

A Strategy for the Correction of Effects of Jitter in AMCW Lidar Images

Lee Streeter, Michael J. Cree and Adrian A. Dorrington

School of Engineering

University of Waikato

Hamilton

Email: cree@waikato.ac.nz

Abstract—Full-field range imaging cameras, which are an example of amplitude modulated continuous wave (AMCW) lidar, are subject to multiple error sources. In this paper we consider the possibility that a global jitter and/or drift occurs between the modulation signals to the active illumination source and the camera sensor. A model for jitter is presented and is used to characterise the jitter in a real camera system. Correction of the effects of jitter in real acquired range images is demonstrated by a reduction of the standard deviation of random noise by up to a factor of five, with the best reductions occurring in pixels that receive a strong signal.

I. INTRODUCTION

Full-field range imaging cameras are a relatively new camera technology for the simultaneous full-field acquisition of range data. Until recently commercial cameras were expensive and use was primarily restricted to researchers with a specific interest in range data. The release of the SoftKinetic DepthSense in 2012 has brought full-field range imaging technology to the consumer market at affordable prices.

Range imaging cameras are an example of amplitude modulated continuous wave (AMCW) lidar. Radio frequency amplitude modulated light illuminates the scene and by means of a modulated shutter on the camera sensor, the time-of-flight of the illumination from light source to the scene and back to the camera is measured. As the speed of light is a well known constant (that is barely modified by the presence of the atmosphere) the range to the scene is straightforward to calculate from the time-of-flight.

AMCW lidar is known to be adversely affected by a number of error sources. This includes non-linearities of the illumination and shutter modulations which generates harmonics that can alias on to the fundamental frequency, the mixing of light return from adjacent objects visible in a pixel leading to erroneous ranges, and light taking multiple paths from illumination source via the scene back to a pixel. These problems have been reasonably well analysed and techniques for mitigation described [1]. Now that the dominant error sources are better understood and can be corrected for to a large degree, smaller error sources are more noticeable.

In this paper we consider the possibility that jitter or drift is present between the modulation signals for the illumination source and the camera shutter [2], [3]. The model we present

compensates for a global mistiming of acquisition of a single phase image and for a small drift in modulation frequency. By ‘global’ it is meant that the mistiming or modulation frequency drift is equally applicable to every pixel across the sensor in a single phase acquisition. Jitter between pixels, say due to separate analogue to digital conversion pathways for subsections of the sensor, is not considered. We are not aware of any prior published studies on the effects of jitter or modulation frequency drift to full-field range image data.

We present in Sect. II a summary of the theory behind AMCW lidar acquisition as employed by typical full-field range imaging cameras. We extend that theory to include two extra terms: one effectively models jitter in the phase difference between the light source and sensor modulation and the other a drift or jitter in modulation frequency. This model is reformulated as a factor analysis problem which enables a linear regression to be applied to recover the model (jitter/drift) parameters and a correction applied to recover a de-jittered range image. The model is applied to range imaging data as described in Sect. III to test the effect jitter can have in a practical camera system. Results of testing on raw data, and data that have been first spatially filtered to remove random noise due to other error sources, is presented in Sect. IV.

II. THEORY

The full-field range cameras of interest operate by throwing radio-frequency amplitude modulated light into the scene and measuring the phase shift of the modulation of light reflected back from the scene. With the modulation phase shift one can calculate the time of flight of the light, hence the distance the light travelled from light source via the scene to the camera sensor. Assuming the light source is coincident with the camera sensor, the range to an object is given by half the distance the modulated light travelled.

It is typical to electronically shutter the camera sensor with a signal the same as the light source but phase shifted by θ_n for the n -th phase acquisition. As the sensor integration period τ_s is very long compared to a period of the modulation ($1/f$) the received signal for a time of flight phase shift of ϕ at a pixel of the sensor is described by

$$I_n = \int_0^{\tau_s} A \sin(2\pi ft + \phi + \theta_n) \sin(2\pi ft) dt \quad (1)$$

$$\approx \frac{A\tau_s}{2} \cos(\phi + \theta_n), \quad (2)$$

which is a good approximation when $\tau_s \gg 1/f$. Let the θ_n be evenly stepped, that is $\theta_n = 2\pi n/N$, for a sequence of N phase acquisitions to evenly sample the cosine function, then the time-of-flight induced phase shift ϕ can be recovered by Fourier transform of the I_n .

