

Compressed Gated Range Sensing

Grigorios Tsagkatakis^a, Arnaud Woiselle^b, George Tzagkarakis^c,
Marc Bousquet^b, Jean-Luc Starck^c, Panagiotis Tsakalides^a

^a ICS - Foundation for Research & Technology - Hellas (FORTH), Crete, Greece

^b Sagem (Safran), 27 rue Leblanc, 75512 Paris Cedex 15, France.

^c Commissariat à l'Énergie Atomique (CEA), Centre de Saclay, F-91191 Gif-Sur-Yvette cedex,
France

ABSTRACT

Active Range Imaging (ARI) has recently sparked an enthusiastic interest due to the numerous applications that can benefit from the high quality depth maps that ARI systems offer. One of the most successful ARI techniques employs Time-of-Flight (ToF) cameras which emit and subsequently record laser pulses in order to estimate the distance between the camera and objects in a scene. A limitation of this type of ARI is the requirement for a large number of frames that have to be captured in order to generate high resolution depth maps. In this work, we introduce *Compressed Gated Range Sensing* (CGRS), a novel approach for ToF-based ARI that utilizes the recently proposed framework of Compressed Sensing (CS) to dramatically reduce the number of necessary frames. The CGRS technique employs a random gating function along with state-of-the-art reconstruction in order to estimate the timing of a returning laser pulse and infer the depth map. To validate our method, software simulations were carried out using a realistic system model. Simulated results suggest that low error reconstruction of a depth map is possible using approximately 20% of the frames that traditional ToF cameras require, while 30% sampling rates can achieve very high fidelity reconstruction.

1. INTRODUCTION

Active Range Imaging systems employ an active illumination source, typically a laser or a light emitting diode (LED), and an imaging sensor in order to generate a 2D depth map of a scene. Time-of-Flight cameras, a very efficient type of an ARI system, perform depth extraction by measuring the time it takes for light pulses, traveling at the speed of light, to propagate from the source to an object and back to the sensor. ToF systems enjoy large measurement ranges and offer high resolution depth images. The extracted depth map can be subsequently used for target detection, robot navigation, and surface modeling in numerous applications including remote sensing, gaming, security, search and rescue, and medical applications^{1,2}.

Currently, there are two broad categories of ToF cameras implementations, namely Continuous Wave Modulation (CWM)³ and Range Gated Imaging (RGI)^{4,5}. CWM range imaging operates by illuminating the scene with an appropriately modulated light source and recording the reflected light. In contrast to continuous wave, in RGI the light source emits a series of pulses and the sensor records the reflected pulses that correspond to a specific range of distances. For each specific range of distances, a 2D depth profile is reconstructed according to the time interval that the sensor was recording reflected pulses. This interval is independently controlled for each pixel by an electronic gate. To obtain a full depth map, multiple frames, equal to the required depth resolution have to be recorded. This technique is commonly referred to as the *sliding gate* approach.

The large number of frames that are necessary in order to extract a full resolution depth map is one of the major limitations of sliding gate based RGI technology. In this work, we propose a novel RGI technique that is able to significantly reduce the number of necessary frames without sacrificing the quality of the depth map.

Further author information: (Send correspondence to G.Ts.) E-mails: G.Ts.: greg@ics.forth.gr, A.W.: arnaud.woiselle@sagem.com, G.Tz.: georgios.tzagkarakis@cea.fr, M.B.: marc.bousquet@sagem.com, J-L.S.: jstarck@cea.fr, P.T.: tsakalid@ics.forth.gr

This work was funded by the IAPP CS-ORION (PIAP-GA-2009-251605) grant within 7th Framework Program of the European Community.

This goal is achieved by exploiting the sparsity of the reflected laser pulses with respect to the full resolution depth signals. More specifically, instead of employing a predefined gating scheme, we utilize a random gating mechanism. Reconstruction of the depth map is achieved by formulating the problem under the Compressed Sensing (CS) framework and employing tools from the CS reconstruction theory. To validate the merits of the proposed RGI scheme, extensive experiments and comparison with standard techniques were carried out on data from a high quality simulated system. The simulation was carefully designed in order to account for the numerous parameters that affect the behavior of an RGI system.

The rest of the paper is organized as follows. Section 2 provides an overview of current state-of-the-art methods in RGI. In section 3, we present an overview of Compressed Sensing while Section 4 discusses the theoretical modeling of depth signals. Section 5 presents our prototype system termed *Compressed Gated Range Imaging* (CGRI). Experimental results are presented in Section 5 and the paper concludes in Section 6.

