Novel Compression Algorithm Based on Sparse Sampling of 3-D Laser Range Scans

OĞUZCAN DOBRUCALI AND BILLUR BARSHAN*

Department of Electrical and Electronics Engineering, Bilkent University, TR-06800 Bilkent, Ankara, Turkey
*Corresponding author: billur@ee.bilkent.edu.tr

Three-dimensional models of environments can be very useful and are commonly employed in areas such as robotics, art and architecture, facility management, water management, environmental/industrial/urban planning and documentation. A 3-D model is typically composed of a large number of measurements. When 3-D models of environments need to be transmitted or stored, they should be compressed efficiently to use the capacity of the communication channel or the storage medium effectively. We propose a novel compression technique based on compressive sampling applied to sparse representations of 3-D laser range measurements. The main issue here is finding highly sparse representations of the range measurements, since they do not have such representations in common domains, such as the frequency domain. To solve this problem, we develop a new algorithm to generate sparse innovations between consecutive range measurements acquired while the sensor moves. We compare the sparsity of our innovations with others generated by estimation and filtering. Furthermore, we compare the compression performance of our lossy compression method with widely used lossless and lossy compression techniques. The proposed method offers a small compression ratio and provides a reasonable compromise between the reconstruction error and processing time.

1. INTRODUCTION

Many techniques have been developed to build 3-D models of indoor and outdoor environments. Three-dimensional modeling techniques allow us to describe environments including objects with indefinite shapes or patterns, although these techniques can be complex and computationally expensive [1]. The main advantage of using 3-D models of environments is that such models are more descriptive and have richer information content than 2-D models in terms of the features extracted from the environments, resulting in less ambiguity in distinguishing features [2]. Three-dimensional models are used in a wide range of fields such as robot motion planning and navigation [1, 3–5], art and architecture [6–10], environmental/industrial/urban planning, water management and forestry documentation [11–13]. These models can be obtained using a variety of sensors measuring range or intensity. A common approach in constructing these models is to employ laser range finders (LRFs) that measure the range between the sensor and the objects along the path of the beam emitted by the sensor. These sensors can supply range measurements within their field of view, as the laser beam is rotated by the sensor. There are several approaches to obtaining 3-D models with LRFs; the first uses a conventional 3-D laser scanner. However, since these products are very expensive, this approach is not frequently employed. Another approach is acquiring 3-D range measurements by translating a 2-D LRF that horizontally or vertically scans a field of view of 180°. A third alternative is to acquire the 3-D range information by rotating the 2-D LRF around a fixed axis [4]. In the latter two, multiple 2-D LRFs can be employed, where each sensor scans either the horizontal or the vertical axis [4]. On the market today, various LRFs, suitable for both indoor and outdoor applications, are provided by several companies such as SICK [14], RIEGL [15], FARO [16], Zoller + Fröhlich [17], Leica [18], MENSI [19] and Velodyne [20].
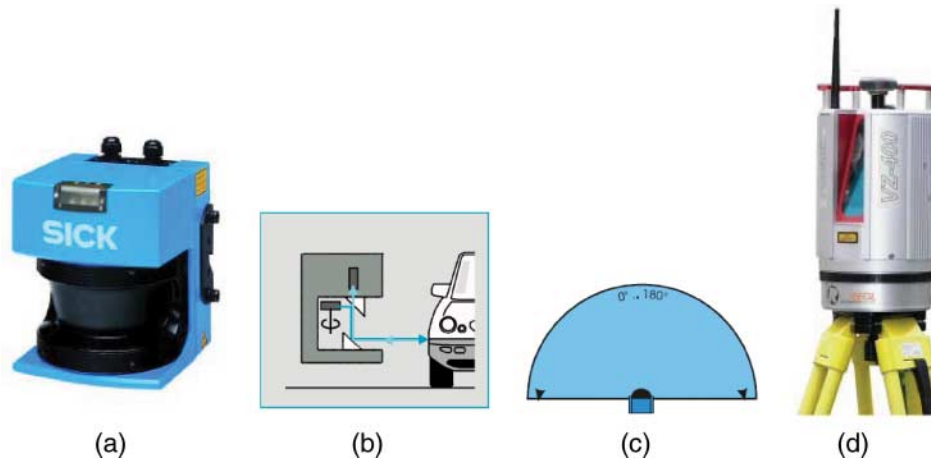
FIGURE 1. (**a**) The front view of the SICK LMS200 LRF, (**b**) its measurement principle, (**c**) its field of view (reprinted from [36]) and (**d**) the front view of the RIEGL VZ-400 LRF.

As an example of the use of 3-D models in the field of robotics, Brenneke *et al.* [1] proposed a technique for simultaneous localization and mapping (SLAM) in outdoor environments. They applied the existing 2-D mapping algorithms to one horizontal layer of a 3-D model. Maps are also obtained in other ways and used in 3-D SLAM applications [21–23]. To build 2-D or 3-D maps from sequentially acquired scans, an iterative closest point (ICP) algorithm is employed, integrated with odometry measurements [4]. This ICP algorithm is not just used in the registration of scans of planar surfaces, but also for curves and non-planar surfaces, as in [24]. In addition to ICP, *semantic information* of the range measurements, which is the gradient between the neighboring measurements, is used for the same purpose [5]. Apart from deterministic methods for the registration of 3-D objects, parametric methods such as *expectation-maximization* [2] and *maximum-likelihood estimation* [25] are employed. Moreover, non-parametric methods, such as the *k-means clustering* algorithm [26], are also used for the same purpose. Devices such as panoramic cameras can be used in conjunction with laser scanners [27]. Besides the techniques for modeling indoor environments, terrains are modeled using airborne laser scanners as in [28, 29], especially for navigation of land vehicles. Apart from modeling environments above the sea level, 3-D models of the sea floor can also be acquired, using autonomous underwater vehicles equipped with a camera, sonar and oceanographic sensors [30]. Building 3-D urban models is another area of application. In summary, many techniques for acquiring and processing 3-D measurements have been developed in a variety of fields, using sensors such as LRFs and cameras as in [31–33]. New techniques are continuously being introduced such as those proposed in [34, 35].

In this study, we consider both indoor and outdoor environments scanned in 3-D with an LRF. Scans of indoor environments have been acquired with the SICK LMS200 illustrated in Fig. 1a. This 2-D device measures the range between itself and the objects within its field of view, based on the *time-of-flight* principle. The sweeping laser beam is aligned by the rotating mirror, as illustrated in Fig. 1b. The laser has a maximum range of 80 m, field of view of 180° (Fig. 1c), range resolution as low as 1 mm and a selectable angular resolution of either 0.25°, 0.5° or 1°. The measurements have a systematic error of ±4 cm, as well as some statistical error that changes with the measurement range, the ambient temperature, illumination and the reflectivity of the objects in the environment. The acquision frequency of the 2-D scans is 75 Hz [36]. In this study, scans of outdoor environments, acquired with the RIEGL VZ-400 illustrated in Fig. 1d, are also used. The measurement principle of this device is similar to that of SICK LMS200. This device is dedicated to outdoor modeling with 100° vertical and 360° horizontal scanning field of view. The angular resolution is at least 0.0024° in both horizontal and vertical directions. The laser has a maximum range of 350 m in the high-speed mode in which 60° is scanned horizontally in one second [37].

The advantages of using a laser beam is reliable detection of the presence of an object and independence of the measurements from the amount of ambient light and the color of the objects. A major disadvantage is that, for proper operation of the sensor, the environment should not contain highly reflective or transparent materials, such as glass. LRFs are used in various tasks such as estimating volumes and positions of objects and their classification, collision prevention for vehicles and surveillance.

The scan data collected by any sensor usually need to be transmitted to a station where the data are processed and analyzed. The scan data of a 3-D model, likely to be composed of hundreds of thousands of range measurements, also need to be stored in a medium, where the amount of allocated memory is required to be as small as possible. Thus, the data must be recorded to the medium efficiently in terms of the memory

space, as must the speed in reading and writing operations for the data. This way, fast and accurate autonomous search and scan systems can be developed.

To satisfy all of the requirements mentioned above, the scan data must be compressed before it is transmitted or stored. By decreasing the size of the data, the amount of data stored can be increased and the elapsed time required to transmit the data through the communication channel can be reduced. Although many compression techniques have been developed for various types of data (see, for instance [38–41]), determining the optimum data compression technique with respect to the following criteria is still an open research field.

An important aspect of data compression is the compression ratio (CR), which is the ratio of the size of the compressed output to the size of the original data. The CR is between zero and one (or 0 and 100%) for a compression operation, and greater than one for an expansion operation. The closer the CR is to zero, the larger is the amount of compression [42].

Salomon [42] points out that no data compression method is perfect; thus compressing any number of bits into one bit, which may be a fictional case, is such a success that even compressing two bits into one bit can be considered as 'perfect'. Therefore, a compression method can be considered efficient when the size of the original data is reduced by more than one half. In other words, an efficient compression method, at least halves the storage and communication costs [43].

The amount of *distortion* is the second aspect in data compression [44]. The size of the data is lowered by employing either *lossless* compression techniques, in which the whole information in the data is encoded, or *lossy* compression techniques, in which the essential part of the information is encoded. Although distortion, which is the difference between the data and its reconstruction using the compressed data, is observed in lossy compression, lossy compression methods are usually preferred since they result in lower CRs than lossless compression methods. Distortion can be measured in various ways depending on the type of data, and is required to be as low as possible. Let $\mathbf{x} = \{x_i\}_{i=1}^{N}$ and $\hat{\mathbf{x}} = \{\hat{x}_i\}_{i=1}^{N}$ represent the data sequence and its reconstruction, respectively. A common measure of distortion, which is also used in this study, is the root mean squared error (RMSE) between $\mathbf{x}$ and $\hat{\mathbf{x}}$, calculated as $\sqrt{(1/N) \sum_{i=1}^{N} (x_i - \hat{x}_i)^2}$.

Speed is another aspect in data compression. It is a measure of how fast the data are compressed (encoding speed) and reconstructed from the compressed data (decoding speed) by using a given compression technique. Speed is inversely proportional to the time required for encoding and decoding the data, and required to be as high as possible.

In the literature, there are various compression techniques dedicated to 3-D range measurements. Some of them are aimed to fit geometrical structures expressed with a few parameters explicitly to the separate clusters of range measurements. For instance, Obaid fits line segments to measurements after arranging them by applying Peano scanning to the whole 3-D scan [45]. Kaushik *et al.* [46] fit polygons to planar surfaces detected in the 3-D scan in an iterative manner. Rivest and Siddiqi [47] attempt to fit rectangles to the consecutive measurements having values close to the average of the measurements inside the same rectangle. In these algorithms, the parameters of the geometrical structures fitted to the measurements are encoded using a lossless compression technique. Furthermore, some well-known algorithms generally employed to encode different types of data are also considered in compressing 3-D scans. For instance, Differential Pulse Code Modulation, widely used in compressing speech signals, is employed in [47]. In a study closely related to our work [48], the authors focus on real-time compression of laser data acquired on board a mobile robot platform. A 3-D scan is handled as a gray-scale image data and different lossless and lossy compression techniques are compared in terms of the transmission time of laser-generated 3-D maps. It is shown that JPEG-LS compression performs the best with a CR of about 19.4% and a maximum RMSE of 4 mm.

In this paper, we propose an effective compression method that can be applied to 3-D laser range measurements as the data are being acquired. The main contribution of this study is to provide a method to generate sparse representations of laser range measurement sequences. These representations include an incredibly small number of nonzero values compared with the number of measurements in the original sequences. Then the sparse representations are compressed by applying sparse sampling techniques that have been applied in sampling parametric signals [49], and are based on *compressive sensing*. The proposed method is similar to *difference encoding* and is a causal system because it generates sparse representations based on current and previous measurements. Therefore, in theory, it can compress even an infinite number of range measurement sequences.

The rest of this paper is organized as follows: Compressive sensing is reviewed in Section 2. The proposed method is described in detail in Section 3 and compared with widely used compression techniques in terms of the CR, distortion and speed in Section 4. Three experimental datasets, independently acquired at different institutions with two different LRF brands, are used for this purpose. Conclusions and directions for future work are provided in the last section.

2. BACKGROUND ON COMPRESSIVE SENSING

Compressive sensing enables signals to be successfully reconstructed with fewer samples than the Shannon/Nyquist sampling theorem requires, by using a linear sampling model with an optimization procedure for reconstructing the original signal [50]. To achieve this, compressive sensing relies on *sparsity* and *incoherence* properties. The sparsity property requires signals to have sparse representations in proper

domains. Sparse signals can be represented with a lower sampling frequency than the Nyquist rate. Furthermore, sparsity enables discrete-time signals to be represented more briefly when expressed using a proper basis $\mathbf{\Psi}$. The incoherence property states that the sparse representation of the signal on basis $\mathbf{\Psi}$ must be extended in the domain in which the signal is sampled [51].

