# U-Star: An Underwater Navigation System based on Passive 3D Optical Identification Tags

Xiao Zhang, Hanqing Guo, James Mariani, Li Xiao
Michigan State University
{zhan1387,guohanqi,mariani4}@msu.edu,lxiao@cse.msu.edu

## ABSTRACT

Underwater optical wireless communication techniques are promising due to a broad bandwidth with a long communication range compared with existing expensive acoustic and RF-based underwater communication techniques. For underwater navigation assistance during dive and rescue, it is more practical to adopt passive optical tags for objects/human identification and location-based services. However, existing optical tags (bar/QR codes) employ one/two dimensional designs, which lack significant element/symbol distance for robust decoding and full-directional localization capabilities for underwater navigation tasks. This paper investigates opportunities to increase the element distance in passive low-order optical tags by exploiting 3D spatial diversity. Specifically, we design **U-Star**, a system that consists of Underwater Optical Identification (UOID) tags and commercial camera-based tag readers for underwater navigation. Our UOID tags embed rich location and guidance information. Additionally, because our UOID tags employ a three-dimensional design, they can also determine the relative location of a user in real-time based on the perspective principles. We design AI based mobile algorithms for underwater denoising, relative positioning, and robust data parsing for tag readers. Finally, we evaluate **U-Star** on real UOID tag prototypes under different underwater scenarios. Results show that our 3-order **UOID** tag can embed 21 bits with a BER of 0.003 at 1m and less than 0.05 at up to 3m, which is sufficient for underwater navigation guidance with backup database.

## CCS CONCEPTS

• **Networks** → *Mobile networks*; **Wireless personal area networks**; • **Human-centered computing** → *Mobile devices*; **Ubiquitous and mobile computing**.

## KEYWORDS

Underwater Optical Wireless Communication, Passive 3D Optical Tag, Underwater Navigation System.

**ACM Reference Format:**
Xiao Zhang, Hanqing Guo, James Mariani, Li Xiao . 2022. U-Star: An Underwater Navigation System based on Passive 3D Optical Identification Tags. In *The 28th Annual International Conference On Mobile Computing And Networking (ACM MobiCom '22), October 17–21, 2022, Sydney, NSW, Australia.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3495243.3517019


*ACM MobiCom '22, October 17–21, 2022, Sydney, NSW, Australia*

ACM ISBN 978-1-4503-9181-8/22/10…$15.00
https://doi.org/10.1145/3495243.3517019

## 1 INTRODUCTION

The ocean, other natural and man-made water areas (e.g. lakes, rivers, ponds, pools, reservoirs, etc) account for more than 71% of the surface area of Earth. Although sea exploration has been undertaken throughout history, much of the underwater world remains a mystery that still needs to be explored by humans[28, 29]. Nowadays, there has been a growing research interest in numerous water-based applications such as climate change monitoring, oceanic animals study, oil rigs exploration, lost treasure discovery, unmanned operations, scuba diving, search/rescue, and underwater navigation assistance[39]. Additionally, it is reported by Market Reports that the Global Scuba Diving Equipment market was valued at USD 1127 million in 2020 and is projected to reach USD 1503 million by 2027[26]. Most of these applications require reliable, flexible, and fast underwater communication to provide a safe and comfortable experience. However, despite the rapid development and progress of terrestrial and space communication, high-speed underwater wireless communication (**UWC**) is still not fully explored[6, 17, 24, 28].

There are significant differences between underwater and terrestrial scenarios, such as a harsh environment and lack of infrastructure deployment. When signals propagate in water, wireless communication faces challenges: water turbulence, limited power supply, unusable GPS, marine animal block issues, etc. Today's most popular UWC techniques adopt acoustic, radio frequency (RF), and optical waves as wireless mediums. However, acoustic signals are generated by high-power sonar (sound navigation and ranging) equipment with a long communication range, but with the cost of high communication latency. As for RF-based UWC techniques, they have low latency but still face high energy consumption issues with a minimal communication range due to severe interference of seawater with the electromagnetic waves[14, 16, 18, 28, 39, 43].

Underwater Optical Wireless Communication (**UOWC**) has shown significant potential due to its longer propagation range, lower propagation delay, and lower power consumption compared with acoustic and RF-based techniques[29, 36, 38, 39, 44, 47]. Moreover, UOWC systems based on passive optical tags, which utilize natural light sources, are more practical because they do not rely on finite battery power in underwater scenarios where it is not feasible to perform frequent battery replacement.

Similar to terrestrial navigation procedures, underwater navigation systems need to be able to answer these two fundamental questions: (1) *Where am I now?* and (2) *How do I get to where I am going?* For GPS-based navigation, systems first determine the user's

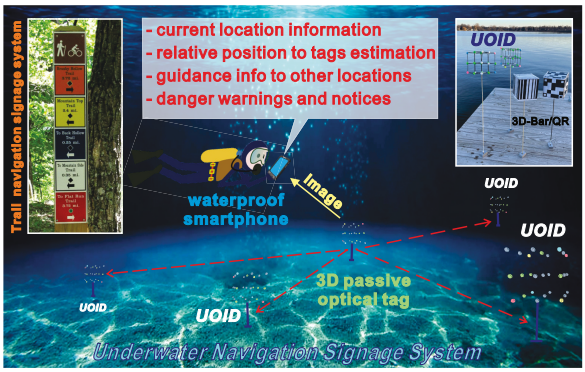



**Figure 1: U-Star based underwater navigation illustration.**

current location by GPS localization and then provide terrestrial navigation guidance based on a pre-established location database. Another common method of terrestrial navigation guidance utilizes signage systems such as visitor guidance boards in a museum, campus, or trail. The tour map on the board has a notation (star/dot) to show the user's current location, and then the user can find the way to somewhere based on the map[22].

Because GPS is unusable underwater and other underwater acoustic/RF-based localization methods are expensive [28], divers traditionally navigate underwater using a portable waterproof compass and location information provided by their guide before diving, which is not smart, reliable, or flexible[13, 20, 37]. Inspired by terrestrial navigation, we can adopt waterproof signage systems to show users rich location information for underwater navigation. This, however, has many challenges, as it is hard to find and read a finite-sized map image or messages underwater due to the harsh optical environment.

Alternatively, we can use passive tags and a portable tag reader for more embedded and clear navigation information. In our daily life, passive optical tags such as barcodes and QR (Quick Response) codes are popular[25, 41], but their short communication range makes underwater navigation impossible because users cannot even find the tags to scan them. A larger tag size could provide a more extended communication range, but at a higher price and would cause a greater disturbance to the original ecological environment.

When discussing passive tags we define a high-order tag as containing more than five elements per dimension. For example the barcode in the left of Figure 2 contains 16 columns, or 16 elements in its one dimension. We also define a low-order tag as having five or less elements per dimension. High-order tags, however, are not feasible for underwater navigation because as the number of elements increases the error rate also increases due to the necessity for elements to be physically closer to each other. On the other side, the amount of embedded data in a low-order barcode or QR code is not rich enough for underwater navigation.

**Motivation:** (1) Acoustic and RF-based UWC is not feasible because of drawbacks such as high latency, low communication range, or need for an external power source. (2) High-order optical tags cannot be reliably used for underwater navigation because of their error rates and short communication range. (3) Existing optical tags only utilize 1D/2D spatial diversity for data embedding[33]. Even the 3D versions of Bar/QR codes shown in Figure 1 have limited element distances and ignore 3D spatial diversity. As a result, there will be more error bits in decoding, especially in muddy underwater scenarios. (4) Existing bar/QR codes, even in 3D, have limited scanning angles and require the user to move to directly face the surface of the codes, which is inconvenient for underwater navigation activities. (5) We can use 3D spatial features to provide underwater positioning based on the perspective principle, which states that objects such as cubes are observed differently at different distance and angles. This is discussed in more depth in Section 2.3.

**Approach:** In this paper, we design **U-Star**, an underwater signage system based on passive 3D optical identification tags for underwater navigation, as illustrated in Figure 1. U-Star consists of UOID tags and the AI-based mobile tag reader. UOID tags are hollowed-out cubes which consist of data elements and positioning elements. The data elements are positioned with proper non-Line-of-Sight spacing on the UOID tag. The positioning elements are set in different clockwise color sequences along the six faces of the UOID. The U-Star tag reader is built on waterproof mobile devices with standard, commercial cameras. The detailed navigation procedures are illustrated in Section 3.

Our **contributions** can be summarized as follows:

**(1)** This is the first work to employ passive **3D** optical identification tags for underwater navigation. We model 3D spatial diversity and utilize it to increase the distance of data elements in our proposed UOID tags for simple and robust underwater navigation.

**(2)** We propose a passive 3D optical identification tag based positioning scheme for underwater navigation. Our UOID tag can help user to determine their current orientation by the arc of clockwise positioning elements and estimate the underwater distance due to perspective principles.

**(3)** We propose AI-based mobile algorithms at the tag reader for robust UOID decoding. We design CycleGAN based underwater denoising, CNN-based relative positioning, and real-time data parsing algorithms without significant computation overhead, latency or energy concerns.

**(4)** We implement U-Star and evaluate its performance on UOID tag prototypes in different underwater scenarios. Our experiment results show that a 3-order UOID tag can embed 21 bits of data with a BER of 0.003 at 1m and less than 0.05 at a distance of up to 3 m. We also make fair comparison with existing optical tags (Bar, QR) to show the superiority of our UOID tags in underwater navigation. U-Star also achieves over 90% accuracy for both optical ranging at up to 7m and orientation guidance.

The rest of the paper is organized as follows. Section 2 introduces background and related work. Section 3 gives the U-Star system overview. Section 4 to 6 illustrate U-Star design: passive 3D optical tag, underwater relative positioning, and mobile AI based tag reader. Section 7 presents U-Star implementation. Section 8 reports the performance evaluation of U-Star. Finally, we have some discussion in Section 9 and conclude the paper in Section 10.