2. STATE-OF-THE-ART

Traditional RGI systems extract the scene’s depth information by recording a large number of frames where each frame encodes reflected pulses, and thus depth information, from a specific range of distances, a technique generally known as *bracketed gating* or time-slicing⁶. Depth information extraction is achieved by opening the electronic gate of the camera only during the time it takes the laser pulse to propagate from the camera to an object at a specific range and back to the sensor. Thus, to obtain a full range sweep, multiple frames, in numbers equal to the depth resolution, are necessary. Although this technique is intuitive in its operation, it is limited by that fact that there is a need for collecting a large number of frames, many of which may be “empty” in the sense that no objects are located at the corresponding distance ranges. Figure 1(left) provides a visual illustration of the gate timing by indicating in red the distance ranges that correspond to each captured frame (measurement).

An alternative and more efficient approach for obtaining a depth map via an RGI technique is *gate coding*.⁷ In gate coding, the gate opening and closing pattern follows a specific sequence such that efficient encoding of the signals is achieved. In one such implementation, the depth signal is modeled as a binary valued vector where only a single element is non-zero, the one corresponding to the depth associated with that pixel. Given the depth resolution one wishes to achieve, the binary encoding scheme requires in the ideal, noise-free case, $k = \log_2(n)$ measurements to encode depth information with n -bin resolution. Despite the fact that this approach requires significantly less frames to be captured, it cannot support the case where multiple objects at different depths are associated with each pixel due to semi-transparent objects in the scene and it is more susceptible to noise. Figure 1(right) provides the corresponding visual illustration of the sampling pattern utilized by gate coding.

In our work we propose a novel approach in RGI design that exploits the characteristics of CS for the sampling and the recovery of the depth map. Utilization of the CS principles for depth map reconstruction has been also investigated under different design paradigms. CoDAC⁸ is an example of a prototype ARI system where a spatial light modulation is responsible for randomly encoding the emitted light which is subsequently captured by a single photodetector. The idea is an extension of the “single-pixel camera”⁹ concept into range imaging and offers numerous benefits such as the use of a single time-resolved photo-diode. However, unlike our proposed scheme, CoDAC requires a specific hardware setup and does not rely on existing ToF ranging hardware systems.

3. COMPRESSED SENSING

Compressed sensing (CS) is a novel approach in signal representation and sampling that was introduced by Donoho¹⁰ and by Candés et al.¹¹ The main concept of CS is that a signal can be recovered from a small number of random measurements, far below the Nyquist-Shannon limit, provided that the signal is sparse, *i.e.*, contains a small number of non-zero elements. Formally, let $\mathbf{x} \in R^n$ be a k -sparse signal such that $\|\mathbf{s}\|_0 < k$, where the zero norm counts the non-zero elements. This signal can be fully recovered from a low-dimensional representation $\mathbf{y} = \Psi\mathbf{x} \in R^m$, where Ψ is the $m \times n$ measurements matrix with $m \ll n$ by solving the following l_0 minimization:

$$\min \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \Psi\mathbf{x}. \quad (1)$$

An important issue regarding the formulation in Eq. (1) is that l_0 minimization is an NP-hard problem and therefore inefficient to solve for even moderate sized problems. To address this issue, greedy approaches such