The first step in compressive sensing is to represent the signal using a proper basis, i.e. one on which the representation is sparse. The basis should contain a set of orthonormal vectors that form a set of waveforms such as a wavelet basis [51]. Let $\mathbf{x} = [x_1, \ldots, x_N]^T$ be the column vector that represents the N samples of the signal in $\Re^N$, where N is a large integer; $\mathbf{\Psi} = [\Psi_1, \ldots, \Psi_N]$ stands for the basis matrix with orthonormal basis vectors $\{\Psi_i\}_{i=1}^N$. Here, it is assumed that the basis vectors are column vectors in $\Re^N$ so that $\mathbf{\Psi}$ is an $N \times N$ matrix. Thus, we can represent the signal as $\mathbf{x} = \sum_{i=1}^N s_i \Psi_i = \mathbf{\Psi s}$, where $\mathbf{s} = [s_1, \ldots, s_N]^T$, $s_i = \langle \mathbf{x}, \Psi_i \rangle$ with $i = \{1, \ldots, N\}$ [50] and $\langle \cdot, \cdot \rangle$ denotes the *inner product* of two vectors. Note that $\mathbf{x}$ and $\mathbf{s}$ are different representations of the same signal in different domains: the time domain and the $\mathbf{\Psi}$ domain, respectively. If the projection of the signal onto the basis $\mathbf{\Psi}$ is sparse, only a small number of coefficients in $\mathbf{s}$, denoted by K, will have large values, whereas the majority, denoted by $(N - K)$, will be close to zero. When $K \ll N$, $\mathbf{s}$ is referred to as K-*sparse*. The sparsity property defined here is motivated by the assumption that most signals are *compressible* with the proper choice of a basis $\mathbf{\Psi}$. The approximation of signals with K-sparse representations is the basis of *transform coding* [50].

After the first step, the signal is sampled using a linear measurement model that computes M measurements, where $M \ll N$. We assume that the measurement model $\mathbf{\Phi} = [\Phi_1^T, \ldots, \Phi_M^T]^T$ is an $M \times N$ matrix, and is composed of basis vectors $\{\Phi_j\}_{j=1}^M$, each of which is a column vector in $\Re^N$. Let the measurement vector be denoted as $\mathbf{y} = [y_1, \ldots, y_M]^T$ composed of $\{y_j\}_{j=1}^M$, where $y_j = \langle \mathbf{x}, \Phi_j \rangle$. Thus, the measurement vector can be defined as $\mathbf{y} = \mathbf{\Phi x} = \mathbf{\Phi \Psi s} = \mathbf{\Theta s}$, which has fewer dimensions than the original signal, referred to as the *undersampled* case [51]. The objective of compressive sensing can be briefly summarized as determining a measurement model $\mathbf{\Phi}$ and a sparsifying basis $\mathbf{\Psi}$ that allow the reconstruction of the signal $\mathbf{x}$, which is not damaged despite the reduction in dimensionality. More briefly, the objective of compressive sensing is determining $\mathbf{\Theta}$.

The solution to the determination of $\mathbf{\Theta}$ must satisfy two important properties: *Restricted Isometric Property* (RIP) and *incoherence*. RIP requires that ζ, a constant between 0 and 1, should be close to 0 by the following statement:

$$(1 - \zeta)\|\mathbf{x}\|_2^2 \leq \|\mathbf{\Theta x}\|_2^2 \leq (1 + \zeta)\|\mathbf{x}\|_2^2, \qquad (1)$$

where $\| \cdot \|_2$ is the two-norm of the corresponding vector. The above statement expresses that any vector multiplied by $\mathbf{\Theta}$ cannot be in the null space of $\mathbf{\Theta}$, and so $\mathbf{\Theta}$ must

preserve the length of the vectors multiplied by itself. The second requirement for $\mathbf{\Theta}$ is the incoherence property, which indicates uncorrelatedness between the sparsifying basis $\mathbf{\Psi}$ and the measurement model $\mathbf{\Phi}$ [51]. Incoherence states that basis vectors in the measurement model cannot *sparsely* represent the basis vectors in the sparsifying basis [50]. Coherence (i.e. the opposite of incoherence) between $\mathbf{\Phi}$ and $\mathbf{\Psi}$ can be referred to as a measurable quantity, computed by

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \sqrt{N} \max_{1 \leq i \leq N, \, 1 \leq j \leq M} |\langle \Psi_i, \Phi_j \rangle|, \qquad (2)$$

where μ indicates coherence, varying between 1 and $\sqrt{N}$ [51]. Low levels of coherence are always preferable for constructing $\mathbf{\Theta}$, so that we have maximal incoherence when μ is 1.

One remaining issue in the design of the compressive sensing structure is determining a lower bound for M, which is the number of measurements obtained by the measurement model. Since the dimension of the sampled signal N and the number of nonzero entries in the sparse representation K are both known, the minimum value of M can be computed from either

$$M \geq c_1 \, K \ln\left(\frac{N}{K}\right) \text{ [50]}, \qquad (3)$$

$$\text{or } M \geq c_2 \, \mu^2(\mathbf{\Phi}, \mathbf{\Psi}) \, K \ln(N) \text{ [51]}, \qquad (4)$$

where c_1 and c_2 are small positive constants. No information is lost after sampling as soon as a set of M samples that satisfies Equation (3) or (4) is acquired in the measurement model. It is seen that fewer samples in the measurement model are claimed to be sufficient as the coherence decreases.

After a measurement vector $\mathbf{y}$, which has far smaller dimensions than the original signal $\mathbf{x}$, is obtained, the next step is the reconstruction of the original signal and its sparse representation $\mathbf{s}$ from the measurement vector. At the end of the sampling process, we have $\mathbf{y} = \mathbf{\Theta s}$, where $\mathbf{s}$ is to be estimated given $\mathbf{y}$ and $\mathbf{\Theta}$. Since $\mathbf{\Theta}$ is an $M \times N$ matrix with $M \ll N$, there are infinitely many $\tilde{\mathbf{s}}$ that satisfy $\mathbf{y} = \mathbf{\Theta}\tilde{\mathbf{s}}$. The optimal solution to $\mathbf{s}$ is stated as

$$\hat{\mathbf{s}} = \arg \min \|\tilde{\mathbf{s}}\|_1 \quad \text{such that } \mathbf{y} = \mathbf{\Theta}\tilde{\mathbf{s}}, \qquad (5)$$

where $\| \cdot \|_1$ is the one-norm of the corresponding vector. Furthermore, if the sparse representation is reconstructed from noisy measurements, the following optimization can be considered:

$$\hat{\mathbf{s}} = \arg \min \|\tilde{\mathbf{s}}\|_1 \quad \text{such that } \|\mathbf{y} - \mathbf{\Theta}\tilde{\mathbf{s}}\|_2 \leq \rho, \qquad (6)$$

where ρ is the bound on the noise in the measurement vector $\mathbf{y}$ [52, 53]. Apart from the one-norm solution, a two-norm solution is available in *regularized minimization* for reconstruction [54]. In this case, $\hat{\mathbf{s}} = \arg \min \|\mathbf{y} - \mathbf{\Theta}\tilde{\mathbf{s}}\|_2^2 + c_0\|\tilde{\mathbf{s}}\|_1$, where c_0 is a small positive constant. One way to solve the given optimization problems is to apply *basis pursuit* algorithms [55]. As soon as the sparse representation of the

signal **s** is estimated as **ŝ**, the original signal **x** is reconstructed as $\hat{\mathbf{x}} = \mathbf{\Psi}\hat{\mathbf{s}}$, with a small distortion between **x** and **x̂**. In this study, we use the solution given by Equation (5) because the measurement model **Φ** employed here provides a noiseless measurement vector **y**.

2.1. Determining the sparsifying basis

As stated above, the first step in compressive sensing is to determine the best sparsifying basis **Ψ** for efficient representation of the original signal **x**. Thus, the projection of **x** onto this basis should represent **x** with fewer values than **x** has, and allow the reconstruction of **x** with small error. Any sparsifying basis is composed of a set of basis vectors that are actually waveforms. In the literature, these waveforms are called *atoms*, and the set of atoms that comprises the sparsifying basis is called a *dictionary*. Although there are some readily available dictionaries, such as wavelet and cosine packets, Gabor and Fourier dictionaries, chirplets and warplets, dictionaries can also be designed and tailored according to the signal features. In this study, Fourier, Gabor and Haar dictionaries [55] are tested with the experimental scan data to acquire sufficiently sparse representations.

2.2. Determining the measurement model

An interesting application of compressive sensing is the *single pixel camera*, which is used to sample sparse images, as reported in [56]. In the implementation of sampling, light is projected onto a digital micromirror device having an array of N micromirrors, as shown in Fig. 2. According to the measurement model used, a random number generator selects a set of micromirrors to focus the reflected light onto a photodiode. The measurement model is generated in three different ways: *raster scan*, *basis scan* and *compressive sampling* (CS). As a result, different combinations of M pixels out of N are measured by the photodiode. In the raster scan, the photodiode measures N pixels one at a time (i.e. $M = N$), where **Φ** is the $N \times N$ identity matrix. In the basis scan, the photodiode measures M pixels, which are determined according to the Walsh basis, one at a time. In this model, **Φ** is the $N \times N$ Walsh matrix that includes binary

coefficients [57]. In CS, the photodiode measures M different linear combinations of N pixels, using random test functions. It is shown that the smallest distortion of images occurs with the smallest number of measurements, which is achieved when CS is used as the measurement model.

In this study, to construct a measurement model using random test functions, the number of basis vectors in the model is determined by taking $c_1 = 1$ in Equation (3) such that

$$M = \left\lceil K \ln\left(\frac{N}{K}\right) \right\rceil, \tag{7}$$

where $\lceil \cdot \rceil$ denotes the *ceiling* function. All elements in the model matrix are selected independently from a Gaussian distribution with zero mean and $1/N$ variance to obtain a model incoherent with any sparsifying basis with high probability [50]. Then the row vectors in **Φ** are orthonormalized by the Gram–Schmidt process. Using this measurement model without a sparsifying basis, where $\mathbf{\Theta} = \mathbf{\Phi}$ and **Ψ** is an $N \times N$ identity matrix, is advantageous in reconstructing the original signal because RIP is satisfied, since **Φ** has no null space with high probability, where ζ is zero in Equation (1). Moreover, the incoherence property is satisfied, where $\mu(\mathbf{\Phi}, \mathbf{\Psi})$ is found likely to be around $\sqrt{2 \ln N}$.

3. THE PROPOSED METHOD

We can use compressive sensing to compress any signal using a suitable sparsifying basis and an incoherent measurement model. This approach is commonly applied in various fields, such as magnetic resonance imaging in medicine [58], interferometric imaging in astronomy [59] and advanced techniques in image processing [60]. Although forming the measurement model is a straightforward process, forming the sparsifying basis is a more challenging problem. The signals considered in this study are range measurement sequences taken within the sensor's field of view, as column vectors in $\mathfrak{R}^N$, where N can be very large. The main objective in this problem is to find a representation of the signal that contains sufficiently sparse critical information to recover the signal with small error [52].
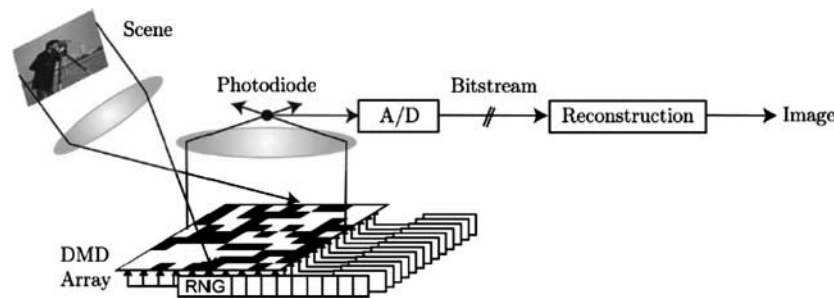

FIGURE 2. The operation scheme of the single pixel camera (reprinted from [50]).

In this section, we present the proposed method using one of the benchmarking datasets described in more detail in Section 4. The dataset used here is composed of 29 3-D scans collected at different locations in the University of Osnabrück's AVZ building in Osnabrück, Germany [61]. Each 3-D scan in the dataset is acquired by collecting 2-D scans from a sensor rotated in small steps around a horizontal axis above the ground level. Each 2-D scan is obtained as the laser beam emitted by the sensor is swept within the sensor's field of view in 0.5° intervals. A 3-D scan in this set is composed of 471 2-D scans that are vectors of 361 consecutive range measurements (i.e. $N = 361$). Different features such as a mannequin, a human being, banisters at the top of stairs and chairs are observed while they are stationary in these 3-D scans, as illustrated in Fig. 3a–d. Gray levels in these images are directly proportional to the range measurements such that white indicates the maximum and black indicates the minimum range measurement.

