2016 Fourth International Conference on 3D Vision

# Single-shot Time-of-Flight Phase Unwrapping Using Two Modulation Frequencies

Changpeng Ti Ruigang Yang
University of Kentucky
{cti222, r.yang}@uky.edu

James Davis
University of California, Santa Cruz
davis@cs.ucsc.edu

## Abstract

*We present a novel phase unwrapping framework for the Time-of-Flight sensor that can match the performance of systems using two modulation frequencies, within a single shot. Our framework is based on an interleaved pixel arrangement, where a pixel measures phase at a different modulation frequency from its neighboring pixels. We demonstrate that: (1) it is practical to capture ToF images that contain phases from two frequencies in a single shot, with no loss in signal fidelity; (2) phase unwrapping can be effectively performed on such an interleaved phase image; and (3) our method preserves the original spatial resolution. We find that the output of our framework is comparable to results using two shots under separate modulation frequencies, and is significantly better than using a single modulation frequency.*

## 1. Introduction

The standard Time-of-Flight (ToF) sensor detects depth by measuring the phase offset between sensor and scene, at a fixed modulation frequency. As such, there is an inherent ambiguity caused by the periodicity of the phase measurement. As a result, the phase measurement produced by the ToF sensor is "wrapped" into a fixed interval, i.e., $[0, 2\pi]$, such that all phase values corresponding to the set $\{\Phi|\Phi = 2k\pi+\phi, k \in Z\}$ become $\phi$. In terms of depth measurement, all depths are "wrapped" into an interval defined by the modulation frequency.

To recover the correct ToF measurement, one needs to "unwrap" the phase image, i.e., find the correct wrapping number for each pixel. One class of phase unwrapping methods involve capturing two shots at different modulation frequencies, and solving for the wrapping number by investigating the relation between the two measurements at each pixel. While each pixel could theoretically be solved independently, these methods require two shots to solve each frame, which cuts down the effective frame rate, and introduces inconsistency issues for dynamic scenes due to the time difference between the two shots. Another class of phase unwrapping methods works with the available one-shot image, and uses global optimization guided by factors other than phase, such as the intensity image. This area is still work in progress, as the performance is not on par with the two-shot methods.

Therefore, it would be desirable to combine the strengths of the two areas, i.e., have a system that preserves the redundancy of phase measurements as in the two-shots methods, while not requiring additional exposure time.

In this paper, we propose a novel phase unwrapping framework as an attempt to achieve this goal. We employ a special pixel arrangement at the capturing sensor, so that pixels adopting separate modulation frequencies are spatially interleaved. We then develop a new unwrapping algorithm that solves for the number of times each pixel has been wrapped. With capable hardware, this innovation will allow us to take advantage of the spatial coherence in the phase image within a single frame.

Main contributions of this paper include: (1) a novel pixel arrangement for the ToF sensor such that two different modulation frequencies are interleaved on a single phase image; (2) a phase unwrapping algorithm that produces state-of-the-art unwrapped result using the frequency-interleaved phase image.

We hope our work could inspire future generations of ToF cameras that support simultaneous multi-frequency and customizable pixel arrangement.

## 2. Related Work

Phase unwrapping is the procedure of recovering the correct wrapping number for signal given wrapped input. It is a problem well known in signal processing techniques, such as synthetic aperture radar (SAR), and magnetic resonance imaging (MRI) [7].

Phase unwrapping methods can be categorized into single-shot and multi-shot methods. Single-shot methods work with available data captured at a single modulation

978-1-5090-5407-7/16 $31.00 © 2016 IEEE
DOI 10.1109/3DV.2016.74

frequency by means of global optimization, such as [4], [3], while multi-shot methods work with data captured at multiple frequencies to solve for the number of wraps, such as [14], [5], [8], [6], [2], [11].

In particular, [4] takes into account intensity information when solving with a single frequency, and is the state-of-the-art in single-frequency unwrapping methods. [1] is a single-frequency work and it does not directly target ToF cameras, yet the graph cuts optimization framework [10] that it uses could be modified to work with multiple frequencies. The illumination demultiplexing method in [9] can also be used to capture multiple light sources. Instead of using custom coded ToF pixels, we use standard ToF pixels with interleaved modulation frequencies. As a result, the exposure time of one frame is entirely used for one depth measurement, allowing for higher SNR.

Our method organizes the input data such that it can be captured in a single shot, and then performs unwrapping following the general guideline of multi-frequency methods. This idea could be traced back to that of multi-sampled imaging [12], however unlike traditional pixel arrangement methods such as the Bayer pattern, our pixel layout does not cause any loss of resolution: wrapping information is recovered at every sensor pixel.

# 3. Framework Design

## 3.1. Spatial Pixel arrangement

As mentioned in Section 1, to exploit spatial coherence among pixels, we use an interleaved pixel arrangement, in which neighboring pixels use different modulation frequencies. A most straightforward layout is a checker pattern, where all pixels corresponding to the "white" blocks adopt one modulation frequency, while all pixels corresponding to the "black" blocks adopt another frequency. As illustrated in Fig. 1.

