I t is a novel method which solves a difficult problem in T O F cameras. More research is needed to make it more practical, however the direction is quite promising.

2) Clarity of Exposition
The exposition is clear.

3) Quality of References
Yes

4) Reproducibility
Yes, if you have a T O F (PMD) camera.

5) Rating
4.0

7) Explanation of Rating
It is a novel method which solves a difficult problem in T O F cameras. More research is needed to make it more practical, however the direction is quite promising.

Reviewer #5:

This paper presents a technique to modify time-of-flight cameras to capture multiple reflection events along a line-of-sight. In comparison, conventional cameras assume that each camera pixel observes a single reflection/scattering event.

There are two main ideas:
1) Make the assumption that each camera pixel receives illumination corresponding only to a sparse and discrete set of light source rays.

In this case, the temporal irradiance profile at the camera can be expressed as a sum of a few discrete shifted profiles (alternatively, as authors explain it, convolution with a sparse vector).

2) Design illumination and camera modulation (temporal) codes that allow separation of the discrete profiles via regularized deconvolution.

The ideas presented in this paper are interesting, and I believe the community will find them useful. T O F imaging is becoming popular these days in many application areas, and multi-path interference is an important problem in T O F.

However, in the current form, there are several limitations that need to be addressed. Also, the applications are not well motivated (see details below). If the paper can address these issues within the time between the review and final submission, it can be accepted. If not, it should be revised and resubmitted.

Reviewer #90:

Title:

Id:

The paper presents a new technique to compute the contributions of multipath rays for time-of-flight depth (range) cameras. It thus enables to measure depth of near-transparent objects, which is a hard task for T O F technology. It also enables to resolve depth of edge pixels (where two objects with different depth are in a single pixel’s field-of-view) and, as a byproduct, to measure the illumination amplitude beyond diffused objects.

The technique is based on measuring the cross correlation function and deconvolving it to recover multiple time-shifts. In the paper only two parallel time-shifts are shown, which are sparse (delta-like, cannot recover scattered multi-path). The deconvolution uses a variant of orthogonal-matching-pursuit algorithm with additional non-negativity and proximity constraints, which fit best the physical problem. The technique seems to work well and outperform other deconvolution methods.

The advantages of the technique is that it is based on off-the-shelf T O F technology, without any significant assumptions or additional complex hardware.

Disadvantages are the current practicality of the method for real applications:

The spatial resolution is only 120x160, the depth resolution is 2cm and the time for a single frame is 4 seconds. Therefore it is not practical to be used for any depth-measurements involving motion and neither for high quality still measurements.

1) Description

The paper presents a new technique to compute the contributions of multipath rays for time-of-flight depth (range) cameras. It thus enables to measure depth of near-transparent objects, which is a hard task for T O F technology. It also enables to resolve depth of edge pixels (where two objects with different depth are in a single pixel's field-of-view) and, as a byproduct, to measure the illumination amplitude beyond diffused objects.

The technique is based on measuring the cross correlation function and deconvolving it to recover multiple time-shifts. In the paper only two parallel time-shifts are shown, which are sparse (delta-like, cannot recover scattered multi-path). The deconvolution uses a variant of orthogonal-matching-pursuit algorithm with additional non-negativity and proximity constraints, which fit best the physical problem. The technique seems to work well and outperform other deconvolution methods.

The advantages of the technique is that it is based on off-the-shelf TOF technology, without any significant assumptions or additional complex hardware.

Disadvantages are the current practicality of the method for real applications:

The spatial resolution is only 120x160, the depth resolution is 2cm and the time for a single frame is 4 seconds. Therefore it is not practical to be used for any depth-measurements involving motion and neither for high quality still measurements.

2) Clarity of Exposition
The exposition is clear.

3) Quality of References
Yes

4) Reproducibility
Yes, if you have a TOF (PMD) camera.

5) Rating
4.0

7) Explanation of Rating
It is a novel method which solves a difficult problem in TOF cameras. More research is needed to make it more practical, however the direction is quite promising.
Following are the main issues /suggestions:

1) Sparse time-profiles: The presented technique relies on the assumption that each camera pixel receives illumination originating at only a few (2 in scenarios shown in the paper) light source rays. The supplementary video states that non-sparse profiles occur only for specialized cases. That is not true.

In fact, sparse profiles happen for very specialized cases - in scenes involving "only" specular reflections, i.e., a scene made only of mirrors or polished glass. Almost all real-world scenes have diffuse interreflections, which will result in non-sparse profiles.

It is not clear how the technique will perform in such a scenario. This is a serious limitation, and is not discussed in the paper. Without addressing this, the technique remains limited to a very restricted set of scenarios.

2) Design of codes: Perhaps the main contribution of the paper is designing codes that allow separation of different profiles. While there is a discussion about different choices of the codes, I believe the paper will be strengthened by discussing the two obvious codes: (a) sinusoidal code, (b) delta code.

