COMPRESSION FOR STORAGE AND TRANSMISSION OF LASER RADAR MEASUREMENTS

Joseph C. Dagher, Michael W. Marcellin, and Mark A. Neifeld

Optical Sciences Center and,
Department of Electrical and Computer Engineering,
The University of Arizona, Tucson, AZ 85721

ABSTRACT

We develop novel methods for compressing volumetric imagery that has been generated by single platform (mobile) range sensors. We exploit the correlation structure inherent in multiple views in order to improve compression efficiency. We show that, for lossy compression, three-dimensional volumes compress more efficiently than two-dimensional (2D) images. In fact, our error metric suggests that accumulating more than 9 range images in one volume before compression yields up to a 99% improvement in compression performance over 2D compression.

Keywords: Compression, Laser RADAR, Volumetric Imagery, Range Imagery

1. INTRODUCTION

The existence of volumetric imagery is a natural consequence of distributed imaging and is a significant potential benefit of the sensor-network paradigm. Volumetric image data has become important for many commercial, scientific, and military applications. There are numerous sensor modalities (e.g., PET, MRI, Interferometry) from which it is possible to produce volumetric imagery.

With the increasing importance of three-dimensional (3D) imaging, efficient storage and transmission of the large resulting data sets is necessary. There exist several two-dimensional (2D) compression algorithms that perform well when treating the volume as multiple 2D slices. These include JPEG2000 [1–3] and the context-based adaptive lossless image codec (CALIC) [4] . These algorithms exploit the correlation structure inherent in the 2D slices. However, it is also important to consider the volumetric correlation structure that is naturally available in 3D. Several methods that utilize dependencies in all three dimensions have been proposed [5–8] . Unfortunately however, most efforts have been devoted to the development of algorithms for compressing "intensity volumes" while little work has been done for the compression of range imagery. Moreover, these efforts are generally based on highly iterative geometric and topological lossy compression schemes, which makes them less suitable for real-time applications [9,10] .

This paper is concerned with efficient methods for the compression of volumetric imagery. We will focus primarily on volumetric data that have been generated by single platform range sensors (e.g., ladar); however, many of the approaches will be extensible to the more traditional intensity sensor data. We analyze the compression performance of two-dimensional range images as well as three-dimensional volumes. We show that 3D volumes are easier to compress than a collection of 2D range images.

The paper is organized as follows: Section 2 presents an overview of laser range imaging as well as a description of the physical setup used in our simulations. Compression performance of 2D range images using JPEG2000 is evaluated in Section 3, while Section 4 analyzes the compression performance of volumetric imagery. Concluding statements are presented in Section 5.

Further author information: (Send correspondence to J.C.D.)
E-mail: joseph@ece.arizona.edu, Telephone: 1 520 621 9769, Address: Department of Electrical and Computer Engineering, 1230 E. Speedway Blvd, Tucson, Az, 85721

2. SENSOR GEOMETRY AND PHYSICAL SETUP

2.1. Laser Range Imaging

Laser radar (ladar) is an imaging modality that employs pulsed laser illumination. The light scattered back from objects in the scene is detected and a time-of-flight measurement determines the distance at which reflection occurs according to $D = cT/2$, where T is the measured time-of-flight and c is the speed of light. (Here we assume that objects are non-transparent to the laser illumination.) Compact solid-state laser systems are becoming capable of sufficient pulse energy to facilitate laser ranging to distances of more than 150 m, resulting in a range record of roughly 150m/15cm = 1,000 bins. Such a system might scan over an array of more than 1000×1000 angles in order to obtain an acceptable field of view. With a laser repetition frequency of 1MHz, such an image measurement will take 1 second and will generate more than 10^9 data values. The need for data compression within such a sensor environment is clear.

2.2. Physical Setup

Figure 1 shows an overview of the sensor geometry that we will study. We are interested in a single platform ladar sensor and the sensor is assumed to make a series of range measurements at ninety equally spaced locations along a straight line 30.48 m above the scene (i.e., the sensor follows the line $y = 0$, $z = 30.48$m). The distance between the first and last measurement location is 60.96 m and the distance between two consecutive measurement locations is constant and is equal to 0.684 m. The sensor operates in a scanning configuration and, at each of the ninety measurement locations, generates a 1500×1500 array of range measurements (a range image). Only the strongest return is retained along each line-of-sight.