## 2 BACKGROUND AND RELATED WORK

### 2.1 Underwater Navigation

Underwater navigation is important for human-related underwater activities, such as scuba diving and underwater accident rescue. Natural underwater navigation requires the diver to utilize physical contours and characteristics of dive sites and combine basic compass skills to find the path to their destination[5, 11, 13]. Natural

underwater navigation is similar to terrestrial navigation shown in Figure 1, the diver first needs to known his/her current location based on the site map or underwater physical features of dive sites and then guide him/herself to their destination based on the information on map or prior knowledge. However, natural underwater navigation relies highly on diver's familiarity with dive sites. If unfamiliarity or any confusion with dive sites, it is very dangerous for divers, to the point that many have lost their lives.

Many researchers have made efforts to improve underwater navigation[14, 15, 19, 27]. However, these are based on acoustic and RF techniques that incur significant drawbacks, including high power consumption, expensive price, long latency, or short communication ranges. To combat these issues, we explore setting underwater, on-site visible signage tags to provide site location information and navigation guidance. Our approach is inspired by traditional terrestrial navigation techniques such as tour maps and location marks in hiking trails[21, 22] and offers new and innovative techniques for underwater navigation.

However, it is not practical to just place the signage tags underwater in a similar fashion to terrestrial navigation. This is because it is not as easy for users to move to directly face the tags as it is on land, the hostile underwater optical environment, and that the lengthy communication distance [12, 15] makes effectively reading the signage impossible. The optical tags used in underwater navigation need three features: (1) **Easily observed**. The color and brightness are striking enough to be observed by users at long distances (10m-20m) and the content on the tag should be visible from practically every angle. (2) **Enough data capacity**. The data embedded in the tags needs to be large enough to record both the location information and guidance advice. (3) **Positioning ability**. The tag needs to provide relative position information to the user. Feature (1) is more based on material and color choice, specifically, to suit the underwater scenario, we discuss this in Section 4.2.2. Feature (1) also relies on the hollowed-out structure of the tag design for the real 3D passive optical tag discussed in Section 2.2. Features (2) and (3) are in the category of optical wireless communication and we discuss the related work below.

## 2.2 Optical Data Embedding

Barcodes and QR codes, as seen in Figure 2, are very popular machine-readable optical tags used in our daily life. Barcodes represent data by parallel lines with varying widths and spacing invented in 1951[35]. The barcode became commercially successful in supermarket checkout systems. Later, two-dimensional (2D) variants (matrix codes) were developed, which can represent more data per unit area[32, 33]. The QR (Quick Response) code is one of these matrix codes and is common in many aspects of life, such as mobile payment, social E-cards, electronic tickets, access control, etc. High-order QR codes, such as the version 40 QR code (177x177), can embed 23,648 bits[32]. However, in underwater navigation scenarios, high-order bar/QR codes are not suitable to be seen and scanned due to their limited scanning angles, limited data element distance, and the quality of the optical environment.

These bar/QR codes only focus on 1D and 2D spatial diversity and ignore the potential opportunity of three-dimensional spatial diversity in optical tag data embedding. Even with the 3D version of Bar/QR codes (six planes of the cube are covered with the same

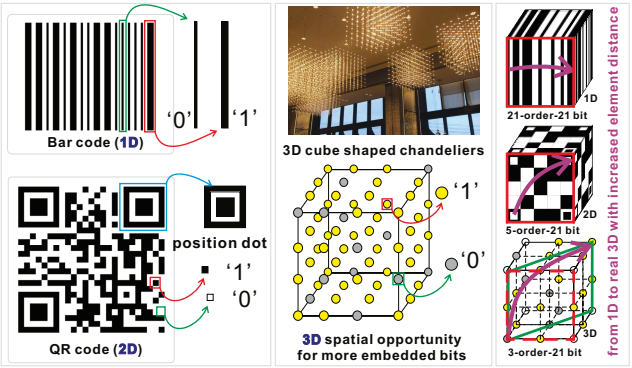



**Figure 2: Existing optical tags and 3D spatial diversity.**

bar/QR codes to ensure consistent content at various angles), the user can record up to three repeat bar/QR codes, which does not increase data element distances and does not fully take advantage of 3D spatial diversity in data embedding. Our 3D optical tag design is inspired by 3D cube-shaped chandeliers shown in Figure 2, but improved and modified for the data and communication needs of underwater scenarios. Each element inside of a 3D light cube can denote bits **1** and **0** via its **On** and **Off** status, as opposed to linear or matrix dots on a surface in a bar/QR code. Although images of the 3D optical tag captured by our tag reader is a 2D pixel matrix, we can restore the 3D optical tag based on perspective principles (discussed in Section 2.3 and Section 6.3). When compared to optical tags with the same tag size and the same amount of embedded data (e.g., 1D, 2D codes, and surface 3D tags with 1D/2D codes attached), our proposed 3D hollowed-out cube improves data element distance by leveraging 3D spatial diversity in data embedding, as shown in the right of Figure 2. In our U-Star system, we design UOID, passive 3D optical identification tags, to utilize the 3D spatial diversity to increase the distances among data elements for robust and full-directional underwater decoding.

## 2.3 Optical Positioning and Perspective

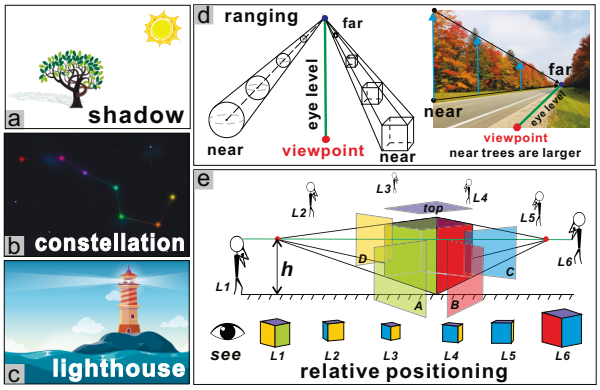



**Figure 3: Perspective principle for positioning**

It is very common for humans to utilize natural or human-made luminous objects for positioning, as shown in Figure 3. For example, we can determine orientation by observing the direction of the shadows during the day time due to the sun's movement and

the direction of the Big Dipper at the night because the orientation of the Big Dipper is unchanged and always pointing to the Earth's North Pole[7, 8]. In addition to orientation and localization based on natural optical objects, lighthouses are an example of human-made optical object based positioning. The basic functions of lighthouses are to guide ships, indicate dangerous areas and help ships to determine their positions[30]. For underwater scenarios, researchers have also made many efforts to design optical-based underwater positioning mechanisms and systems. Akhoundi et al. design RSS (Received Signal Strength) based optical positioning systems that calculate location based on the received optical signal from multiple anchors[4]. In other work[40], the authors proposed a ToA and RSS-based underwater localization system. However, these works require a significant power supply and expensive devices with high-accuracy sensors.

Perspective principles are traditionally used in vision and art[23, 31]. Creatively, we can utilize the perspective principles for ranging and relative positioning. The perspective principle simply describes the visual relationship between the observer and the observed object: (1) increasing the distance between the observer and the object results in a reduced size of the observed object, as shown in Figure 3 (d); (2) varying the angle from the view point to the object results in a variable shape and observed content of the observed object, as shown in Figure 3 (e). Our U-Star design also utilizes UOID tags as fixed underwater beacons utilizing 3D spatial diversity for optical ranging and orientation guidance besides its data embedding.

Compared with existing work, our UOID tags are based on passive optical wireless communication and therefore utilize natural light sources to present data and provide relative positioning without energy consumption concerns. The tag readers are also commercial camera-based devices instead of expensive sensors.

## 3 SYSTEM OVERVIEW

Our proposed underwater navigation system consists of two parts, as shown in Figure 4: (1) 3D passive optical tags: UOID tags, and (2) AI-based mobile tag reader.

**UOID tags.** UOID tags are anchored underwater with fixed facing direction. They are made of fluorescent materials and can absorb light from natural underwater environment or users' flashlight. There are data elements and positioning elements in UOID tags, which are assigned with proper spacing to eliminate LoS blockage in the tag's 3D spatial domain to present data.

**Tag reader.** The tag readers are based on commercial smart devices such as smartphones or sports cameras. These devices can capture images of UOID tags and perform underwater, robust, and real-time data parsing and relative positioning by its onboard computation abilities. The U-Star tag readers have three key modules: (1) CycleGAN denoising based pre-processing, (2) CNN based relative positioning, (3) 3D restoring based decoding.

**User operation and navigation procedure.** The detailed U-Star underwater navigation procedures are: (1) The diver, equipped with a tag reader, looks for luminous UOID tags. (2) The diver uses waterproof tag reader to take pictures of a specific UOID tag at current location. (3) The tag reader performs image style transformation for denoising, then the tag reader can determine diver's relative position including the distance estimation and orientation guidance. (4) The diver knows where he/she is now and can navigate him/herself to new sites based on the pre-recorded data from the backup database that the tag reader can query with the embedded data from the UOID tag (which we call a query code). The user operation is simple and quick and can be performed at different distances with all directions in different environments and time.

**Challenges.** (1) **LoS blockage.** When capturing tag images, some inside elements are blocked by their front elements due to lights' line-of-sight propagation. We address this by assigning elements with proper spacing and a machine learning based restoration. (2) **Harsh optical environment.** The underwater environment decreases the quality of captured UOID images, and thus makes them hard to decode. We design CycleGAN based algorithms to transfer unclear images into clear images (Unity3D-style images) before decoding. (3) **Underwater relative positioning.** The UOID tag is expected to help determine the distance between the user and the tag as well as user's current orientation for relative positioning. We propose clockwise positioning arc schemes to denote planes and a CNN method to infer relative position. (4) **3D decoding.** The tag reader needs to restore each element to a standard 3D space from a random 2D image during decoding. We utilize the perspective principle to reconstruct the 3D structure for data parsing.