Other possible arrangements include row-interleaved (or column-interleaved) patterns, in which every other row (or column) take the same modulation frequency. This paper mainly discusses the checker-style pixel arrangement. The other two patterns deliver similar performance, as is discussed in the supplementary material.

## 3.2. Simultaneous Capturing at Multiple Modulation Frequencies

In an actual hardware implementation of the above pixel arrangement, there should be two simultaneous light sources at work, each operating at its own modulation frequency. The sensor should also have pixels operating at either frequency. Therefore, it is important to make sure that the light that reaches the sensor, as a mixture of different frequencies, does not cause trouble for pixels set to one of these frequencies.

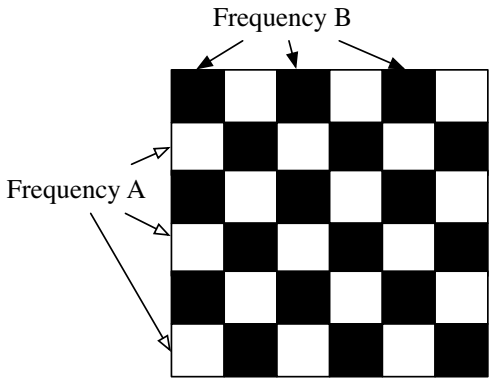



Figure 1. The checker-style interleaved pixel layout, where each pixel is surrounded by pixels operating at a different modulation frequency.

Let us consider a regular case, where light is emitted from the ToF camera's active illumination, and reflected back to a pixel on the camera sensor. Under the sine wave ToF model, the signal of the emitted light could be written as

$$l(t) = I\cos(\omega_l t) + I_0, \tag{1}$$

where $\omega_l$ is the modulation frequency of the light source.

Under the assumption of Lambertian reflection at a distance, and the sensor co-locating with the light source, the return light signal is in the form

$$g(t) = \alpha\cos(\omega_l t + \phi) + k, \tag{2}$$

where $\phi$ is the phase offset of the reflecting object w.r.t. the sensor.

The received signal $g(t)$ undergoes a per-pixel modulation with an on-sensor sinusoid function implemented by periodically shifting charge carriers between the two readout nodes. This sinusoid function could be written as

$$s(t) = \cos(\omega_p t). \tag{3}$$

The modulated signal for each pixel would be

$$\begin{aligned} c(t) &= s(t)\cdot g(t) \\ &= \cos(\omega_p t)\cdot(\alpha\cos(\omega_l t+\phi)+k) \\ &= \alpha\cos(\omega_p t)\cos(\omega_l t+\phi) + k\cos(\omega_p t) \\ &= \frac{\alpha}{2}\cos[(\omega_p+\omega_l)t+\phi] + \\ &\quad \frac{\alpha}{2}\cos[(\omega_p-\omega_l)t-\phi] + \\ &\quad k\cos(\omega_p t). \end{aligned} \tag{4}$$

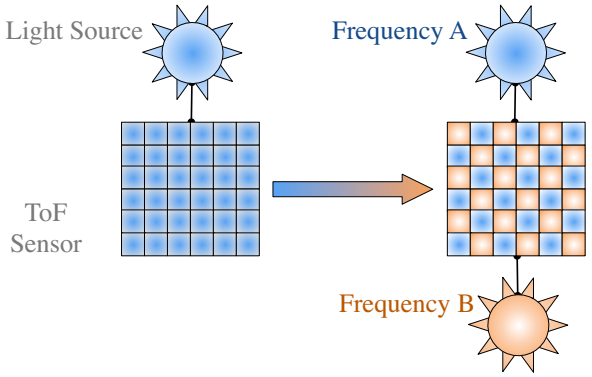



Figure 2. Illustration of our pixel interleaving design (right), comparing with the standard ToF camera (left). Although the two lights operate simultaneously, a pixel only receives the component of the reflected light signal that has the same modulation frequency as itself.

The ToF sensor executes the pixel's exposure by integrating the modulated signal $c(t)$ over an integration time $T$, and the raw phase readout of the ToF pixel would be

$$p(T) = \int_0^T c(t)\,\mathrm{d}t. \tag{5}$$

Since $T$ is normally orders of magnitude longer than the period of any term in Eq. 4, all terms with a non-zero $\omega$ will be nullified. In which case, the phase readout becomes

$$p(T) = \begin{cases} T \cdot \dfrac{\alpha}{2}\cos(-\phi), & \text{if } \omega_p = \omega_l, \\ 0, & \text{if } \omega_p \neq \omega_l, \end{cases} \tag{6}$$

where $\omega_p = \omega_l$ denotes when the light source's modulation frequency agrees with the sensor pixel's modulation frequency.

Eq. 6 reveals that, for a ToF pixel operating at frequency $\omega_p$, only the portion of the received light that has the same modulation frequency is effective towards its phase readout, meaning that having two distinct modulation frequencies do not cause interference for the ToF pixels.

It should be emphasized, however, that for the analysis from Eq. 5 to Eq. 6 to be true, $\omega_p$ and $\omega_l$ need to be sufficiently apart (see "orders of magnitude longer"). In our experiments, $|\omega_p - \omega_l|$ is in the megahertz range, and $T$ is in the millisecond range, which ensures the appropriateness of Eq. 6.