To apply the sampling model described in Section 2, we first consider the projections of a 3-D scan, illustrated in Fig. 3b, onto some of the well-known sparsifying bases. The 2-D scans forming the 3-D scan are projected one at a time onto $N \times N$ sparsifying bases formed by using the Fourier, Gabor and Haar dictionaries as follows:

(i) The Fourier dictionary is formed by N cosine waveforms with frequencies $\omega = l\pi/N$, where $l = \{\frac{1}{2}, \frac{3}{2}, \ldots, N - \frac{1}{2}\}$.

(ii) The Gabor dictionary is formed by N waveforms with no delay, the unit standard deviation of the Gaussian envelope of the waveforms and different frequencies uniformly selected from $[0, \pi)$.

(iii) The Haar dictionary is formed by N wavelets with $\frac{1}{32}$ dilation and $l/32$ translation for $l = \{0, 1, \ldots, N - 1\}$.

In the projections, the average percentages of the number of nonzero values to the total number of values are 74.7, 61.3 and 88.7%, respectively, indicating that the projections onto the bases described above are not sufficiently sparse. Thus, both the CR and distortion would be high if CS were applied to these projections [52].

During the process of data acquisition, the sequences of range measurements are extracted and stored in arrays where each value is stored after being converted from binary to decimal format. 2-D scans acquired consecutively have similarities as well as differences. The differences may be caused by changes taking place in a dynamic environment, as well as by the translational or rotational motion of the sensor; at each step of the sensor, a different cross section of the 3-D environment is being observed. Since we observe that the raw 2-D scans do not have highly sparse representations in the domains listed above, we attempt to represent them with sparse innovations exploiting the correlation between two consecutively acquired scans, when the sensor is rotated by a small amount before acquiring the next scan. Thus, we define the innovations between:

(a) two consecutive scans;

(b) each scan and its estimate using *linear regression* [62] based on the last two scans;

(c) each scan and its estimate using *second-order polynomial fitting* [62] based on the last three scans;

(d) each scan and its estimate adding the previous scan to a difference estimate using a *second-order Wiener filter* [62] under the assumption that the differences between consecutive scans form a stationary random sequence;

(e) each scan and its estimate adding the previous scan to a difference estimate using a 1-D random walk on the previous difference; and

(f) each scan and its estimate using a linear *Kalman filter* with the constant velocity kinematic state model [63], which is also called a polynomial filter, because the mesh points in consecutive scans form piecewise polynomial functions.

The implementation details of methods (a)–(f) are provided in [64]. The average percentages of the number of nonzero values to the total number of values in the innovations are 43.7, 92.5, 99.5, 71.5, 27.3 and 50.9%, respectively, for the 3-D scan in Fig. 3b. On the average, we obtain the most sparse innovations in (e), with 27.3%, but even this is not found to be sufficient.

In this study, we propose a method to generate much more sparse innovations (6.5%). The proposed method is composed of *encoder* and *decoder* parts, where the encoder consists of sparsifying, measurement and reconstruction stages and the decoder involves only the reconstruction stage, as depicted in Fig. 4. The sparsifying model generates sparse innovations for each scan in the sparsifying stage, and the measurement model samples the innovations with the minimum number of samples in the measurement stage. Finally, the reconstruction model rebuilds each scan from the samples encoded by the measurement model in the reconstruction stage. The following subsections provide more details on these three models.

3.1. The sparsifying model

In the sparsifying model, we generate the innovations between consecutive 2-D scans as follows: suppose $\mathbf{r}_n$ is the nth 2-D scan that is currently acquired, and $\mathbf{r}_{n-1}$ is the previous one. First, $\mathbf{r}_{n-1}$ is generated at the encoder by employing the reconstruction procedure in Section 3.3 that the decoder follows, to adapt the sparsifying parameters according to the reconstruction at the decoder. Then $\mathbf{r}_{n-1}$ is approximated to $\mathbf{r}_n$ by shifting $\mathbf{r}_{n-1}$ along the vertical and horizontal axes by amplitude (ϵ) and phase (δ) shifts, respectively. An example illustrating ϵ and δ is shown in Fig. 5a.

Assume that the individual range measurements in $\mathbf{r}_n$ and $\mathbf{r}_{n-1}$ are denoted by $r_n[i]$ and $r_{n-1}[i]$, respectively, where $i = 1, 2, \ldots, N$. We define an error function $\mathcal{E}^2 = \sum_{i=1}^{N} [r_n[i] - (r_{n-1}[i + \delta] + \epsilon)]^2$ and set its partial
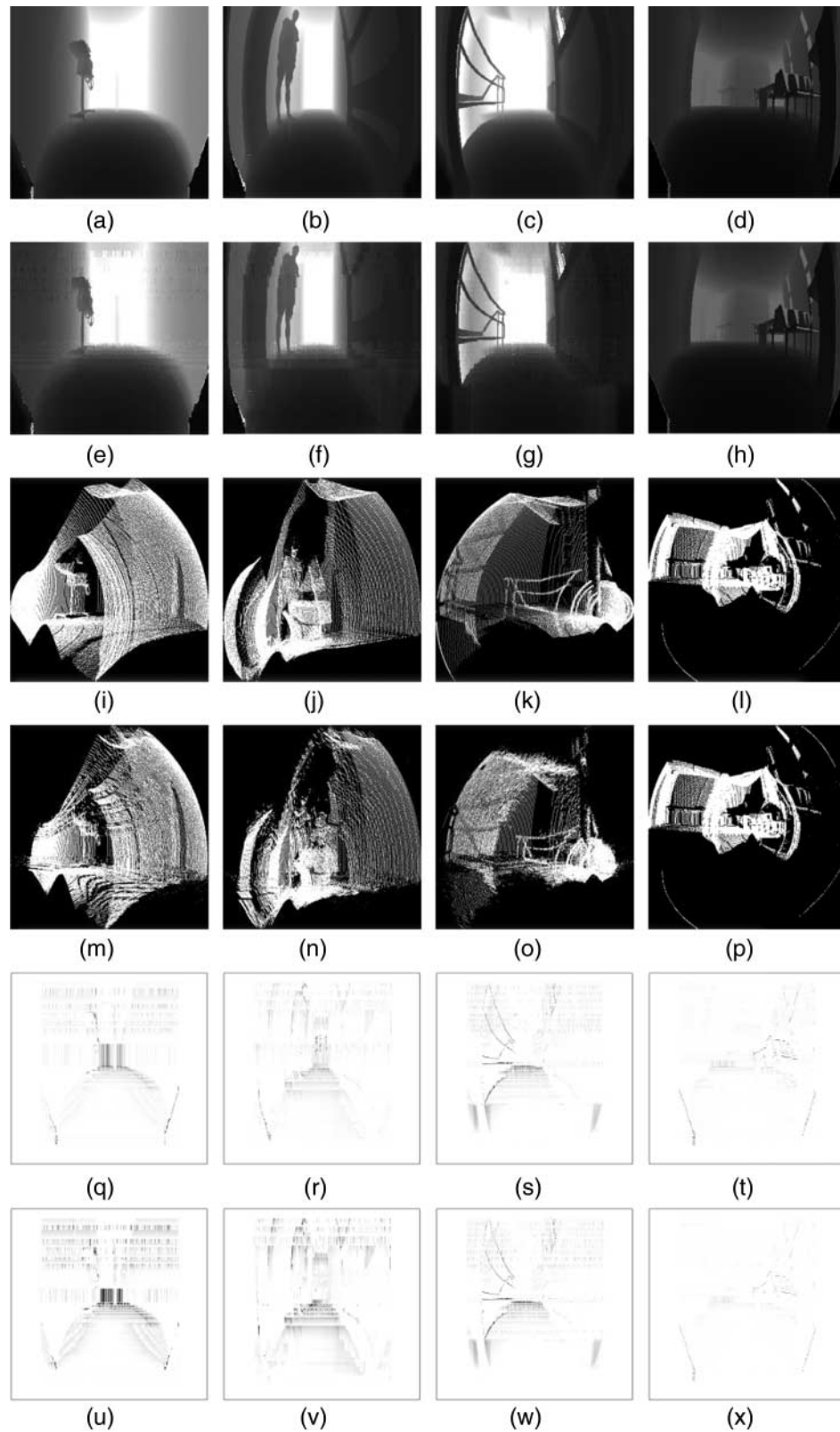
FIGURE 3. (**a**)–(**d**) Sample 3-D scans collected at the University of Osnabrück AVZ building using SICK LMS200; (**e**)–(**h**) their reconstructions; (**i**)–(**l**) point cloud representations of (a)–(d); (**m**)–(**p**) point cloud representations of (e)–(h); (**q**)–(**t**) difference sequences of (a)–(d); and (**u**)–(**x**) the resulting distortion error.
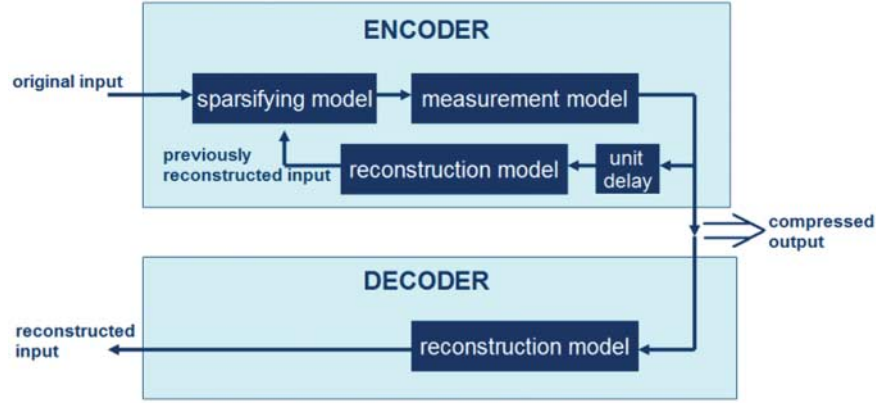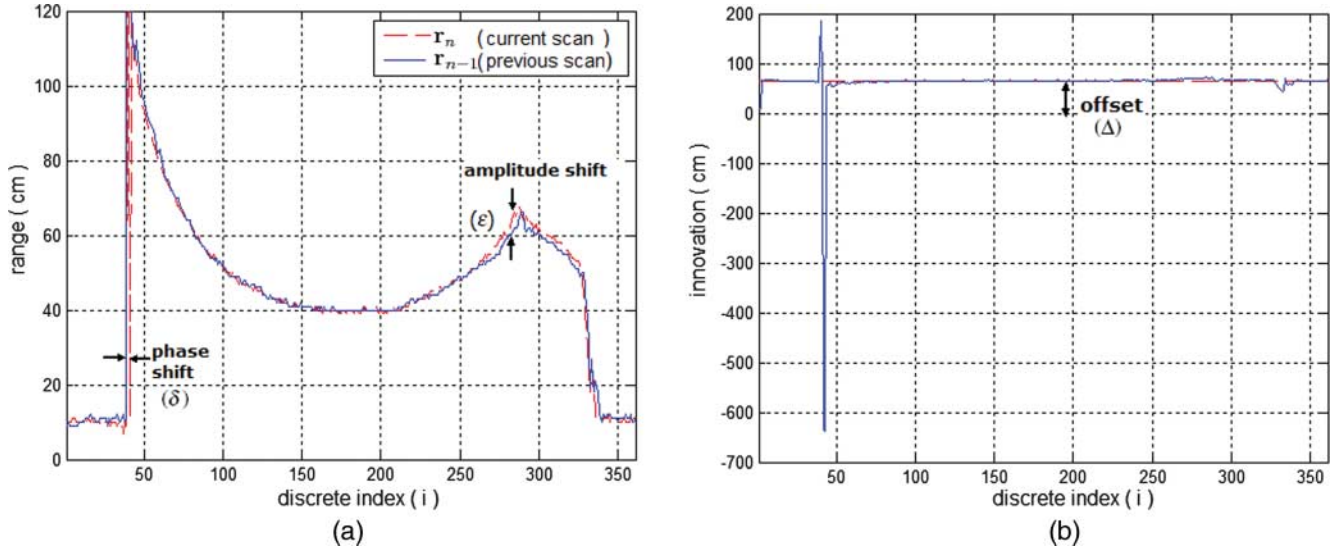
FIGURE 4. The operation scheme of the proposed method.