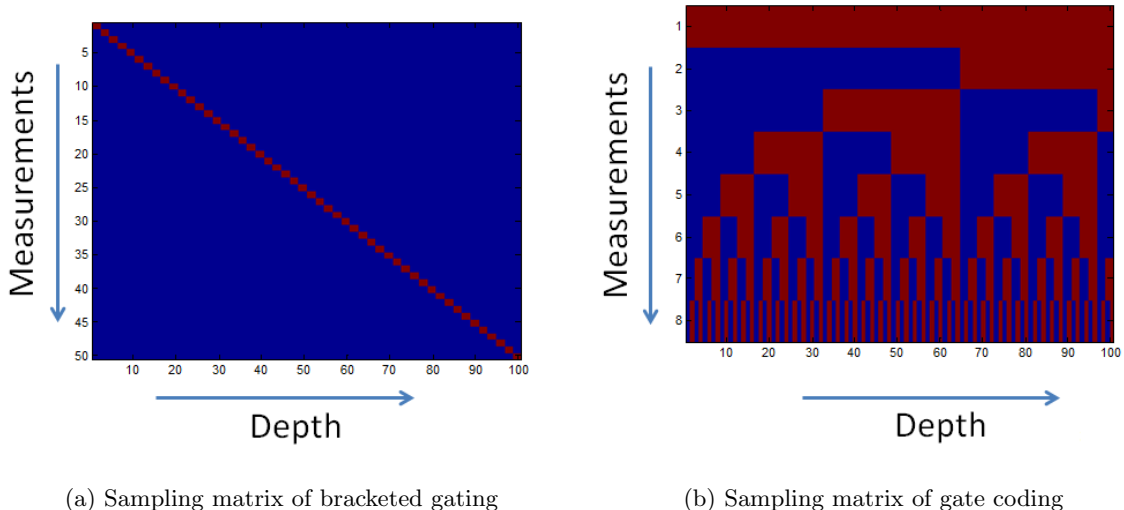


Figure 1: Illustrative examples of the sampling patterns in bracketing (a) and coding (b) for each sampling instance (rows) for multiple frames (columns). In the matrices shown, blue coded entries correspond to gate "close" while red to gate "open". One can see that in bracketing mode the gate is open for a single time interval during each frame, while in coding mode the gate may open and close several times during each frame.

as the Orthogonal Matching Pursuit (OMP)¹² algorithm have been proposed. OMP greedily tries to identify the elements that contain most of the signal energy. At each iteration of the OMP, the element that produces the higher absolute inner product with the residual is selected. This element is added to the collection and a new approximation of the signal is evaluated by projecting the input signal to the linear span of the selected elements. The new residual is found by subtracting the approximated signal from the input signal. The process continues until the residual is within an acceptable approximation limit.

Alternately, the CS theory suggests that for sufficiently sparse signals and for measurements matrices whose elements are drawn from an appropriate distribution and that satisfy the so called RIP property¹³, reconstructing the original signal \mathbf{x} from the measurements \mathbf{y} can be achieved by solving a much more efficient l_1 optimization problem given by:

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y} = \Psi \mathbf{x} \quad (2)$$

where the l_1 norm is defined as the sum of the absolute values of the elements in the vector. In order to acquire stable solutions, the number of random measurements should be greater than $O(k \log(n/k))$. The above optimization is called Basis Pursuit¹⁴ and can be solved in polynomial time. For noisy cases, or for approximated sparse signals, the problem in (2) can be relaxed to:

$$\min \|\mathbf{x}\|_1 + \lambda \|\mathbf{y} - \Psi \mathbf{x}\|_2 < \epsilon_1 \quad (3)$$

where ϵ_1 is a bound on the residual error of the approximation which is related to the amount of noise in the data and λ is the regularization parameter controlling the tradeoff between the sparsity and fidelity of the solution.

The CS model described in Eq. (1) assumes that the signal to be sampled is sparse *i.e.* only a few elements are non-zeros. However, the theory of CS has also been extended to signals that can be sparsely modeled in an appropriate dictionary of prototype elements called atoms^{11,15}. For strictly sparse signals, the minimization is given by:

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y} = \Psi \mathbf{D} \mathbf{x} \quad (4)$$

where \mathbf{D} is the, possibly redundant, dictionary of elements. Similarly to Eq. (2), for the noisy case, an approximate formulation is given by:

$$\min \|\mathbf{x}\|_1 + \lambda \|\mathbf{y} - \Psi \mathbf{D} \mathbf{x}\|_2 < \epsilon_2. \quad (5)$$

The design of the dictionary is a critical issue that is dictated by the signal model and it can be either parametric, such as when using DCT and Wavelet bases, or extracted from the data via a dictionary learning stage.

4. SYSTEM MODEL

In order to obtain a realistic performance evaluation, special care is given to the modeling of the system. In our model, we assume that each pixel is associated with a single reflected pulse, corresponding to a single non-transparent object at a specific depth. We approximate such as depth signal as a temporal delta function with non-zero value only at the time instance that corresponds to the round-trip time of the laser pulse from the source to the target and back and is given by:

$$\mathbf{s}_{init}(z) = \delta\left(t - \frac{2z}{c}\right) = \delta_\tau \quad (6)$$

where $\tau = \frac{2z}{c}$ encodes the depth of the object. This *clean* depth signal propagates and thus interacts with the atmosphere. This interaction between the laser photons and the molecules in the atmosphere also contributes to the reflected pulse and it is manifested as signal attenuation.