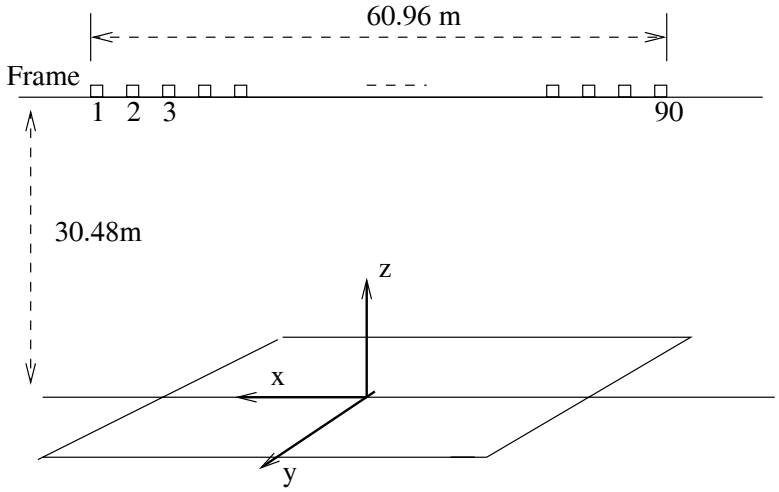


Figure 1. Geometry describing the mobile ladar sensor under study. Sensor scans scene with uniform angular step sizes from 90 positions.

Figure 2 shows four candidate representations of simulated data acquired from a single viewing location. The measured data spans the range of 22m to 116m. Figure 2a is a pictorial representation of this range data. Each pixel in Figure 2a is the range value (a real number that was quantized to an 8-bit value for display purposes) at which the maximum return was received. Thus, large and small range values are denoted by bright and dark pixels, respectively. The number of pixels in this image (1500×1500) is equal to the number of scan angles that were used during the measurement process.

A convenient alternate representation of the data from Figure 2a is provided by the "z-image" of Figure 2b. This image is obtained by using the known sensor geometry to compute the height above the ground (Z-coordinate) of each range measurement depicted in Figure 2a. This representation is more visually pleasing and demonstrates the expected pixel correlation structure that a typical visible image might display.

A third potential range representation is shown in Figure 2c. In this figure we have generated a volumetric representation of a single range image. A black dot is placed at each point in 3-space from which a ladar return was received. This strategy produces the so-called "point-cloud" shown in Figure 2c and facilitates the convenient generation of alternate viewing perspectives as shown in Figure 2d. This representation also enables visualizing accumulated data acquired from multiple sensor locations on the same plot. In subsequent sections we will investigate the utility of these and other data representations toward the goal of efficient volumetric data compression.

