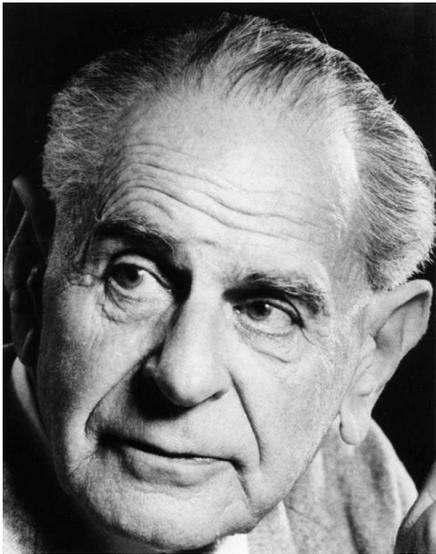




TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

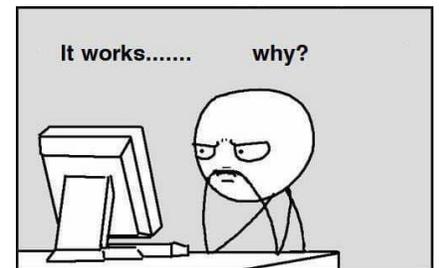
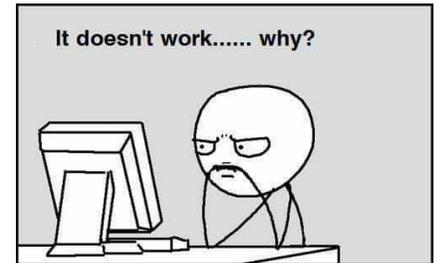
Iuri Frosio, GTC 2019 (San Jose, CA)

THE IMPORTANCE OF NEGATIVE RESULTS



“I shall require that [the] logical form [of the theory] shall be such that it can be singled out, by means of empirical tests, in a negative sense: it must be possible for an empirical scientific system to be refuted by experience” (Karl Popper, The Logic of Scientific Discovery, 1959).

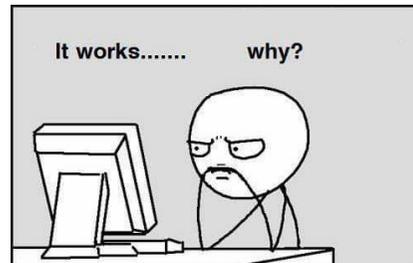
In simple words, “negative results are fundamentals for the advancement of science”.



TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*
- Back to physics
- DNN architecture
- Results
- Conclusion



* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

TIME OF FLIGHT (TOF) CAMERAS & ARTIFACTS

E.g., Kinect 2

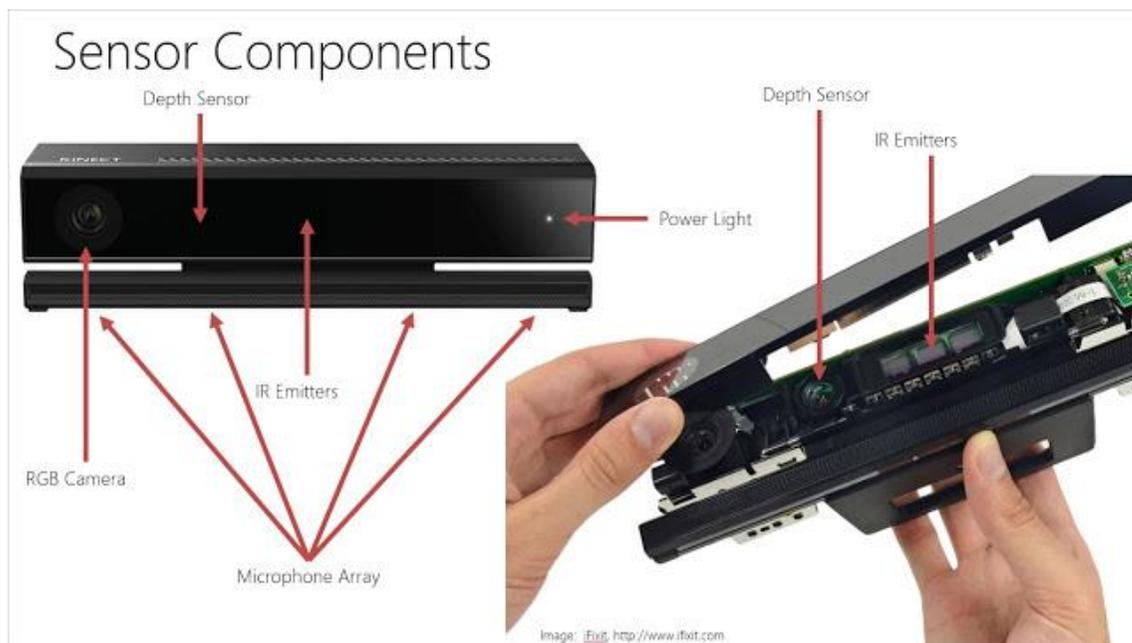


Image from <https://stackoverflow.com/questions/22921390/how-to-scale-a-kinect-depth-image-for-applying-lbp-on-it-in-matlab?rq=1>

TIME OF FLIGHT (TOF) CAMERAS & ARTIFACTS

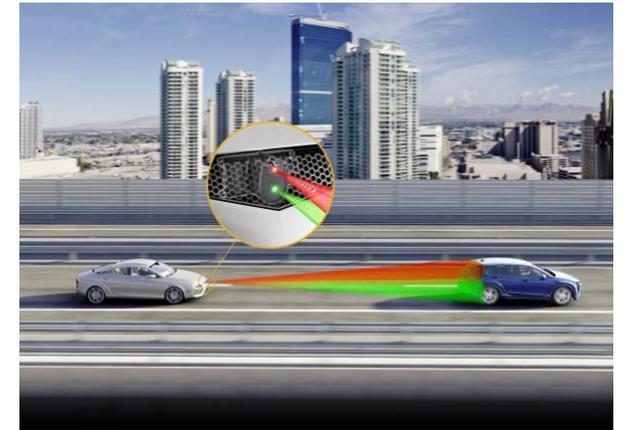
Applications



Image from amazon.com



Image from https://www.physio-pedia.com/The_emerging_role_of_Microsoft_Kinect_in_physiotherapy_rehabilitation_for_stroke_patients



TIME OF FLIGHT (TOF) CAMERAS & ARTIFACTS

Artifact #1: shot noise

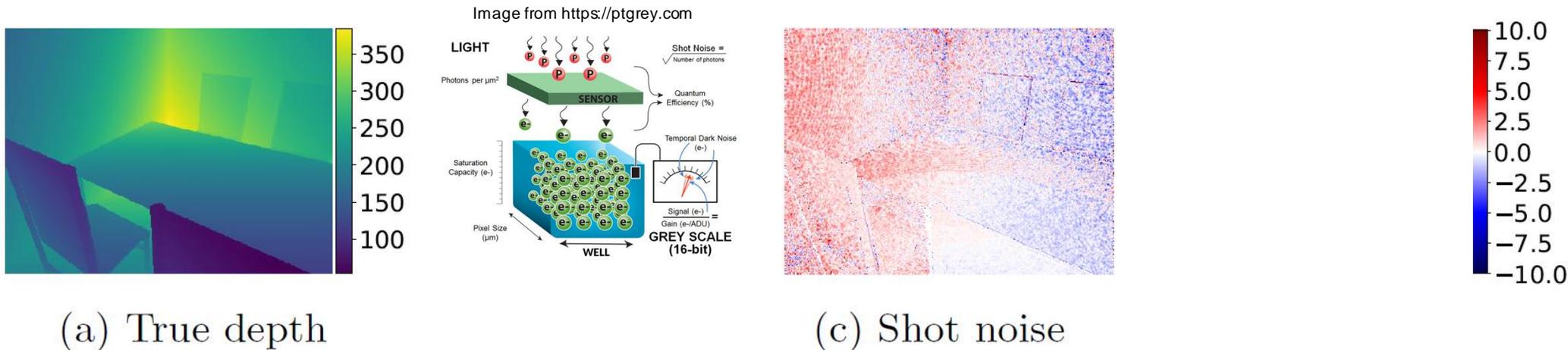
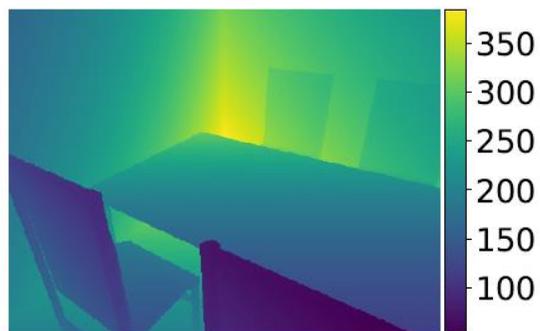


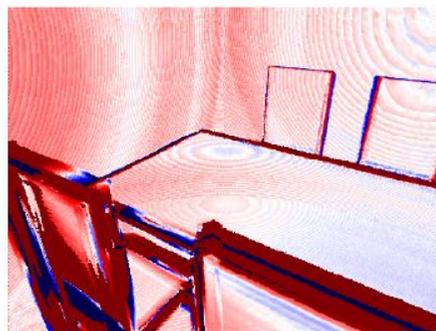
Fig. 3: Effect of non-idealities on the LF2 depth reconstruction error, in cm.