First, we consider the effects of backscatter that describes the amount of light that is scattered by the atmosphere into the backward direction. More specifically, the effect of the backscatter is modeled as a constant value across all distances that is added to the clean depth signal, *i.e.*, $\mathbf{s}(z) = \mathbf{s}_{init}(z) + c\mathbf{I}$. This model corresponds to relatively clear and isotropic atmospheric conditions. Furthermore, part of the laser pulse's power is absorbed by the molecules in the atmosphere resulting in a lower power reflected signal. The atmospheric absorption is modeled following the Lambert-Beer-Bouguer law described by $\mathcal{A}(z) = e^{(-2z/\alpha)}$, where α is the atmospheric transmission coefficient. Last, an additional power reduction caused by the geometry of the laser beam's divergence is modeled as $\mathcal{D}(z) = 1/z^2$.

To realistically model the behavior of the gate, we assume that the gate function can be expressed as the convolution of an ideal gate function, modeled by a rectangular function $\mathbf{g}_{ideal}(t) = \Pi(t + t_k)$, with a filtering function $\mathbf{f}_g(t) = e^{\frac{t}{t_{gate}}}$ ($t \geq 0$) encoding the characteristics of the sampling process, that is $\mathbf{g}(t) = \mathbf{g}_{ideal}(t) * \mathbf{f}_g(t)$. Similarly, the laser pulse is modeled as the convolution of the ideal pulse $\mathbf{p}_{ideal}(z) = \mathbf{s}(z)\mathcal{A}(z)\mathbf{d}(z)$ with a similar filter with parameter t_{pulse} giving $\mathbf{p}(t) = \mathbf{p}_{ideal}(t) * \mathbf{f}_p(t)$.

The noise-free signal that reaches the sensor corresponds to the multiplication of the pulse function with the gate function defined before, *i.e.* $\mathbf{l}_{ideal}(t) = \mathbf{p}(t)\mathbf{g}(t)$. The final noisy signal captured by the sensor also encodes the contribution of natural sources of noise and is given by $\mathbf{l}(t) = \mathbf{l}_{ideal}(t) + \mathbf{n}(t)$. The characteristic differences between an ideal depth signal and the one modeled by our system is graphically presented in figure 2.

5. COMPRESSED GATED RANGE SENSING (CGRS)

Following the discussion in the previous section, in this work we exploit an important characteristic of the depth signal $\mathbf{s}_{init}(z)$, its temporal sparsity, in order to provide an efficient sampling mechanism. More specifically, the depth signal is modeled as a vector where the number of elements is given by the depth resolution. Each element of this vector encodes the amount of light captured by the sensor during a specific time interval, which also corresponds to a specific depth range. In our model, we assume that the full resolution depth signal for each pixel will consist primarily of zeros, expect for a small number of non-zero values, corresponding to the temporal locations of the reflected pulses. Formally, we propose the application of a CS based sampling scheme where the recorded value at each pixel is given by $\mathbf{y} = \mathbf{G}\mathbf{s}$. In this scenario, the gating function $\mathbf{G} = \{\mathbf{g}_t \| t = 0, \dots, m\}$ serves as the sampling matrix Ψ .

A limitation of such a straightforward approach based on CS sampling is that it does not consider the additional information regarding the signal characteristics such as attenuation due to distance and geometry. To introduce prior information regarding the signal characteristics, an appropriate dictionary must be considered. In this work, we propose the construction of such a dictionary that will capture properties such as depth dependencies and backscatter.