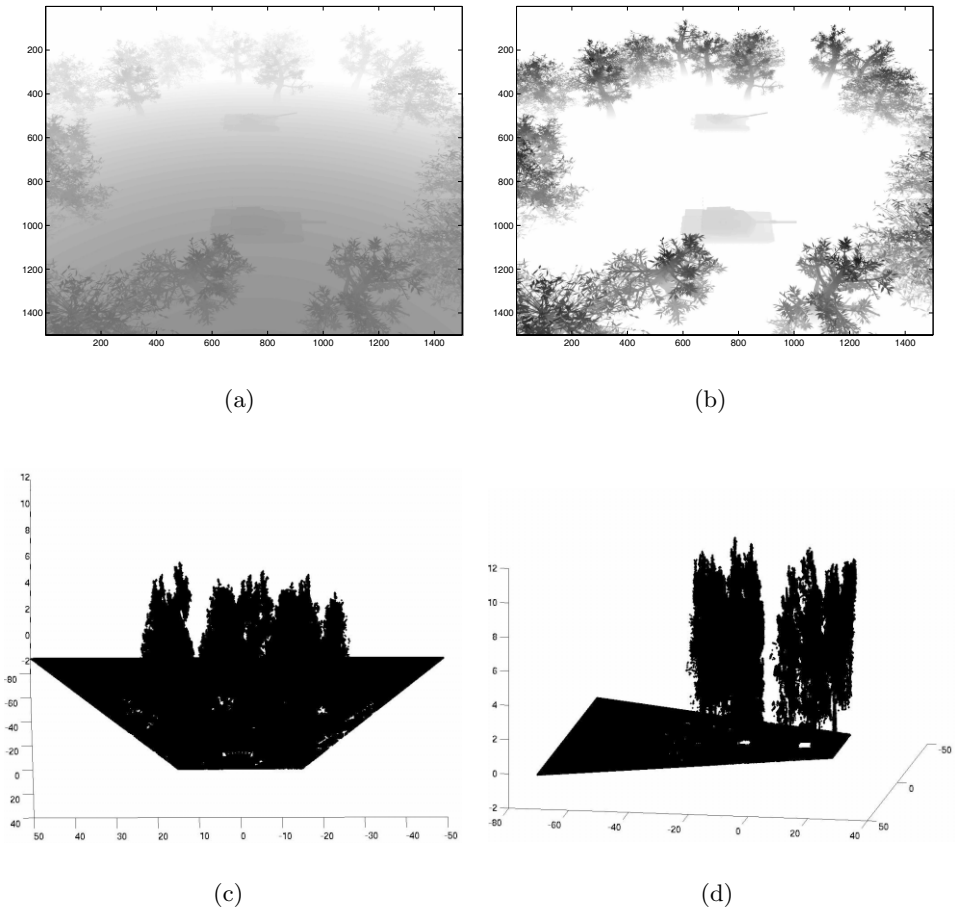


Figure 2. Candidate representations of a single range image. (a) Range of strongest return depicted as grayscale pixels at 8bpp. (b) Elevation-only data extracted from (a). (c) Data from (a) rendered as a volume. (d) Same as (c) but rendered from a different perspective.

3. 2D RANGE IMAGE COMPRESSION

As indicated in Section 2, there exist various ways to represent a range measurement from a ladar sensor. In this section, we investigate the compression performance of 2-space representations of range imagery as depicted in Figure 2a. We propose two approaches for compressing measurements generated by the single platform ladar sensor setup described in Section 2.

In the first approach, each sensor measurement is treated individually, i.e. range images are independently compressed and transmitted. This exploits the correlation structure within each image. The other approach is to treat the set of sensor measurements as a sequence of images. This second approach aims to increase compression efficiency by exploiting the correlation structure between images.

3.1. Independent, Frame-by-Frame

3.1.1. Range Pre-processing and Compression

We start with the grayscale range representation discussed in Section 2, and shown for three sensor positions in Figure 3. Although the images of Figure 3 are printed using 8 bits per pixel (bpp), the floating-point range values were digitized to 16 bits per range pixel for compression purposes. It should be noted that there is some distortion introduced with this digitization process. We will treat this subject in more detail in Section 3.1.2. However, it suffices to mention here that the average RMSE introduced by digitizing to 16 bits per range pixel is approximately 1 millimeter.

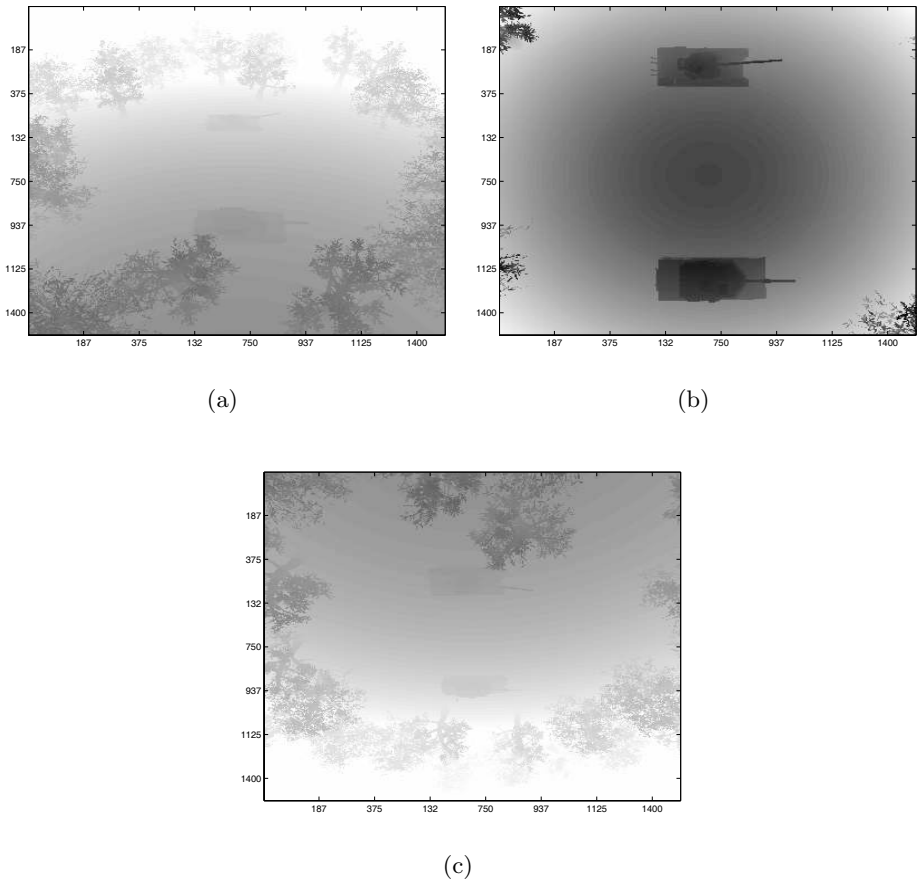


Figure 3. Sample range images associated with different sensor positions. (a) Range Measurement from sensor position 1. (b) Range Measurement from sensor position 45. (c) Range Measurement from sensor position 90.