TIME OF FLIGHT (TOF) CAMERAS & ARTIFACTS

Artifact #2: movement



(a) True depth



(b) Motion

Image from <https://support.xbox.com>

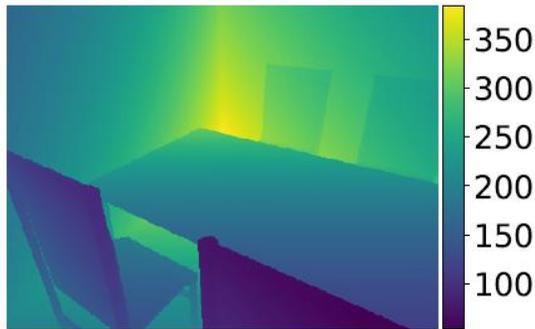


Fig. 3: Effect of non-idealities on the LF2 depth reconstruction error, in cm.

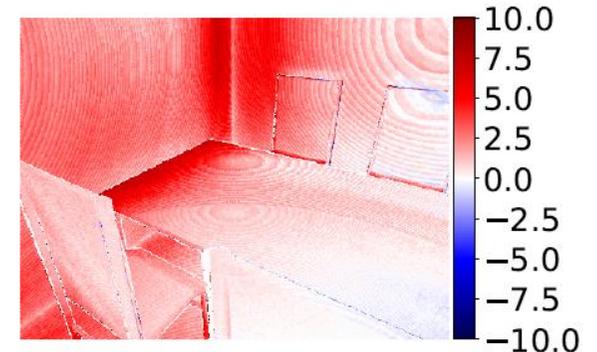
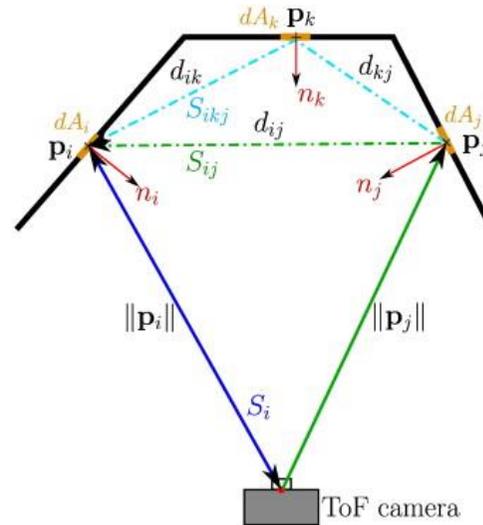
TIME OF FLIGHT (ToF) CAMERAS & ARTIFACTS

Artifact #3: multiple reflections

Image from
<https://www.sciencedirect.com/science/article/pii/S0262885613001650>



(a) True depth



(d) MPI

Fig. 3: Effect of non-idealities on the LF2 depth reconstruction error, in cm.

TIME OF FLIGHT (TIF) CAMERAS & ARTIFACTS

Artifact #3: multiple reflections

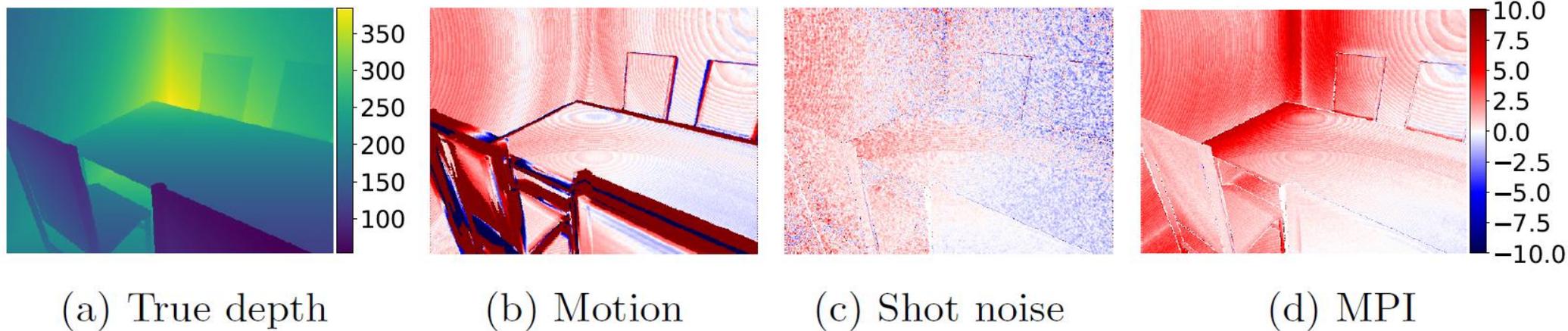


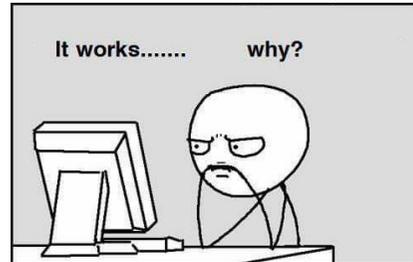
Fig. 3: Effect of non-idealities on the LF2 depth reconstruction error, in cm.

TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*

- Back to physics
- DNN architecture
- Results
- Conclusion



* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

NAÏVE MACHINE LEARNING (ML) FOR TOF RECONSTRUCTION

What do we need?

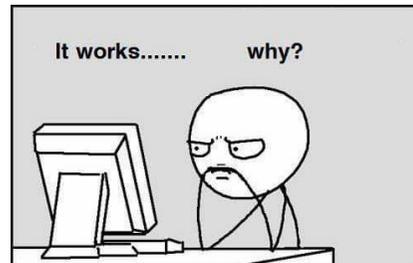


- (1) A large dataset of scenes...
- (2) ... corrupted by:
 - (1.1) photon noise,
 - (1.2) motion,
 - (1.3) multiple reflections...
- (3) ... with clean output data...
- (4) ... And a DNN.

TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*
- Back to physics
- DNN architecture
- Results
- Conclusion



* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

TOF WORKING PRINCIPLES

Time of flight is not time of flight 😊

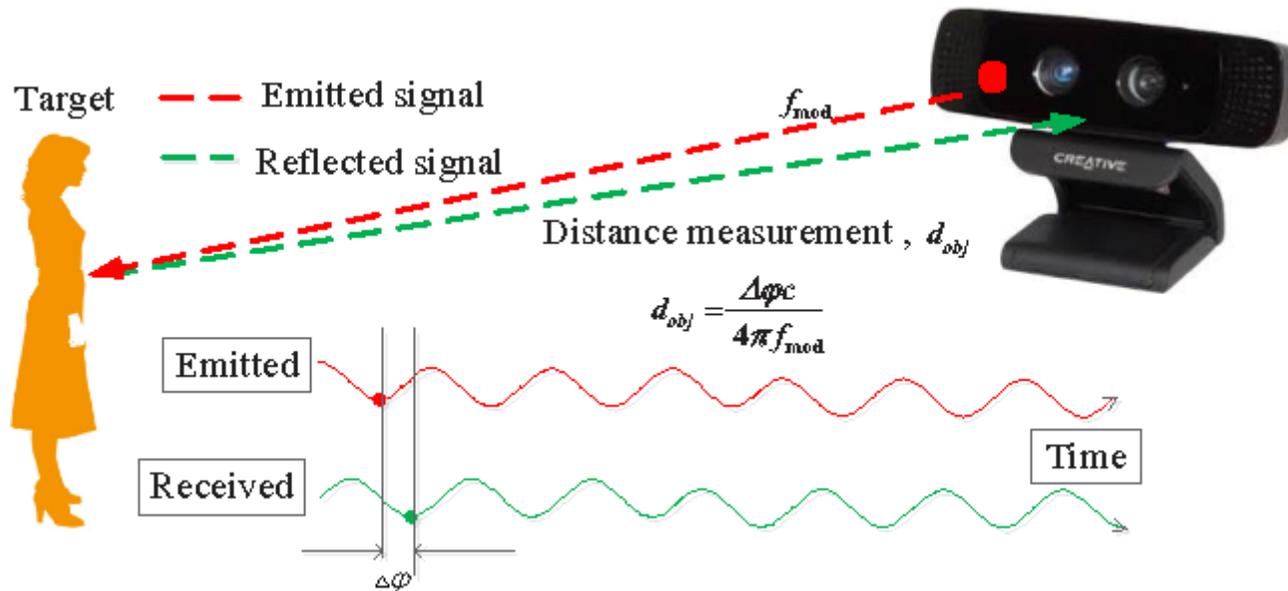
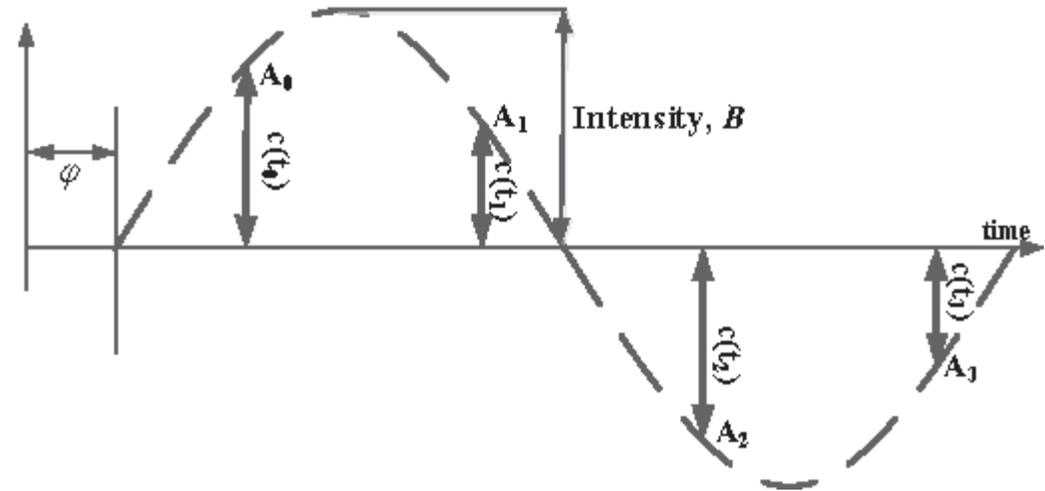
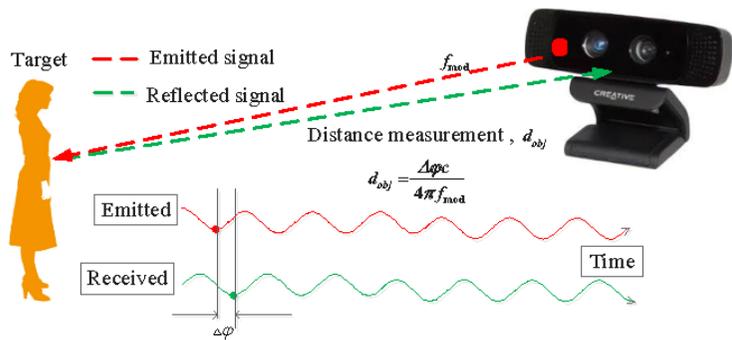


Image from [https://www.semanticscholar.org/paper/Interference-mitigation-technique-for-\(ToF\)-camera-Islam-Hossain/](https://www.semanticscholar.org/paper/Interference-mitigation-technique-for-(ToF)-camera-Islam-Hossain/)