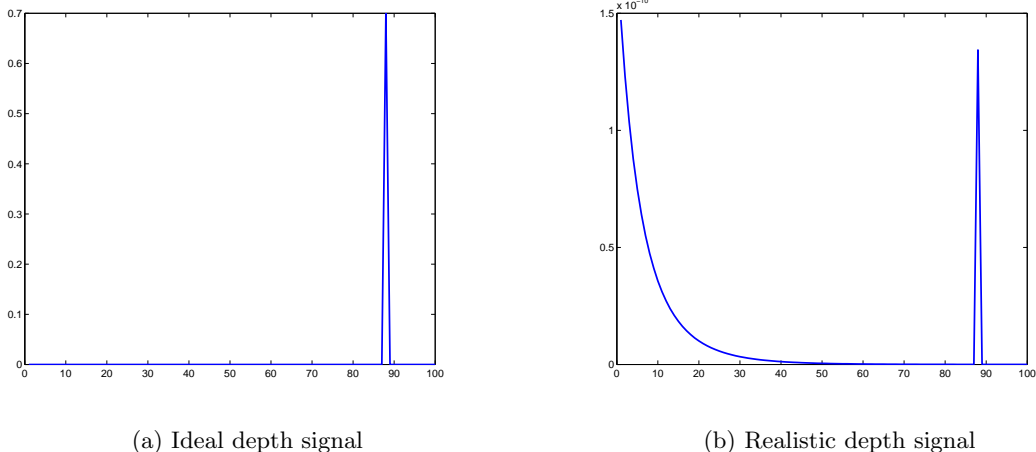


Figure 2: Graphical illustration of an ideal (a) and the corresponding realistic (b) depth signal. In the realistic signal model, one is able to observe phenomena such as back propagation and the effects of the convolution with the gate. Such effects make the process of identification of the true underlying signal model a challenging task.

Formally, this dictionary is modeled by $\mathbf{D} = [\mathbf{I} \ \mathbf{1}] \times \mathcal{A}(z)\mathcal{D}(z)$. For the generation of the dictionary, we consider \mathbf{I} , which is an appropriately sized identity matrix and $\mathbf{1}$, an all-ones vector, in an effort to encode the depth signal and the backscatter effect correspondingly. The matrix composed of the concatenation of the identity matrix with the all-ones vector is further multiplied by $\mathcal{A}(z)\mathcal{D}(z)$ which corresponds to the transmission and beam divergence effects. By employing such a dictionary, the requested sparse vector will only encode the depth information, while other effects are encoded in the dictionary. Reconstruction of the depth signal of pixel (i, j) is given by:

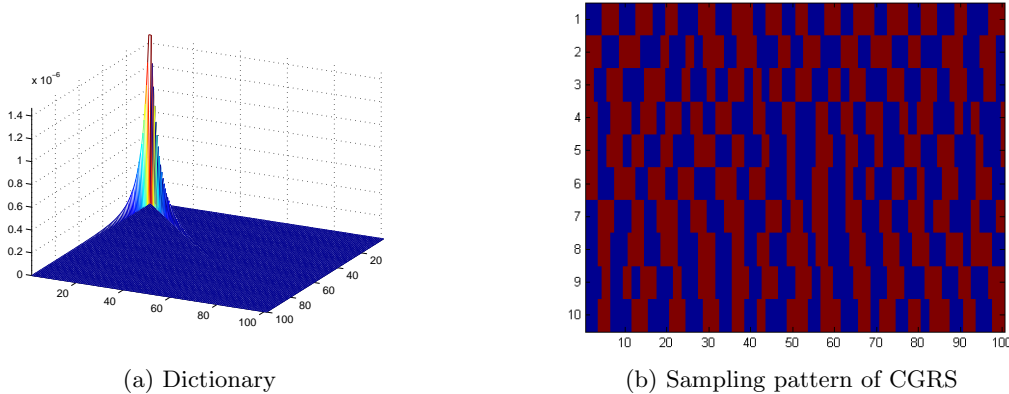
$$\hat{\mathbf{s}}_{i,j} = \arg \min_{\mathbf{s}} \|\mathbf{s}\|_1 + \lambda \|\mathbf{y} - \mathbf{G}\mathbf{D}\mathbf{s}\|_2^2. \quad (7)$$

which can be solved using l_1 or Matching Pursuit type algorithms. Figure 3 provides a graphical illustration of the dictionary (left) and the sampling pattern of the gate (right). Although CS theory suggests that the sampling pattern should be completely random, physical constraints limit the randomness of the gating function. In our model, the gate switch from the open to the close state and vice versa every 100 to 200 milliseconds which corresponds to the duration of 2 to 4 pulses. Furthermore, a 5 milliseconds respond time was also introduced as the t_{gate} parameter in the filtering process that produced the realistic gate signal.

6. EXPERIMENTAL RESULTS

To validate the merits of the proposed *Compressed Gated Range Sensing* (CGRS) scheme, we report results on the reconstruction of a synthetically generated scene and compare the performance with two state-of-the-art techniques, namely the classical bracketed gating and the gate coding method. In all cases, the gate duration was set to 100 ns, while the pulse duration to 50 ns for extracting the depth in the range of 500 to 3500 meters with 300 meters depth resolution. In order to account for noise sources that are involved in the signal sampling, we considered scenarios with high SNRs (30dB) and low SNR (20dB). Furthermore, the backscatter coefficient was set to 10^{-5} , and the transmission coefficient α was set to 10^3 .