FIGURE 5. Illustration of (**a**) the amplitude and phase shifts and (**b**) the offset.

derivatives with respect to ϵ and δ to zero to find the optimal values of ϵ and δ, respectively. First, we determine ϵ from

$$\frac{\partial \mathcal{E}^2}{\partial \epsilon} = \sum_{i=1}^{N} [-2r_n[i] + 2(r_{n-1}[i + \delta] + \epsilon)] = 0. \quad (8)$$

When we neglect δ and assume that $r_{n-1}[i + \delta] \cong r_{n-1}[i]$ in Equation (8), we obtain the following solution for ϵ:

$$\epsilon = \frac{1}{N} \sum_{i=1}^{N} (r_n[i] - r_{n-1}[i]). \quad (9)$$

In other words, ϵ corresponds to the average amplitude difference between $\mathbf{r}_n$ and $\mathbf{r}_{n-1}$. Next, we determine δ from

$$\frac{\partial \mathcal{E}^2}{\partial \delta} = \sum_{i=1}^{N} \left(-2r_n[i] \frac{\partial r_{n-1}[i + \delta]}{\partial \delta} \right.$$
$$\left. + 2 (r_{n-1}[i + \delta] + \epsilon) \frac{\partial r_{n-1}[i + \delta]}{\partial \delta} \right) = 0. \quad (10)$$

The $r_{n-1}[i + \delta]$ term in Equation (10) can be expanded using a *Taylor series* expansion around i, such that $r_{n-1}[i + \delta] = r_{n-1}[i] + r'_{n-1}[i]\delta + (\frac{1}{2})r''_{n-1}[i]\delta^2 + \cdots$, where $r'_{n-1}[i]$ and $r''_{n-1}[i]$ are the first- and the second-order differences of the sequence $\mathbf{r}_{n-1}$ at i, respectively. Assuming that δ is very small compared with N, we use only the first two terms of the expansion and

obtain the following first-order approximation to δ:

$$\delta = \frac{\sum_{i=1}^{N} r'_{n-1}[i](r_n[i] - r_{n-1}[i] - \epsilon)}{\sum_{i=1}^{N} r'_{n-1}[i]^2}. \tag{11}$$

We investigated using a second-order approximation to δ in [64], and have concluded that the first-order approximation is simple and sufficiently accurate.

Shifting $\mathbf{r}_{n-1}$ along the vertical and the horizontal axes by ϵ and δ, respectively, we obtain an approximation $\hat{\mathbf{r}}_n$ to $\mathbf{r}_n$. Then the difference sequence is $\tilde{\mathbf{v}}_n \triangleq \mathbf{r}_n - \hat{\mathbf{r}}_n$. Here, $\tilde{\mathbf{v}}_n$ is a sparse signal representing discontinuities in the scanned environment. To illustrate this fact, the difference sequences obtained for the 3-D scans in Fig. 3a–d are shown in Fig. 3q–t, respectively. In these figures, the darker features correspond to larger differences. Note that most of the dark features in these images occur where there is a sudden change in the measured range.

If there is any remaining offset level in $\tilde{\mathbf{v}}_n$, as in the example given in Fig. 5b, $\tilde{\mathbf{v}}_n$ is further shifted to the zero level either in the positive or the negative vertical direction by the offset value Δ to improve the sparsity. Here, $-\Delta$ is the most frequently appearing value in $\tilde{\mathbf{v}}_n$. After shifting the amplitude of $\tilde{\mathbf{v}}_n$ by Δ, we eventually obtain a highly sparse innovation $\mathbf{v}_n$, where 70% of the values computed for the first dataset are zero [64].

After we obtain the innovation sequence $\mathbf{v}_n$ as described above, we test whether time consecutive innovation sequences are correlated with each other by applying a whiteness test in the autocorrelation domain. It is shown in [64] that $\mathbf{v}_n$ is indeed a white sequence in time.

Consequently, $\mathbf{r}_n$ is represented with ϵ, δ, Δ and $\mathbf{v}_n$. When $\mathbf{r}_n$ and $\mathbf{r}_{n-1}$ are highly correlated, $\mathbf{v}_n$ becomes very small, so $\mathbf{r}_n$ is represented without $\mathbf{v}_n$ in that case. On the other hand, when $\mathbf{r}_n$ and $\mathbf{r}_{n-1}$ are not sufficiently correlated, $\mathbf{v}_n$ does not become a sparse sequence. This time, $\mathbf{r}_n$ is not encoded. The degree of correlation between $\mathbf{r}_n$ and $\mathbf{r}_{n-1}$ is measured by comparing the RMSE between $\mathbf{r}_n$ and $\hat{\mathbf{r}}_n$ with an experimentally determined threshold (200 cm) that is 10 times the maximum allowable distortion (20 cm) that can be visually tolerated in the reconstruction of $\mathbf{r}_n$. (When the distortion is above 20 cm, it is difficult to visually recognize objects in the reconstructed 3-D scans.) Where $\mathbf{r}_n$ is encoded, the flowchart of the algorithm followed in the sparsifying model is briefly delineated in Fig. 6a.

Finally, the performance of the sparsifying model under additive white Gaussian noise is investigated. The 2-D scans in the 3-D scan illustrated in Fig. 3b are sparsified after zero mean white Gaussian noise is added to them. When noise is not added to the 2-D scan, 6.5% of the values in the representations are nonzero on the average. When the standard deviations of the noise are 1, 2, 3, 4, 5, 10, 20 and 30 cm, the average percentages of nonzero values are 5.1, 4.5, 7.3, 11.8, 13.7, 20.3, 73.7 and 81.9%, respectively. As the sparsity of the representations mentioned above are compared with each other, it is observed that, under the presence of noise with a standard deviation

up to 3 cm, the sparsifying model maintains the performance observed in the case without any added noise. In addition to that, the model provides representations with acceptable sparsity under the presence of noise with a standard deviation as much as 10 cm. Beyond this level of noise, the representations can no longer be considered sparse. Thus, we conclude that the proposed method has reasonably good noise tolerance.

Note that the method proposed here has some similarities with *optical flow* techniques used for motion estimation in image and video processing [65, 66]. In optical flow, spatial and temporal shifts are used to estimate the relative motion between the scene and the camera (the observer). The solution of the following partial differential equation is required:

$$\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0, \tag{12}$$

where V_x and V_y are the x and y components, respectively, of the velocity of the optical flow of the image intensity $I(x, y, t)$, and $\partial I/\partial x$, $\partial I/\partial y$ and $\partial I/\partial t$ are the partial derivatives of the image at (x, y, t) in the corresponding directions. In our method, two spatial shifts, δ and ϵ (and Δ), are involved, whose time derivatives correspond to V_x and V_y in the optical flow equation, respectively.

3.2. The measurement model

The measurement model gets the minimum number of samples from $\mathbf{v}_n$ by using either simple coding (SC) or CS. Simple coding encodes $\mathbf{v}_n$ with the pairs of location and amplitude of the nonzero values in $\mathbf{v}_n$. The measurement size M in this case increases proportionally with the number of nonzero values K, where $M = 2K$. Despite this, the reconstruction error is zero when $\mathbf{v}_n$ is rebuilt from the measurements taken with SC. Compressive sampling measures arbitrary linear combinations of the values in $\mathbf{v}_n$. In this case, the measurement model is determined as described in Section 2.2, and M is calculated using Equation (7). Then $\mathbf{v}_n$ is encoded by using the product $\mathbf{\Phi} \mathbf{v}_n$. Furthermore, the resulting reconstruction error, which arises when $\mathbf{v}_n$ is rebuilt from the measurements taken with CS, increases with K.

When a 2-D scan with N range measurements is received, we first determine the number of nonzero elements K of the scan, as an indication of sparsity. The size M of the measurement vector $\mathbf{m}$ acquired by using either SC or CS is illustrated in Fig. 7. The K^* is the value of K at which the two curves intersect and the number of measurements M for SC and CS are equal. If the equation $M = 2K$ for SC and Equation (7) for CS, corresponding to the two curves, are solved simultaneously for K, we find $K^* = N/e^2$ that only depends on the number of range measurements N in a 2-D scan. When $K \leq K^*$, using SC is advantageous over CS because of the smaller M and the reconstruction error. Consequently, we employ SC when $K \leq K^*$, and use CS, otherwise. We include a special character (i.e. π) at the beginning of $\mathbf{m}$ when SC is applied
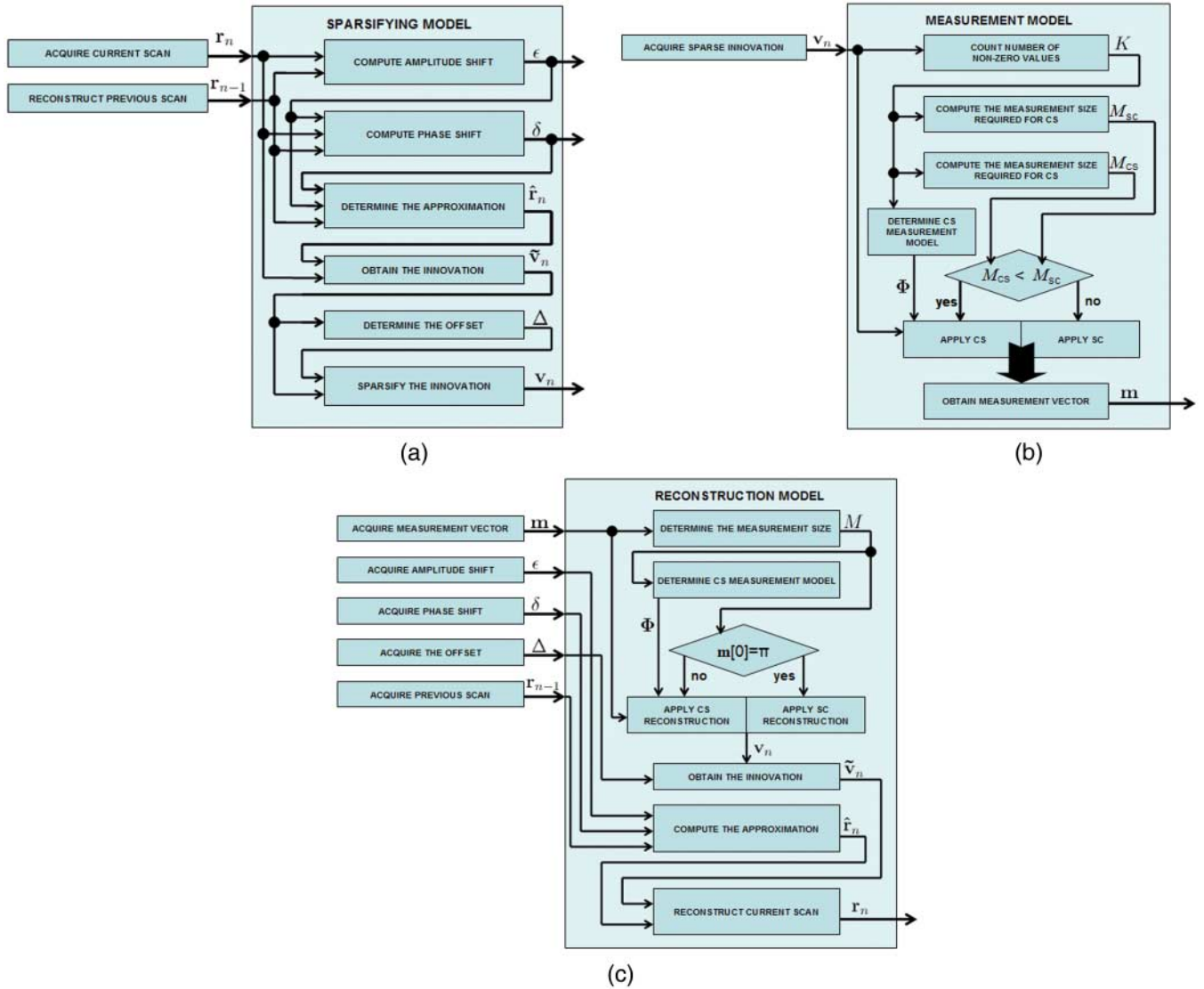
FIGURE 6. The flowcharts of the (**a**) sparsifying, (**b**) measurement and (**c**) reconstruction models.

to inform the decoder that we are using SC instead of CS. Besides, when $K > N/2$ such that more than half of the values in $\mathbf{v}_n$ are nonzero, $\mathbf{v}_n$ cannot be considered sparse, since the reconstruction error would be very high if $\mathbf{v}_n$ were sampled using CS. In that case, $\mathbf{r}_n$ is not encoded. When $\mathbf{r}_n$ is encoded, the flowchart of the algorithm followed in the measurement model is given in Fig. 6b.

If $\mathbf{r}_n$ is encoded, it is represented by $\{\epsilon, \delta, \Delta, \mathbf{m}\}$ at the output of the measurement model. If $\mathbf{r}_n$ is not encoded, the length of the representation remains as N. We investigate whether it is possible to achieve measurements with shorter lengths using a simpler method, such as run length encoding (RLE) [67]. This simple method is commonly used for encoding fax images of typical office documents. RLE first determines the sets in the input data, each of which is formed by the repetition of

a single character. Then it encodes each set with its length and the character repeated in the set. The proposed method provides more efficient measurements than RLE, since the size of the first dataset is reduced, on the average, by 51% using RLE and by 89.1% on the average using the proposed method.

