong, Q., & Faugeras, O. (1996). The fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, *17*(1), 43–75. doi:10.1007/BF00127818

Mair, E., Hager, G. D., Burschka, D., Suppa, M., & Hirzinger, G. (2010). Adaptive and generic corner detection based on the accelerated segment test. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010* (pp. 183–196). Springer. doi:10.1007/978-3-642-15552-9_14

Mertens, D. M. (2018). Ethics of qualitative data collection. In U. Flick (Ed.), *The SAGE handbook of qualitative data collection* (pp. 33–48). doi:10.4135/9781526416070.n3

Norouzi, M., Fleet, D. J., & Salakhutdinov, R. (2012). Hamming Distance metric learning. *Neural Information Processing Systems, 25*, 1061–1069. https://papers.nips.cc/paper/4808-hamming-distance-metric-learning.pdf

Plageras, A. P., Psannis, K. E., Stergiou, C., Wang, H., & Gupta, B. B. (2018). Efficient IoT-based sensor BIG Data collection–processing and analysis in smart buildings. *Future Generation Computer Systems*, *82*, 349–357. doi:10.1016/j.future.2017.09.082

Rauch, S., Wallner, B., Ströhle, M., Dal Cappello, T., & Brodmann Maeder, M. (2020). Climbing accidents—Prospective data analysis from the international alpine trauma registry and systematic review of the literature. *International Journal of Environmental Research and Public Health*, *17*(1), 203. doi:10.3390/ijerph17010203 PMID:31892182

Rivlin, R. S., & Ericksen, J. L. (1997). Stress-Deformation relations for isotropic materials. In G. I. Barenblatt & D. D. Joseph (Eds.), *Collected papers of R.S. Rivlin* (vols. 1-2, pp. 911–1013). Springer. doi:10.1007/978-1-4612-2416-7_61

Sanchez, J., Denis, F., Cœurjolly, D., Dupont, F., Trassoudaine, L., & Checchin, P. (2020). Robust normal vector estimation in 3D point clouds through iterative principal component analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, *163*, 18–35. doi:10.1016/j.isprsjprs.2020.02.018

Scaramuzza, D. (2014). Omnidirectional Camera. In K. Ikeuchi (Ed.), *Computer vision: A reference guide* (pp. 552–560). Springer US. doi:10.1007/978-0-387-31439-6_488

Triggs, B., McLauchlan, P. F., Hartley, R., & Fitzgibbon, A. (2000). Bundle adjustment — A modern synthesis. In B. Triggs, A. Zisserman, & R. Szeliski (Eds.), *Vision algorithms: Theory and practice* (pp. 298–372). Springer. doi:10.1007/3-540-44480-7_21

Weinbruch, S., & Nordby, K. C. (2013). Fatalities in high altitude mountaineering: A review of quantitative risk estimates. *High Altitude Medicine & Biology*, *14*(4), 346–359. doi:10.1089/ham.2013.1046 PMID:24377342

Whymper, E. (1871). *Scrambles Amongst the Alps in the years 1860-69*. https://www.gutenberg.org/files/41234/old/41234-pdf.pdf

Yan, T., Shen, S., Zhou, A., & Chen, X. (2022). Prediction of geological characteristics from shield operational parameters by integrating grid search and K-fold cross-validation into stacking classification algorithm. *Journal of Rock Mechanics and Geotechnical Engineering*, *14*(4), 1292–1303. doi:10.1016/j.jrmge.2022.03.002

Zaliva, V., & Franchetti, F. (2014). Barometric and GPS altitude sensor fusion. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi:10.1109/ICASSP.2014.6855063

Zhang, Z., Deriche, R., Faugeras, O., & Luong, Q. (1995). A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, *78*(1–2), 87–119. doi:10.1016/0004-3702(95)00022-4