Most current time-of-flight cameras use sinusoidal codes. Also, several older ToF cameras used the delta code. These two codes are interesting as they lie at the extremes of the frequency spectrum (sinusoid’s frequency response is a delta, whereas delta’s frequency response is flat).

What is the advantage of the proposed codes over these? I suspect that the sinusoidal codes will not allow the sparse deconvolution because the resulting Toeplitz matrix will not be invertible. It will be good to show this.

But, what about the delta codes? The Toeplitz matrix made of the delta code is an identity matrix, which is obviously invertible. So, what is the advantage of using the proposed codes over the delta codes? Is there an SNR argument (delta codes do not capture sufficient light)? If so, that argument should be made clearly in the paper.

I believe this discussion would make the paper considerably stronger.

3) Proposed applications: While I like the application regarding separating multiple reflections, I am not convinced about the time profile imaging application. Such a sequence of images can be created simply by using a conventional camera, and capturing very short exposure photographs using an external trigger (such as the one used in this paper). The camera’s frame rate need not be very high. The trigger’s temporal location can be moved to create the “light-sweep”.

It is not clear how the presented technique improves over the above mentioned baseline method?

4) Number of measurements: There should be a discussion on the number of measurements required as a function of the sparsity factor. Clearly, as the number of spikes in the environment function increases, the number of measurements should increase accordingly. A theoretical/empirical discussion on this issue will help strengthen the paper.

Minor comment, but it will also be more compelling to show examples with more than two spikes.

The paper is generally clear, but it will benefit from motivating the code design part better. See above for details.

Also, the image formation process can be better explained. What is the camera exposure used? How many images are captured? A lot of important details are missing.

References are adequate.

In the related work section, Narasimhan et al.'s paper is cited under depth imaging of translucent objects. Strictly speaking, that paper deals with the problem of depth imaging "through" translucent media.

For depth imaging of the surfaces of translucent objects, please see:

1) Micro Phase Shifting, Mohit Gupta and Shree Nayar
2) Polarization and Phase Shifting for 3D Scanning of Translucent Materials, Chen et al.
Reviewer #65:

This paper describes a technique for recovering multiple per-pixel depths taken by a time-of-flight camera/light setup that includes amplitude modulated coded illumination and detection systems. By assuming that the multiple depths being recovered occur at hard surfaces, and thus create delta spikes, a numerical method is proposed to effectively deconvolve the cross correlation measurement to recover the delta spikes, and thus the multiple depths. The method works on commercially available hardware, and had several other advantages. Several examples are demonstrated.

The seems practical and the results are pretty good given the hardware.

1) Description

I found 80 - 99 really awkward. There seem to be a lot of incomplete sentences with missing elements. I don't like "multiple" and "multipath" being used as nouns (if that is even what you meant). Are you required to know K (the number of expected depths) before performing the solve?

2) Clarity of Exposition

Nothing missing that I am aware of.

3) Quality of References

Seems fine.

4) Reproducibility

My confidence is very low in this area, but the paper seems to present novel material that is interesting to the community, seems to work quite well given the hardware costs, and the video and images show some compelling data. The depth recovery in scenes containing transparent/translucent interfaces is indeed challenging and this seems to present a solid step forward in this area. The limitations to not necessarily working well with diffusive media seem fair, although I would expect the same issue in a complex environment with many high albedo surfaces, where many orders of interreflection would contribute to the light reflected to any given pixel, smearing out the deltas in a way that would seem unlikely to recover.

5) Rating

Reviewer #10:

The paper demonstrates how per-pixel time profiles can be extracted as a sequence of impulses. The idea is simple: use a custom code that has strong auto-correlation but weak cross-correlation properties. The authors showed that the Maximum Length (Pseudorandom Number) sequence is the most promising. With this idea, the authors are able to show time profile imaging, extracting depths of a transparent object and background, "looking" through a diffuse material, and handling boundary pixel errors. This work is significant, given the new technical contributions and the increasing interest in TOF cameras.

1) Description

The paper is generally well-written. There are typos:
* Line 84: "multiple" -> "multiple paths"
* Equation between lines 143-144: ... {t} -> ... \( \tau \)
* Line 218: "digonialized" -> "diagonalized"
* Figure 5 caption: "Finall" -> "Finally"

Line 166: Is \( \tau \) the same as \( \varphi \)?

2) Clarity of Exposition

The references are adequate. The authors may be interested in commenting on the following paper and how the errors change with their technique:
M. Frank et al., "Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras," Optical Engineering, 48(1), 2009. (They show that the computed range follows an Offset Normal distribution.)

3) Quality of References
4) Reproducibility
The work can be replicated by someone with a good hardware/electronics background.

5) Rating
4.0
This is a very nice paper: it has a great idea, the math is solid, and it shows never-before seen results with a TOF camera. The paper is also generally well-written.

The authors should acknowledge that phase-based TOF is one of two major types; the other is pulse-based (with built-in shutters). How would the technique change for a pulse-based TOF camera?