Figure 4 shows the RMSE resulting from lossy compression of thirteen of the range images using JPEG2000 [1] . The 13 views are chosen in an equally spaced fashion from the possible 90 different positions. For each rate, we indicate in the figure the average, μ, and the variances, σ^2, of the RMSE values. Each image in the figure is independently compressed to the rate denoted. For example, when the first image is compressed/decompressed using 0.2 bits per range pixel, the resulting RMSE is approximately 1.5 meter.

It is interesting to note the symmetry in the results of Figure 4. Specifically, the RMSE for image 1 is roughly equal to that of image 13. The situation for images 2 and 12 is similar, and so on. This is due to the symmetry of the sensor geometry. Specifically, the range sensor acquires each one of those "pairs" of images at the same slant angle, but while at opposite ends of the scene. Image 6 corresponds to nadir viewing conditions. Another point to note from Figure 4 is that, for a given bit-rate, the high slant angle imagery is more difficult to compress

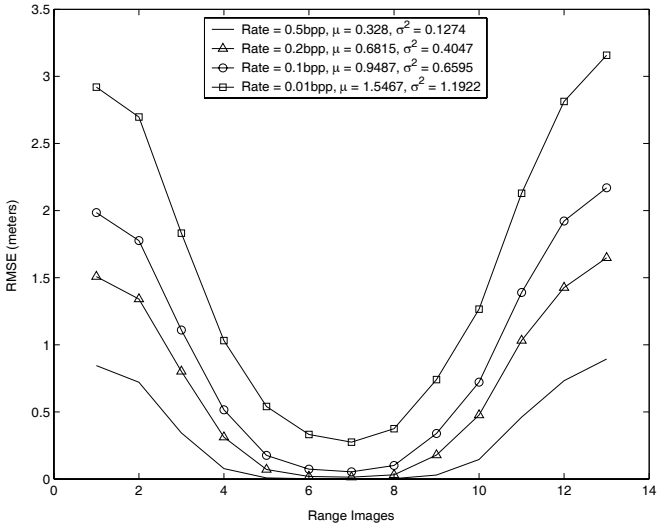


Figure 4. RMSE for independent lossy compression of range images using JPEG2000.

(larger RMSE) than the nadir imagery. A direct consequence of this is the high variances of the RMSE curves, especially at lower rates. To justify this phenomenon, three reasons have been proposed and studied. Due to page limit constraints, we only list the reasons here. For a more thourough discussion, the interested reader is referred to [11] .

Reason #1: The high slant angle imagery of Figure 3 contain more objects that are hard to compress (e.g., trees) compared to the nadir range imagery.

Reason #2: Object positions that are far apart do not exhibit much correlation and thus do not compress as well as near-by object positions.

Reason #3: Looking back at Figure 3 we also notice that images measured from off-nadir locations show a trend in range values going from bright (i.e., farther) to dark (i.e., nearby). This linear trend increases the dynamic range of the pixel values and makes the images harder to compress. In particular, off-nadir images possess range values with a large dynamic range compared to that of nadir images (95 m vs. 10 m).

3.1.2. Resolution Specific Compression

Recall from Section 2 that our ladar sensor collects range measurements that are accurate down to some resolution Δ meters. This range resolution determines the eventual file size that would be required to store the uncompressed image. Specifically, we require B bits to store one measurement, where B is given by

$$B = \lceil \log_2(\frac{r_{max} - r_{min}}{\Delta}) \rceil, \tag{1}$$

r_{max} and r_{min} being the maximum and minimum possible range values that could be detected, respectively.

It is apparent from this equation that the coarser the resolution is, the fewer the number of bits that are required to represent the range value. Now, assume that we are to store the range measurements using a bitdepth b, less than B. This might be achieved by using a coarser resolution, say $\delta > \Delta$.

One implementation would quantize all the range measurements down to resolution, δ using a uniform quantizer. The resulting RMSE between the original and the lower resolution image would be [12] ,

$$RMSE = \sqrt{\delta^2/12}. \tag{2}$$

One final step might be to compress the uniformly quantized image *losslessly*. Let R be the average number of bits per pixel required for such compression.