### 3.3. Practical Verification of Simultaneous Capturing at Separate Modulation Frequencies

Experiments are conducted to verify our theoretical conclusion in Section 3.2 that separate modulation frequencies do not interfere with each other. We set a SwissRanger ToF camera to 29MHz, and use another identical SwissRanger camera as external illumination at 31MHz, as shown in Fig. 3. The scene is a simple flat wall, close enough $(2 \sim 3m)$ so that the entire frame is within wrapping limit for either frequency.


Figure 3. Our experimental setup and scene to verify multi-frequency interference performance.

Consider the two scenarios, 1) capture at 29MHz with the external illumination at 31MHz turned off, and 2) capture with the external illumination turned on. We compare the difference between 1) and 2) against the difference between two captures under 1), to see how the external illumination affects the captured range data. For each capture, we take 100 frames to minimize noise, as well as to check that there are no temporal oscillations common to simultaneous captures at identical frequencies, as will be discussed below.

We take three separate captures, $R_1$, $R_2$ and $R_3$. $R_1$ and $R_2$ are taken by the 29MHz camera alone, $R_3$ is taken by the 29MHz camera, with the 31MHz light also turned on.

As illustrated in Fig. 4, the extra 31MHz light makes little difference in the captured range map, as the difference of measurement with the extra light on is very similar to that with the extra light off.

We also compare the behaviors of measurements over time, for both scenarios. We first calculate

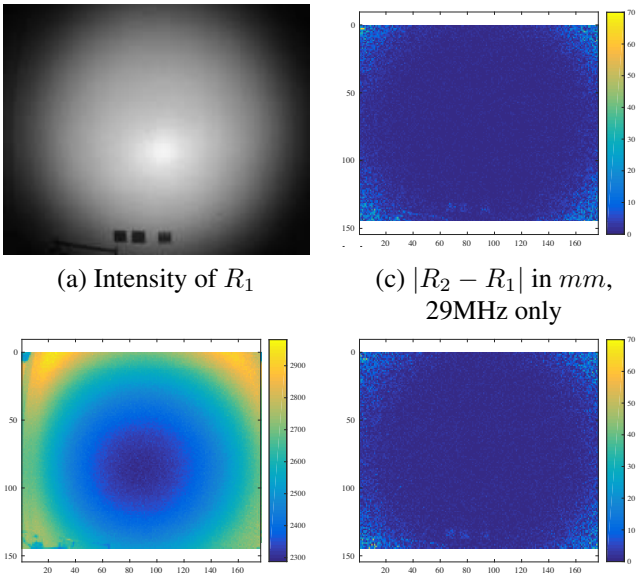
$$M_{21}(i,j) = \mathrm{mean}(|R_2(i,j,t) - R_1(i,j,t)|), \tag{7}$$

$$M_{31}(i,j) = \mathrm{mean}(|R_3(i,j,t) - R_1(i,j,t)|), \tag{8}$$

$$S_{21}(i,j) = \frac{\mathrm{std}(R_2(i,j,t))}{\mathrm{std}(R_1(i,j,t))}, \tag{9}$$

$$S_{31}(i,j) = \frac{\mathrm{std}(R_3(i,j,t))}{\mathrm{std}(R_1(i,j,t))}, \tag{10}$$

for each pixel, where $t$ is the frame index, then compare $\mathrm{mean}(M)$, $\mathrm{std}(M)$, $\mathrm{mean}(S)$ and $\mathrm{std}(S)$ between the two scenarios. The purpose of comparing the standard deviations is to detect temporal oscillations, as any oscillation



(a) Intensity of $R_1$

(c) $|R_2 - R_1|$ in $mm$, 29MHz only

(b) $R_1$ in $mm$, 29MHz only

(d) $|R_3 - R_1|$ in $mm$, 29MHz and 31MHz lights

Figure 4. Range differences between shots. (a) and (b) are reference data captured with the external 31MHz light off, (c) shows the difference between two captures both with the external light off, (d) shows the difference between one capture with the external light off, and another with the external light on. All range images are averaged over 100 frames.

| $(s,t)$ | $(2,1)$ | $(3,1)$ |
|---|---|---|
| $\text{mean}(M_{st}(i,j))$ | $1.8063mm$ | $2.0150mm$ |
| $\text{std}(M_{st}(i,j))$ | $3.1436mm$ | $3.1492mm$ |
| $\text{mean}(S_{st}(i,j))$ | 1.0060 | 1.0802 |
| $\text{std}(S_{st}(i,j))$ | 0.1045 | 0.1163 |

Table 1. Quantitative evaluation on the presence of an external light ($(s,t) = (3,1)$) vs. no external light ($(s,t) = (2,1)$). $s$ and $t$ denote different captures as previously defined. The two scenarios have very similar noise performance.

will make the distribution of measurements spread out more than default. Table 1 lists quantitative comparisons. It can be observed that interference from another light source at a different frequency is fully rejected in practical terms. $0.2mm$ more noise at a range of over $2m$ is completely negligible. There is also no visible temporal oscillation, as the spread of measurements over time with the external light on is very close to that with the external light turned off.