It is useful to described the recovered (complex) phasor by $P_k = A_k e^{i\phi_k}$ where A_k is the magnitude (brightness) of the received signal at the k -th pixel on the sensor and ϕ_k is the phase shift of the modulation due to the time-of-flight at the same pixel. Then by discrete Fourier transform,

$$P_k = \sum_{n=0}^{N-1} I_n e^{-2\pi i n/N} \quad (3)$$

$$= A_k \sum_{n=0}^{N-1} \cos\left(\phi_k + \frac{2\pi n}{N}\right) e^{-2\pi i n/N}, \quad (4)$$

where the extra factor of $\tau_s/2$ has been absorbed into the A_k .

Any high speed signal generation or acquisition is subject to degradation due to jitter [2]. Here we consider jitter primarily in the modulation signal generation leading to jitter in the incremental phases θ_n and also potential frequency fluctuations of the modulation frequency f . Such jitter will manifest errors in acquisition equally in all pixels of the camera. The model presented below does not describe pixel-dependent jitter that might result, for example, because of clock amplifiers specific to subsections of the sensor.

Let ε_n be the error in stepping the phase θ_n , and Δf_n be the error in modulation frequency of the n -th phase acquisition, due to signal jitter. Under such errors the measured P_k is given by

$$P_k = A_k \sum_{n=0}^{N-1} \cos\left((1 + \Delta f_n)\phi_k + \frac{2\pi n}{N} + \varepsilon_n\right) e^{-2\pi i n/N}. \quad (5)$$

Note that the noise due to jitter increases linearly with active brightness, thus increasing brightness of the illumination or a change in the reflectivity of objects in the scene does not reduce the relative jitter error, thus does not affect the signal to noise ratio when considering noise due to jitter. Note also that no assumption has been made as to the form or shape of the jitter distribution profile.

A. An Algorithm for Jitter Correction

We now develop a means of characterising errors due to jitter. Expanding the cosine term of Eqn. 5 with standard trigonometric identities and using the Taylor polynomial expansions about $\Delta f_n = 0$, namely,

$$\cos((1 + \Delta f_n)\phi_k) = \cos \phi_k - \Delta f_n \phi_k \sin \phi_k + O(\Delta f_n^2), \quad (6)$$

$$\sin((1 + \Delta f_n)\phi_k) = \sin \phi_k + \Delta f_n \phi_k \cos \phi_k + O(\Delta f_n^2). \quad (7)$$

leads to the linear model

$$P_k = A_k [\cos \phi_k, \sin \phi_k, \phi_k \cos \phi_k, \phi_k \sin \phi_k] \cdot [c_1, c_2, c_3, c_4], \quad (8)$$

with the coefficients given by

$$c_1 = \sum_{n=0}^{N-1} \cos\left(\varepsilon_n + \frac{2\pi n}{N}\right) e^{-\frac{2i\pi n}{N}}, \quad (9)$$

$$c_2 = -\sum_{n=0}^{N-1} \sin\left(\varepsilon_n + \frac{2\pi n}{N}\right) e^{-\frac{2i\pi n}{N}}, \quad (10)$$

$$c_3 = -\sum_{n=0}^{N-1} \Delta f_n \sin\left(\varepsilon_n + \frac{2\pi n}{N}\right) e^{-\frac{2i\pi n}{N}}, \quad (11)$$

$$c_4 = -\sum_{n=0}^{N-1} \Delta f_n \cos\left(\varepsilon_n + \frac{2\pi n}{N}\right) e^{-\frac{2i\pi n}{N}}. \quad (12)$$

All dependence on Δf_n and ε_n is confined to the coefficients c_1 to c_4 , that is, those coefficients are global range image frame jitter parameters. For multiple range acquisitions of a static scene, the jitter coefficients vary from frame to frame, but are invariant from pixel to pixel within any one frame. In contrast, the active brightness A_k and range phase ϕ_k is invariant across frames for an individual pixel. Therefore over a number of range image frames, Eqn. 8 fits the model of a factor analysis problem, thus can be solved for the jitter coefficients for any one range image frame using linear algebra.