TOF WORKING PRINCIPLES

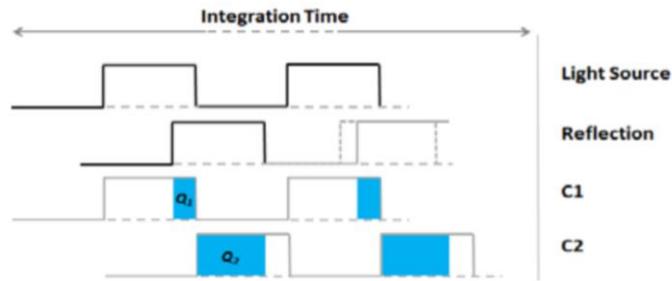
Time of flight is not time of flight 😊



TOF WORKING PRINCIPLES

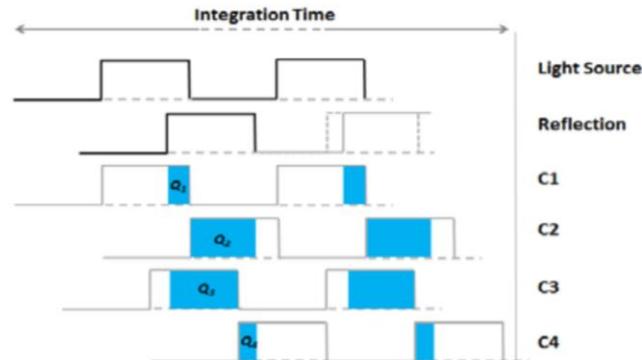
Multiple measurements

Pulse method



$$d = \frac{1}{2} c \Delta t \left(\frac{Q_2}{Q_1 + Q_2} \right).$$

Continuous wave method

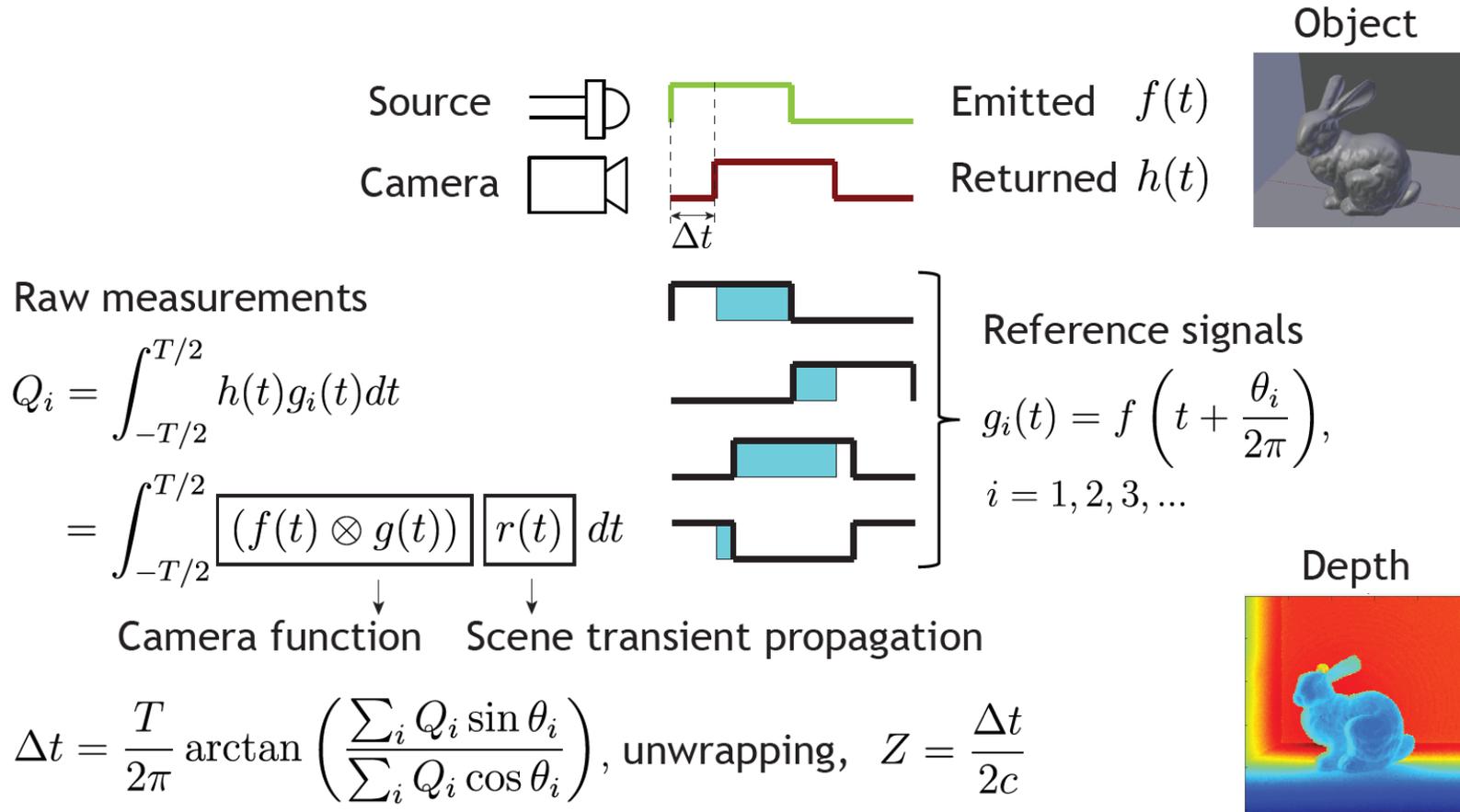


$$\varphi = \arctan \left(\frac{Q_3 - Q_4}{Q_1 - Q_2} \right),$$

$$d = \frac{c}{4\pi f} \varphi.$$

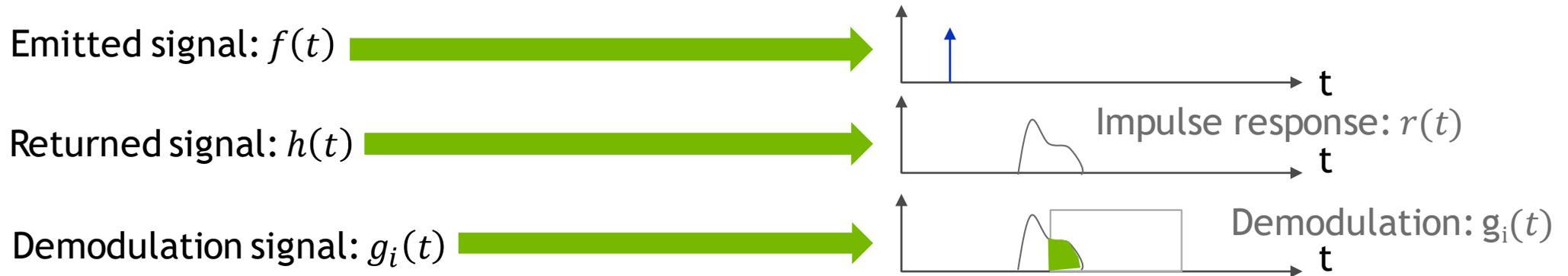
TOF WORKING PRINCIPLES

Camera functions and scene response



TOF WORKING PRINCIPLE

More on scene response



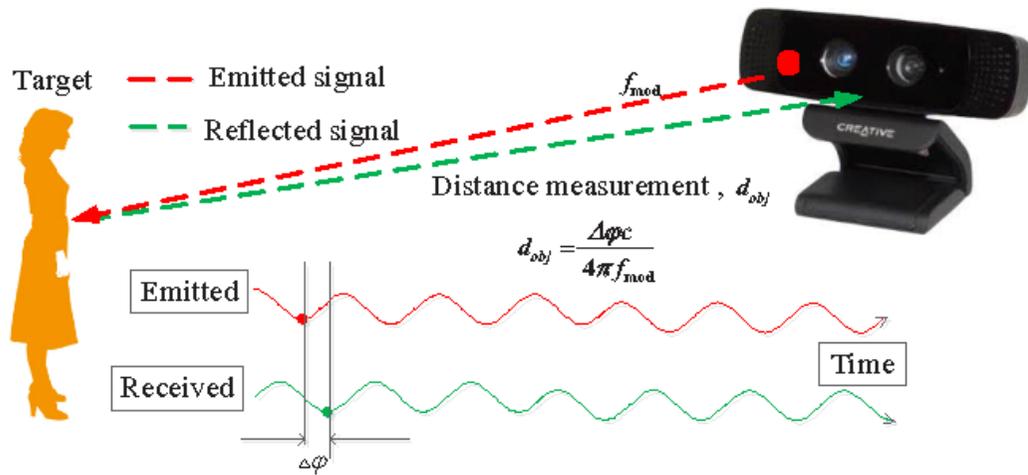
$$\text{Raw measurement: } Q_i(t) = \int_{-\frac{T}{2}}^{\frac{T}{2}} h(t) g_i(t) dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} (f(t) * r(t)) g_i(t) dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} \boxed{(f(t) \otimes g_i(t))} \boxed{r(t)} dt$$

The equation shows the raw measurement $Q_i(t)$ as the integral of the returned signal $h(t)$ multiplied by the demodulation signal $g_i(t)$ over the interval $[-\frac{T}{2}, \frac{T}{2}]$. This is equivalent to the integral of the convolution of the emitted signal $f(t)$ and the impulse response $r(t)$ multiplied by $g_i(t)$. The final form shows the correlation of the emitted signal $f(t)$ and demodulation signal $g_i(t)$ (labeled 'Camera') with the impulse response $r(t)$ (labeled 'Scene').

$$\text{Depth: } Z = \frac{Tc}{4\pi} \arctan \frac{\sum_i \sin \theta_i Q_i(t)}{\sum_i \cos \theta_i Q_i(t)}$$

TOF WORKING PRINCIPLES

Multiple frequencies

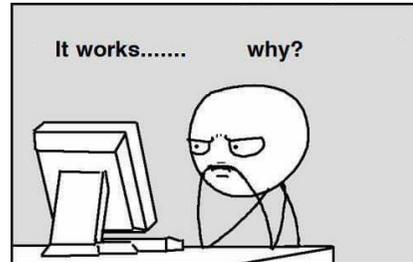
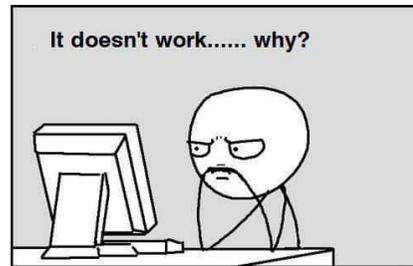


- Different max length (combine them)
- Different resolutions
- Agreement between different measurements

TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

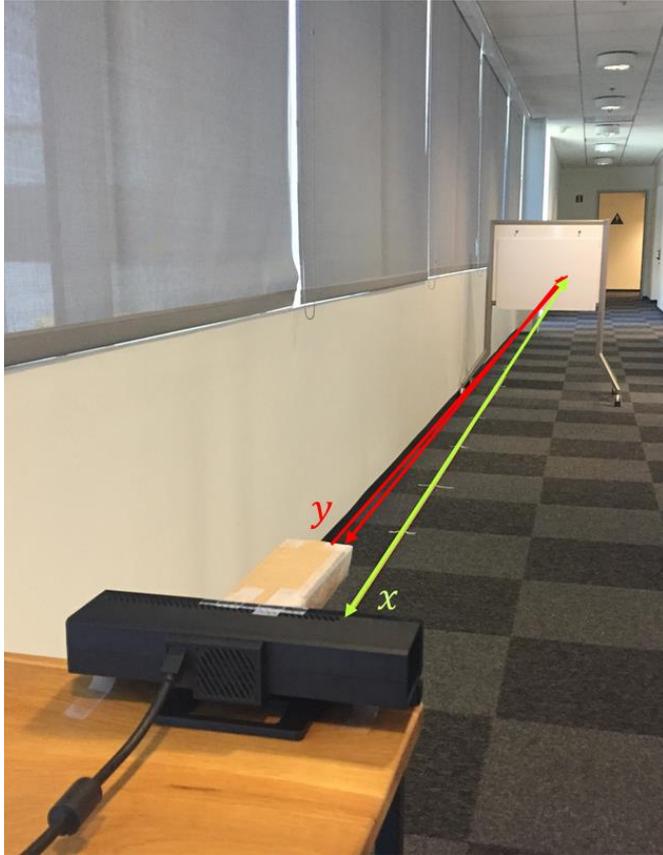
- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*
- Back to physics
- DNN architecture
- Results
- Conclusion



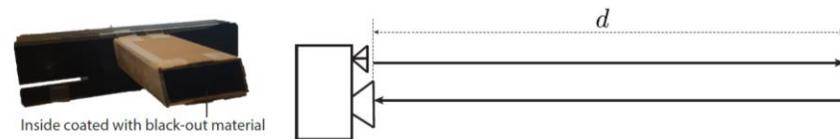
* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

CAMERA CALIBRATION

Camera response functions (flat scene)