## 3.4. The Phase Unwrapping Algorithm

Given the phase readout from the interleaved sensor pixels, our novel phase unwrapping algorithm can be applied. The general idea is to first generate an approximate depth map from the interleaved phase image / depth map, then

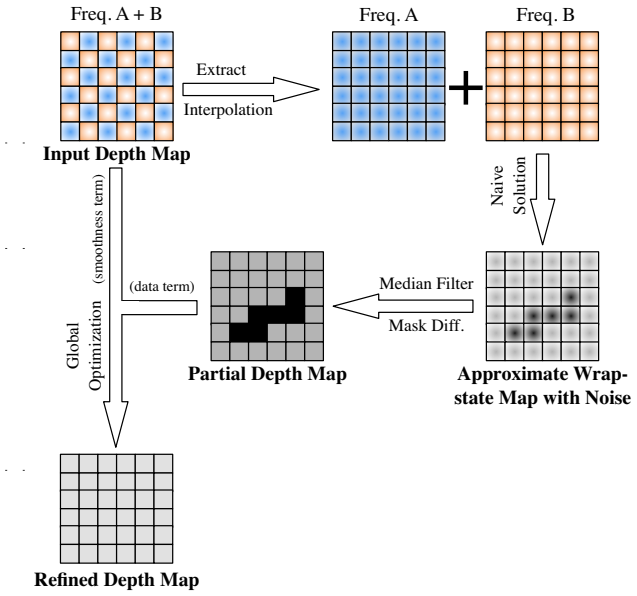



Figure 5. Flow chart of our phase unwrapping algorithm.

use it as a guiding term in the global optimization stage, to obtain a refined result. The main insight of our algorithm is that stable areas in the approximate depth map provide useful cues for the final solution, and that we can rely on global optimization to bridge the unstable areas using coherence from the input depth map.

### 3.4.1 Initial Solution

First, we extract the two sets of pixel depths from the raw depth map $d$, grouped by their operating modulation frequency. Each set of pixel depths constitute an incomplete depth map. Holes are left where pixels of the other frequency used to be at.

We now fill the holes in the incomplete depth maps by interpolation. This will generate two depth maps, $d_1$ and $d_2$, which for the time being we consider independent, akin to those captured consecutively at two different frequencies.

In the next step, we naively solve for phase unwrapping. The interpolated phase images are converted to depth maps, according to the following assumption

$$\begin{cases} d_1 = D \bmod r_1, \\ d_2 = D \bmod r_2, \end{cases} \tag{11}$$

where $r_1$ and $r_2$ are the depth wrapping ranges determined by the corresponding modulation frequency, and $D$ is the unwrapped depth map. In the solver, Eq. 11 is rewritten as

$$D = k_1 \cdot r_1 + d_1 = k_2 \cdot r_2 + d_2, \tag{12}$$

where $D$, and the wrap-state maps, $k_1$ and $k_2$, are unknowns. The equation is solved using the method of [8]. The wrap-state map records the number of wraps for all pixels. $k_1$ and $k_2$ are then assembled into a global $k_0$, according to the pixel arrangement pattern.

This initial method works reasonably well for relatively smooth regions, where no wrappings, natural discontinuities, or high noises occur. However, at those difficult areas this method fails, because the interpolated depth values no longer satisfy the assumption of Eq. 11. As a result, false values appear as high spikes in the wrap-state maps and the unwrapped depth map.

#### 3.4.2 Median Filtering

At this stage, a median filter is applied to $k_1$ and $k_2$ to remove lone spikes and produce a smoothed version, $k_1^s$ and $k_2^s$. We use a $5 \times 5$ filter in our experiments. After this step, $k_1^s$ and $k_2^s$ are reassembled pixel-wise into a single wrap-state map $k_0^s$.

In our algorithm, pixels where $k_0^s \neq k_0$ are considered unstable, and their neighborhood areas are masked out in a stability mask $M$. The size of the neighborhood can be empirically determined. We use $5 \times 5$ in our experiments.

At this point, we produce a partial depth map $d^s$ by the following

$$d^s = d + k_0^s \cdot r, \tag{13}$$

where $r$ is assembled from $r_1$ and $r_2$ according to the pixel arrangement pattern. Note that $d^s$ is only useful where pixels are unmasked in $M$.