## 4 PASSIVE 3D OPTICAL TAG

### 4.1 3D Spatial Diversity Exploration

As shown in Figure 5, we use a 3D cube instead of a 2D matrix to represent more bits in an optical tag. Naturally, there are two methods to embed data in a 3D cube: (1) embed data on its six surfaces, (2) embed data on both its surfaces and inside space (i.e., hollowed-out), which fully utilize the 3D spatial diversity. For method (1), the tag reader can only capture the dots on 1 and up to 3 surfaces due to the line-of-sight (LoS) characteristic of light. Method (1) cannot also guarantee that the embedded data captured at different angles is always the same (unless all 6 planes cover the same content) due to

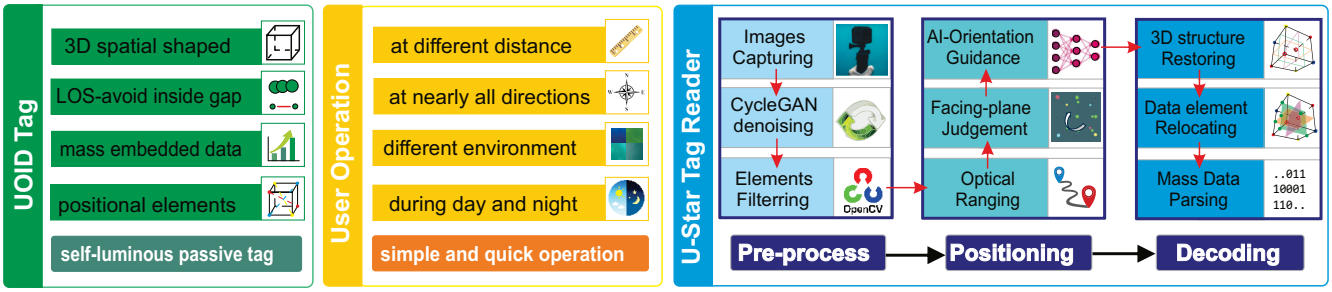



**Figure 4: U-Star system diagram including UOID tag, user operation, and tag reader.**

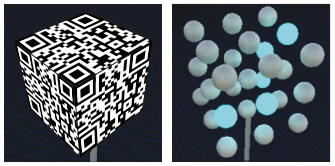



**Figure 5: Surface/real 3D.**

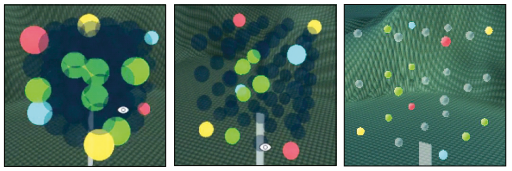



**Figure 6: Proper spacing to combat LoS.**

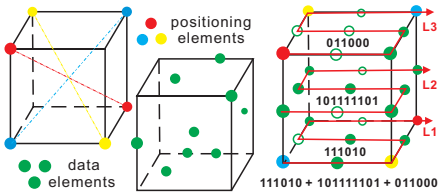



**Figure 7: Two element types in UOID.**

the potential of capturing different surfaces, which means that the tag's decoded data will change without consistency. Additionally, Method (1) results in smaller data element distances and a shorter communication range, which is verified in Section 8.3.

Thus we choose method (2) to embed data in our UOID design. However, the LoS issue can also occur if we embed data inside of a 3D cube due to mutual blockage among elements physically near each other. As shown in Figure 6, the 4-order (4x4x4) tag without proper spacing will have the mutual blockage issue. Three factors affect the blockage: (1) Tag order. As the order of tags increases (3-order, 4-order, 5-order, etc), more and more blockage occurs. Similarly, as the order of tags decreases, so does blockage. (2) Element size. The smaller the element size, the less blockage. (3) Mutual Spacing. The larger the mutual spacing of elements, the less blockage. In this paper, we discuss a 3-order UOID tag with fixed element size and we address the mutual blockage by extending the spacing among nearby nodes to guarantee the tag reader can capture all elements in most cases.

## 4.2 UOID Tag Design

*4.2.1 Positioning and data elements.*

In our UOID tag design, there are two types of elements: positioning elements and data elements, as shown in Figure 7. The positioning elements are on six vertex points with three pairs of colors. The positioning elements help determine the relative position of the user and assist in reconstructing the 3D cube for data parsing. The data elements make up most of the elements in a UOID tag for data embedding. They are located at the two remaining vertex points as well as inside of the tag itself.

**Positioning elements.** As shown in Figure 7, each pair of colored elements are at a pair of vertex points. Thus, each plane of the cube has three different colored positioning elements. They can denote six surfaces based on the generated clockwise arc color sequence (discussed in Section 5.2, Figure 10 (a)). Then the tag can determine which surfaces the user is facing based on captured surfaces of the tag and determine orientation based on the perspective principle to support underwater navigation. Furthermore, these positioning elements can help to reconstruct the 3D structure from captured 2D images based on the perspective principle for data parsing. The reason for using three instead of four positioning elements to denote a plane are: (1) Three dots can already determine a surface. Four dots will sacrifice the positions that could be used for assigning data elements and thus decrease the embedded data. (2) Fewer overall colors is desirable, as more colors will increase the color detection error for decoding due to the fewer hue gaps.

**Data elements.** The data elements of our UOID are assigned to various 3D spatial locations. There are three layers L1, L2, and L3. For each layer, we assign data elements in an '**S**' shape. If the data element is colored green, the embedded bit is **1**, if the data element is not colored, the embedded bit is **0**. As illustrated in Figure 7, L1 embeds bits '111010', L2 embeds bits '101111101' and L3 embeds bits '011000'. This 3-order UOID tag embeds a total of $3^3$-6=21 bits, '111010 101111101 011000'. We set the current angle of view to be the standard coordinate system for data parsing. With the assistance of positioning elements (discussed in Section 6.3, Figure 13), we can map the tag images from any angle of view into the standard coordination system and then conduct the mass data parsing.

*4.2.2 Underwater-specific Tag Design.*

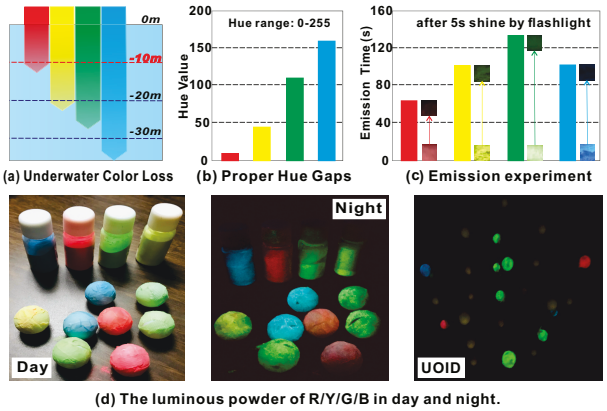



**Figure 8: Color choices and luminous powder.**

**Color Choices.** Light with different wavelengths/color have different absorption rates in water. As shown in Figure 8 (a), the green and blue light have less absorption in deeper underwater environments such as a depth of 20 m[28, 42]. However, considering most commercial underwater activities do not exceed depths greater than 10 m, the color choices (red, yellow, green, and blue) above in the UOID tag are reasonable (for deeper underwater navigation, finer-grained blue and green can be chosen). As shown in Figure 8 (b), these four colors also have sufficient hue value gap to decrease the wrong detection of colors during decoding[45]. The green light has the longest emission time after 5s of being shined by a flashlight as shown in Figure 8 (c). Because data elements are the most numerous and important elements, we set them to green.

**Luminous powder.** Our UOID tags are passive, without any power supply. As illustrated in Figure 8 (d), we coat the elements in luminous powder, which is cheap and nontoxic to marine animals. As shown in Figure 8 (c), the luminous powder with our chosen colors can keep emitting light more than 60 seconds (1 min) after being shined by a flashlight for 5 seconds in our experiments. This ensures that the UOID tags work by absorbing natural underwater light and emitting light in specific colors, allowing us to see and scan UOID tags at any time of day or night.

# 5 UNDERWATER POSITIONING

## 5.1 Optical Ranging

For underwater navigation, the perception and estimation of distance is very important. Our UOID tags can give the user a rough feeling of the distance between themselves and the tag. We use the rough size of the captured tag to infer the current distance from the user to the tag. The estimated relative distance has no relation with the angle of capturing images by the user.

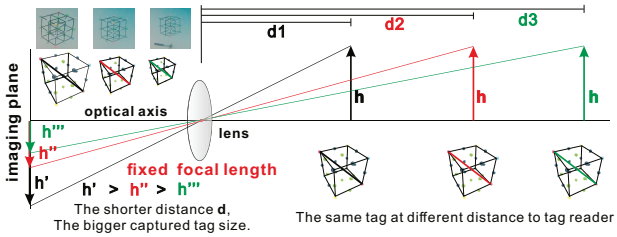



**Figure 9: 3D spatial perspective based optical ranging.**

As shown in Figure 9, We can estimate the distance based on the captured tag size because the tag size increases when the user is getting closer to the tag due to the spatial perspective principle. We first collect the captured images (camera is set with fixed focal length) at different distances and use this dataset to train the CNN model for classification offline. Then we can use the trained CNN model to predict and estimate the current distance from the user to the tag in real-time.

## 5.2 Orientation Guidance

We map the six planes of the UOID tag onto six different clockwise color arcs which start from the non-positioning element: Yellow(Y)-Blue(B)-Red(R) maps to Plane 1, BRY to Plane 2, RBY to Plane 3, YRB to Plane 4, RYB to Plane Top, and BYR to Plane Bottom as shown in Figure 10 (a). The UOID tag is fixed underwater (i.e., a specific plane of the UOID tag always faces in a specific direction), and thus the user/tag reader can know his/her orientation based on the plane of the UOID the user is currently facing. For example, as shown in Figure 10 (b), Plane 1 is facing South. That means if the user is facing Plane 1, the user can know his/her current orientation is directed North.

For underwater navigation, the Plane Top and Bottom faces do not provide value to orientation decisions. Additionally, North, East, South, and West are not sufficiently descriptive for navigation. Therefore, we define 8 user facing orientations: North (facing Plane 1), Northwest (facing Plane 1 & 2), West (facing Plane 2), Southwest (facing Plane 2 & 3), South (facing Plane 3), Southeast (facing Plane 3 & 4), East (facing Plane 4), Northeast (facing Plane 4 & 1) as shown in Figure 10 (b). Naturally, we can determine the plane the user is facing based on the color arc detected in images. However, due to the small size of elements in captured images, it is hard to judge which plane the user is facing. Thus, we employ CNN models to learn plane features offline and then predict the plane in the captured image in real-time, similar to the AI method used in the optical ranging procedure.