Assume that we have a reference image, such as the previous range image acquired, or an average over a number of range images. The following steps serve to estimate and remove the effect of jitter using the above described model:

- 1) Compute the jitter parameters of the current frame with respect to the reference frame by solving Eqn. 8 by linear regression.
- 2) Having found the jitter parameters for the current frame, recalculate the current frame, namely $P_k = A_k e^{i\phi_k}$, using the real component of c_1 and the imaginary component of c_2 .

III. MATERIALS AND METHODS

150 range images were taken of a static scene consisting of flat targets of varying reflectivity, however due to transmission errors between camera and computer five range images were found to contain errors and were discarded leaving 145 range images for analysis. Each range image was acquired with $N = 5$ phase acquisitions as a reasonable trade-off between keeping N small, but also a high enough odd value to avoid excessive aliasing of harmonics to the fundamental. As the scene is static, a reference image (shown in Fig. 1) was constructed by taking the mean over time of the 145 images. Since the noise sources including but not limited to jitter are zero mean centred, the reference image is used as the noise-free image for comparison.

The data were processed with the algorithm outlined in Sect. II-A. The effect of the algorithm on the recovered phases ϕ_k , thus the ranges, is of most interest. We do calculate the root mean square error (RMSE) of the phase of a restored image against the reference image phase, however this measure must be interpreted with care as very large errors in phase occur in areas of low intensity for reasons other than jitter, and can adversely influence the RMSE values.

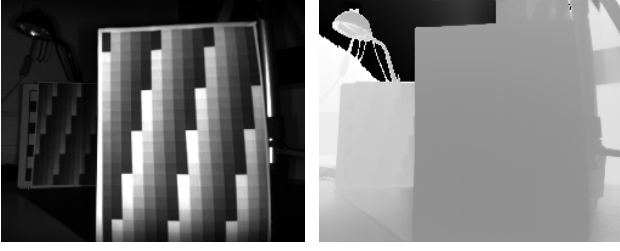


Fig. 1. The reference range image obtained by averaging over many images. The scene consists of two test patterns. Left: active brightness of the reference frame. Right: phase of the reference frame; increasing brightness reflects increasing range.

As there are 145 range images, each of which is processed by the de-jitter algorithm, the standard deviation of variation in the recovered phase at a pixel is calculated over time (that is, over the 145 range images). It is expected that the de-jitter algorithm should reduce the fluctuations in restored phase and this should be reflected in the calculated standard deviation.

For this analysis no attempt was made to account or correct for errors in ranging due to non-linearities, mixed pixels or multipath interference, as these sources are deterministic in nature and should result in exactly the same error to every one of the acquired 145 range images. While such errors in ranging have a greater effect than any random jitter or drift, their deterministic nature means that they act in a sufficiently different manner such that their presence does not affect the analysis below.

The acquired images are subject to random spatial noise independent of the jitter noise. Spatial filtering of the range images, with well understood spatial noise suppression filters, should reduce zero mean random spatial noise enabling the effect of jitter to be better observed. We therefore also test spatial filtering of the individual range images before applying the de-jitter algorithm to better observe the effect of jitter. Standard noise suppression filters such as 2D Gaussian and median filtering can be straightforwardly applied, but must be applied to the real and imaginary components of the complex range phasors separately. Also note that Gaussian filtering tends to blur edges, and that median filtering is not optimal for Gaussian distributed noise.

There are some noise filtering algorithms designed for colour image processing that 1) have good noise suppression properties without unfortunate blurring of edges and loss of detail, and 2) can easily be adapted to operate on complex data. We therefore also test a number of filters (Tbl. I) including the Bilateral Filter [4], Local Adaptive Means [5], Guided Image Filter [6], the Fuzzy Vector Filter (FVF) [7], [8] and the Non-local Means (NL-means) Filter [9]. The NL-means filter requires the computation of a sum of square of differences, which we modify to be calculated as the sum of the square of moduli of the difference between complex values. A similar change is made for the Guided Means filter; the minimisation of the sum of squares of an energy function in its derivation is replaced by a minimisation of the sum of squares of moduli of complex values. The FVF is a method arising from analysis of colour images, and is adapted herein to treat complex values as 2D vectors.

TABLE I. THE SPATIAL FILTERING METHODS TESTS; THE MEANING OF THE PARAMETERS IS EXPLAINED IN THE TEXT AND THE ABILITY TO INHERENTLY PROCESS COMPLEX DATA IS NOTED.