7) Explanation of Rating
Would the technique help resolve the multi-path problem when a concave object (e.g., inside of a bowl, even a face) is used? Cases like this should be discussed.

Transparent objects tend to have internal reflections. This hasn't been observed to be a problem?

I would have given the paper a higher score if the real-time performance was explored with success.

Reviewer #93:

1) Description
This paper introduces a method for resolving TOF traces in the case of multiple (non-scattering) interactions. The last interactions can be scattering I guess. The key observation is that in such conditions the trace is basically a sparse sum of (near-)dirac functions. Hence, the problem now comes down to computing the phase-shifts (i.e., location) of each dirac (and the number diracs) which involves a deconvolution (the hardware introduces a low-pass blur on the incoming signal). A single frequency wave is used a lighting signal, and various different coding (i.e., at the pixel) signals are investigated. Not surprisingly, a simple block signal (cf dirac train) is not a good signal when deblurring the incoming signal.

I am not an expert on TOF hardware, hence I do not know how novel it is to look at different control signals. However, mathematically, there not that much novelty. The applicability is fairly limited as very few (mostly uninteresting) scenes yield sparse profiles.

2) Clarity of Exposition
The exposition is not good. There are numerous typos. I could not parse some sentences. Section 3 does not help to explain the principle behind TOF cameras (the video does a much better job). The transition between consecutive subsections is abrupt, and often one wonders what is now being discussed. It feels as if the paper was written in a rush -- it misses coherency. The mathematical derivation is too vague and too short on the critical parts, while being overly lengthy on trivial parys.

3) Quality of References
OK. Maybe some deblurring references would be in order?

4) Reproducibility
No, the method cannot be reproduced. The exposition is just too vague that it is not clear what the authors have exactly done.

The limitations are fairly well discussed.

5) Rating
2.9
I find repurposing TOF cameras an interesting research topic. It is clear that there is still a lot of low hanging fruit in this area.

This paper tries to answer an interesting question: "what if the trace is sparse? can we exploit this?"

Unfortunately, sparse time traces do not occur in many interesting situations -- all but the last bounce should be 'specular'. The light sweeps do not make much sense (basically you are showing a temporal encoding of the depth image) for the examples shown. It would have been much more interesting to show this on a scene with more interreflections.

What I do not really understand is why this paper only focuses on sparse traces in the canonical domain? One could enforce sparsity in any domain -- cf compressive sensing. It would have been more interesting to explore this in the context of compressive sensing; the only difference is the additional convolution, but that could be solved on the fly by just taking it in account when solving the L1 minimization.

7) Explanation of Rating
The exposition is very confusing. It is not clear which algorithm was used to compute the solution (It seems like a variant of OMP, but I am not sure since the line before states that another algorithm was used). Certain steps are supposedly justified (e.g., why binary control signals are used), but to me the connection was not always immediately clear.

It is not clear why exactly those binary control signals were used and why not some others (why not a random signal as in compressive sensing?).
I think the authors are on to something interesting, but the current state of the manuscript do not meet the quality-bar expected for SIGGRAPH Asia. Furthermore, the novelty is at this point limited (the whole mathematics can be explained more concisely and more clearly). However, it form a great starting point for more generalizations.

**Reviewer #51:**

<table>
<thead>
<tr>
<th>1) Description</th>
<th>The authors present a new signal processing method for time-of-flight cameras, based on compressed sensing (specifically, l1 norm optimization).</th>
</tr>
</thead>
<tbody>
<tr>
<td>2) Clarity of Exposition</td>
<td>I think the exposition is adequate.</td>
</tr>
<tr>
<td>3) Quality of References</td>
<td>I don't have enough background knowledge in the field of time-of-flight cameras.</td>
</tr>
<tr>
<td>4) Reproducibility</td>
<td>L1 norm optimization has been widely used in recent years, some solvers are available on the Internet. I would assume it is not difficult to implement the paper once a solver is obtained. Time-of-flight cameras are available on the market, but I would assume it is slightly difficult to customize.</td>
</tr>
<tr>
<td>5) Rating</td>
<td>2.75</td>
</tr>
<tr>
<td>7) Explanation of Rating</td>
<td>The authors use l1 norm optimization to retrieve correct response from the reflected profile. It is more or less similar to a compressed sensing problem. I believe technically it is a good paper. However, I would assume this paper will be a perfect candidate for computer vision conferences instead of SIGGRAPH, which is a graphics-motivated conference. I can’t find how this technique can be related to real problems in 3D graphics, or I can’t find a good graphics application that will actually be benefited from this technology. For example applications such as digitizing 3D object or scanning a large environment such as a movie set. I would suggest the authors show one or more captured scene with actual 3D geometry. Looking through a diffuser is kind cool, but still lack of practical use in graphics community. If the authors are able to address more discussion regarding the application side then I would love to raise my score.</td>
</tr>
</tbody>
</table>

**Submission Information:**

Logged in as: achoo@mit.edu
You must have session cookies enabled to use this system.