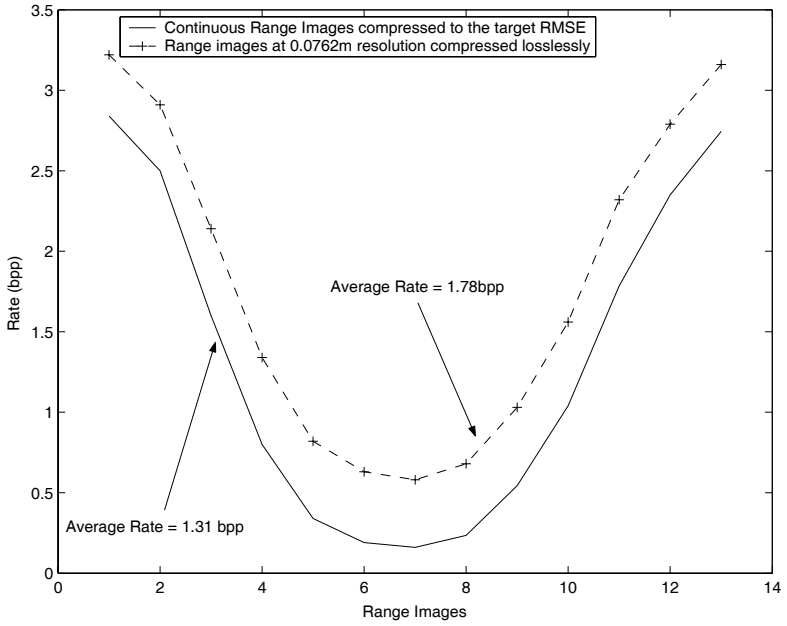


Figure 5. Comparison of "lossless" vs. lossy compression of range imagery.

Another implementation might apply *lossy* compression directly to the original data to achieve the same $RMSE = \sqrt{\delta^2/12}$. We illustrate a comparison of these two approaches in Figure 5. For this figure, continuous simulated range data was directly quantized to a resolution of $\delta = 0.0762$ m, yielding 11-bit integer range representations. Each of the the range images was so treated, and then compressed losslessly. The bit-rate required (in bits per range pixel) is shown for each image by the dotted curve in the figure. Next, each continuous simulated range image was compressed using sufficient rate to achieve the same *average* RMSE = 0.022 m, obtained using Equation (2). The required bit rates for this procedure are shown using the solid line. It can be seen that the rates (and eventually the file sizes) obtained using lossy compression are on average 26.5% smaller than those obtained by the lossless compression method for the same RMSE. Also we can see that for the target spatial resolution, a lossless compression efficiency up to 70:1 can be achieved (9:1 on average). It is worth noting however, that in the uniform case, the maximum range error corresponding to any given pixel is limited to $\pm\delta/2 = 0.0381$ m. This is not necessarily true in the lossy case.

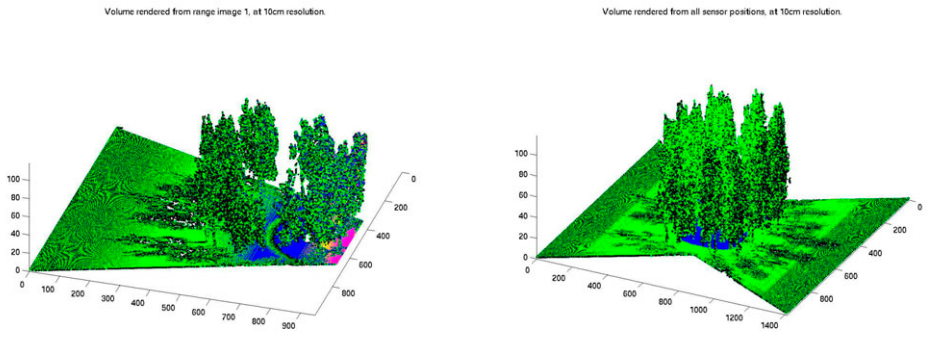
4. 3D VOLUMETRIC LADAR COMPRESSION

Throughout this paper, we deal with multiple range images acquired and processed by a single mobile sensor. Such a sensor may acquire many views of a single scene. Of course, the simplest strategy is just to compress and downlink each range image independently as it is acquired. Alternately, multiple views may be collected prior to compression and downlink. In this section, we present 3D representations of range measurements that yield significant compression improvements compared to the traditional 2D approaches that we discussed in previous sections.

4.1. 3D Representations

As noted in Section 2, a range image may be converted to a 3D volume by plotting a "point" at every location in space that produces a ladar return. Examples of such "point cloud" representations are given in Figures 2c and 2d. One of the advantages of such a representation is the ability to plot multiple range images on the same 3D volume thus generating a point cloud containing the aggregate information from all corresponding views. Several representations of point clouds are possible. For example, an unstructured list of all (x, y, z) triples, resulting from all views, will suffice. Another representation is the histogram 3D volume shown in Figure 6. This representation is suggested naturally by the point cloud figures discussed above. When plotting point clouds, (x, y, z) can be

(at least conceptually) continuous. However, for the purpose of representation or storage/transmission of 3D data cubes, discretization of (x, y, z) is often required. In other words, for a given 3D "bin" or "voxel" size of $(\Delta_x, \Delta_y, \Delta_z)$, the value of the histogram $h(x_0, y_0, z_0)$ is set equal to the number of returns occuring within the voxel centered at (x_0, y_0, z_0). Thus, a histogram volume will tend to a point cloud equivalent representation as the voxel size tends to zero. Figure 6a shows a plot of the 3D histogram volume corresponding to Figure 2d for a voxel size of 10 cm on each side.