Filter	Parameters	Complex
Gaussian Filtering [5]	$\sigma_d = 3$, width = 9	No
Bilateral Filter [4]	$\sigma_d = 1$, $\sigma_r = 0.01$, width = 7	No
NL-means [9]	$M = 11$, $k = 3$, $\sigma_r = 0.01$	Yes
Complex FVF [8]	$p = 10$, width = 5	Yes
Median Filter [5]	width = 5	No
Guided Filter [6]	$\varepsilon = 0.001$, width = 11	Yes
Local Adaptive Means [5]	$\sigma_n = 0.01$, width = 5	No

The filtering methods tested, with programmable parameters specific to each filter's operation, are listed in Tbl. I. The choice of the filters tested was based on 1) speed, namely the existence of a fast implementation, or the potential for a fast implementation seemed feasible, and 2) the ability of the filter to suppress noise without excessive damage to image structure. The 'complex' column in the table refers to whether the implementation of the filter operates on complex data directly (thus 'yes'), or whether the filter operates on the real and imaginary parts separately and independently (thus 'no' in the column).

With the exception of the Gaussian filter, the parameters for filtering were chosen to trade off between error introduced and noise reduction. The full meaning of the parameters used for each algorithm can be discerned from the relevant references, nevertheless we give a brief description here. 'Width' refers to the filtering window width and height. The standard deviation in domain, σ_d , pertains to filtering algorithms that incorporate a spatial filtering kernel with a Gaussian profile. The range standard deviation σ_r is for algorithms that compute a Gaussian of some distance metric in image pixel values. (Note that σ_r has been specified in units where the complex phasors have been normalised to unit magnitude for maximum signal.) For the local adaptive means algorithm, σ_n is an estimate of the noise level in the image. In the NL-means algorithm: M is the search window size and k is the matching window size. For the guided filter ε is a regularisation parameter that is analogous to the ridge parameter in ridge regression. In performing guided filtering, increasing ε increases the smoothing effect.

IV. RESULTS

The effect of spatial filtering in terms of RMSE is shown in Tbls. II and III. All RMSE values are calculated with respect to the reference image, which is formed by taking the mean of the unfiltered frames over time. An increase in RMSE (as can be seen for the RMSE calculated over the phase for the Guided Filter) over that of the raw data, indicates that the filter has introduced unwanted structural effects, such as blurring of edges, in excess of the improvement in noise filtering. The Non-local Means filter provides the best improvement in phase (see Tbls. II and III) due to its better ability to adapt the filter about image structure.

The relative performance of the filters is as expected. The Gaussian filtering introduces substantial blurring of the images and loses image detail, and even though random noise is almost eliminated the RMSE does not improve. Filters that better adapt noise suppression to image structure perform significantly better, with the Non-local Means filter the best performing on the RMSE measure.

TABLE II. RMS ERROR (RMSE) OVER THE 145 FRAMES BEFORE APPLICATION OF THE DE-JITTER ALGORITHM.

Method	Phase (rad)	Real (DN)	Imaginary (DN)
Raw Data	0.304	30	29
Gaussian Filtering [5]	0.283	582	375
Raw, Gaussian Params.	0.304	30	29
Bilateral Filter [4]	0.128	31	29
NL-means [9]	0.115	41	31
Complex AFVF [8]	0.125	311	241
Median Filter [5]	0.155	316	242
Guided Filter [6]	0.475	67	153
Local Adaptive Means [5]	0.141	48	43

TABLE III. RMS ERROR (RMSE) OVER THE 145 FRAMES AFTER APPLYING THE DE-JITTER ALGORITHM.

Method	Phase (rad)	Real (DN)	Imaginary (DN)
Raw Data	0.304	25	23
Gaussian Filtering [5]	0.283	582	375
Raw, Gaussian Params.	0.304	25	23
Bilateral Filter [4]	0.128	26	22
NL-means [9]	0.115	38	26
Anisotropic Image Filter	0.171	102	91
Complex AFVF [8]	0.125	311	240
Median Filter [5]	0.156	315	241
Guided Filter [6]	0.478	65	151
Local Adaptive Means [5]	0.141	45	39

The RMSE values in Tbl. II are before applying the de-jitter algorithm. Applying the de-jitter algorithms results in the RMSE values as listed in Tbl. III. It appears disappointing that the de-jitter algorithm has not reduced the RMSE in phase, however this measure is dominated by two factors: a small number of very large errors in phase from pixels that received almost no signal and a bias towards increased RMSE due to the blurring effect of certain spatial filters. We therefore seek a measure that better captures the improvement that de-jittering makes.