# 6 AI-BASED MOBILE TAG READER

## 6.1 CycleGAN based Denoising

CycleGAN is a popular deep learning method and is mostly used for image style transforming which can convert images between Style X and Style Y. For example, to generate a monet-style image from a real world picture or vice versa[46]. We adopt a lightweight CycleGAN to convert the real underwater images taken of real, physical UOIDs created for U-Star (Style X) into clear Unity3D-style images created in the Unity3D game engine (Style Y) for further processing. The images in real underwater scenarios have a random and different background (i.e., with noise) for UOID tags. The images in the Unity 3D version have clear and pure backgrounds (i.e., there is no noise from the background in these images). Thus, we can utilize CycleGAN to convert real-world images with noise (Style X) to Unity3D-version images without noise (Style Y) to perform underwater denoising as shown in Figure 11.

In our CycleGAN-based denoising, instead of the typical unpaired datasets, we create the partial-paired datasets, the Real UOID tags (60 images) and the Virtual UOID tags (60 images), for each underwater environment setting in the CycleGAN training procedure, as shown in Figure 11 (a). Partial-paired means the positioning elements are paired between the real UOID tag images and the Unity3D version images of the training datasets, while the inside data elements are not paired. Partial-paired CycleGAN denoising guarantees mostly correct conversion of both the tag structure, data elements and the color of positioning elements.

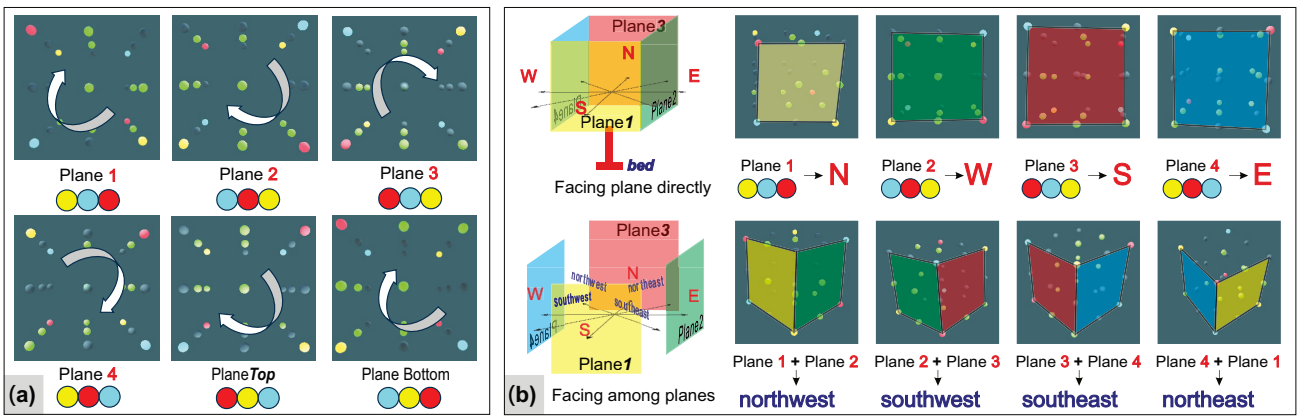
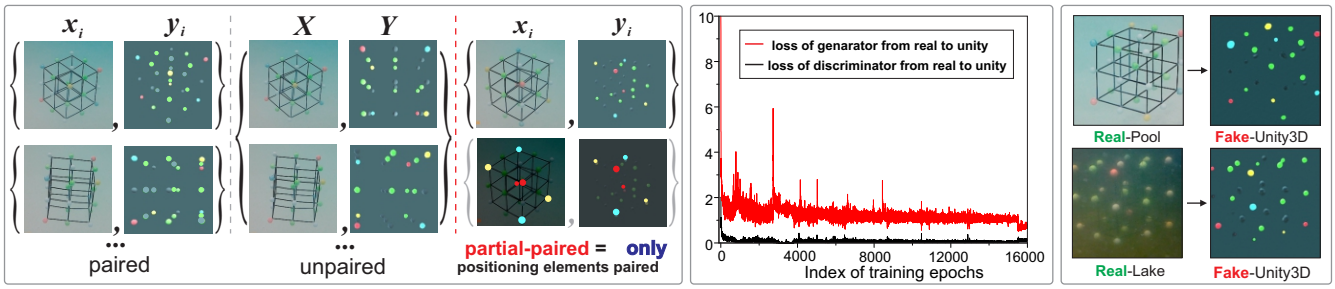



**Figure 10: Positioning elements for the plane decision and the orientation guidance principle illustration.**




(a) Partial-paired training dataset (b) Loss curve in training (c) Result samples

**Figure 11: CycleGAN based denoising from real underwater tag images to the Unity3D version tag images.**

To train the CycleGAN efficiently, we use three different types of losses to train our Cycle-GAN. More specifically, we apply an identity loss ($\mathcal{L}_{id}$) for generator network, a GAN loss ($\mathcal{L}_{GAN}$) for the Discriminator, and a cycle loss ($\mathcal{L}_{cycle}$) for the cycle step.

$$\mathcal{L}_{CycleGAN} = \mathcal{L}_{id} + \lambda_1 \mathcal{L}_{GAN} + \lambda_2 \mathcal{L}_{cycle} \quad (1)$$

Both identity loss and GAN loss are using L1 loss, while the cycle loss is applied by a MSE loss. We summed those three losses together with different prior assigned weights ($\lambda_1$ and $\lambda_2$) to help the model converge. The value of $\lambda_1$ and $\lambda_2$ are selected empirically, in our case, we use 10 and 5 for $\lambda_1$ and $\lambda_2$ respectively. By integrating the three losses together, we feed the pairwise training images to the CycleGAN and train the generators and discriminators. The loss curve in the training of the generator and discriminator (from real images to Unity3D-style images) are shown in Figure 11 (b). The varying trend of the loss curves show the conversion from the real underwater UOID tag images to the Unity3D-style UOID tag images converges successfully.

The original captured underwater images and the denoised images are shown in Figure 11 (c). We can see that underwater images from both a pool and lake can be successfully denoised and converted to Unity3D-style images with a mostly correct tag structure, color, and element positioning. The CycleGAN denoising also removes the physical UOID frame components to reduce the LoS blockage. Although there are a few elements with unmatched colors, we can correct them based on the original image easily. The next steps of relative position determination and data parsing can then be based on these converted Unity3D-style UOID tag images to lessen the influence of harsh underwater optical environment.

## 6.2 CNN based Relative Positioning

We adopt CNN-based deep learning methods to determine the relative position instead of non-deep, traditional computer vision methods to simplify the task and decrease the computation overhead. It is difficult to calculate relative distance directly with different underwater backgrounds, which requires several steps: (1) locate the tag in the image using AI or CV methods, (2) calculate the tag size, and (3) utilize the distance estimation relation to calculate the estimated distance. In comparison, we choose a CNN model because it does not necessitate detecting the tag in the image or calculating the tag size. Instead, we directly output the prediction distance in different underwater environments using the trained CNN model and captured images of UOID tags.

We create two datasets (1) optical ranging dataset (280 images of Unity3D version and 280 images of real underwater), and (2) orientation dataset (320 images of Unity3D version and 320 images of real underwater) for the offline CNN training. The reason using both real underwater tag images and Unity3D version tag images in training is to increase the generality of the prediction model. Then we use CycleGAN denoised tag images for real-time relative position determination. As shown in Figure 12, our CNN models, ORM (optical ranging model) and OM (orientation model), adapt the ResNet-18 architecture. ResNet-18 is a neural network architecture that adds a skip connection between disconnected layers, such that the input of deep layers will not only take the output from its preceding layer, but also from its former layers which may contain original data. Such design effectively copes with gradient vanishing problem in DNNs[10], and further increases the depth of network with fewer additional parameters. ResNet has demonstrated superior performance on image classification tasks [1–3], which is particularly suitable for our goal that distinguishing relative position both optical ranging and orientation guidance. We follow the ResNet-18 design due to its efficiency and high accuracy on image classification tasks. Specifically, we retain all of the convolutional and pooling layers, and modify the output feature of the last fully connected layer to match the number of possible options (i.e., 7 for ORM and 8 for OM).

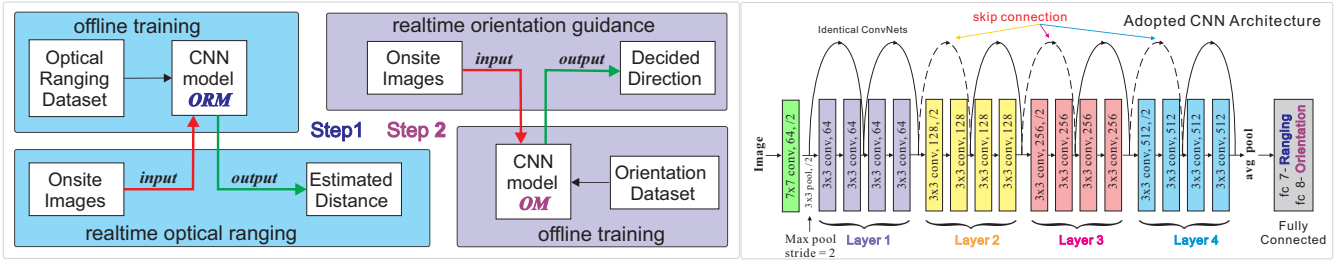



**Figure 12: CNN based relative positioning of optical ranging, orientation guidance and adopted ResNet-18 network architecture.**