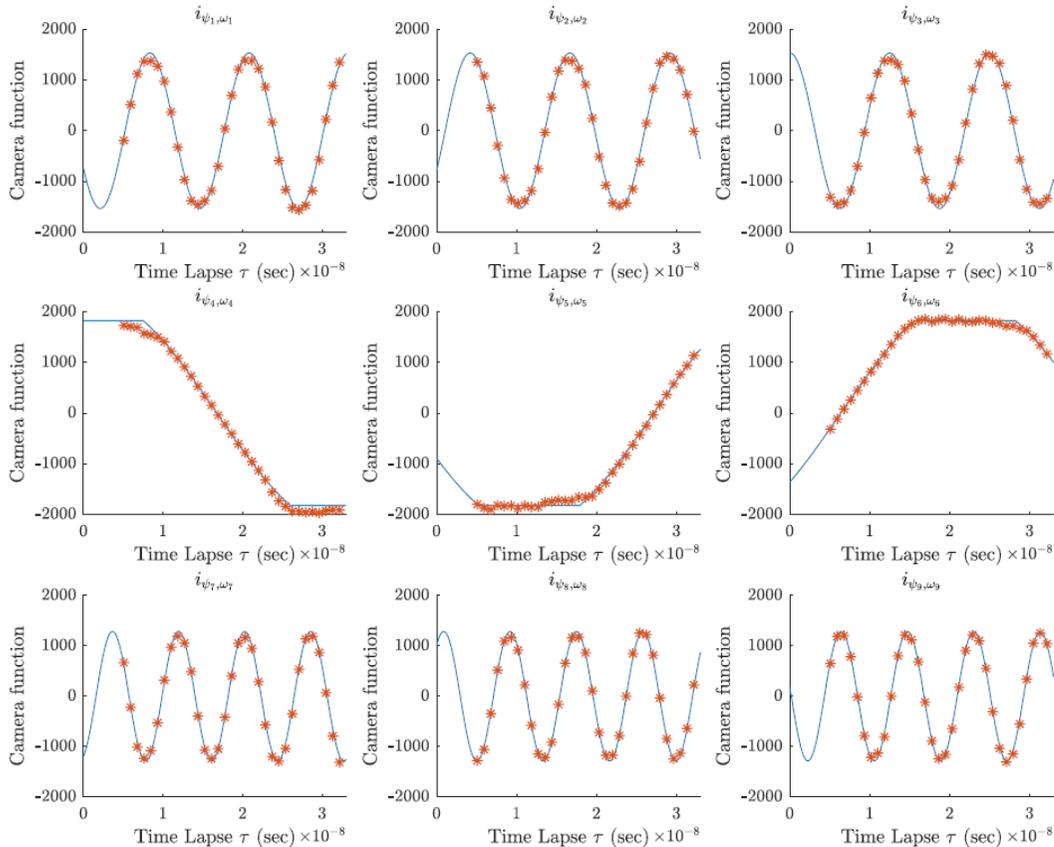


Inside coated with black-out material



CAMERA CALIBRATION

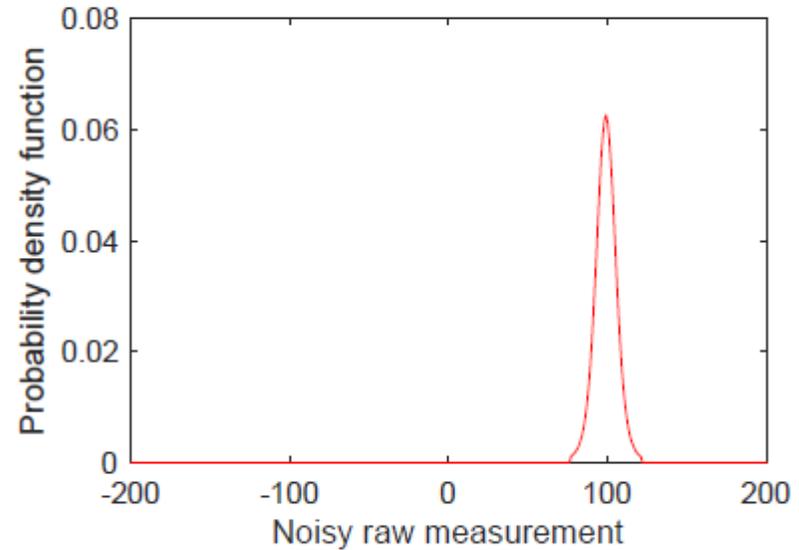
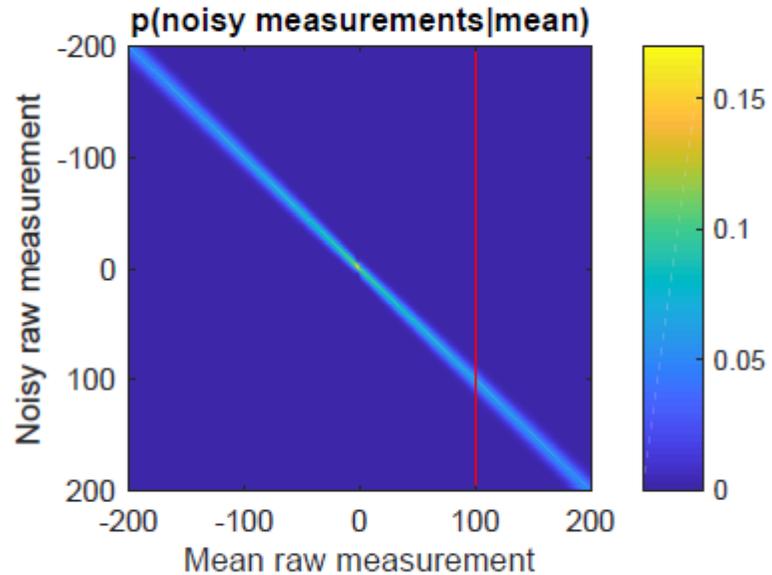
Camera response functions (flat scene)



- Three “frequencies”
- Three measurements per frequency

CAMERA CALIBRATION

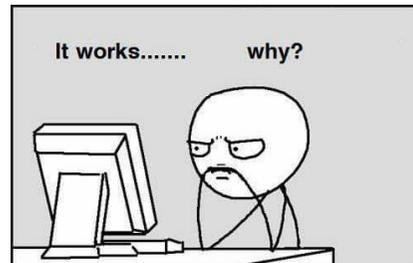
Photon noise



TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*
- Back to physics
- DNN architecture
- Results
- Conclusion



* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

THE FLAT DATASET

Flexible, Large, Augmentable, ToF (FLAT)

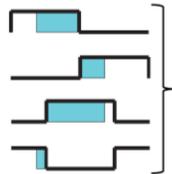
Raw measurements

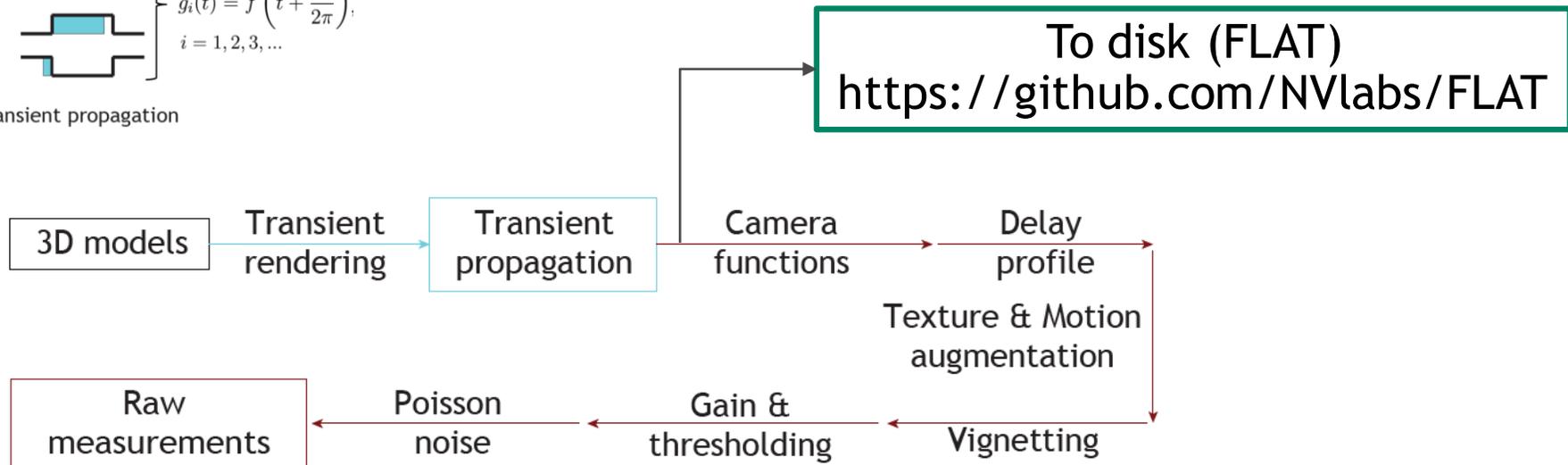
$$Q_i = \int_{-T/2}^{T/2} h(t)g_i(t)dt$$
$$= \int_{-T/2}^{T/2} \underbrace{(f(t) \otimes g(t))}_{\text{Camera function}} \underbrace{r(t)}_{\text{Scene transient propagation}} dt$$