#### 3.4.3 Global Optimization

We modify the graph cuts procedure in [1] to globally refine the wrap-state map. The energy function is defined as

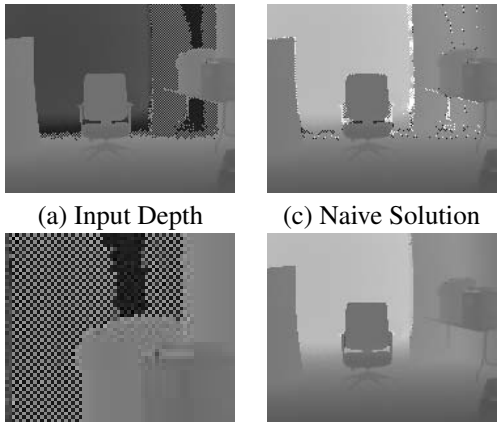
$$E(k|d, r, d^s, M) = E_s(k|d, r) + \lambda E_d(k|d, r, d^s, M), \tag{14}$$

with smoothness term

$$E_s(k|d, r) = \sum_{i,j} \left( w_{i,j}^h \cdot V\left(\frac{2\pi \Delta D_{i,j}^h}{r_{i,j}}\right) + w_{i,j}^v \cdot V\left(\frac{2\pi \Delta D_{i,j}^v}{r_{i,j}}\right)\right), \tag{15}$$

and data term

$$E_d(k|d, r, d^s, M) = \sum_{i,j} M_{i,j} \cdot \left|d^s - (d_{i,j} + k_{i,j} \cdot r_{i,j})\right|, \tag{16}$$



(a) Input Depth (c) Naive Solution

(b) Detail in (a) (d) Final Result

Figure 6. Illustration of data at different stages in our algorithm.

where $k = \{k_{i,j} \in Z\}$ is our solution wrap-state map. $\lambda$ is a balancing factor between the two terms that can be determined empirically.

$E_s(k|d, r)$ has been adapted to accommodate our pixel arrangement. $\Delta D^h$ and $\Delta D^v$ denote pixel distance differences where

$$\begin{cases} \Delta D_{i,j}^h = (d_{i,j} - d_{i,j-1}) + (k_{i,j} \cdot r_{i,j} - k_{i,j-1} \cdot r_{i,j-1}), \\ \Delta D_{i,j}^v = (d_{i,j} - d_{i-1,j}) + (k_{i,j} \cdot r_{i,j} - k_{i-1,j} \cdot r_{i-1,j}), \end{cases} \tag{17}$$

and the clique potential $V$ is defined as

$$V(x) = \begin{cases} \theta^{-1.9} x^2, & \text{if } |x| \leqslant \theta. \\ |x|^{0.1}, & \text{if } |x| > \theta. \end{cases} \tag{18}$$

We find in our experiments that $\theta = 2.5\pi$ yields good results.

The optional $w^h$ and $w^v$ convey discontinuity information, where lower values suggest more discontinuity and higher values otherwise. One may use the normalized directional gradient maps of the intensity image as the discontinuity maps. No discontinuity maps are used in our experiments.

The refined depth map $D$ is given by

$$D = d + k \cdot r. \tag{19}$$

Fig. 6 shows the input data, intermediate depth map after naive solution, and final result, for a sample scene.

In summary, the above algorithm recovers the wrapping number of all pixels, i.e., how many times the depth value has gone over the wrapping limit. To this end, we interpolate the initial interleaved (and incomplete) depth maps to generate correct results for smooth areas, and use median filter to find unstable areas that need special attention (i.e.,

global optimization). The median filter and the global optimization steps do not smooth the data; they are used to bridge the unstable areas by exploiting coherence from the input depth map. Finally, as Eq. 19 indicates, there is no loss in spatial resolution.

# 4. Experiments

In our experiments we use the SwissRanger [13] SR-4000 ToF camera for data collection. Because we do not have the actual sensor that can simultaneously capture at two modulation frequencies, we capture two real frames, then interleave the pixels according to our pixel arrangement. Since separate modulation frequencies do not interfere with each other, as demonstrated in Section 3.3, our data can be considered identical in quality to data captured by the actual hardware.

For our captures we use the frequency combination of 29MHz and 31MHz. Within the frequencies available to the camera, these are the largest co-prime pair that are not so far from each other as to behave discernibly different in noise performance. We also discuss other frequency pairs in the supplementary material.

We demonstrate the performance of our phase unwrapping framework on scenes captured in an indoor environment. Additionally, we use the dataset in [4] to quantitatively verify our method, as well as compare with the state-of-the-art algorithm in single-shot phase unwrapping. Furthermore, we compare our results with conventional multi-frequency phase unwrapping, which uses two modulation frequencies in the form of two separate frames.

## 4.1. Single-Frequency Unwrapping

Single-frequency unwrapping is performed using the method in [4]. Source code and the associated dataset are provided by the author. For our data, we use the phase images captured at the modulation frequency of 31MHz.

## 4.2. Conventional Multi-Frequency Unwrapping

We implemented a two-frequency unwrapping algorithm that globally optimizes the discontinuity term and frequency term as described in [5]. We used graph cuts as the optimization framework. To ensure performance in the presence of many discontinuities, we have also incorporated our data term, as described in Equation 16, into the algorithm.