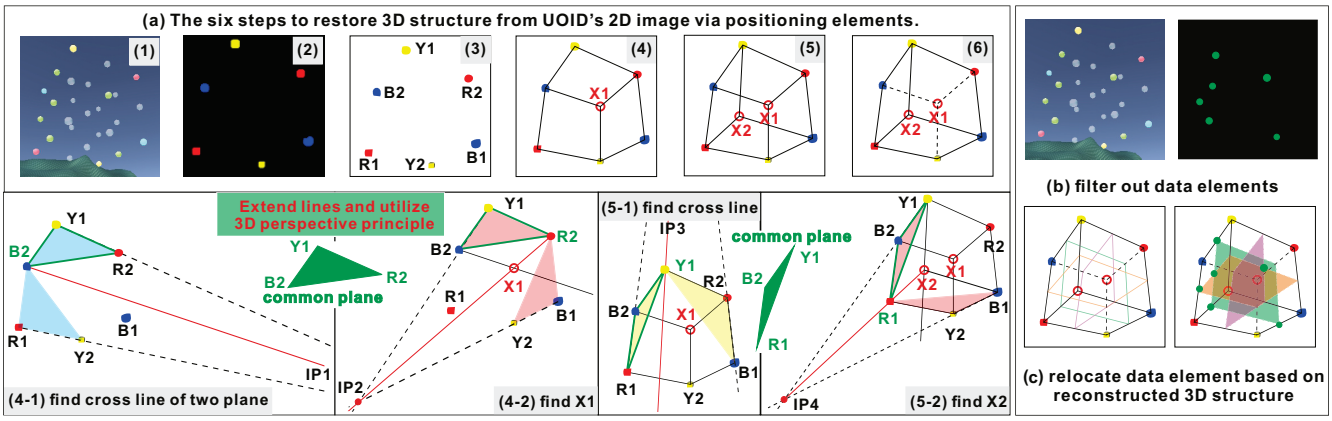



**Figure 13: The illustration of 3D structure restoring and data parsing is based on the perspective principle.**

## 6.3 Data Parsing via Perspective Principle

The data elements in captured images are different when the user is at different relative positions to the UOID tag. To decode the embedded data in the tag, the tag reader needs to know the 3D locations of data elements in a standard coordinate system to then perform decoding.

**Restore 3D structure.** Based on three pairs of positioning elements, the tag reader can restore the 3D structure of UOID tag based on captured 2D images in six steps shown in Figure 13 (a): (1) obtain Unity3D-style UOID image after CycleGAN based denoising, (2) filter out three pairs of positioning elements via computer vision tools, (3) decide which positioning element for each pair is in the front or rear based on element size, (4) find one of the two remaining vertices, (5) find the other remaining vertex, and (6) decide which of remaining vertices is front or rear based on the element size of nearby positioning elements. Finally, we can reconstruct the 3D structure based on the total of 8 vertices of the 3D cube.

For step (4), there are two sub-steps: **(4-1)** Extend line Y1R2 and R1Y2 to find the intersection point IP1(not shown in the figure). Then connect B2 with IP1, which is the cross line of plane Y2R1**B2** and Y1R2**B2**. **(4-2)** Extend line Y1B2 and B1Y2 to find the intersection point IP2. Then connect R2 with IP2, which is the cross line of plane B2Y1**R2** and B1Y2**R2**. Then we can find the first vertex, which is the intersection point of B2IP1 and R2IP2. The sub-steps for step (5) are similar to the sub-steps in step (4).

**Data element location restoration.** As shown in Figure 13 (b) and (c), we can restore the location of data elements by matching the filtered data element and locations of each element calculated based on the positioning elements. If the specific filtered data element is near or at the specific calculated location from the restored 3D structure, it signifies a match. Then we can denote that this location has a data element as bit **1** while other vacant calculated data element locations will be decoded as bit **0**. Then the tag reader decodes the embedded data and generates the bitstream based on the data assignment rule illustrated in Section 4.2 and Figure 7.

# 7 U-STAR SYSTEM IMPLEMENTATION

## 7.1 UOID Tags

We implement two versions of UOID tags. One is a virtual NxNxN UOID tag created in the Unity3D cross-platform game engine to

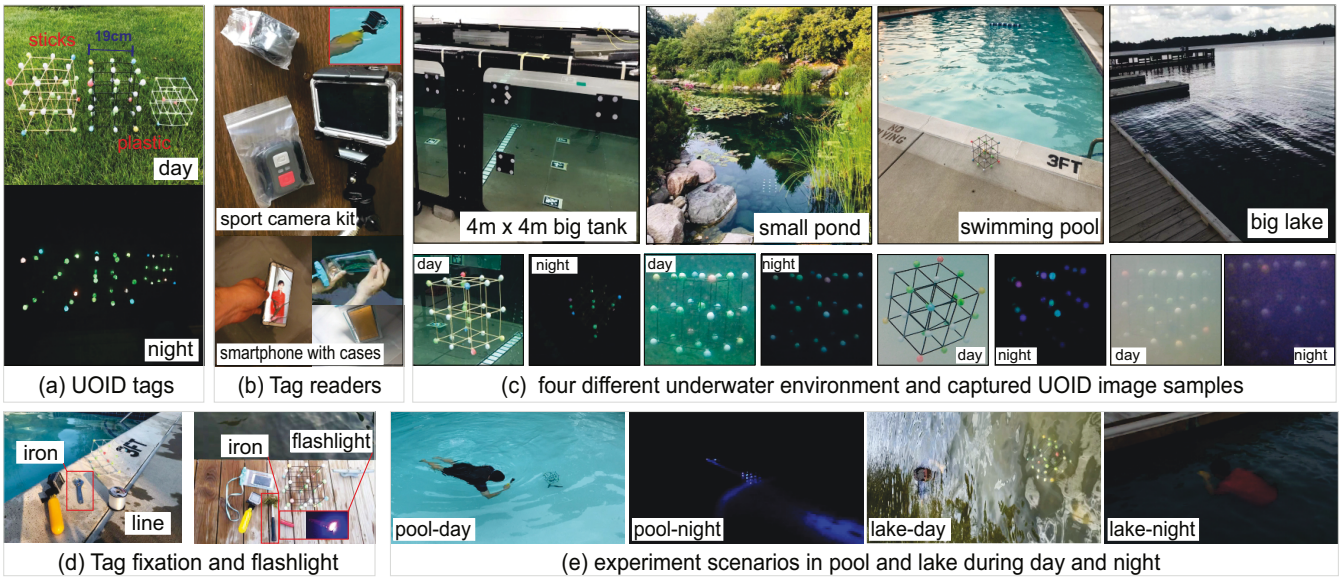



**Figure 14: U-Star system implementation, setup and experiment scenarios in day and night.**

simulate UOID tags of various order and also different permutations of embedded data within tags of the same order. We also implement multiple physical 3x3x3 UOID tags for use underwater.

**Virtual UOID tag.** The elements in our virtual UOID tags are translucent with fluorescent effects and are assigned with the proper spacing, as shown in Figure 10 and Figure 13.

**Real UOID tag.** As shown in Figure 14 (a), the UOID tags can be observed well during both the day and night because they absorb natural light and emit light, as discussed in Section 4.2.2. For the elements of our physical UOID tags we employ soft plastic balls ($\phi$ = 2cm) glazed with fluorescent powder and attach them on 3 types of cube structure frames for exploration (sticks, black and transparent plastic). Finally we choose the black plastic frame-based UOID tags ( edge: 19cm, weight: 14g) for evaluation in Section 8.

## 7.2 Tag Reader

There are many commercial smart devices that can be adopted for use in our U-Star system. Some of these include underwater sports cameras and smartphone with transparent, waterproof cases, as shown in Figure 14 (b). These commercial camera devices are popular and cheap. In our experiment, we use the Campark sport camera, which costs less than $50 and set it at a fixed focal length.

## 7.3 Setup

**Different underwater environment.** Figure 14 (c) shows four underwater environment (indoor big tank, outdoor small pond, swimming pool, and big lake) and captured images of UOID tags.

**Tag fixation and flashlight.** As discussed in Section 5.2, we fix the UOID tags at the bottom of a body of water, i.e., a specific UOID plane always faces a specific direction. We use iron and connection pole to sink and fix the UOID tag underwater, as shown with Figure 14 (d) and Figure 18. As discussed in Section 4.2.2, during the night, the user can use a flashlight for underwater lighting to activate the UOID tags, as shown in Figure 14 (c), (d) and (e).

# 8 PERFORMANCE EVALUATION

In this section, we evaluate three performance aspects of our U-Star system: (1) relative positioning, (2) data parsing, (3) comparison with existing optical tags. In addition, we conduct an underwater navigation case study in a 4m x 10m indoor pool with 4 UOID tags. Finally, we evaluate other aspects such as cost/price, computing overhead, and latency.

## 8.1 Accurate Relative Positioning.

We evaluate the relative positioning performance in three aspects: optical ranging accuracy, orientation guidance accuracy (both at $100^{th}$ epoch), and their training loss in [5, 200] epochs.

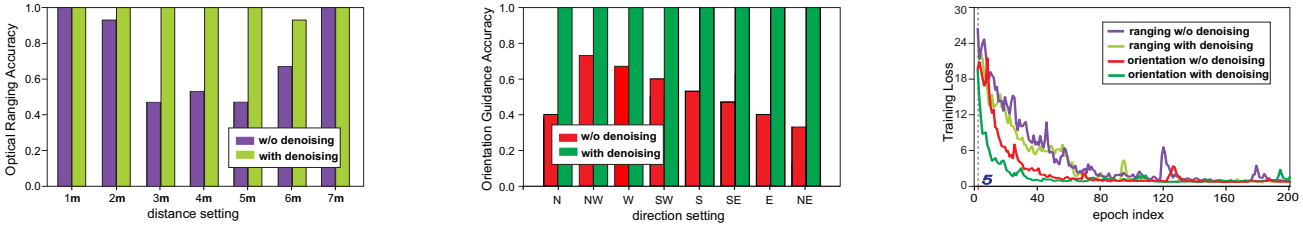
**Optical ranging.** We have 7 different distance settings: 1m, 2m, 3m, 4m, 5m, 6m, and 7m. As shown in Figure 15 (a), due to the considerable tag size difference, the ranging accuracy of 1m and 7m distance settings are 100% for both with and without CycleGAN denoising. After CycleGAN denoising, the ranging accuracy improves significantly and reaches nearly 100% for other distance settings. The results show that the trained CNN model for optical ranging performs well to estimate the distance from the user to the tag with CycleGAN denoising. The results show our current U-Star prototypes can provide up to 7 meters of optical ranging with average accuracy nearly 100%.