Reference signals

$$g_i(t) = f\left(t + \frac{\theta_i}{2\pi}\right),$$

$i = 1, 2, 3, \dots$





Flexible: Separately models **scenes** and **cameras**

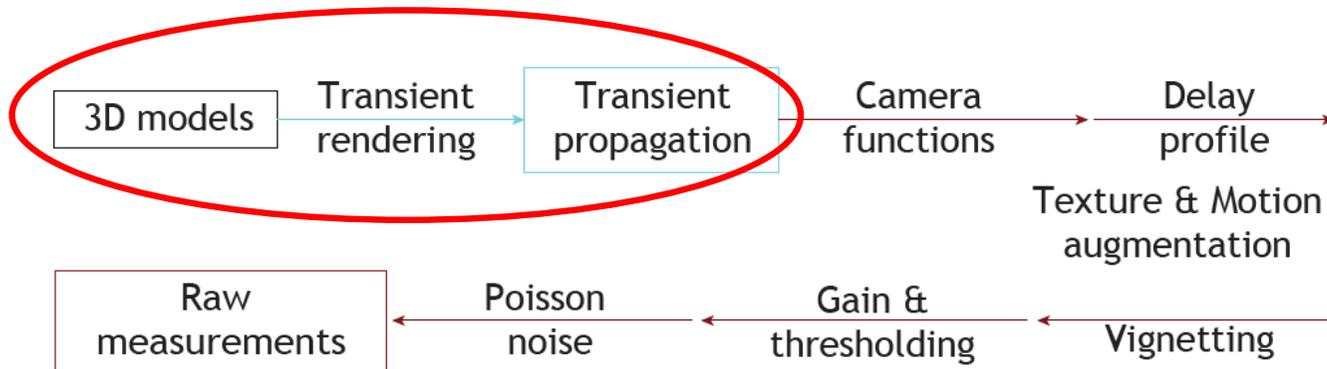
Large: 2061 static and 29 dynamic scenes for now, numbers are increasing

Augmentable: approximate textures and motions can be added on the fly

THE FLAT DATASET

Flexible, Large, Augmentable, ToF (FLAT)

Transient rendering (scene response function) based on Jarabo, A., Marco, J., Muñoz, A., Buisan, R., Jarosz, W., Gutierrez, D.: A framework for transient rendering. In: ACM Transactions on Graphics (SIGGRAPH ASIA), (2014).



Flexible: Separately models **scenes** and **cameras**

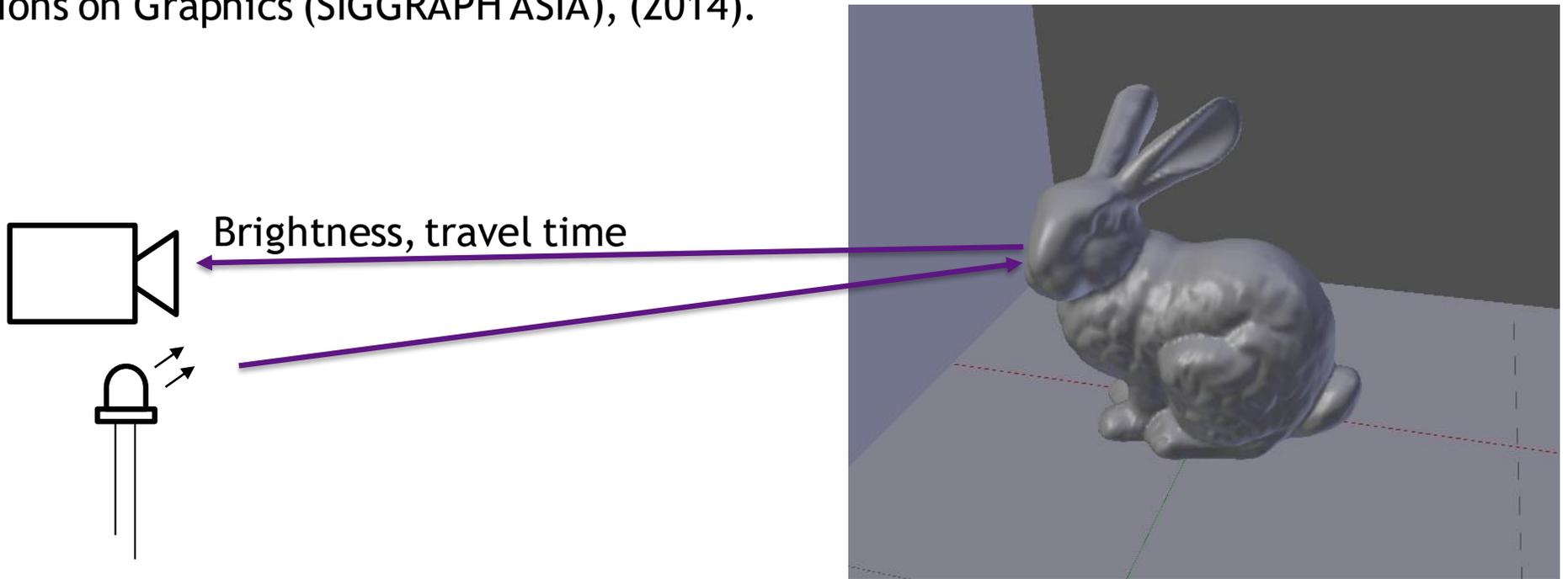
Large: 2061 static and 29 dynamic scenes for now, numbers are increasing