The ground truth range image is presented in Fig. 4 (left), while Fig. 4 (right) provides a numerical evaluation of the performance, expressed by the RMSE, as a function of sampling rate, where the sampling rate is the normalized ratio between the number of acquired images and the number of images necessary for full resolution depth sampling (100 frames). The figure presents the performance characteristics for the bracketing coding, the gate coding and the proposed CS based approach. The plot clearly indicates the superior performance of CS over the other competing methods, especially in the low sampling ratio regime. Furthermore, we observe that



(a) Dictionary

(b) Sampling pattern of CGRS

Figure 3: Visualization of the dictionary (a) and the sampling pattern (b) of *CGRS*. The two main components of the dictionary are evident, while we observe the random sampling pattern and its distinction from the deterministic designs in Fig. 1.

while CS and gate coding offer a graceful degradation with lower sampling rates, bracketing coding experiences a more dramatic reduction in performance.

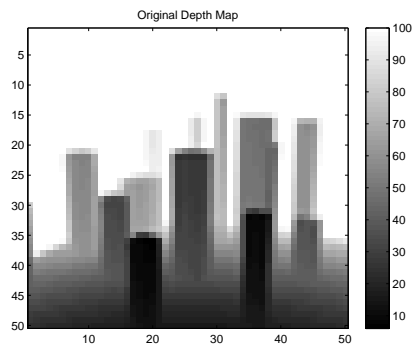
The reconstruction performance can also be seen in Figures 5-8 where some indicative reconstructed depth maps are presented. In these results one can make various observations regarding the performance of each method on different conditions. First, we observe that the classical time-slicing method is more heavily influenced by the low sampling rate and exhibits a confusion on the distance estimation by identifying longer distances, such as the one in the background, as shorter distances.

7. CONCLUSIONS

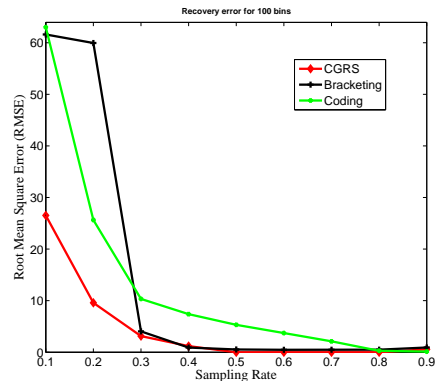
In this paper, we presented a novel design for Time-of-Flight based Active Range Imaging systems that employs the concepts of the recently developed framework of Compressed Sensing. The proposed scheme offers high quality reconstruction while it requires a significantly less number of frames to be captured compared to traditional bracketed gating. This goal is achieved by employing a random gating function that captures multiple reflected laser pulses from multiple distances during each integration period. The true depth signal is obtained by performing a decoding of the sampled measurements via the sparse coding framework. Furthermore, in order to capture light propagation effects such as backscattering, we employ an appropriately designed dictionary of prototype elements. Experimental results on realistically designed simulations suggest that Compressed Sensing based sampling and reconstruction is more efficient than classical bracketed gating and the gate coding scheme and offers higher robustness to noisy environments.

REFERENCES

- [1] Sansoni, G., Trebeschi, M., and Docchio, F., “State-of-the-art and applications of 3d imaging sensors in industry, cultural heritage, medicine, and criminal investigation,” *Sensors* **9**(1), 568–601 (2009).
- [2] Blais, F., “Review of 20 years of range sensor development,” *Journal of Electronic Imaging* **13**(1) (2004).
- [3] Lange, R. and Seitz, P., “Solid-state time-of-flight range camera,” *Quantum Electronics, IEEE Journal of* **37**(3), 390–397 (2001).
- [4] Busck, J. and Heiselberg, H., “High accuracy 3d laser radar,” in [*Defense and Security*], 257–263, International Society for Optics and Photonics (2004).
- [5] Steinvall, O., Andersson, P., Elmqvist, M., and Tulldahl, M., “Overview of range gated imaging at foi,” in [*Defense and Security Symposium*], 654216–654216, International Society for Optics and Photonics (2007).
- [6] Bonnier, D. and Larochelle, V., “A range-gated active imaging system for search and rescue, and surveillance operations,” in [*Proc. SPIE*], **2744**, 134–145 (1996).