# 5. Results

Figures 7∼9 showcase the performance of our unwrapping framework with respect to conventional two-frequency unwrapping and single-frequency unwrapping, using our data. From the simplest scene in Fig. 7 to the more complex scenes that follow, our method achieves results that significantly outperform the state-of-the-art single-frequency unwrapping algorithm, and are extremely close to results of conventional multi-frequency unwrapping. This is not surprising, as we manage to obtain a better initial estimation by exploiting spatial coherence in our rearranged pixels, and thus making the global optimization better conditioned. This is not possible with the single-frequency solution. We continue to observe that the single-frequency method is sensitive to the choice of parameters and illumination calibration, whereas our method requires no intensity calibration, and the same parameter settings can be applied across multiple images, even data captured by different cameras, with no apparent ill effect. Also, the virtually non-existent performance gap with conventional multi-frequency unwrapping demonstrates that, just as the multi-frequency method is theoretically capable of solving for each pixel independently, our method is practically capable of extracting coarse-scale information (i.e., wrapping) through careful arrangement of fine-scale information (i.e., interleaving different modulation frequencies).

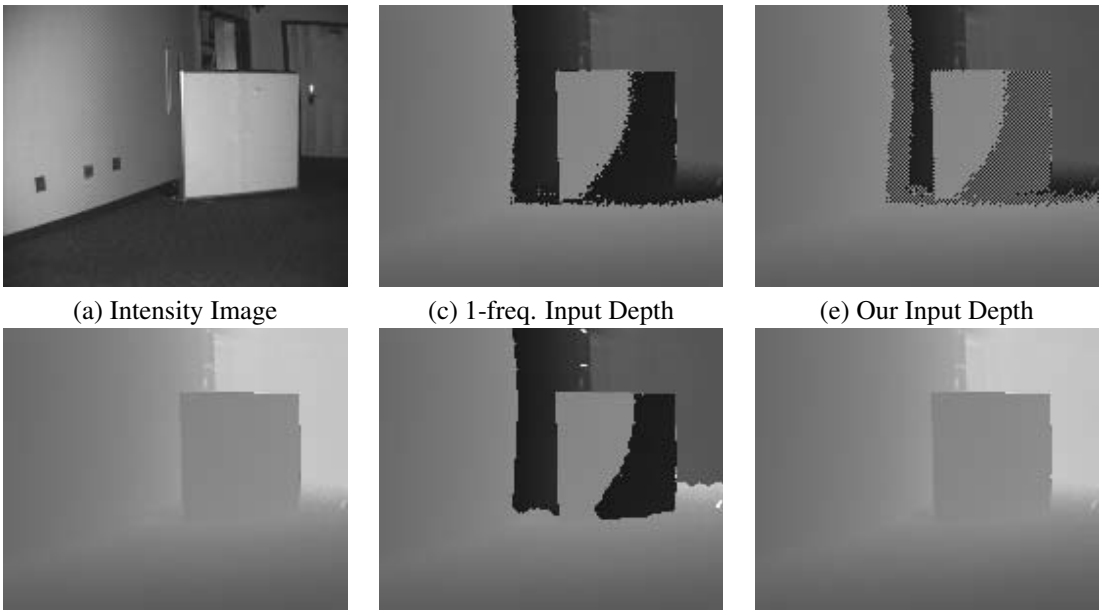
| Category | I | II | III |
|---|---|---|---|
| Number of scenes | 5 | 9 | 31 |
| Number of phase wraps | 1 | 2 | 3 |
| Percent correct (1-freq.) | 99.4% | 91.4% | 83.3% |
| Percent correct (Ours) | 99.9% | 99.8% | 97.7% |

Table 2. Quantitative evaluation on the dataset of [4]. Wrap count maps generated by [4] and our method are compared for the percentage of correct pixels. An unwrapped pixel is considered correct if its wrap count is equal to the wrap count of the same pixel in the ground truth.

Table 2 lists the performance comparisons on the dataset from [4] in terms of the percentage of correctly unwrapped pixels. Notice the performance drop of the single-frequency method at higher wrap numbers, as the algorithm cannot keep up with scene complexity. In contrast, given the combination of the guidance of depth knowledge in stable areas and continuity constraints in unstable areas, there is no significant performance deterioration with our algorithm. A sample scene where both algorithms are compared w.r.t. ground truth is shown in Fig. 10.