3.3. The reconstruction model

The reconstruction model rebuilds $\mathbf{r}_n$ from the output generated by the encoder. When $\mathbf{r}_n$ is encoded, the output is composed of $\{\epsilon, \delta, \Delta, \mathbf{m}\}$, and its length is $(M + 3)$, which is less than N. If $\mathbf{r}_n$ is not encoded, the output is $\mathbf{r}_n$ with length N. Therefore, the reconstruction process begins by checking the length of the encoder output. If the length is N, the output is stored directly as the reconstruction of $\mathbf{r}_n$. If not, the output is
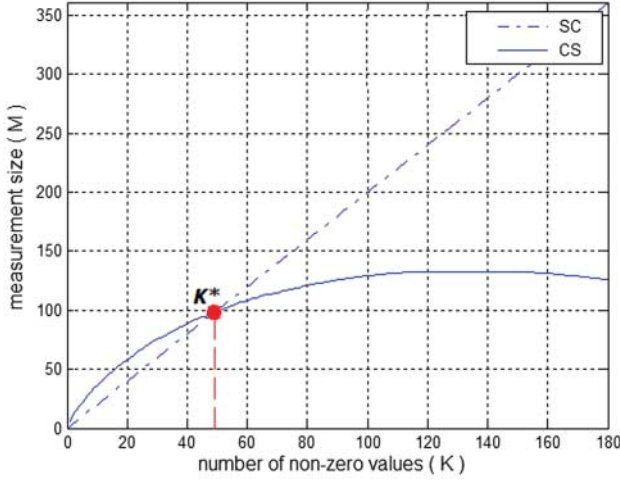
FIGURE 7. The measurement size M in SC and CS with respect to the number of nonzero values K of a signal. In drawing the logarithmic curve for CS, the value of $N = 361$ is used.

decomposed into ϵ, δ, Δ and $\mathbf{m}$. After this step, $\mathbf{r}_{n-1}$, which has been previously reconstructed, is shifted along the vertical and horizontal axes by ϵ and δ, respectively. The resultant signal $\hat{\mathbf{r}}_n$ is the approximation to $\mathbf{r}_n$. Then $\tilde{\mathbf{v}}_n$ is rebuilt from $\mathbf{m}$ and Δ. In this step, if the first value of $\mathbf{m}$ is π, then $\mathbf{v}_n$ is rebuilt, decoding the rest of $\mathbf{m}$ with respect to the SC scheme, which involves filling an empty signal in $\Re^N$ with the values of location and amplitude pairs provided in the measurements. Otherwise, $\mathbf{v}_n$ is rebuilt, decoding $\mathbf{m}$ with respect to the CS scheme, which involves solving Equation (5), where $\mathbf{y} = \mathbf{m}$, $\Theta = \Phi$ and $\hat{\mathbf{s}} = \mathbf{v}_n$, following the procedure in [52]. In our implementation, $\mathbf{v}_n$ is determined using the MATLAB function *perform_l1_recovery* written by Peyre [68]. Then $\tilde{\mathbf{v}}_n$ is obtained by shifting the amplitude of $\mathbf{v}_n$ by $-\Delta$. Eventually, $\mathbf{r}_n$ is reconstructed by adding $\tilde{\mathbf{v}}_n$ to $\hat{\mathbf{r}}_n$. The algorithm followed in the reconstruction process is summarized in Fig. 6c.

The reconstruction model is used at the decoder, as well as at the encoder to estimate the input reconstructed by the decoder.

4. COMPARING THE COMPRESSION PERFORMANCE OF THE PROPOSED METHOD WITH SOME WELL-KNOWN COMPRESSION TECHNIQUES

Three different experimental datasets are used for benchmarking in this study. The datasets are composed of numerous 3-D scans. The first two datasets contain 3-D scans of indoor environments acquired using the SICK LMS200 LRF. The 3-D scans are acquired by collecting 2-D scans from a sensor on board a mobile robot. The sensor is rotated in small steps around a horizontal axis above the ground level while the position of the robot is fixed. Each 2-D scan in the datasets is obtained as the laser beam emitted by the sensor is swept within the sensor's field

of view in 0.5° steps. The first dataset contains 29 3-D scans collected at different locations in the University of Osnabrück's AVZ building in Osnabrück, Germany [61]. The sensor rotates 471 steps to acquire the 2-D scans forming a 3-D scan in this set. The second dataset is collected, while a mobile robot platform equipped with SICK LMS200 follows a parallelogram-shaped path in one of the halls of Dagstuhl Castle in Wadern, Germany. It is composed of 82 3-D scans taken at different locations in the hall [69]. Each 3-D scan in this set is acquired by rotating the sensor in 225 steps. As a consequence, each 3-D scan from the first and the second dataset constitutes 471 and 225 2-D scans, respectively. The 2-D scans are sequentially acquired as vectors in $\Re^{361}$ (i.e. $N = 361$). The third dataset contains 3-D scans of outdoor environments acquired using the RIEGL VZ-400 LRF. The 3-D scans in this dataset are acquired at the city center of Bremen. Each of the two 3-D scans used from this set contains 2250 horizontal 2-D scans, sequentially acquired as vectors in $\Re^{3000}$ (i.e. $N = 3000$) [70].

In this section, we compare the compression performance of the proposed method with some well-known and widely used lossless and lossy compression techniques. The 3-D scans are compressed by applying each technique to the 2-D scans comprising the 3-D scans individually. For each technique in the comparison, we compare the overall CR, the average distortion ($\mathcal{D}$), that is, the average RMSE between the 2-D scans and their reconstructions defined in Section 1, and the time required for encoding (t_{enc}) and decoding (t_{dec}) the 3-D scans. These values are found by averaging over the values obtained for all 3-D scans, including 4 930 899 (= 29 3-D scans ×471 2-D scans ×361 measurements) range measurements in the first dataset, 6 660 450 (= 82 3-D scans ×225 2-D scans ×361 measurements) range measurements in the second dataset and 13 500 000 (= two 3-D scans ×2250 2-D scans ×3000 measurements) range measurements in the third dataset.

The following implementations are executed on a computer platform with a 2 GHz Intel Core2 Duo processor and 2 GB RAM. All executable tasks are run in the MATLAB environment installed on a Microsoft Windows Vista operating system.

4.1. Implementation and comparison with well-known lossless techniques

First, the 3-D scans are compressed using four different lossless techniques, which are *Huffman*, *arithmetic*, *ZLIB* and *GZIP* coding techniques. Huffman coding maps every character in the input data to distinct binary patterns based on the frequency of appearance of the characters. It is the optimal lossless coding technique since the characters that appear more frequently are mapped to shorter patterns than the characters that appear less frequently, and the two characters that appear least frequently are mapped to two different patterns having the same length [44]. Arithmetic coding maps *blocks* of characters, instead of single characters, to distinct binary patterns based on how frequently the blocks appear. It is observed that arithmetic

coding can sometimes be more efficient than Huffman coding, depending on the nature of the signal to be encoded [44]. The 3-D scans are encoded by Huffman and arithmetic coding using the *huffmanenco* and *arithenco* functions in the MATLAB Communications Toolbox, respectively.

ZLIB and GZIP are two popular compression techniques that are variations of *LZ77* [71], which is a widely used compression method that encodes repeated strings in the input data with pairs of *distance* and *length*. Distance is the separation between the beginning of the last location and the previous location of the repeated string in the data. Length is the size of the corresponding repeated string. Two independent Huffman trees are used in compressing the distance and length information, respectively. ZLIB is a general purpose coding library, and can be used in any operating system. ZLIB is reported to provide satisfactory compression on various types of data with optimum use of system resources. ZLIB is also claimed to be able to compress the input data by up to 99.9% in theory [72]. GZIP is a coding technique that is designed to be used instead of *compress*, which is a compression utility used in UNIX operating systems [73]. The 3-D scans are encoded by ZLIB and GZIP using the functions written by Kleder [74] and Hopkins [75], respectively.

4.2. Implementation and comparison with well-known lossy techniques

Besides the lossless compression techniques, we also apply two lossy compression methods to the 3-D scans: *JPEG* and the *wavelet transform.*

Since the 2-D scans forming a 3-D scan are basically cross-sectional intensity images of the scanned environment where the intensity values represent depth information, JPEG compression is applied first. JPEG refers to a family of image compression standards including both lossless and lossy techniques. Lossy JPEG techniques are based on the discrete cosine transform (DCT) applied on 8×8 blocks of pixels in the image data. The CR for lossy JPEG is claimed to be as low as about 5% in compressing colored images, when the distortion in the reconstructed images is not visually recognizable [76]. In colored images, each pixel is represented with three channels, each of which holds an 8-bit intensity value that corresponds to an unsigned integer between 0 and 255. On the other hand, range measurements in the 2-D scans of the experimental datasets are represented in the binary format using 16 bits. Therefore, before encoding a 2-D scan using JPEG, each range measurement is encoded with three channels such that the most significant 8 bits are placed in the first channel, the least significant 8 bits are placed in the second channel and the third channel is left blank. Then the 2-D scan, which is now a $1 \times N$ image with three channels, is duplicated eight times to form an $8 \times N$ image, since JPEG divides the image into 8×8 blocks before applying the DCT. Eventually, the resultant image data are encoded by lossy JPEG using the MATLAB *imwrite* function. Without these operations on the raw data prior to encoding with JPEG, the CR and the distortion will be high [77].

Besides JPEG, the wavelet transform, which is also widely used in image compression, is applied to the raw scan data. The wavelet transform analyzes the input signal at separate bandwidths by applying the input signal to a specific *filterbank*, that is, a set of low-pass and high-pass filters connected in a network [78]. Every 2-D scan in the 3-D scans is compressed using up to a three-level wavelet transform with the Haar filterbank, for which the function *dwt* in the MATLAB Wavelet Toolbox is employed. The raw scan data are decomposed into a number of frequency components, ranging from low frequencies to high frequencies and denoted by x_L and x_H in a one-level transform; x_{LL}, x_{LH}, x_{HL} and x_{HH} in a two-level transform; and x_{LLL}, x_{LLH}, x_{LHL}, x_{LHH}, x_{HLL}, x_{HLH}, x_{HHL} and x_{HHH} in a three-level transform. For the reconstruction, only the lowest frequency components, which are x_L, x_{LL}, x_{LLL} in one-level, two-level and three-level transforms, respectively, are used. Therefore, some distortion on the reconstructions is expected.

4.3. Implementation and comparison with the proposed method

Finally, the 3-D scans are encoded using the proposed method, which is a lossy technique. In the implementation, small fluctuations in the compression performance are observed, such that at most $\pm 2\%$ variations appear in the CR, since the measurement model in CS is determined arbitrarily in each trial. Recall that in the proposed method, each 2-D scan in a 3-D scan is encoded with one of the following:

(i) ϵ, δ, Δ and $\mathbf{m}$ acquired using CS;
(ii) ϵ, δ, Δ and $\mathbf{m}$ acquired using SC;
(iii) ϵ, δ and Δ;
(iv) itself (no coding).

The number of occurrences of each type of code for a given 3-D scan is denoted by k_1, k_2, k_3 and k_4, respectively. Thus, these numbers change as the CR fluctuates. Moreover, the distortion also changes in this situation. Therefore, every 3-D scan is encoded using the proposed method ten times, then the average values of the CR, $\mathcal{D}$, t_{enc}, t_{dec}, k_1, k_2, k_3 and k_4 are obtained.

The distortion $\mathcal{D}$ is dependent on the information provided by the code of $\mathbf{r}_n$. At first, $\mathbf{r}_n$ is defined with the relative vertical and horizontal shifts (ϵ and δ) to $\hat{\mathbf{r}}_n$. When the RMSE between $\mathbf{r}_n$ and $\hat{\mathbf{r}}_n$ is >20 cm for indoor environments and >1 m for outdoor environments, $\mathbf{r}_n$ is defined with additional information provided by sampling $\mathbf{v}_n$ through the measurement model. Therefore, $\mathcal{D}$ is restricted to either 20 cm or 1 m depending on the type of the range of measurements, which is the maximum distortion that can be tolerated in the reconstruction. For instance, when the 3-D scans illustrated in Fig. 3a–d from the first dataset are compressed using the proposed method, the resulting average distortions are about 15, 13, 14 and 11 cm,

respectively. Their reconstructions are shown in Fig. 3e–h for comparison with their originals. In parts (i)–(l) and (m)–(p), the original scans and their reconstructions are, respectively, illustrated as point clouds, where range measurements are represented as discrete points in the 3-D Cartesian space with the LRF located at its origin. In these figures, the point cloud is viewed from a different perspective than that of the LRF. These illustrations are obtained using *CloudCompare*, which is a free software for 3-D point and mesh processing [79]. Difference sequences of (a)–(d) are shown in parts (q)–(t). Furthermore, the distortion errors between these 3-D scans and their reconstructions are illustrated in Fig. 3u–x to provide a visual comparison. According to the figures, the distortion becomes significant in those 2-D scans that are encoded with only $\{\epsilon, \delta, \Delta\}$, as indicated by the darker horizontal stripes in the difference images.