Augmentable: approximate textures and motions can be added on the fly

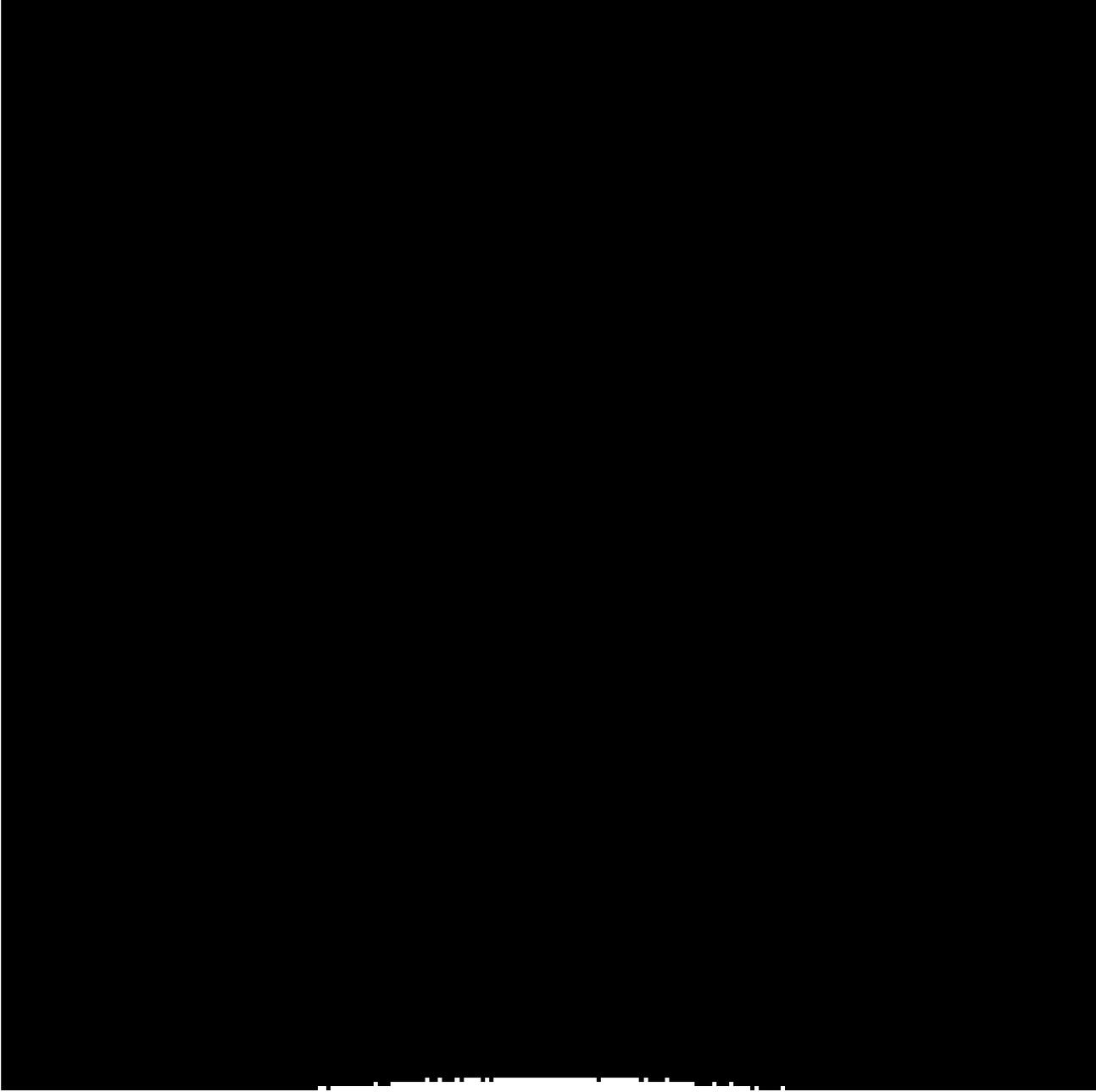
THE FLAT DATASET

Flexible, Large, Augmentable, ToF (FLAT)

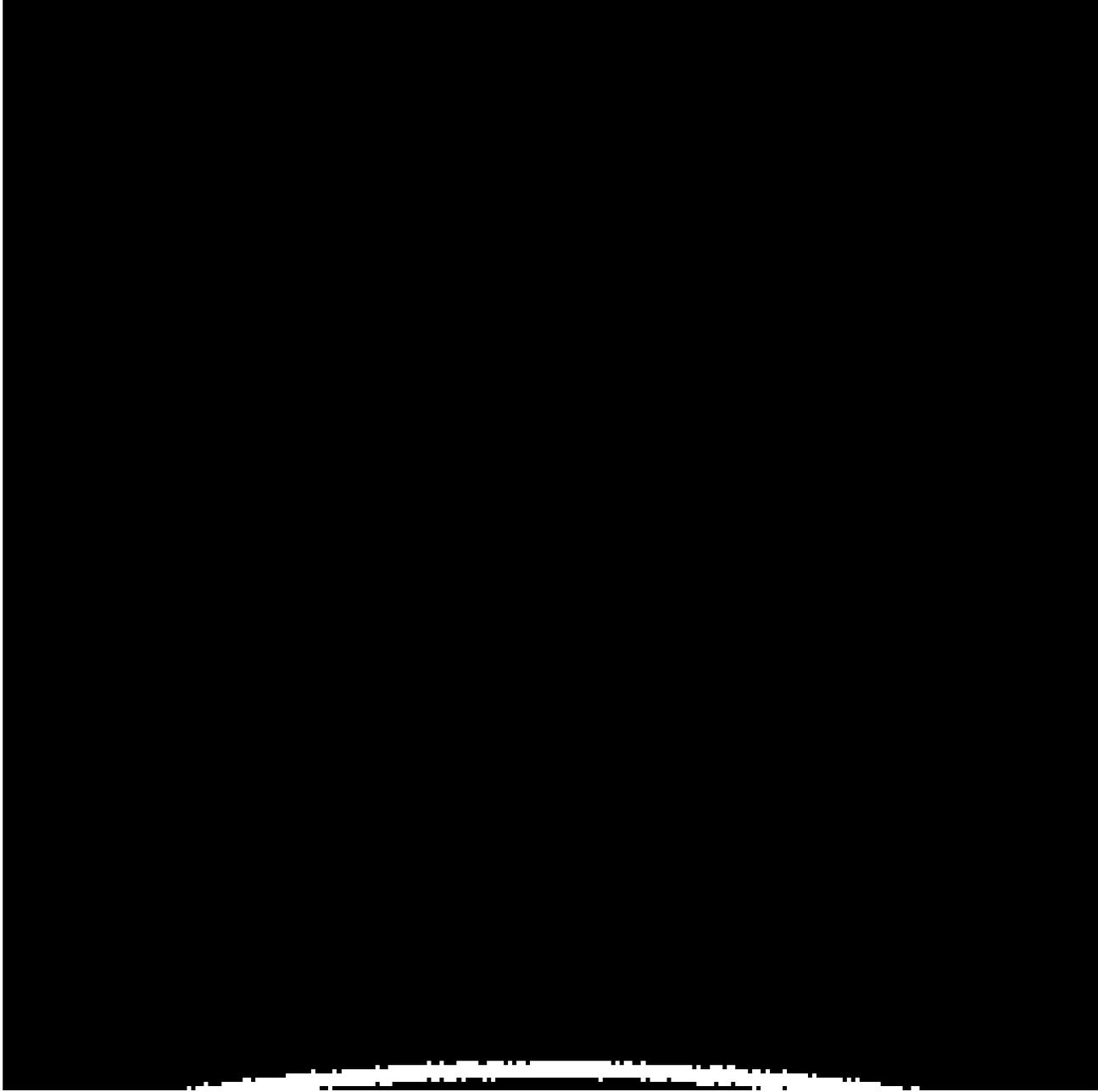
Transient rendering (scene response function) based on Jarabo, A., Marco, J., Muñoz, A., Buisan, R., Jarosz, W., Gutierrez, D.: A framework for transient rendering. In: ACM Transactions on Graphics (SIGGRAPH ASIA), (2014).