# 6. Conclusion

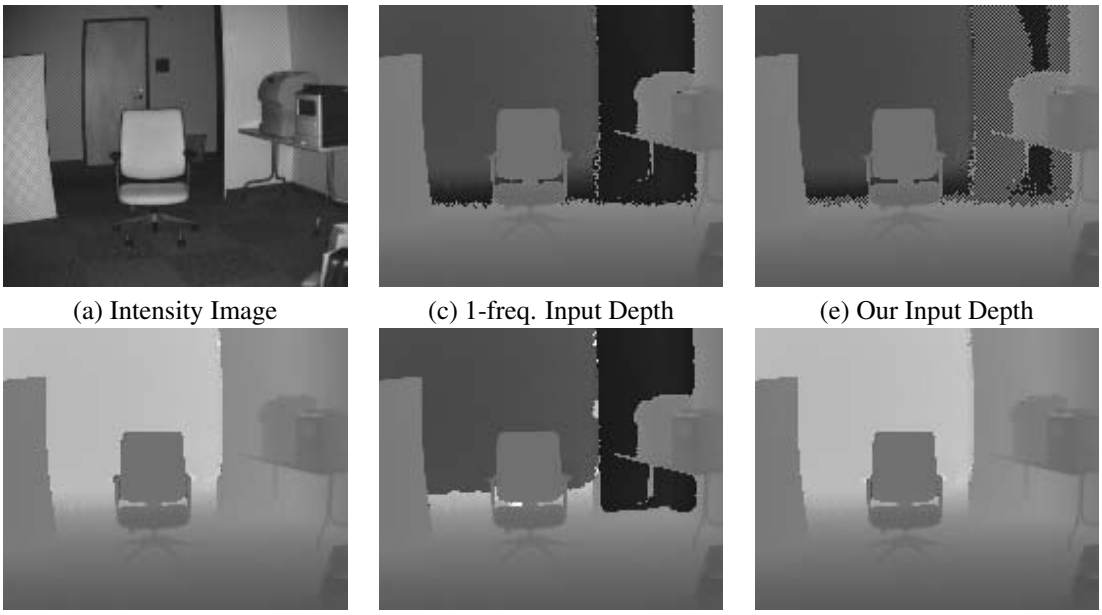
In this paper, we propose a novel single-shot phase unwrapping framework for the ToF camera. We introduce a special pixel arrangement, where spatially interleaved pixels capture simultaneously at different modulation frequencies. We have proved that having two modulation frequencies on the same sensor does not cause interference as long as the frequencies are sufficiently apart. In our unwrapping algorithm, interpolating the partial depth maps provides good guesses in smooth areas, and the continuity constraint in the global optimization helps maintain performance in ar-



(a) Intensity Image (c) 1-freq. Input Depth (e) Our Input Depth

(b) Conventional 2-freq. Result (d) 1-freq. Result (f) Our Result

Figure 7. Our scene #1. From (a) to (f) are: (a) Intensity image; (b) Reference unwrapping result, using two frequencies, two frames; (c) Raw depth map, single frequency; (d) Single frequency unwrapping solution from (c); (e) Our interleaved depth map, assembled from two raw depth maps; (f) Solution of our unwrapping algorithm. Lighter shades imply greater distance in the depth maps. 99.98% of all pixels in (b) and (f) have the same wrap count.

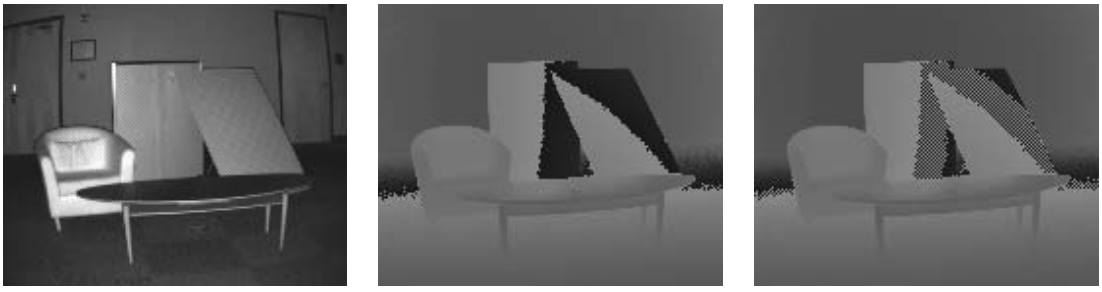


(a) Intensity Image (c) 1-freq. Input Depth (e) Our Input Depth

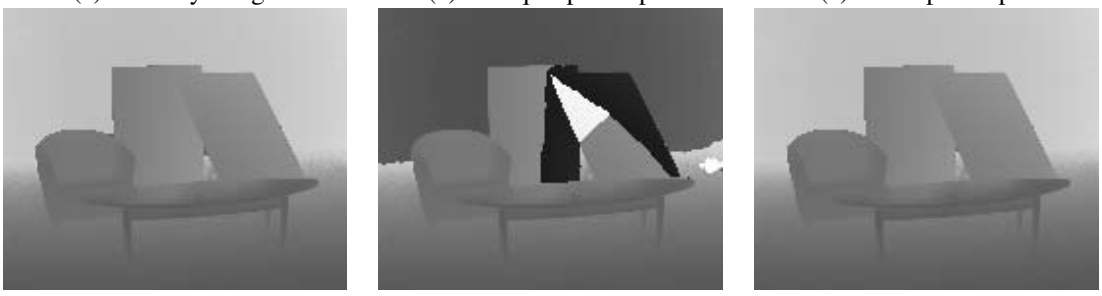
(b) Conventional 2-freq. Result (d) 1-freq. Result (f) Our Result

Figure 8. Our scene #2. From (a) to (f) are: (a) Intensity image; (b) Reference unwrapping result, using two frequencies, two frames; (c) Raw depth map, single frequency; (d) Single frequency unwrapping solution from (c); (e) Our interleaved depth map, assembled from two raw depth maps; (f) Solution of our unwrapping algorithm. Lighter shades imply greater distance in the depth maps. 99.94% of all pixels in (b) and (f) have the same wrap count.

eas with difficulties such as wraps, discontinuities, and high noise. We demonstrate that our method achieves compara-