The average compression performances of the methods described so far are summarized in Table 1 for the three datasets. For the performances of the lossless methods, arithmetic coding can be said to be efficient in terms of the CR for both datasets; however, it is slow compared with ZLIB and GZIP. On the other hand, these methods compress less than arithmetic coding for the first two datasets. ZLIB and GZIP seem better than other lossless methods for the third dataset. Despite the high average CR and t_{enc} when Huffman coding is applied to the raw scan data, they can be lowered by encoding the differences between consecutive 2-D scans, since the range of distinct characters in the differences is narrower than in the raw scan data. However, coding the differences instead of the raw scan data may not lower the CR for other compression techniques because their compression performance is not directly related to the range of distinct characters in the input data, as in Huffman coding. With this approach, a considerable amount of compression is observed for the first dataset, but not for the other datasets. For the performances of the lossy methods, JPEG can be considered a fast and efficient technique in terms of the CR, however, it results in intolerable distortion. As the level of the wavelet transform increases, the CR decreases exponentially; meanwhile, the distortion increases and becomes intolerable. The wavelet transform is very fast; the time required for both encoding and decoding does not exceed a few seconds. When the proposed method is compared with the lossless methods considered here, it is found to be faster than the variations of Huffman and arithmetic coding, but slower than ZLIB and GZIP for the first two datasets. On the other hand, the proposed method compresses much more than ZLIB and GZIP. For the third dataset, the proposed method compresses less than ZLIB and GZIP, but is much faster than them. The performance of the proposed method is remarkable when the lossless methods fail in compressing, as observed for the second and the third datasets. When the proposed method is compared with the lossy methods, the proposed method does not provide the least CR, but provides acceptable low CR with very low distortion and high speed. For lossy compression, there always exists a trade-off

TABLE 1. CR, $\mathcal{D}$, t_{enc} and t_{dec} when (a) the first, (b) the second and (c) the third datasets are compressed by different methods.

Method	CR (%)	$\mathcal{D}$ (cm)	t_{enc} (s)	t_{dec} (s)
(a)				
Lossless methods				
HC on $\mathbf{r}_n$	41.7	0	165.6	610.6
HC on $(\mathbf{r}_n - \mathbf{r}_{n-1})$	12.0	0	49.0	14.0
AC on $\mathbf{r}_n$	11.1	0	37.6	48.9
AC on $(\mathbf{r}_n - \mathbf{r}_{n-1})$	16.4	0	45.6	58.4
ZLIB	65.3	0	0.4	0.2
GZIP	76.7	0	0.5	0.3
Lossy methods				
JPEG	9.0	164.6	4.2	7.3
One-level WT	50.2	21.3	0.3	0.3
Two-level WT	25.2	29.8	0.5	0.6
Three-level WT	12.7	37.3	0.8	0.9
Proposed method	10.9	12.9	15.3	14.5
(b)				
Lossless methods				
HC on $\mathbf{r}_n$	683.3	0	101.8	363.1
HC on $(\mathbf{r}_n - \mathbf{r}_{n-1})$	253.6	0	12.6	34.2
AC on $\mathbf{r}_n$	27.1	0	21.8	25.7
AC on $(\mathbf{r}_n - \mathbf{r}_{n-1})$	35.8	0	16.4	19.1
ZLIB	140.8	0	0.3	0.1
GZIP	143.8	0	0.3	0.2
Lossy methods				
JPEG	10.0	743.6	1.5	3.6
One-level WT	50.2	204.4	0.2	0.2
Two-level WT	25.2	283.0	0.4	0.3
Three-level WT	12.7	353.8	0.5	0.5
Proposed method	32.0	4.8	1.9	1.7
(c)				
Lossless methods				
HC on $\mathbf{r}_n$	427.7	0	1042.1	5820.0
HC on $(\mathbf{r}_n - \mathbf{r}_{n-1})$	135.9	0	1862.3	2756.2
AC on $\mathbf{r}_n$	21.1	0	416.1	1160.8
AC on $(\mathbf{r}_n - \mathbf{r}_{n-1})$	25.0	0	175.1	592.3
ZLIB	15.2	0	3.3	4.2
GZIP	15.4	0	4.8	5.7
Lossy methods				
JPEG	4.5	8690	36.3	314.4
One-level WT	50.0	440	3.0	1.2
Two-level WT	25.0	550	4.6	3.3
Three-level WT	12.5	620	6.0	5.0
Proposed method	37.0	63	2.1	0.2

HC, Huffman coding; AC, arithmetic coding.

between reducing the size of the input data and minimizing the distortion on the reconstructions [44]. Consequently, being a lossy method, the proposed method provides a reasonably good compromise between the CR, accuracy of the reconstructions
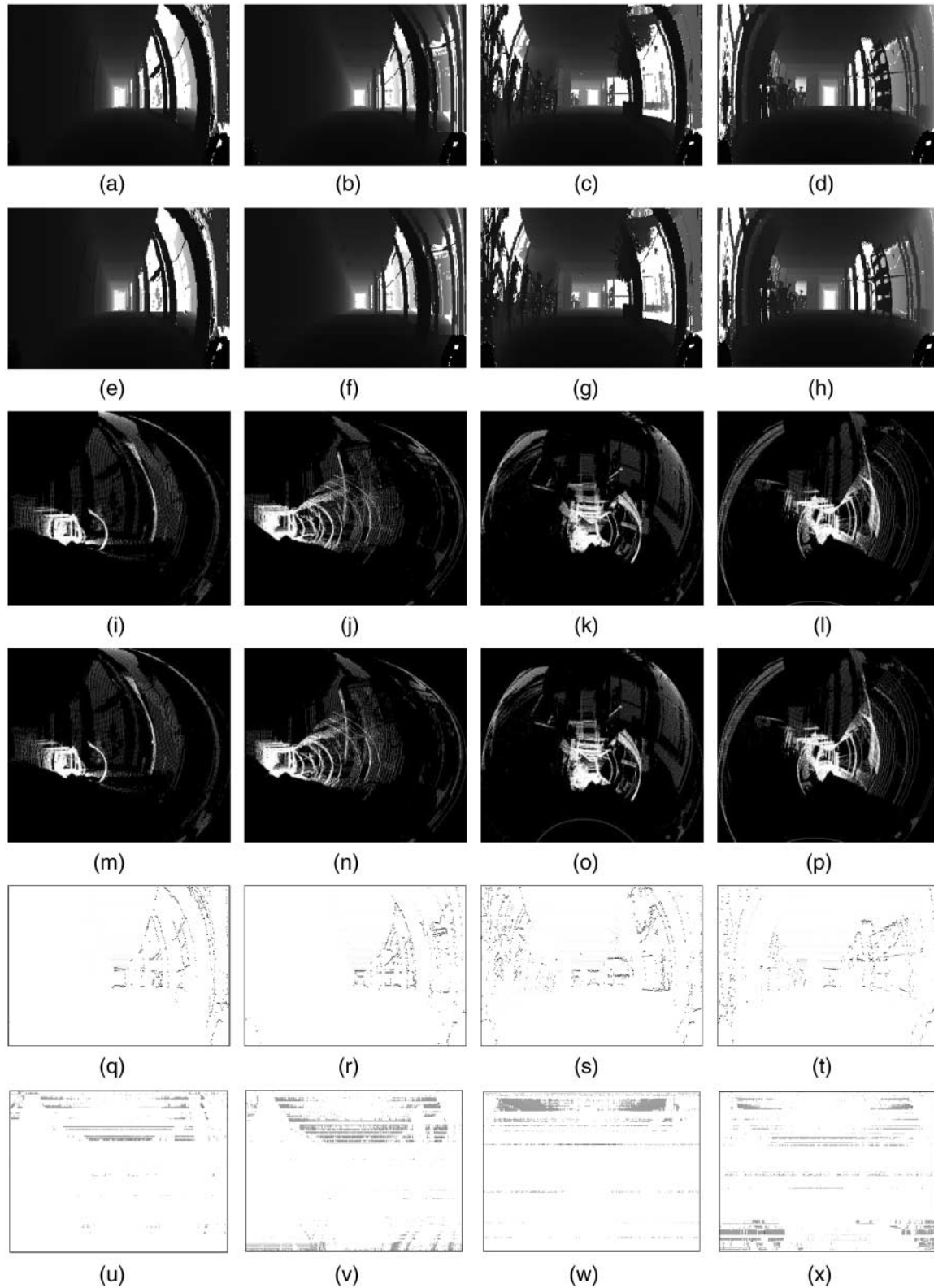
FIGURE 8. (**a**)–(**d**) Sample 3-D scans collected at Dagstuhl Castle using SICK LMS200; (**e**)–(**h**) their reconstructions; (**i**)–(**l**) point cloud representations of (a)–(d); (**m**)–(**p**) point cloud representations of (e)–(h); (**q**)–(**t**) difference sequences of (a)–(d); and (**u**)–(**x**) the resulting distortion error.

TABLE 2. The average percentages of k_1, k_2, k_3 and k_4, when the three datasets are compressed using the proposed method.

Dataset	k_1 (%)	k_2 (%)	k_3 (%)	k_4 (%)
#1 (University of Osnabrück)	16.0	24.0	57.0	3.0
#2 (Dagstuhl Castle)	4.9	17.3	13.8	64.0
#3 (City Center of Bremen)	0.3	31.0	63.0	37.0

and speed when its performance is compared with the performances of the well-known techniques considered in this study.

In the second dataset, besides stationary features such as furniture (also present in the first dataset), buildings outside the hall are visible through the glass windows along both sides of the hall, as illustrated in Fig. 8a–d. Parts (e)–(h) depict the reconstructions of the original scans. In parts (i)–(l) and (m)–(p) of the same figure, the original scans and their reconstructions are represented as point clouds, respectively. Because of the additional detailed features in this dataset, more discontinuities appear as illustrated in the difference sequences in Fig. 8q–t, compared with that of the first dataset. Therefore, the complexity of the scenes in the second dataset is higher, and the similarity between consecutive 2-D scans in the second dataset is lower than in the first. The average correlation coefficients of the range measurements in the first and the second datasets are calculated as 0.9521 and 0.9102, respectively [64]. Because of this difference in complexity between the datasets, the compression performance of the proposed method demonstrates variations, as indicated in Table 2, in the average percentages of the number of the 2-D scans according to how they are encoded. The table demonstrates how the degree of similarity between the 2-D scans in a 3-D scan affect the compression performance of the proposed method; the 3-D scans in the first and the second datasets are compressed by 89.1% (Table 1a) and 68% (Table 1b), on the average, respectively. Although the size of the second dataset is not reduced as much, the distortion on its reconstructions is lower, as shown in Fig. 8e–h. The resulting distortion errors of the corresponding reconstructions are illustrated in Fig. 8u–x.

The third dataset, acquired at the city center of Bremen, contains 3-D scans of outdoor environments in contrast to the first two datasets. Therefore, the upper bound on the range measurements (350 m) and the size of the 2-D scans ($N = 3000$) are much larger than that of the first two datasets. This makes compressing the third dataset more difficult, and increases the complexity of the scenes. Two sample 3-D scans, illustrated in Fig. 9a and b, are used in the experiments. The average correlation coefficient of the range measurements is calculated to be 0.9122, which is closer to that of the second dataset. In parts (e, f) and (g, h), the original scans and their reconstructions are represented as point clouds, respectively.
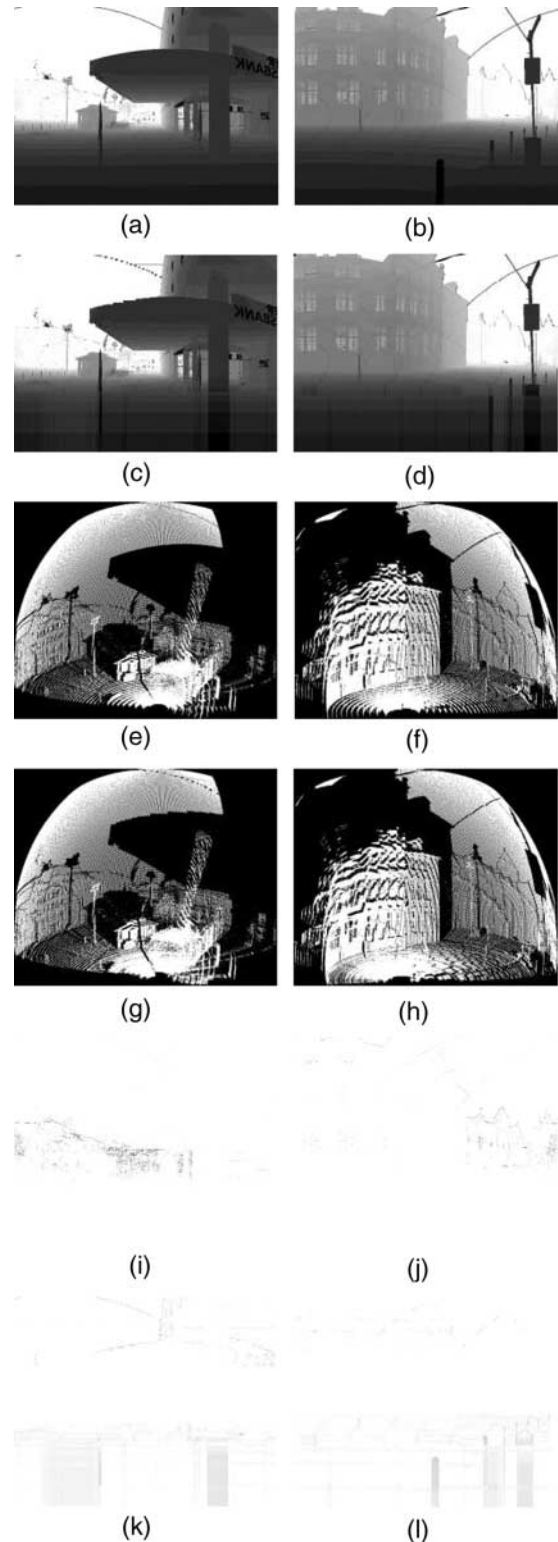

FIGURE 9. (**a**, **b**) Sample 3-D scans collected at the city center of Bremen using RIEGL VZ-400; (**c**, **d**) their reconstructions; (**e**, **f**) point cloud representations of (a, b); (**g**, **h**) point cloud representations of (c, d); (**i**, **j**) difference sequences of (a, b); and (**k**, **l**) the resulting distortion error.