**Orientation guidance.** We provide eight recognized orientations for underwater navigation: North(N), North West(NW), West(W), South West(SW), South(S), South East(SE), East(E), and North East(NE). As shown in Figure 15 (b), no matter what was the user is facing (any of the eight recognized orientations) the accuracy of our orientation classification is always 100% when performing orientation guidance with CycleGAN based denoising. We also present orientation guidance performance without CycleGAN based denoising for comparison. The results show that the performance with CycleGAN denoising is better than without CycleGAN denoising. This shows that the CycleGAN based denoising helps the CNN model to improve the orientation guidance performance by decreasing the impact of harsh water conditions. The results show that our U-Star system can provide accurate orientation guidance amongst all eight orientations.

**Training loss.** For relative positioning, we also measure the loss in CNN based training for optical ranging and orientation guidance separately. As shown in Figure 15 (c), the optical ranging training loss curves both with/without denoising are above the orientation training loss curves during the training process. This means that features (tag size) in the optical ranging dataset are not as rich as the features (positioning elements and their various permutations) in the orientation dataset. The curves with CycleGAN denoising are beneath those without CycleGAN denoising during the entire training process no matter the optical ranging training or orientation training. That means that using the CycleGAN denoising can help decrease training loss more quickly and limit the impact of harsh underwater optical conditions for relative positioning.

## 8.2 Robust Data Parsing

We use our tag reader to capture images of four real UOID tags A1, A2, B1, B2 with random capturing poses in different distances, water conditions, and time of day to evaluate the decoding performance of U-Star. A1 and B1 embed raw bits without error correction codes. A2 has 3, 5, and 3 common data bits with A1 in layers 1, 2, and 3 respectively. A2 also has 3, 4, and 3 Hamming ECC parity bits in layers 1, 2, and 3. B2 has 3, 5, and 3 common data bits with B1 in layers 1, 2, and 3 respectively, and also has 3, 4, and 3 Hamming ECC parity bits in layers 1, 2, and 3. Hamming ECC[9] can correct




(a) optical ranging performance (b) orientation guidance performance (c) training loss curve

**Figure 15: Relative positioning performance in aspects of optical ranging, orientation guidance and training loss.**

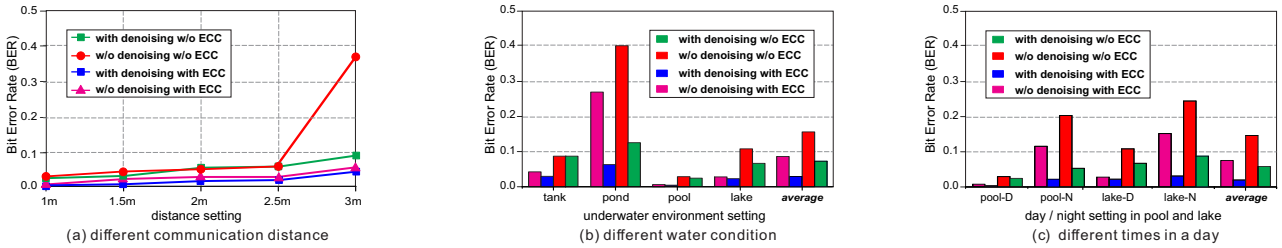



**Figure 16: Decoding performance with different communication distance, water condition, and time (day/night).**

1 error bit per bitstream, thus, for a total of 3 error bits correction capability for a tag. The bits in A1, A2, B1, B2 are shown in Table 1.

| Tag | bits in 1st layer | bits in 2nd layer | bits in 3rd layer |
|---|---|---|---|
| A1 | 101101 | 110010001 | 001011 |
| A2 | 101101 | 111110011 | 010101 |
| B1 | 001101 | 100111010 | 101001 |
| B2 | 010101 | 101100111 | 101101 |

common data bits of A1 & A2 or B1 & B2 | data bits without ECC | Hamming ECC parity bit | Valid data bit

**Table 1: Embedded bits in 4 UOID tags: A1, A2, B1 and B2.**

We define the BER as the average bit error ratio of the entire embedded valid data bits in two UOID tags with different data embedding (i.e., two tags: A1 and B1 or two tags: A2 and B2). Each BER value is calculated using 30 captured images and we use it as a metric to evaluate the decoding performance of our U-Star system. Besides the difference between UOID tags with and without Hamming ECC codes[34], we also compare BER performance with and without CycleGAN based denoising as comparison.

**In different communication distances.** We adjust the distance of the tag to the tag reader to be 1m, 1.5m, 2m, 2.5m, and 3m in clean water (pool) during the daytime. As shown in Figure 16 (a), the BER remains low, consistently less than 0.09 after CycleGAN denoising in all distance settings. We found that the best data parsing distance for current U-Star prototypes is 1m, as the BER is 0. The BER performance without CycleGAN denoising is significantly worse than with CycleGAN denoising at 3m. This confirms that the CycleGAN denoising works well, especially at longer distances. Both with and without CycleGAN denoising, the BER with ECCs is lower than for without ECCs. The BER is 0.003 at 1m and continues to be less than 0.05 up to 3m with Hamming ECC and CycleGAN denoising simultaneously.

**In different water conditions.** We explore four water conditions during the day in experiments: indoor tank with clean water, small pond, swimming pool, and big lake, as shown in Figure 14 (c). We conduct experiments at a distance of 1m (the best capturing distance for data parsing of the current U-Star prototype mentioned above). As shown in Figure 16 (b), without CycleGAN denoising, our data parsing performs best in the pool and worst in the pond. This is because the pool is clean enough for data parsing without the denoising process and the small pond makes the color of the elements change too much. After CycleGAN based denoising, the BER decreased significantly in all four water conditions. The Hamming ECC codes decreased the BER even further, resulting in a BER lower than 0.07 for all four water conditions. Notably, the tank, pool, and lake situations show a BER approaching 0. The average BER decreases from 0.16 to 0.03 after CycleGAN denoising and Hamming error correction. In summary, the BER in four different water conditions is all low enough with CycleGAN based denoising and Hamming error correction for robust data parsing.

**In different times of the day.** We conduct experiments during both day and night at a distance of 1m in the swimming pool and lake. As shown in Figure 16 (c), the BER in the daytime is lower than in the night for both the pool and lake. Even with a flashlight shining to activate the UOID tag, the current UOID tag only has luminous powder covering the element surface, which is not as bright as in the day time. Moreover, at night, the BER without denoising in the lake is worse than the clean pool, because the emitted light from the UOID tag is too weak to go through more muddy water in the lake. After CycleGAN based denoising and Hamming error correction, the BER in all four settings decreased significantly and is lower than 0.03. The results show that the current U-Star system performs data parsing well with CycleGAN based denoising and Hamming error correction both day and night.

## 8.3 Comparison with Existing Optical Tags

We implement the 3D version of existing Bar/QR codes with the same 21 embedded data bits (101101 110010001 001011) and the same tag size (cube edge: 19cm) as our UOID tag for a fair comparison across various aspects. The data alignment, implemented tags and the comparison experiment scenarios are shown in Figure 17 (a). We conduct experiments and make comparisons in the five aspects below to demonstrate the superiority and necessity of our designed UOID tags over existing optical tags for underwater navigation.

**(1) Same tag order with more embedded bits.** Despite the fact that the user can capture the information of one and up to three surface planes of a 3D version of existing Bar/QR codes in N-order, the decoded bits are the same as the bits in one plane. The embedded bits in an N-order barcode are roughly $N$. The embedded bits in an NxN QR code are roughly $N^2$-4 bits. The embedded bits in an NxNxN UOID tag are $N^3$-6 bits. As shown in Figure 17 (b), the amount of embedded bits in a UOID tag increases exponentially compared to the same order 1D/2D optical tags. Even their 3D versions cannot compare to the UOID tags (e.g., 3-order UOID embeds 7x and 4.2x bits of the same order Bar and QR code).

**(2) Same tag size & data with larger element distance.** The larger the average element distance and the broader the distribution of element distances (shown in the right of Figure 2), the better the detection performance and the less the error bits. We measure distances between all 21 data elements in 3D versions of the Bar/QR, and UOID tag that have the same embedded bits and tag size (edge is 19cm). As shown in Figure 17 (a) and (c), data element distances in Bar and QR codes are all smaller than 20cm, however the data element distances in UOID tags are completely distributed in a greater range of [5, 30] cm.

**(3) Same tag size & data with longer communication range.** We also investigate the goodput performance of three tags mentioned above in two different underwater scenarios: clean creek and

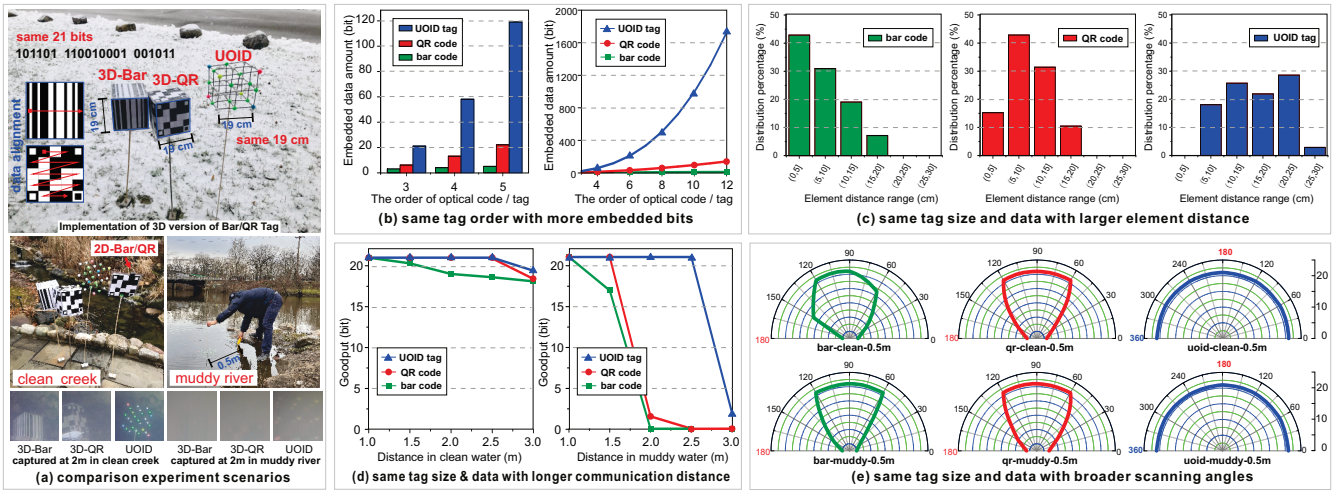



**Figure 17: Comparison between UOID tags with existing optical tags. (a) Experiment scenarios, (b) Data improvement, (c) Broader element distance, (d) Better goodput performance, and (e) Full-directional scanning.**

muddy river at varying distances from 1m to 3m. In clean creek, all three tags perform well and produce more than 17 bits of goodput at up to 3m, as illustrated in Figure 17 (a) and (d). However, in the muddy river, the goodput of Bar and QR codes in the 3D version drops dramatically after 1.5m, whereas the UOID tag maintains its high goodput until 2.5m.