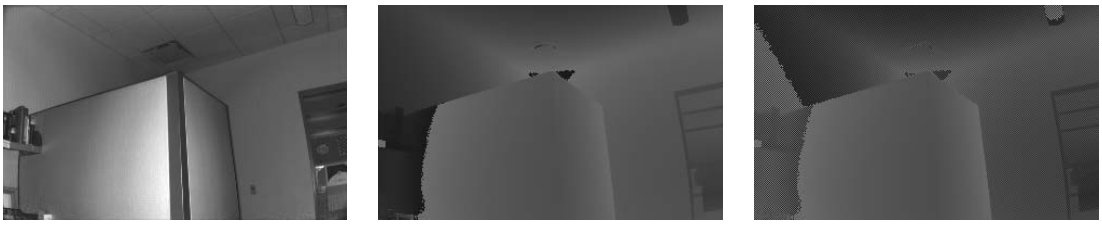


(a) Intensity Image (c) 1-freq. Input Depth (e) Our Input Depth



(b) Conventional 2-freq. Result (d) 1-freq. Result (f) Our Result

Figure 9. Our scene #3. From (a) to (f) are: (a) Intensity image; (b) Reference unwrapping result, using two frequencies, two frames; (c) Raw depth map, single frequency; (d) Single frequency unwrapping solution from (c); (e) Our interleaved depth map, assembled from two raw depth maps; (f) Solution of our unwrapping algorithm. Lighter shades imply greater distance in the depth maps. 99.96% of all pixels in (b) and (f) have the same wrap count.



(a) Intensity Image (c) 1-freq. Input Depth (e) Our Input Depth

(b) Ground Truth (d) 1-freq. Result (f) Our Result

Figure 10. Scene #9 from the dataset in [4]. From (a) to (f) are: (a) Intensity image; (b) Ground truth depth map in the dataset; (c) Raw, single frequency depth map; (d) Single frequency unwrapping solution of (c); (e) Our interleaved depth map, assembled from two raw depth maps; (f) Solution of (e) using our unwrapping algorithm. Lighter shades imply greater distance in the depth maps. The ratio of correctly unwrapped pixels is 99.83% in our method v.s. 80.59% in the single-frequency unwrapping method.

ble performance as the conventional two-shot unwrapping algorithm, and significantly outperforms the state-of-the-art in single-shot unwrapping methods.

## 7. Acknowledgements

This work is supported in part by US NSF grants IIS-1208420 and IIP-1543172.

## References


[1] J. M. Bioucas-Dias and G. Valadão. Phase unwrapping via graph cuts. *IEEE Transactions on Image Processing,*, 16(3):698–709, 2007.

[2] O. Choi, S. Lee, and H. Lim. Interframe consistent multifrequency phase unwrapping for time-of-flight cameras. *Optical Engineering*, 52(5):057005–057005, 2013.

[3] O. Choi, H. Lim, B. Kang, Y. S. Kim, K. Lee, J. D. Kim, and C.-Y. Kim. Range unfolding for time-of-flight depth cameras. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 4189–4192. IEEE, 2010.

[4] R. Crabb and R. Manduchi. Probabilistic phase unwrapping for single-frequency time-of-flight range cameras. In *3D Vision (3DV), 2014 2nd International Conference on*, volume 1, pages 577–584. IEEE, 2014.

[5] D. Droeschel, D. Holz, and S. Behnke. Multi-frequency phase unwrapping for time-of-flight cameras. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1463–1469, 2010.

[6] D. Falie and V. Buzuloiu. Wide range time of flight camera for outdoor surveillance. In *Microwaves, Radar and Remote Sensing Symposium, 2008. MRRS 2008*, pages 79–82. IEEE, 2008.

[7] D. C. Ghiglia and M. D. Pritt. Two-dimensional phase unwrapping: theory, algorithms, and software. 4, 1998.

[8] S. B. Gokturk, H. Yalcin, and C. Bamji. A time-of-flight depth sensor-system description, issues and solutions. In *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*, pages 35–35. IEEE, 2004.

[9] A. Kadambi, A. Bhandari, R. Whyte, A. A. Dorrington, and R. Raskar. Demultiplexing illumination via low cost sensing and nanosecond coding. In *2014 IEEE International Conference on Computational Photography, ICCP 2014, Santa Clara, CA, USA, May 2-4, 2014*, pages 1–10, 2014.

[10] V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, 2004.

[11] J. Mei, A. Kirmani, A. Colaço, and V. K. Goyal. Phase unwrapping and denoising for time-of-flight imaging using generalized approximate message passing. In *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pages 364–368. IEEE, 2013.

[12] S. K. Nayar and S. G. Narasimhan. Assorted pixels: Multisampled imaging with structural models. In *Computer VisionECCV 2002*, pages 636–652. Springer, 2002.

[13] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc. An all-solid-state optical range camera for 3d real-time imaging with sub-centimeter depth resolution (swissranger). In *Optical Systems Design*, pages 534–545. International Society for Optics and Photonics, 2004.

[14] W. Xu, E. C. Chang, L. K. Kwoh, H. Lim, W. Cheng, and A. Heng. Phase-unwrapping of sai interferogram with multi-frequency or multi-baseline. 1994.