TABLE 3. The results of adding different amounts of white Gaussian noise to the first dataset.

Noise st. dev.	SNR (dB)	CR (%)	$\mathcal{D}$ (cm)	t_{enc} (s)	t_{dec} (s)	k_1	k_2	k_3	k_4
								(c o u n t)	
–	–	10.9	12.9	15.3	14.5	76	113	267	15
1	52.7	10.8	12.9	12.1	10.9	74	115	268	15
2	46.6	10.8	12.9	14.6	13.5	73	115	268	14
3	43.1	10.5	12.8	11.3	10.2	70	117	270	14
4	40.6	10.3	12.6	11.1	9.9	66	120	271	14
5	38.7	10.2	12.6	10.5	9.4	64	121	272	14
10	32.7	11.2	12.5	10.2	9.1	62	117	276	16
20	26.6	20.0	13.4	23.4	22.3	147	0	286	38
30	23.1	27.5	13.5	14.2	13.3	88	0	287	96
50	18.7	88.1	6.7	2.5	0.4	0	0	56	415
100	12.7	100.0	0.0	2.9	0.4	0	0	0	471

When the third dataset is used, the majority of the 2-D scans are encoded with $\{\epsilon, \delta, \Delta\}$ as indicated in Table 2, because dense discontinuities do not appear all over the 3-D scans as illustrated in Fig. 9i and j. The 3-D scans in the third dataset are compressed by 63% (Table 1c) with a distortion of about 63 cm per measurement. The reconstructions and distortion errors are given in Fig. 9c, d and k, l, respectively. The proposed method maintains its performance as observed with the other two datasets.

Finally, the compression performance of the proposed method under the presence of additive white Gaussian noise is investigated. In this part, the 3-D scans in the first dataset are compressed after zero mean white Gaussian noise is added to them. For different noise levels, each 3-D scan is compressed ten times using the proposed method, then the average signal-to-noise ratio (SNR), CR, $\mathcal{D}$, t_{enc}, t_{dec}, k_1, k_2, k_3 and k_4 values are obtained, as given in Table 3. Here, the SNR is the ratio of the power of the 3-D scan to the noise power. According to the table, it is observed that the proposed method maintains its performance, with the CR around 10% and $\mathcal{D}$ about 13 cm, under the presence of noise, with a standard deviation up to 10 cm. Note that the sparsifying model also performs properly when the standard deviation of the noise is below the same level. The method provides acceptable compression when the standard deviation remains below 30 cm. Beyond that level, the method cannot perform effective compression. In other words, the method can operate with an SNR larger than 23 dB, and works properly with an SNR larger than 30 dB.

5. CONCLUSIONS AND FUTURE WORK

In this study, we consider the task of efficient representation and transmission of 3-D laser range scans of indoor and outdoor environments, which has applicability in a variety of fields. The task involves compressing 2-D range scans forming a 3-D model simultaneously with the acquisition of the scans, to be able to use the capacity of the communication channel or the data storage medium effectively during transmission or storage of the scan data. From this perspective, we propose a compression technique based on compressive sensing for sequentially acquired 2-D scans that are correlated. Then we demonstrate the superiority of the proposed method over some well-known lossy and lossless compression techniques in encoding the scan data based on the previous scans.

The proposed technique involves sampling the sparse signal efficiently. Each sparse signal is obtained from the difference between the current scan and its estimate, which is generated by shifting the previous scan along the horizontal and vertical axes by certain amounts. The amount of displacements along these axes are formulated with respect to the current and previous scans. In this sense, the proposed method is similar to difference encoding. Then the amplitude of the difference signal is offset to improve the sparsity. Compression is achieved by sampling the sparse signal using either SC or CS.

The compression performance of the proposed method relies on the similarity between consecutive 2-D scans in the input data. The higher the correlation between consecutive 2-D scans and the lower the complexity and detail level of the scanned environment, the lower is the CR. For instance, the proposed method compresses a 3-D scan in the first dataset that includes ∼170 000 range measurements within 15 s and by about 89% on the average, with an average distortion of about 13 cm per measurement. This is about 1.36 MB data since the storage of each range measurement requires 8 bytes in the MATLAB environment. Moreover, the proposed method maintains this performance under the presence of zero mean white Gaussian noise added to the scan data when the SNR is larger than 30 dB. However, for the second dataset where the similarity is somewhat lower than in the first one, the proposed method compresses a 3-D scan from this set that contains ∼81 000 range measurements (about 650 KB data) within two seconds and by about 68%

on the average, with the average distortion per measurement being about 5 cm. A 3-D scan from the third dataset contains 6 750 000 range measurements (about 54 MB data). Using the proposed method, such a 3-D scan is compressed within about 2 s and by about 63% on the average, with the average distortion per measurement being about 0.63 m. It can be stated that distortion on the reconstructions becomes significant when the correlation between the adjacent horizontal scans is large, resulting in a high compression rate. The amount of distortion can be limited in the algorithm by setting the threshold for the maximum distortion tolerated in the reconstructions. Therefore, the proposed method is fast and efficient according to the criteria described in Section 1. The proposed method is recommended for applications where both the CR and speed are crucial. However, a lossless compression technique, such as arithmetic coding, can be used in applications where the accuracy of the range measurements is more important.

In summary, the proposed method provides an acceptable CR compared with the alternative compression techniques that we have considered, and as it provides a reasonably good compromise between reconstruction accuracy and speed, it can be effectively used for 3-D representation of both indoor [80] and outdoor environments as shown in the experimental results. The performance of the method can be improved by further coding the encoder output via lossless coding techniques such as arithmetic coding or LZ77 [71], which would reduce the CR even more. Despite the existence of several compression techniques dedicated to bit encoding (e.g. Huffman Coding, RLE, etc.), it would be interesting to adapt the proposed method to compress the sequence of 2-D range measurements in binary form. Another possible future work direction would be the hardware implementation of the proposed method on FPGAs so that it can be used in real-time applications. We also plan to extend its application to other types of datasets composed of measurement sequences, such as video data.

FUNDING

REFERENCES

[1] Brenneke, C., Wulf, O. and Wagner, B. (2003) Using 3-D Laser Range Data for SLAM in Outdoor Environments. *Proc. IEEE/RSJ Int. Conf. Intelligent Robots Syst.*, Las Vegas, NV, USA, October 27–31, Vol. 1, pp. 188–193. IEEE, NJ, USA.

[2] Liu, Y., Emery, R., Chakbarti, D., Burgard, W. and Thrun, S. (2001) Using EM to Learn 3-D Models of Indoor Environments with Mobile Robots. *Proc. 18th Int. Conf. Mach. Learn.*, Williamstown, MA, USA, June 28–July 1. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[3] Surmann, H., Nüchter, A. and Hertzberg, J. (2003) An autonomous mobile robot with a 3-D laser range finder for 3-D exploration and digitalization of indoor environments. *Robot. Auton. Syst.*, **45**, 181–198.

[4] Borrmann, D., Elseberg, J., Lingemann, K., Nüchter, A. and Hertzberg, J. (2008) Globally consistent 3-D mapping with scan matching. *Robot. Auton. Syst.*, **56**, 130–142.

[5] Nüchter, A., Wulf, O., Lingemann, K., Hertzberg, J., Wagner, B. and Surmann, H. (2006) 3-D Mapping with Semantic Knowledge. In Bredenfeld, A. *et al.* (eds), *Lecture Notes in Computer Science (RoboCup 2005), 4020/2006*, pp. 335–346. Springer, Berlin, Heidelberg, Germany.

[6] Levoy, M. *et al.* (2000) The Digital Michelangelo Project: 3-D Scanning of Large Statues. *Proc. 27th Annual Conf. Comp. Graphics Interactive Tech. (ACM SIGGRAPH 2000)*, New Orleans, LA, USA, July 23–28, pp. 131–144. ACM Press/Addison-Wesley Publishing Co., New York, USA.

[7] Bernardini, F., Rushmeier, H., Martin, I.M., Mittleman, J. and Taubin, G. (2002) Building a digital model of Michelangelo's Florentine Pieata. *IEEE Comput. Graph.*, **22**, 59–67.

[8] Chevrier, C. and Perrin, J.P. (2001) Interactive 3-D Reconstruction for Urban Areas—An Image Based Tool. In Vries, B. de, Leeuwen, J. van and Achten, H. (eds), *CAAD Futures: Proc. Ninth Int. Conf. Comput. Aided Archit. Des. Futures*, July 8–11, pp. 753–765. Kluwer Academic Publishers, Eindhoven, The Netherlands.

[9] Guidi, G., Beraldin, J.A. and Atzeni, C. (2004) High-accuracy 3D modeling of cultural heritage: the digitizing of Donatello's 'Maddalena.' *IEEE T. Image Process.*, **13**, 370–380.

[10] Spagnolo, G.S., Majo, R., Carli, M. and Neri, A. (2004) 3-D Scanner and Virtual Gallery of Small Cultural Heritage Objects. In Corner, B.D., Li, P. and Pargas, R.P. (eds), *Proc. Soc. Photo-Opt. Inst. (SPIE)*, Vol. 5302: Three-Dimensional Image Capture and Applications VI, San Jose, CA, USA, January 19, pp. 148–155. SPIE, Bellingham, WA, USA.

[11] Ergun, B., Sahin, C., Baz, I. and Ustuntas, T. (2010) Creating a 3-D model by terrestrial laser scanners and photogrammetry techniques: a case study on the historical peninsula of Istanbul. *Environ. Monit. Assess.*, **165**, 595–601.

[12] Arayici, Y. (2007) An approach for real world data modelling with the 3D terrestrial laser scanner for built environment. *Automat. Constr.*, **16**, 816–829.

[13] Watt, P.J. and Donoghue, D.N.M. (2005) Measuring forest structure with terrestrial laser scanning. *Int. J. Remote Sens.*, **26**, 1437–1446.

[14] SICK AG (2011) *SICK partner portal*. http://www.mysick.com/.

[15] RIEGL Laser Measurement Systems GmbH (2011) *Mobile laser scanning*. http://www.riegl.com/nc/products/mobile-scanning.

[16] FARO Technologies Inc. (2011) *FARO laser tracker*. http://www.faro.com/lasertracker/home.

[17] Zoller + Fröhlich GmbH (2011) *Z+F laser scanner*. http://www.zf-laser.com/e_z_f-laserscanner.html.

[18] Leica Geosystems AG (2011) *Laser tracker systems*. http://www.leica-geosystems.com/en/Laser-Tracker-Systems_69045.htm.

[19] MENSI Inc. (2011) *3-D modeling systems and 3D modeling software*. http://mensi.free.fr/english/techno.htm.

[20] Velodyne (2011) *High definition LIDAR*. http://velodynelidar. com/lidar/lidar.aspx.

[21] Ellekilde, L.P., Huang, S., Miro, J.V. and Dissanayake, G. (2007) Dense 3-D map construction for indoor search and rescue. *J. Field Robot.*, **24**, 71–89.