(a) (b)

Figure 6. 3D volumetric representation. (a) data from Figure 2d quantized to $10cm \times 10cm \times 10cm$ spatial resolution, viewed from a different angle. (b) Same as (a) but rendered using all data from 13 viewing positions.

Histogram volumes give the distribution of returns within the space of the scene, and may be useful in many applications. An area of interest could be determined by knowledge of the reflectivity of a region. This is important in applications such as target recognition, multi-view compression in a multiple sensor framework, etc. Moreover, this reflectivity could be integrated across multiple viewing positions. Figure 6b shows a 3D histogram volume generated from 13 different views. These are the same equally spaced views used in previous sections.

Another potentially useful representation is the binary 3D volume. This representation can be derived from the histogram 3D volume described above. It simply consists of replacing every non-zero voxel by 1, which is equivalent to placing a 1 at any location where a return is indicated in one or more range images.

4.2. Volumetric Compression

Lossy compression of 3D histogram volumes is also of interest. However, in the case of volumetric compression, the choice of quality metric is not obvious. One choice might be to compute the RMSE on voxel counts, as we did for range pixels in Section 3. However, the interpretation is quite different. Errors in range induce errors which have units of distance. On the other hand, errors in the 3D volumetric representation are measured in "number of returns." Nevertheless, the RMSE on voxel counts may be a useful metric for applications that need information about local reflectivity within a scene. However, for the purpose of facilitating comparison between 2D and 3D compression, we introduce another performance measure based on "remeasuring" range data from 3D histogram volumes.

Figure 7 illustrates the steps involved in computing this metric. We start with a 3D volume computed from one or more *original* range images. From that volume, we create a number of "simulated" range images (shown on the left in Figure 7). The number of these simulated range images is the same as the number of original range images used to create the volume. Furthermore, the sensor locations used to create the simulated range images are the same as the locations used to gather the original range images. We refer to these simulated range images as "initial images." Note that because of discretization in the 3D volume, the initial images are not generally identical to the original range images.

The 3D histogram volume is then compressed and decompressed. The process described above is repeated using the resulting lossy volume to generate simulated range images. We refer to these as "final images" (as shown on the right side of Figure 7). The RMSE is then computed for each initial/final pair.

For the sake of comparison, a similar process is carried out for the 2D case. For a given original 2D range image, we render a 3D volume. From this volume, we create a single simulated range image (as shown on the left side of Figure 8). As before, the location used to generate this "initial" simulated range image is the same as the location used to gather the original 2D range image. The original range image is next compressed/decompressed (as a 2D image) as described previously in Section 3. The resulting lossy range image is then subjected to the same rendering and measurement process as the original range image. The result of this process is the "final" simulated range image (as shown on the right side of Figure 8). Finally, the RMSE between the initial and final simulated range image is computed.

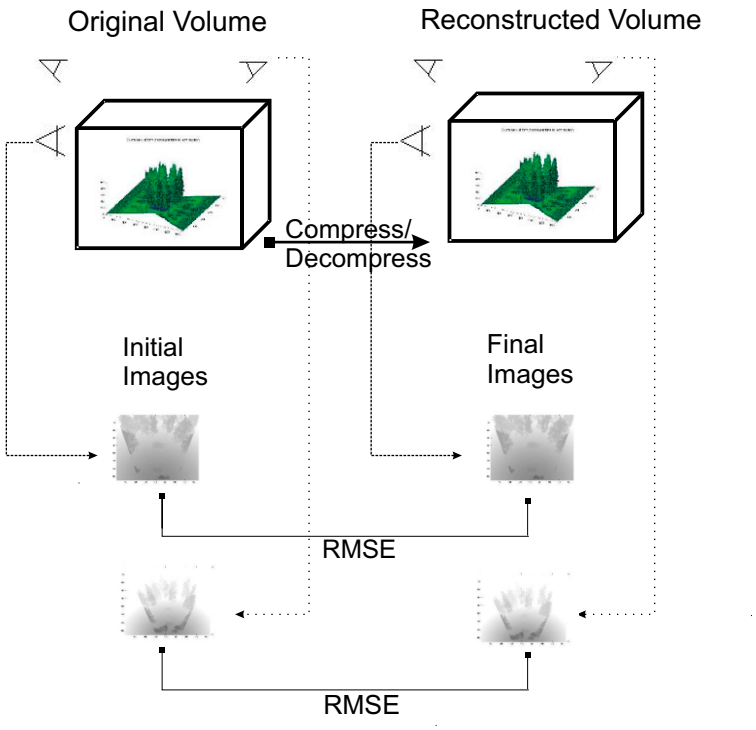


Figure 7. Computation of lossy compression metric for 3D Volumes.