(a) Original depth map



(b) Reconstruction error

Figure 4: Original depth map (a) and reconstruction error as a function of sampling rates (number of acquired frames) for the three techniques (b). Results suggest that CGRS achieves the best performance across all sampling rates while for lower sampling rates the gate coding approach is more reliable than classical bracketed gating.

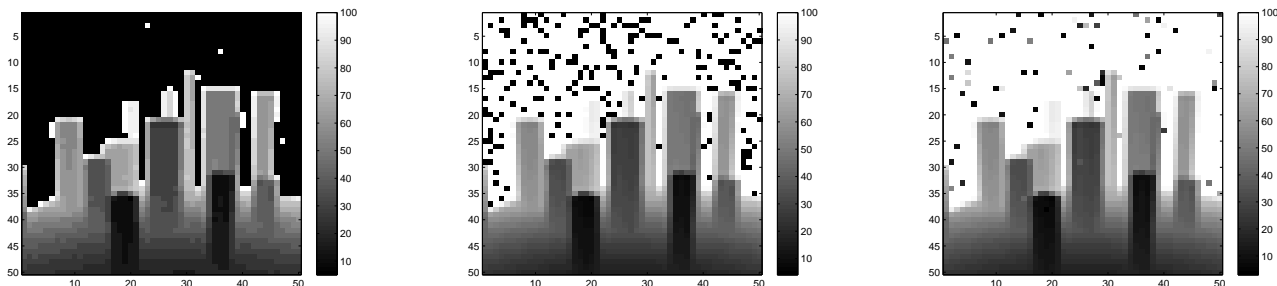


Figure 5: Illustrative example of the classical (left), coded (middle) and random (right) gating based reconstruction from 20 frames (20% sampling rate) in a high SNR scenario (30dB). The reconstruction performance is given by the RMSE equal to 63.8, 27.6 and 11.2 correspondingly. We observe that propagation effects result in low SNR for long range signals that create a depth folding (longer distances appearing as shorter distances) in classical and coded gating.

[7] Laurenzis, M. and Bacher, E., “Image coding for three-dimensional range-gated imaging,” *Applied Optics* **50**(21), 3824–3828 (2011).

[8] Kirmani, A., Colaço, A., Wong, F., and Goyal, V., “Codac: A compressive depth acquisition camera framework,” in [*Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*], 5425–5428, IEEE (2012).

[9] Duarte, M., Davenport, M., Takhar, D., Laska, J., Sun, T., Kelly, K., and Baraniuk, R., “Single-pixel imaging via compressive sampling,” *Signal Processing Magazine, IEEE* **25**(2), 83–91 (2008).

[10] Donoho, D., “Compressed sensing,” *Information Theory, IEEE Transactions on* **52**(4), 1289–1306 (2006).

[11] Candes, E., Eldar, Y., Needell, D., and Randall, P., “Compressed sensing with coherent and redundant dictionaries,” *Applied and Computational Harmonic Analysis* **31**(1), 59–73 (2011).

[12] Tropp, J. and Gilbert, A., “Signal recovery from random measurements via orthogonal matching pursuit,” *Information Theory, IEEE Transactions on* **53**(12), 4655–4666 (2007).

[13] Baraniuk, R., Davenport, M., DeVore, R., and Wakin, M., “A simple proof of the restricted isometry property for random matrices,” *Constructive Approximation* **28**(3), 253–263 (2008).