[22] Zhou, W., Miro, J.V. and Dissanayake, G. (2008) Information-efficient 3-D visual SLAM for unstructured domains. *IEEE T. Robot.*, **24**, 1078–1087.

[23] Pervölz, K., Nüchter, A., Surmann, H. and Hertzberg, J. (2004) Automatic Reconstruction of Colored 3-D Models. *Proc. Deutsche Fachtagung Robotik 2004*, Munich, Germany, 17–18 June, pp. 215–222. VDI, Germany.

[24] Besl, P.J. and McKay, N.D. (1992) A method for registration of 3-D shapes. *IEEE T. Pattern Anal.*, **14**, 239–256.

[25] Hähnel, D. and Burgard, W. (2002) Probabilistic Matching for 3-D Scan Registration. *Proc. VDI-Conf. Robotik*, Ludwigsburg, Germany, June 19–20, Vol. 1679, pp. 137–143. VDI, Düsseldorf, Germany.

[26] Zhou, H. and Liu, Y. (2006) 3-D Modelling from Multi-view Registered Range Images using K-Means Clustering. *Proc. IEEE Int. Conf. Industrial Technology*, Bombay, India, December 15–17, pp. 722–727. IEEE, NJ, USA.

[27] Biber, P., Fleck, S., Busch, F., Wand, M., Duckett, T. and Strasser, W. (2005) 3D Modeling of Indoor Environments by a Mobile Platform with a Laser Scanner and Panoramic Camera. *Proc. 13th European Signal Processing Conf.*, Antalya, Turkey, September 4–8, European Association for Signal Processing (EURASIP).

[28] Vosselman, G. and Dijkman, S. (2002) 3D Building Model Reconstruction from Point Clouds and Ground Plans. *Int. Arch. Photogrammetry Remote Sens.*, Annapolis, MD, USA, October 22–24, Vol. XXXIV-3/W4, pp. 37–43. International Society for Photogrammetry and Remote Sensing, Enschede, The Netherlands.

[29] Filin, S. (2004) Surface classification from airborne laser scanning data. *Comput. Geosci.*, **30**, 1033–1041.

[30] Johnson-Roberson, M., Pizarro, O., Williams, S.B. and Mahon, I. (2010) Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys. *J. Field Robot.*, **27**, 21–51.

[31] Allen, P., Stamos, I., Gueorguiev, A., Gold, E. and Blaer, P. (2001) AVENUE: Automated Site Modeling in Urban Environments. *Proc. Third Int. Conf. 3-D Digital Imaging and Modeling*, Quebec City, Canada, May 28–June 1, pp. 357–364. IEEE Computer Society, NJ, USA.

[32] Zhao, H. and Shibasaki, R. (2001) Reconstructing Urban 3-D Model using Vehicle-Borne Laser Range Scanners. *Proc. Third Int. Conf. 3-D Digital Imaging and Modeling*, Quebec City, Canada, May 28–June 1, pp. 349–356. IEEE Computer Society, NJ, USA.

[33] Nüchter, A., Surmann, H. and Hertzberg, J. (2003) *Proc. Fourth Int. Conf. 3-D Digital Imaging and Modeling*, Banff, Canada, October 6–10, pp. 394–401. IEEE Computer Society, NJ, USA.

[34] Masuda, T. (2001) Generation of Geometric Model by Registration and Integration of Multiple Range Images. *Proc. Third Int. Conf. 3-D Digital Imaging and Modeling*, Quebec City, Canada, May 28–June 1, pp. 254–261. IEEE Computer Society, NJ, USA.

[35] Lanman, D., Crispell, D. and Taubin, G. (2007) Surround Structured Lighting for Full Object Scanning. *Proc. Sixth Int. Conf. 3-D Digital Imaging and Modeling*, Montreal, Canada, August 21–23, pp. 107–116. IEEE Computer Society, NJ, USA.

[36] 8008970/QI72/2006-12 (2006) *Technical description: LMS 200/ 211/221/291 laser measurement systems. SICK AG*. https://mys ick.com/saqqara/pdf.aspx?id=im0012759.

[37] Datasheet, RIEGL VZ-400 (2011) *3-D terrestrial laser scanner with online waveform processing: RIEGL VZ-400. RIEGL Laser Measurement Systems GmbH*. http://riegl.com/uploads/tx_pxpri egldownloads/10_DataSheet_VZ400_12-09-2011.pdf.

[38] Marcelloni, F. and Vecchio, M. (2009) An efficient lossless compression algorithm for tiny nodes of monitoring wireless sensor networks. *Comput. J.*, **52**, 969–987.

[39] Klein, S.T. and Ben-Nissan, M.K. (2010) On the usefulness of Fibonacci compression codes. *Comput. J.*, **53**, 701–716.

[40] Kiely, A., Xu, M., Song, W., Huang, R. and Shirazi, B. (2010) Adaptive linear filtering compression on realtime sensor networks. *Comput. J.*, **53**, 1606–1690.

[41] Popa, I. (2011) SE-Compression: a generalization of dictionary-based compression. *Comput. J.*, **54**, 1876–1881.

[42] Salomon, D. (2002) *A Guide to Data Compression Methods*. Springer, NY, USA.

[43] Lelewer, D.A. and Hirschberg, D.S. (1987) Data compression. *ACM Comput. Surv.*, **19**, 261–296.

[44] Sayood, K. (2000) *Introduction to Data Compression*. Academic Press, San Diego, CA, USA.

[45] Obaid, A. (1994) Range image compression. Master's Thesis, Univ. of Ottawa, Dept. of Electrical Eng., Ottawa, Canada.

[46] Kaushik, R., Xiao, J., Joseph, S.L. and Morris, W. (2010) Fast Planar Clustering and Polygon Extraction from Noisy Range Images Acquired in Indoor Environments. *Proc. IEEE Int. Conf. Mech. Autom.*, Xi'an, China, August 4–7, pp. 483–488. IEEE, NJ, USA.

[47] Rivest, J.-F. and Siddiqi, M. (1996) Range Image Compression. In Stevenson, R.L., Drukarev, A.I. and Gardos, T.R. (eds), *Proc. Soc. Photo-Opt. Inst. (SPIE)*, Vol. 2669: Still-Image Compression II, San Jose, CA, USA, January 30, pp. 95–105. SPIE, Bellingham, WA, USA.

[48] Kaess, M., Arkin, R.C. and Rossignac, J. (2003) Compact Encoding of Robot-Generated 3D Maps for Efficient Wireless Transmission. *Proc. IEEE Intl. Conf. Advanced Robotics*, Coimbra, Portugal, June 30–July 3, pp. 324–331. IEEE, NJ, USA.

[49] Blu, T., Dragotti, P.L., Vetterli, M., Marziliano, P. and Coulot, L. (2008) Sparse sampling of signal innovations. *IEEE Signal Proc. Mag.*, **25**, 31–40.

[50] Baraniuk, R.G. (2007) Compressive sensing. *IEEE Signal Proc. Mag.*, **24**, 118–121.

[51] Candès, E.J. and Wakin, M.B. (2008) An introduction to compressive sampling. *IEEE Signal Proc. Mag.*, **25**, 21–30.

[52] Candès, E.J. and Romberg, J.K. (2005) Signal Recovery from Random Projections. In Bouman, C.A. and Miller, E. L. (eds), *Proc. Soc. Photo-Opt. Inst. (SPIE)*, Vol. 5674: Computational Imaging III, San Jose, CA, USA, January 17, pp.76–86. SPIE, Bellingham, WA, USA.

[53] Candès, E.J., Romberg, J.K. and Tao, T. (2006) Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.*, **59**, 1207–1223.

[54] Zhu, C. (2008) Stable recovery of sparse signals via regularized minimization. *IEEE T. Inform. Theory*, **54**, 3364–3367.

[55] Chen, S.S. (1995) Basis pursuit. PhD Thesis, Stanford Univ., Dept. of Statistics, CA, USA.

[56] Duarte, M.F., Davenport, M.A., Takhar, D., Laska, J.N. Sun, T., Kelly, K.F. and Baraniuk, R.G. (2008) Single-pixel imaging via compressive sampling. *IEEE Signal Proc. Mag.*, **25**, 83–91.

[57] Beauchamp, K.G. (1975) *Walsh Functions and Their Applications*. Academic Press, Inc., NY, USA.

[58] Lustig, M., Donoho, D.L., Santos, J.M. and Pauly, J.M. (2008) Compressed sensing MRI. *IEEE Signal Proc. Mag.*, **25**, 72–82.

[59] Wiaux, Y., Jacques, L., Puy, G., Scaife, A.M.M. and Vandergheynst, P. (2009) Compressed sensing imaging techniques for radio interferometry. *Mon. Not. R. Astron. Soc.*, **395**, 1733–1742.

[60] Fadili, M.J., Starck, J.-L. and Murtagh, F. (2009) Inpainting and zooming using sparse representations. *Comput. J.*, **52**, 64–79.

[61] Nüchter, A. (June 2011) *University of Osnabrück and Jacobs University Robotic 3D Scan Repository, Knowledge-Based Systems Research Group, Institute of Computer Science, University of Osnabrück*. http://kos.informatik.uni-osnabrueck.de/3Dscans/ (17th dataset as of June 2012).

[62] Hayes, M.H. (1996) *Statistical Digital Signal Processing and Modeling*. John Wiley &Sons, NY, USA.

[63] Bar-Shalom, Y. and Fortmann, T.E. (1987) *Tracking and Data Association*. Academic Press, San Diego, CA, USA.

[64] Dobrucalı, O. (2010) A novel compression algorithm based on sparse sampling of 3-D laser range scans. Master's thesis, Bilkent Univ., Dept. of Electrical and Electronics Eng., Ankara, Turkey.

[65] Beauchemin, S.S. and Barron, J.L. (1995) The computation of optical flow. *ACM Comput. Surv.*, **27**, 433–467.

[66] Fleet, D.J. and Weiss, Y. (2006) Optical Flow Estimation. In Paragios, N., Chen, Y. and Faugeras, O. (eds), *Handbook of Mathematical Models in Computer Vision*, ch. 15, pp. 239–258. Springer, Berlin, Heidelberg, Germany.

[67] Bose, R. (2008) *Information Theory, Coding and Cryptography*. Tata McGraw-Hill, New Delhi, India.

[68] Peyre, G. (2009) *Toolbox sparsity*. http://www.mathworks.com/matlabcentral/fileexchange/16204.

[69] Nüchter, A. and Lingemann, K. (June 2011) *University of Osnabrück and Jacobs University Robotic 3D Scan Repository, Knowledge-Based Systems Research Group, Institute of Computer Science, University of Osnabrück*. http://kos.informatik.uni-osnabrueck.de/3Dscans/ (12th dataset as of June 2012).

[70] Borrmann, D. and Elseberg, J. (June 2011) *University of Osnabrück and Jacobs University Robotic 3D Scan Repository, Knowledge-Based Systems Research Group, Institute of Computer Science, University of Osnabrück*. http://kos.informatik.uni-osnabrueck.de/3Dscans/ (7th dataset as of June 2012).

[71] Ziv, J. and Lempel, A. (1977) A universal algorithm for sequential data compression. *IEEE T. Inform. Theory*, **23**, 337–343.

[72] Roelofs, G. and Adler, M. (2009) *The ZLIB Homepage*. http://www.zlib.net/.

[73] Gailly, J. and Adler, M. (2003) *The GZIP Homepage*. http://www.gzip.org/.

[74] Kleder, M. (2005) *Rapid lossless data compression of numerical or string variables*. http://www.mathworks.com/matlabcentral/fileexchange/8899.

[75] Hopkins, J. (2009) *Compression routines*. http://www.mathworks.com/matlabcentral/fileexchange/25656-compression-routines.

[76] Pennebaker, W.B. and Mitchell, J.L. (2004) *JPEG Still Image Data Compression Standard*. Kluwer Academic Publishers, The Netherlands.

[77] Dobrucalı, O. and Barshan, B. (2010) A Compression Method based on Compressive Sampling for 3-D Laser Range Scans of Indoor Environments. In Gelenbe, E., Lent, R., Sakellari, G., Sacan, A., Toroslu, H. and Yazici, A. (eds), *Lecture Notes in Electrical Engineering (Proc. ISCIS 2010: 25th Int. Symp. Comp. Inf. Sciences)*, Vol. 62, ch. 7. Springer, Berlin, Germany.

[78] Strang, G. and Nguyen, T. (1997) *Wavelets and Filterbanks*. Wellesley-Cambridge Press, Wellesley, MA, USA.

[79] EDF Research and Development (2012) *CloudCompare—3-D point and mesh processing software*. http://www.danielgm.net/cc/ [www.danielgm.net].

[80] Dobrucalı, O. and Barshan, B. (2011) A compression method for 3-D laser range scans of indoor environments based on compressive sensing. *Proc. 19th European Signal Processing Conf.*, Barcelona, Spain, pp. 1115–1119, August 29–September 2.