Figure 9 shows the results of these performance measures applied to lossy compression of 2D and 3D data. All of these data points correspond to 0.44 bits per range pixel used. The family of non-solid curves represents the computed RMSE for 3D volumetric compression. As in the lossless case, different 3D volumes contain different numbers of accumulated range images. From these results, we can observe how accumulating range images in a 3D volume affects compression performance. For example, the square symbol represents the RMSE resulting from compressing a volume rendered from a single range image. That value is obtained by computing the RMSE of a single initial/final pair. The two circle symbols depict the RMSE resulting from compressing a volume rendered using the first two range images. Note that there are now 2 RMSE values corresponding to 2 initial/final pairs. This continues for 3 pairs, 4 pairs, etc., until the dashed bold line which represents the RMSE values resulting from compressing the 3D volume rendered from all 13 original range images. For comparison, the solid bold line represents the RMSEs from performing 2D compression on individual original range images.

One can readily see from Figure 9 that, after accumulating 9 (or more) out of the 13 range images in one 3D volume, 3D lossy compression is always better than 2D lossy compression for this performance criteria. Eventually, after accumulating all 13 range images in one 3D volume, 3D lossy compression yields an average RMSE value that is 99% lower than that obtained by 2D lossy compression of the same 13 range images at the

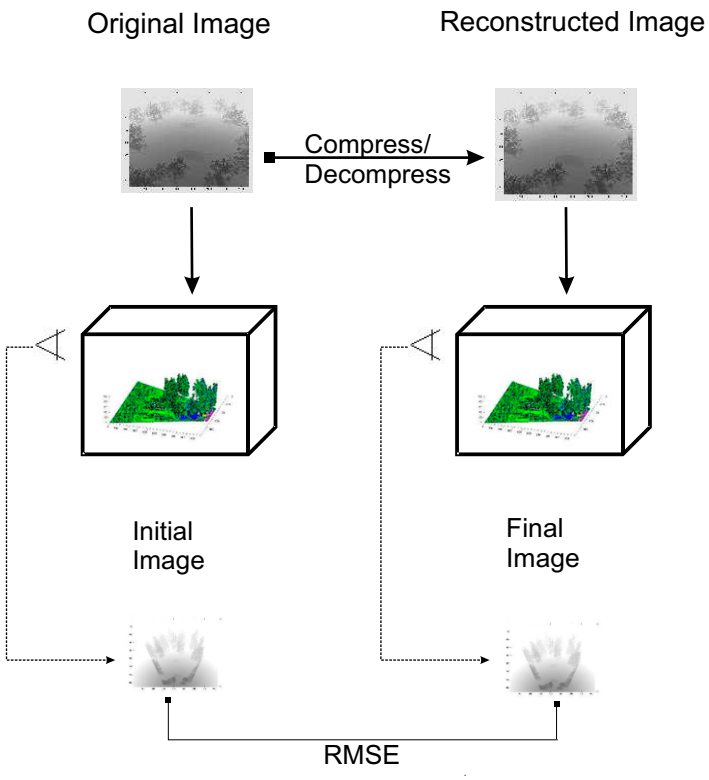


Figure 8. Computation of lossy compression metric for range imagery.

same nominal target rate. Finally, we should note from Figure 9, the RMSE curves start to "rise" back up again after accumulating 11 range images. This is expected. As we noted earlier in Section 3, high slant angle images are harder to compress compared to near nadir imagery. This fact manifests itself here again. With the addition of high slant angle imagery, the compression performance of the entire volume suffers: bits are taken away from the "easy" to compress data and given to the "harder" to compress data.

5. CONCLUSION

The goal of the work presented in this paper is the efficient compression of volumetric image data, generated by a single platform mobile ladar sensor. We exploit the correlation structure inherent in multiple views in order to improve compression efficiency. Our error metric suggests that accumulating more than 9 range images in one volume before compression yields up to a 99% improvement in compression performance over 2D compression.

It may be possible to enhance the compression performance of 3D volumetric data by designing a special purpose compression algorithm with optimized 3D context models. This is the subject of future work. Finally, it should be noted that all algorithms presented in this paper relate to compression data acquired from a single (mobile) sensor. Results pertaining to compressing data acquired from multiple sensors is the subject of future research.