**(4) Same tag size & data with broader scanning angles.** Furthermore, for all three of the aforementioned tags, we evaluate the goodput performance with varying scanning angles at 0.5m under the clean creek and muddy river. As shown in Figure 17 (a) and (e), the usage view range of the existing optical tags has also been increased from less than 120° to 360° of UOID tags for both clean creek and muddy river.

**(5) Other benefits of UOID design.** When compared to the 2D plane (the version of 1D Bar and 2D QR codes in our daily life, shown in the left middle of Figure17 (a)) and confined 3D cube (3D version of Bar/QR) to maintain tag's location and orientation in flowing water or current (i.e., creek, river, tide and etc.), the hollowed-out UOID can lessen influence of water current to allow it to flow through the tags and maintain stabilization.

## 8.4 Case Study with Multiple UOID tags

The usage of our U-Star signage system is similar to barcode/ QR code adopted in auto-supermarket systems. The data embedded in codes are the query codes used for searching a backup database with records for all offered goods. Due to the large enough storage ability on the mobile device, the ability to embed more query codes will result in better navigation. Our 3-order UOID tag can embed $2^{3x3x3-6}$ = 2,097,152 possible query codes. Even with Hamming ECC parity bits that sacrifice 10 (3+4+3=10) bits, there are still 11 data bits available for embedding $2^{11}$ = 2,048 query codes. As shown in Figure 18, we implement four UOID tags with Hamming error correction codes in the case study, and their 11 valid data bits match to distinct query codes in range of [0, 2047] in the backup database. The database stores the current absolute location information, the guidance information, and risk warnings such as "shark near" which can queried via the related query codes. Our demo in a 4m x 10m indoor pool, the user dives at the start site of B and plan to go to the destination site of C and then back.

When the user scans Tag B at the start location, the user will be given the current absolute location (i.e., facing North and at (2m, 0.5m) in the coordinate system) as well as information about its nearby nodes (i.e., D is the nearest tag with 4.5m relative distance to B's NorthEast direction, A is 5.3m away from B to B's EastSouth direction, and C is 9m away from B to B's East direction) to help navigate himself to other spots.

The user intends to visit Tag D first. He looks for a bright dot around 4.5m away (the optical ranging of UOID provides him a sense of underwater distance) at the NorthEast direction of Tag B. If he cannot find his way, he will travel to another nearby node such as Tag A.

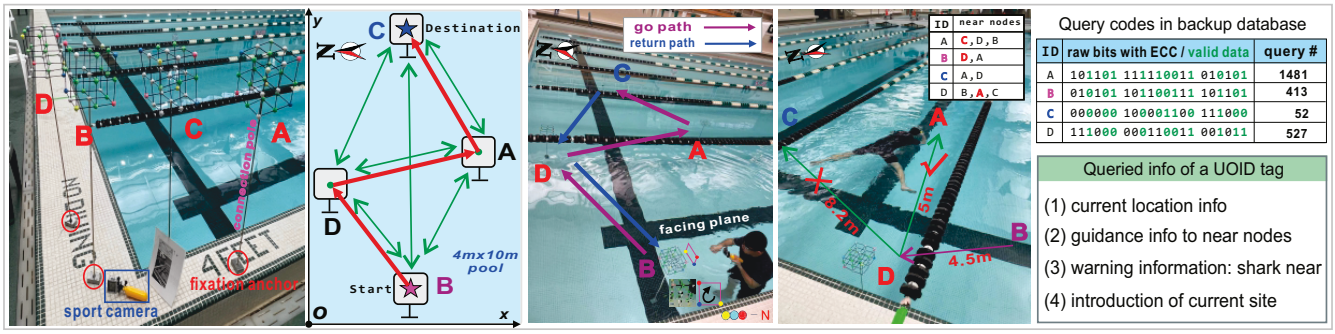



**Figure 18: Underwater navigation case study of U-Star in a 4mx10m indoor pool with 4 UOID tags and backup database.**

After confirmation of D's existence, he moves to Tag D and repeats the similar procedure to go to Tag A first (compared with 8.2m to C, the distance to A is 5m and A is the nearest uncovered node to D). And next, from A, he finally reaches at destination C.

His path (locally optimal) is B-D-A-C and return path is C-D-B (effective path) while globally optimal path C-B may not work due to he may not confirm B's existence from C. By following the procedures above, he achieves self-guided underwater navigation easily and effectively, regardless of the start and destination tags.

## 8.5 Other Concerns

**Cost and price.** As shown in Figure 19, the main cost of the U-Star system is the tag reader, while the UOID tag is very cheap (less than $3 for each). For practicality, the tag reader can be replaced with the user's own smartphones covered with a waterproof case, which is less than $4. Considering multiple UOID tags deployed underwater, the U-Star system with 20 UOID tags costs less than $100 for an underwater site with an area of $1km^2$ ($7m$ x $7m$ x 20).

| Device | Material | Cost ($) |
|---|---|---|
| One 3x3x3 UOID tag | element balls | < 1 |
| | stick / plastic | < 0.5 |
| | hot melt glue | < 0.5 |
| | double-side tap | < 0.5 |
| | luminous powder | < 0.5 |
| | Total for a tag | ≈ 3 |
| One tag reader | sport camera | 30 (basic) |
| | smart phone | self-contained |
| | waterproof case | 3.5 (Amazon) |
| U-Star with 20 tags | | < 100 |

**Figure 19: Cost & price.**

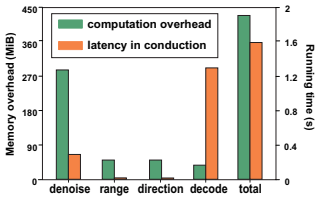



**Figure 20: Overhead.**

**Computation overhead.** For underwater situations, battery is limited and not easy to replace. The tag reader should not conduct complex computations that consume energy too fast. The training processes are offline, the real-time tasks are denoising, optical ranging, orientation guidance, and decoding. As shown in Figure 20, the denoising requires the most memory resources and decoding required the fewest memory resources. For all four tasks, they require a combined 430 MiB of memory and is not a computational burden for a commercial smart device.

**Latency.** For underwater navigation tasks, time can be important to improve the user experience and even save people's lives. Compared with state-of-art underwater navigation systems, including audio-based systems, U-Star has nearly no time delay in signal propagation due to the fast propagation of light. Thus we only consider the computational latency. As shown in Figure 20, optical ranging and orientation guidance have the lowest running time of 0.002 s, while decoding has the longest running time at 1.25 s. All four tasks consume 1.59 s total, which is still quick enough for a good user experience.

# 9 DISCUSSION

**Usage instruction of scanning UOID.** Even with appropriate spacing between data elements in UOID tags, there is some LoS blockage at certain scanning angles. However, by slightly adjusting capturing poses without moving the user's location, it is simple to avoid blockages and capture all data elements.

**The number of guidance directions.** Our current U-Star prototype can provide user orientation guidance in 8 directions, which is sufficient for practical underwater navigation as illustrated in Section 8.4. U-Star, however, may be updated to finer-grained orientation guidance using a same CNN training with more directions (e.g., 16 directions).

**UOID deployment.** Because GPS is unusable for underwater scenarios, the positions of deployed UOID tags are identified and saved in a backup database on shore at a one-time deployment cost. We can use spring installation techniques to fix UOID tags on the underwater floor with little regard for location and orientation fluctuation caused by tide and flow. They can make the tag flexible when subjected to tide power and automatically resume its suspected position when it becomes static, much like how tall building dampers maintain stability and extend tag usage lifetime.

**System robustness and potential side-effect on marine animals.** (1) moss/scum removing: Because moss grows slowly, we can periodically (e.g., every month) remove the accumulated moss and maintain UOID tags as part of underwater infrastructure maintenance. We can utilize an ultrasonic technique to remove moss touchlessly while causing no harm to the UOID tags or other marine life. (2) luminous powder: To prevent pollution and harm to marine life, we apply non-toxic, non-radioactive, and long-lifespan (more than 15 years) luminous powder wrapping with waterproof glues. (3) marine debris: We can use integrated molding technology and 3D printing techniques in the future to produce recycled, solid and not easily damaged UOID tags to avoid marine debris.

**Applications benefited by U-Star.** (1) Recreation scuba diving as illustrated in Section 1. (2) Underwater rescuing. In addition to using fixed UOID tags as infrastructure for safe underwater activities, we can attach smaller size UOID tags (which store people's identifying information) on top of underwater helmets as mobile UOID tags for persons participating in underwater activities. As a result, rescuers can scan UOID tags to identify people and learn about on-site situation (how many people and who are in danger or need rescue). The trapped people, on the other hand, can scan larger UOID tags on rescuers to actively seek help and instructions from rescuers. (3) Future directions combined with Augmented Reality. We can update the tag reader side from current sport camera/smart phone to AR goggles to show guidance info in more direct and visual manner instead of small display on smartphone for user experience of WYSIWYG, "see UOID, see INFO".

# 10 CONCLUSION

In this paper, we implement the U-Star system for simple and robust underwater navigation. We investigate 3D spatial diversity for data embedding with wider element distances and additionally use it for relative positioning. We address challenges in system design and implementation, e.g, combating harsh underwater environments and 3D structure restoration for data parsing. Finally, we conduct experiments based on virtual and real UOID tags in multiple underwater scenarios. Our 3-order UIOD prototype can embed 21 bits and achieves a BER of 0.003 at 1m and less than 0.05 at up to 3 m with approaching 100% relative positioning precision.