To better characterise the improvement in phase due to the de-jitter correction the standard deviation of the phase (σ_{ϕ_k}) is calculated over time (i.e. across the 145 frames) for each of the pixels in a frame. The σ_{ϕ_k} are then plotted against the mean active brightness of the pixel A_k as a scatter graph. This enables one to distinguish between pixels of good SNR (where most of the error is due to jitter and the de-jitter correction is expected to have a useful effect) and those of poor SNR (where random errors due to sources other than jitter dominate). Two sets of data are plotted on each graph: one for the spatially filtered data before application of the de-jitter correction (labelled *raw* in the graphs), and one for the spatially filtered data after application of the de-jitter algorithm (respectively labelled *dejittered*).

The application of the de-jitter correction to the spatially unfiltered (raw) data is shown in Fig. 2. A decrease of pixel variation, σ_{ϕ_k} , of up to 60% is observed for pixels of high SNR (i.e. those with large A_k). As A_k increases the curve tends to flatten out and it appears that it might be asymptotically approaching a non-zero positive value; this is evidence that noise sources that cannot be described by a Poisson or Gaussian model are present. Jitter is such a noise source.

In the following we present results for only some of the spatial filters listed in Tbl. I. Since it is well known and understood we start with the Gaussian filter with results presented in Fig. 3. The two images at the top of the figure show the active brightness and phase of the mean of the 145 frames after applying spatial Gaussian filtering and time-based

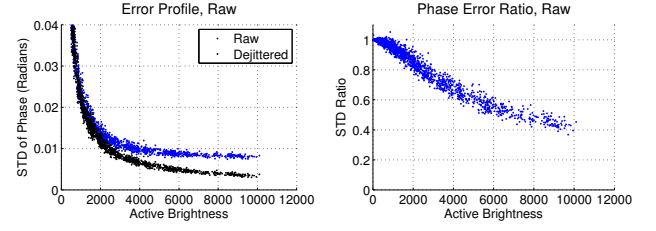


Fig. 2. Jitter correction on the spatially unfiltered (raw) data. Left: graph of the standard deviation of a pixel plotted against its active brightness for the raw data, and the de-jitter corrected data. Right: The ratio of the two standard deviations plotted in the graph at left.

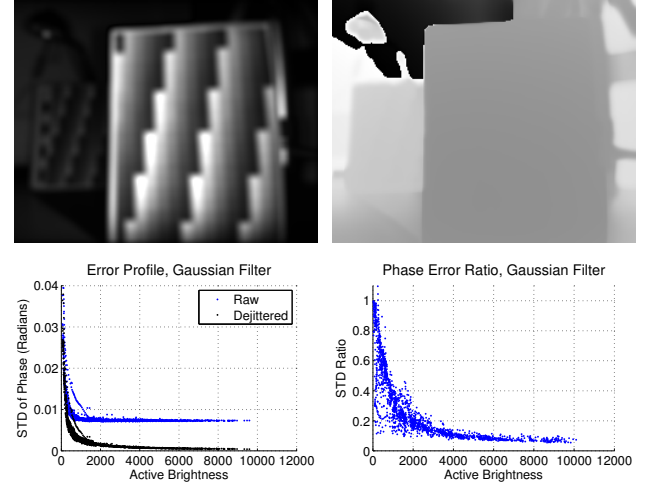


Fig. 3. Jitter correction on the spatially Gaussian filtered data. Top-Left: Active brightness of the mean over time of the Gaussian filtered and de-jitter corrected data. Top-Right: Phase of mean over time of the Gaussian filtered and de-jitter corrected data. Bottom-Left: graph of the standard deviation of a pixel plotted against its active brightness for the Gaussian filtered data, and the Gaussian jitter corrected data. Bottom-Right: The ratio of the two standard deviations plotted in the graph at left.

de-jitter filtering. While random noise is substantially removed there is very noticeable, thus unacceptable, blurring of the image.