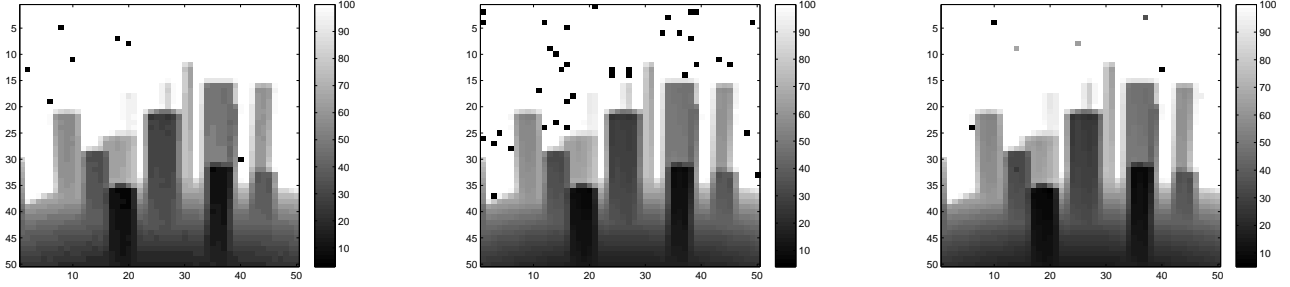


Figure 6: Illustrative example of the classical (left), coded (middle) and random (right) gating based reconstruction from 30 frames (30% sampling rate) in a high SNR scenario (30dB). The reconstruction performance by the RMSE is equal to 5.2, 11.6 and 3.6 correspondingly. We observe that higher sampling rates typically provide a relatively high quality reconstruction.

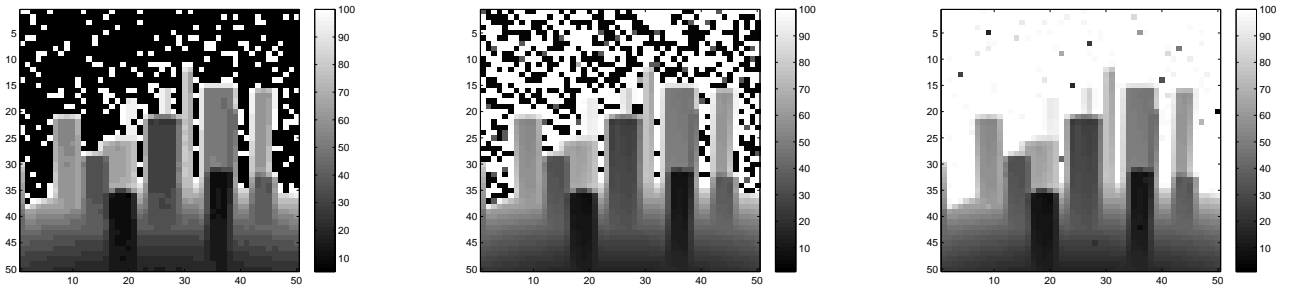


Figure 7: Illustrative example of the classical (left), coded (middle) and random (right) gating based reconstruction from 20 frames (20% sampling rate) in a noisy environment (20dB). The corresponding reconstruction errors are 56.7, 41.3 and 4.5 correspondingly.

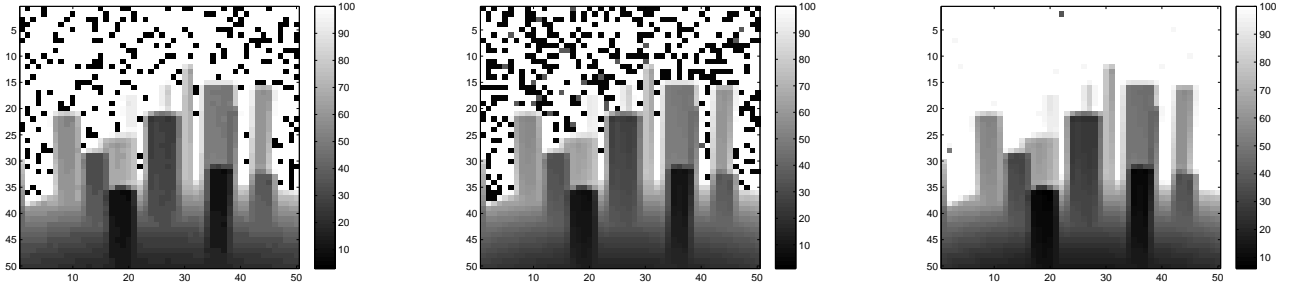


Figure 8: Illustrative example of the classical (left), coded (middle) and random (right) gating based reconstruction from 30 frames (30% sampling rate) in a noisy environment (20dB). The corresponding reconstruction errors are 29.3, 36.7 and 1.5 correspondingly.

- [14] Chen, S. S., Donoho, D. L., and Saunders, M. A., “Atomic decomposition by basis pursuit,” *SIAM journal on scientific computing* **20**(1), 33–61 (1998).
- [15] Rauhut, H., Schnass, K., and Vandergheynst, P., “Compressed sensing and redundant dictionaries,” *Information Theory, IEEE Transactions on* **54**(5), 2210–2219 (2008).