# ACKNOWLEDGMENTS

This work was partially supported by the U.S. National Science Foundation under Grants CNS-2226888 and CCF-2007159.

# REFERENCES


[1] 2009. *CIFAR10.* https://paperswithcode.com/sota/image-classification-on-cifar-10

[2] 2009. *CIFAR100.* https://paperswithcode.com/sota/image-classification-on-cifar-100

[3] 2014. *COCO.* https://paperswithcode.com/dataset/coco

[4] Farhad Akhoundi, Amir Minoofar, and Jawad A Salehi. 2017. Underwater positioning system based on cellular underwater wireless optical CDMA networks. In *2017 26th Wireless and Optical Communication Conference (WOCC)*. IEEE, 1–3.

[5] Azzedine Boukerche and Peng Sun. 2020. Design of Algorithms and Protocols for Underwater Acoustic Wireless Sensor Networks. *ACM Computing Surveys (CSUR)* 53, 6 (2020), 1–34.

[6] Charles J. Carver, Zhao Tian, Hongyong Zhang, Kofi M. Odame, Alberto Quattrini Li, and Xia Zhou. 2021. AmphiLight: Direct Air-Water Communication With Laser Light. *GetMobile: Mobile Comp. and Comm.* 24, 3 (Jan. 2021), 26–29. https://doi.org/10.1145/3447853.3447862

[7] Darren Caulfield and Kenneth Dawson-Howe. 2004. Direction of Camera Based on Shadows. In *Proceedings of the Irish Machine Vision and Image Processing Conference*. Citeseer, 216–223.

[8] Jazmine Gaona and Ray Oltion. 2013. Natural Navigation. (2013).

[9] Richard W Hamming. 1950. Error detecting and error correcting codes. *The Bell system technical journal* 29, 2 (1950), 147–160.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[11] Tariq Islam and Seok-Hwan Park. 2020. A comprehensive survey of the recently proposed localization protocols for underwater sensor networks. *IEEE Access* (2020).

[12] Mohammad Jahanbakht, Wei Xiang, Lajos Hanzo, and Mostafa Rahimi Azghadi. 2021. Internet of underwater Things and big marine data analytics—A comprehensive survey. *IEEE Communications Surveys & Tutorials* (2021).

[13] Fahad Jalal and Faizan Nasir. 2021. Underwater Navigation, Localization and Path Planning for Autonomous Vehicles: A Review. In *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*. IEEE, 817–828.

[14] Junsu Jang and Fadel Adib. 2019. Underwater backscatter networking. In *Proceedings of the ACM Special Interest Group on Data Communication*. 187–199.

[15] Ruhul Khalil, Mohammad Babar, Tariqullah Jan, and Nasir Saeed. 2020. Towards the Internet of underwater things: Recent developments and future challenges. *IEEE Consumer Electronics Magazine* (2020).

[16] Chenning Li, Hanqing Guo, Shuai Tong, Xiao Zeng, Zhichao Cao, Mi Zhang, Qiben Yan, Li Xiao, Jiliang Wang, and Yunhao Liu. 2021. NELoRa: Towards Ultra-low SNR LoRa Communication with Neural-enhanced Demodulation. In *Proceedings of ACM SenSys*.

[17] Chi Lin, Yongda Yu, Jie Xiong, Yichuan Zhang, Lei Wang, Guowei Wu, and Zhongxuan Luo. 2021. Shrimp: a robust underwater visible light communication system. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 134–146.

[18] Huaiyin Lu, Ming Jiang, and Julian Cheng. 2020. Deep Learning Aided Robust Joint Channel Classification, Channel Estimation, and Signal Detection for Underwater Optical Communication. *IEEE Transactions on Communications* 69, 4 (2020), 2290–2303.

[19] Chengcai Lv, Binjian Shen, Chuan Tian, Shengzong Zhang, Liang Yu, and Dazhen Xu. 2020. Signal Design and Processing for Underwater Acoustic Positioning and Communication Integrated System. In *2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP)*. IEEE, 89–93.

[20] Nino E Merencilla, Alvin Sarraga Alon, Glenn John O Fernando, Elaine M Cepe, and Dennis C Malunao. 2021. Shark-EYE: A Deep Inference Convolutional Neural Network of Shark Detection for Underwater Diving Surveillance. In *2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*. IEEE, 384–388.

[21] Olga Mirgorodskaya, Olesya Ivanchenko, and Narine Dadayan. 2020. Using Digital Signage Technologies in Retail Marketing Activities. In *Proceedings of the International Scientific Conference-Digital Transformation on Manufacturing, Infrastructure and Service*. 1–7.

[22] András J Molnár. 2020. TRAILSIGNER: A Conceptual Model of Hiking Trail Networks with Consistent Signage Planning and Management. In *Information Modelling and Knowledge Bases XXXII*. IOS Press, 1–25.

[23] Zhang Nan, Zhang Fan, and Enmao Liu. 2020. Design of A Shared Platform for Interactive Public Art from Perspective of Dynamic Vision. In *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 37–42.

[24] Ibrahima N'Doye, Ding Zhang, Mohamed-Slim Alouini, and Taous-Meriem Laleg-Kirati. 2021. Establishing and Maintaining a Reliable Optical Wireless Communication in Underwater Environment. *IEEE Access* 9 (2021), 62519–62531. https://doi.org/10.1109/ACCESS.2021.3073461

[25] Hao Pan, Yi-Chao Chen, Lanqing Yang, Guangtao Xue, Chuang-Wen You, and Xiaoyu Ji. 2019. mQRCode: Secure QR Code Using Nonlinearity of Spatial Frequency in Light. In *The 25th Annual International Conference on Mobile Computing and Networking*. 1–18.

[26] Market Reports. 2021. *Globle Scube Diving Equipment Industry Research Report, Growth Trends and Competitive Analysis 2021-2027.* https://www.marketreportsworld.com/global-scuba-diving-equipment-industry-18271751

[27] Aleksandr Rodionov, Petr Unru, and Aleksandr Golov. 2020. Long-Range Underwater Acoustic Navigation and Communication System. In *2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*. IEEE, 60–63.

[28] Nasir Saeed, Abdulkadir Celik, Tareq Y Al-Naffouri, and Mohamed-Slim Alouini. 2019. Underwater optical wireless communications, networking, and localization: A survey. *Ad Hoc Networks* 94 (2019), 101935.

[29] Giuseppe Schirripa Spagnolo, Lorenzo Cozzella, and Fabio Leccese. 2020. Underwater optical wireless communications: Overview. *Sensors* 20, 8 (2020), 2261.

[30] Truman R Strobridge. 1974. *Chronology of Aids to Navigation and the Old Lighthouse Service, 1716-1939.* Public Affairs Division, United States Coast Guard.

[31] Witold Szymański and Maurycy Kin. 2019. The perspective transformation in illusionistic ceiling painting of late Baroque. *Teka Komisji Architektury, Urbanistyki i Studiów Krajobrazowych* 15, 1 (2019), 104–112.

[32] Sumit Tiwari. 2016. An introduction to QR code technology. In *2016 international conference on information technology (ICIT)*. IEEE, 39–44.

[33] WiKi. 2021. *Barcode.* https://en.wikipedia.org/wiki/Barcode#Matrix_(2D)_barcodes

[34] WiKi. 2021. *Hamming code.* https:en.wikipedia.orgwikiHamming_code

[35] Norman J Woodland and Silver Bernard. 1952. Classifying apparatus and method. US Patent 2,612,994.

[36] Yue Wu, Purui Wang, Kenuo Xu, Lilei Feng, and Chenren Xu. 2020. Turboboosting Visible Light Backscatter Communication. In *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication* (Virtual Event, USA) *(SIGCOMM '20)*. Association for Computing Machinery, New York, NY, USA, 186–197. https://doi.org/10.1145/3387514.3406229

[37] Yue Yang, Yang Xiao, and Tieshan Li. 2021. A Survey of Autonomous Underwater Vehicle Formation: Performance, Formation Control, and Communication Capability. *IEEE Communications Surveys & Tutorials* 23, 2 (2021), 815–841.

[38] Zhice Yang, Zeyu WANG, Jiansong Zhang, Chenyu Huang, and Qian Zhang. 2019. Polarization-Based Visible Light Positioning. *IEEE Transactions on Mobile Computing* 18, 3 (2019), 715–727. https://doi.org/10.1109/TMC.2018.2838150

[39] Zhaoquan Zeng, Shu Fu, Huihui Zhang, Yuhan Dong, and Julian Cheng. 2016. A survey of underwater optical wireless communications. *IEEE communications surveys & tutorials* 19, 1 (2016), 204–238.

[40] Bo Zhang and Hoi Dick Ng. 2016. An experimental investigation of the explosion characteristics of dimethyl ether-air mixtures. *Energy* 107 (2016), 1–8.

[41] Kai Zhang, Yi Zhao, Chenshu Wu, Chaofan Yang, Kehong Huang, Chunyi Peng, Yunhao Liu, and Zheng Yang. 2019. Chromacode: A fully imperceptible screen-camera communication system. *IEEE Transactions on Mobile Computing* (2019).

[42] Weidong Zhang, Lili Dong, Xipeng Pan, Peiyu Zou, Li Qin, and Wenhai Xu. 2019. A survey of restoration and enhancement for underwater images. *IEEE Access* 7 (2019), 182259–182279.

[43] Xiao Zhang and Li Xiao. 2020. Effective Subcarrier Pairing for Hybrid Delivery in Relay Networks. In *2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE, 238–246.

[44] Xiao Zhang and Li Xiao. 2020. Lighting Extra Data via OWC Dimming. In *Proceedings of the Student Workshop*. 29–30.

[45] Xiao Zhang and Li Xiao. 2020. RainbowRow: Fast Optical Camera Communication. In *2020 IEEE 28th International Conference on Network Protocols (ICNP)*. IEEE, 1–6.

[46] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*. 2223–2232.

[47] Shilin Zhu, Chi Zhang, and Xinyu Zhang. 2017. Automating visual privacy protection using a smart led. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. 329–342.