The graph at the bottom-left of the figure is of the standard deviation in phase σ_{ϕ_k} versus the active brightness before and after applying the de-jitter correction to the Gaussian filtered data. Note that the *raw* curve (which is the Gaussian filtered data before application of the de-jitter correction) reveals a reduced noise level compared to the raw data of Fig. 2 particularly for pixels of low active brightness. We presume that spatially variant noise has been reduced to the point that jitter is the dominant noise source. The Gaussian filter has also failed to improve the noise level of some pixels of very low active brightness (the sub-curve of larger σ_{ϕ_k} for very low A_k). These are likely evidence of errors due to multipath effects in imaging, that is, where multiple returns are incident on a single pixel leading to erroneous range values; these errors can be somewhat systematic in nature across regions of pixels thus are not corrected by standard spatial noise-suppression filters. Even though Gaussian filtering blurs the resultant range images to an unacceptable degree it does provide a useful baseline for comparison of other filtering techniques.

The results with the Bilateral filter, the Non-Local Means

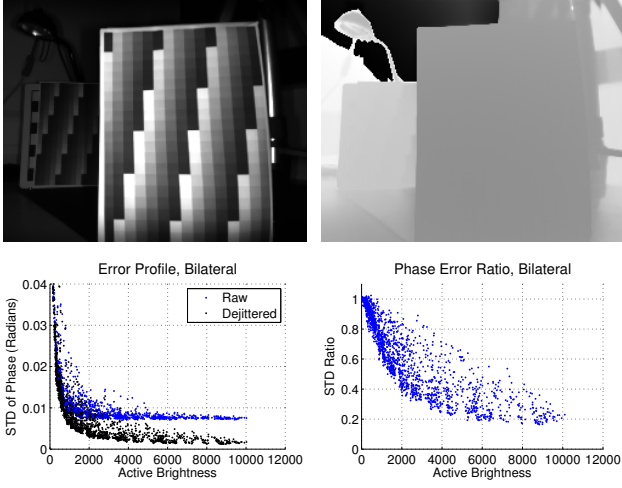


Fig. 4. Jitter correction on the spatially Bilateral filtered data. Top-Left: Active brightness of the mean over time of the Bilateral filtered and de-jitter corrected data. Top-Right: Phase of mean over time of the Bilateral filtered and de-jitter corrected data. Bottom-Left: graph of the standard deviation of a pixel plotted against its active brightness for the Bilateral filtered data, and the Bilateral jitter corrected data. Bottom-Right: The ratio of the two standard deviations plotted in the graph at left.

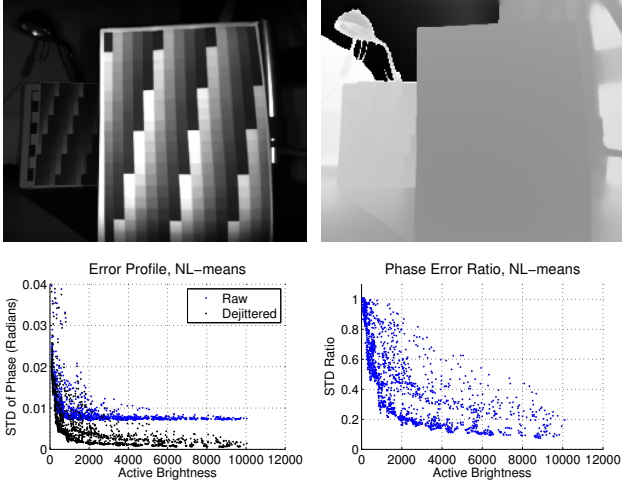


Fig. 5. Jitter correction on the spatially Non-local Means filtered data. Top-Left: Active brightness of the mean over time of the Non-local Means filtered and de-jitter corrected data. Top-Right: Phase of mean over time of the Non-local Means filtered and de-jitter corrected data. Bottom-Left: graph of the standard deviation of a pixel plotted against its active brightness for the Non-local Means filtered data, and the Non-local Means jitter corrected data. Bottom-Right: The ratio of the two standard deviations plotted in the graph at left.

filter and the Guided filters are shown in Figs. 4 to 6. The Bilateral and Non-Local Means performs best. The filtering parameters (see Tbl. I) were chosen to minimise both the standard deviation σ_{ϕ_k} and the ratio of σ_{ϕ_k} calculated before and after de-jitter correction, while retaining image structure. The Guided filter performed particularly poorly in terms of the σ_{ϕ_k} measure and errors are visible in the phase image (Fig. 6). The Non-Locals Means filter exhibits an appearance of two curves which indicates that this filter applies a better noise suppression at edges in the range frame data.