REFERENCES

1. D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Practice and Standards*, Kluwer Academic Publishers, Massachusetts, 2002.
2. M. Marcellin, M. Gormish, A. Bilgin, and M. Boliek, "An overview of JPEG-2000," *Data Compression Conference, Institute of Electrical and Electronics Engineers, Snowbird, UT*, pp. 523–541, 2000.
3. "Jpeg 2000 part i final draft international standard," *ISO/IEC JTC 1/SC 29/ WG1, Doc. No. N1855* , Aug. 2000.

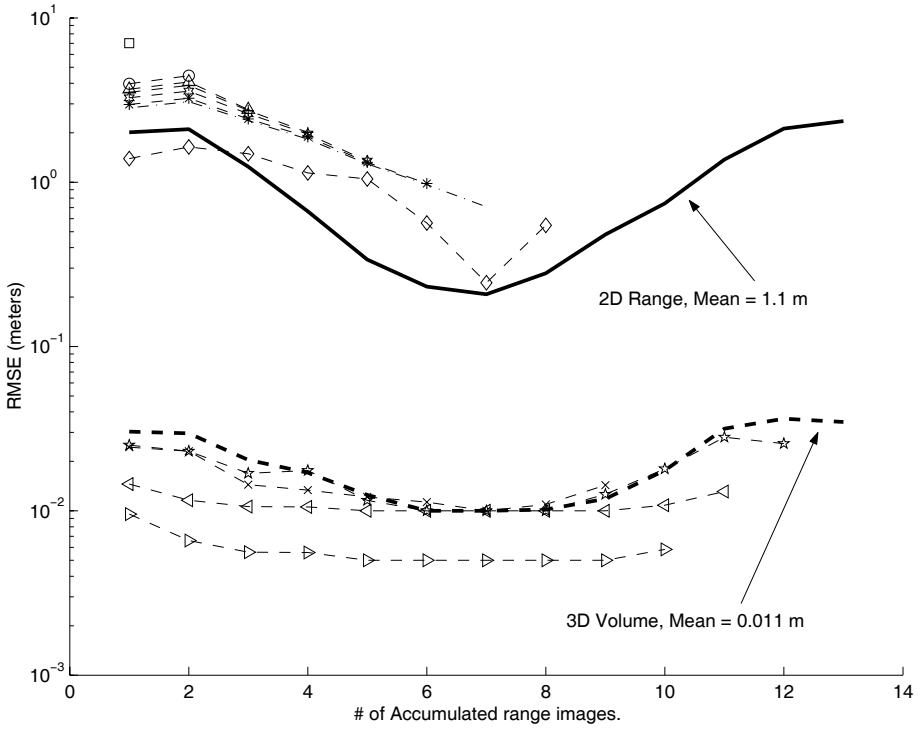


Figure 9. Comparison of lossy compression of range imagery vs. 3D volumes at 0.44 bits per pixel per range image.

4. W. Xiaolin and N. Memon, "Calic)-(a) context based adaptive lossless image codec," in *International Conference on Acoustics, Speech, and Signal Processing*, *Institute of Electrical and Electronics Engineers, New York*, pp. 1890–1893, 1996.
5. J. Luo, X. Wang, C. W. Chen, and K. J. Parker, "Volumetric medical image compression with three-dimensional wavelet transform and octave zerotree coding," in *Visual Communications and Image Processing*, *Proc. SPIE* **2727**, pp. 579–590, 1996.
6. B. Aiazzi, P. S. Alba, S. Baronti, and L. Alparone, "Three-dimensional lossless compression based on a separable generalized recursive interpolation," in *International Conference on Image Processing*, *Institute of Electrical and Electronics Engineers*, pp. 85–88, 1996.
7. A. Baskurt, H. Benoit-Cattin, and C. Odet, "On a 3-D medical image coding method using a separable 3-D wavelet transform," in *Medical Imaging*, *Proc. SPIE* **2431**, pp. 184–194, 1995.
8. K. L. Lau, W. K. Vong, and W. Y. Ng, "Lossles compression of 3-D images by variable predictive coding," *Proceedings of 2nd Singapore International Conference on Image Processing*, pp. 161–165, 1992.
9. M. Soucy and D. Laurendeau, "Multi-resolution surface modeling from multiple range views," in *Conference on Computer Vision and Pattern Recognition*, *Institute of Electrical and Electronics Engineers, CA*, pp. 348–353, 1992.
10. A. D. Sappa, M. A. Garcia, and B. X. Vintimilla, "Geometric and topological lossy compression of dense range images," in *Conference on Image Processing*, *Institute of Electrical and Electronics Engineers, NJ*, pp. 423–426, 2000.
11. J. C. Dagher, M. W. Marcellin, and M. A. Neifeld, "Efficient transmission/storage of ladar imagery," *Submitted to Applied Optics, Information Processing* .
12. N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall, Englewood Cliffs, Massachusetts, NJ, 1984.