A holistic view is required to identify the best performing

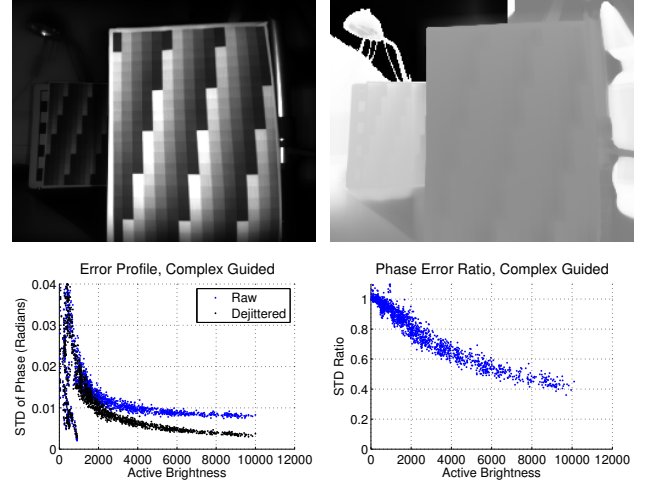


Fig. 6. Jitter correction on the spatially Guided filtered data. Top-Left: Active brightness of the mean over time of the Guided filtered and de-jitter corrected data. Top-Right: Phase of mean over time of the Guided filtered and de-jitter corrected data. Bottom-Left: graph of the standard deviation of a pixel plotted against its active brightness for the Guided filtered data, and the Guided jitter corrected data. Bottom-Right: The ratio of the two standard deviations plotted in the graph at left.

spatial filters. When the ability to de-jitter is considered we find that the Gaussian, Bilateral filter and the NL-Means filters all perform well, however the standard deviation measurement reported above only confirms that random noise in time is reduced; it does not mean the result is actually *unbiased* and thus correct. Tbls. II and III provide a measure of correctness and reveal that the Gaussian filter performs poorly in this regard. We conclude then that the Bilateral filter and the NL-Means filters both perform well in combination with the de-jitter algorithm.

V. CONCLUSION

It is believed that variations in phase and frequency, whether due to jitter or drift, occur between the light source and camera sensor modulations. The model presented herein appears to successfully model the variations in phase and frequency seen in a range imaging sensor. The algorithm presented for jitter characterisation provided a real improvement in range precision, and the success of the algorithm depends only on the assumptions in the model that jitter affects all pixels in the scene.

For the purpose of characterisation the analysis presented above was calculated against a reference image that was the composite of many acquired range images. While such a reference image may not be available for jitter correction in a practical system, we nevertheless expect that it would be straightforward to adapt the algorithm for practical implementation on dynamic scenes.

ACKNOWLEDGEMENT

The authors thank John Godbaz for helpful discussions in the early stages of this research.

REFERENCES

- [1] J. P. Godbaz, M. J. Cree, and A. A. Dorrington, "Understanding and ameliorating non-linear phase and amplitude responses in AMCW lidar," *Remote. Sens.*, vol. 4, pp. 21–42, 2012.
- [2] M. Shinagawa, Y. Akazawa, and T. Wakimoto, "Jitter analysis of high-speed sampling systems," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, pp. 220–224, 1990.
- [3] M. J. M. Pelgrom, *Analog-to-Digital Conversion*. Dordrecht, Netherlands: Springer, 2010.
- [4] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth International Conference on Computer Vision*, Bombay, India, January 1998, pp. 839–846.
- [5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2001.
- [6] K. He, J. Sun, and X. Tang, "Guided image filtering," in *11th European Conference on Computer Vision Location*, vol. 6311, Heraklion, Greece, September 2010, pp. 1–14.
- [7] K. N. Plataniotis and A. N. Venetsanopoulos, "Vector filtering," in *Colour Image Processing Handbook*, S. J. Sangwine and R. E. N. Horne, Eds. London, U.K.: Chapman & Hall, 1998, pp. 188–209.
- [8] M. J. Cree, "Observations on adaptive vector filters for noise reduction in color images," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 140 – 143, 2004.
- [9] J. Wang, Y. Guo, Y. Ying, Y. Liu, and Q. Peng, "Fast non-local algorithm for image denoising," in *International Conference on Image Processing*, Atlanta, GA, October 2006, pp. 1429–1432.