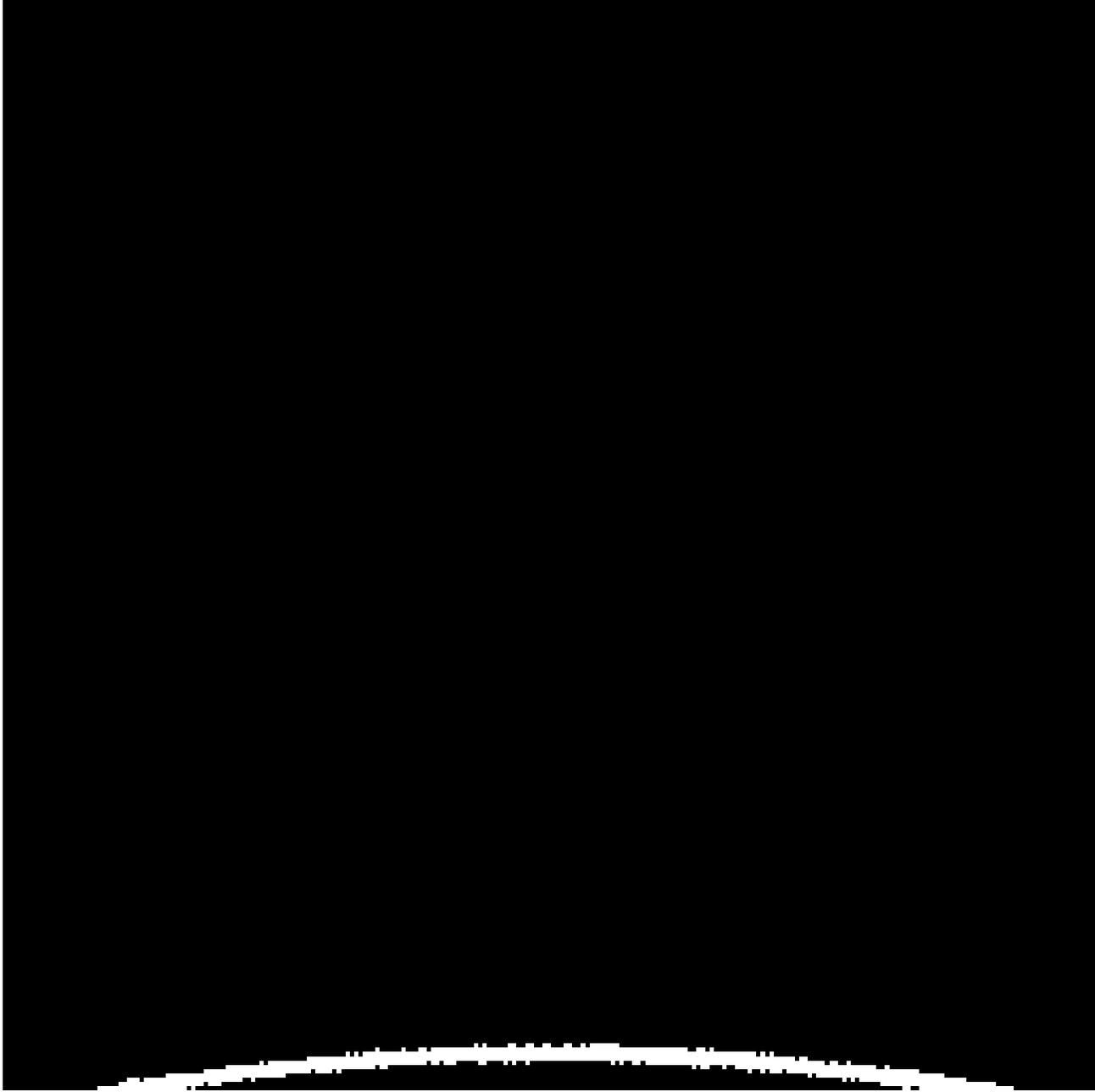
Impulse response



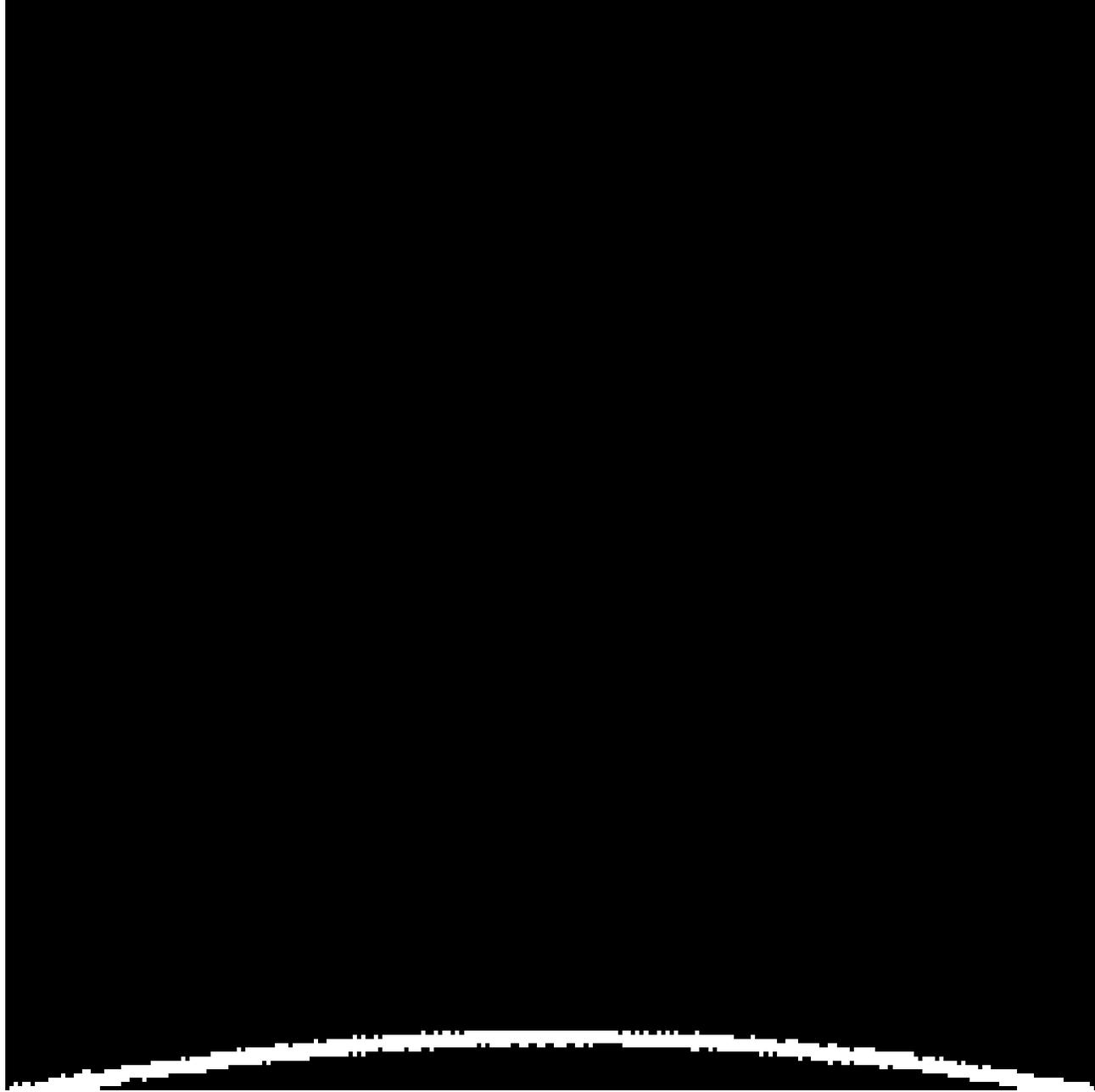
Impulse response



Impulse response



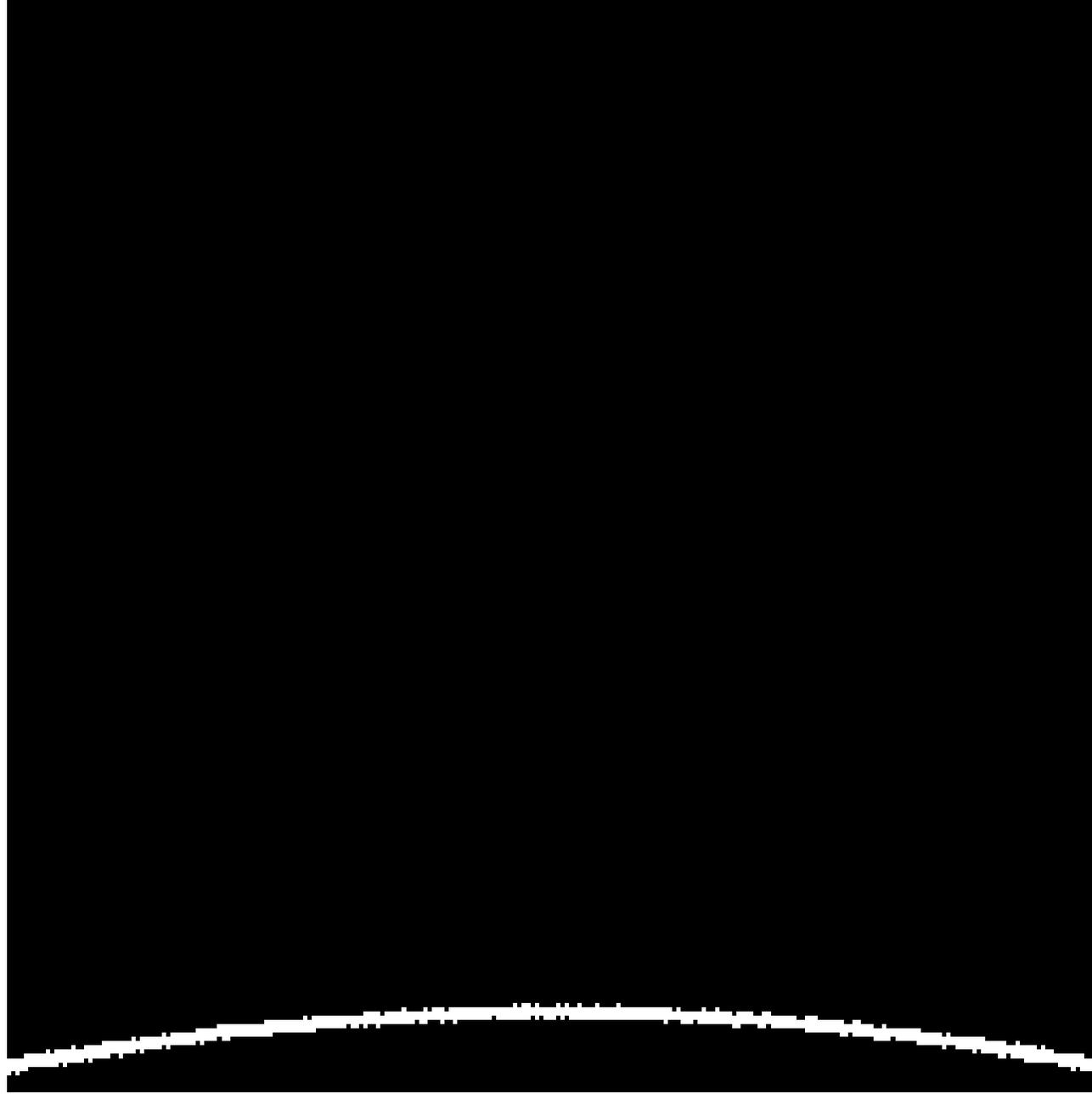
Impulse response



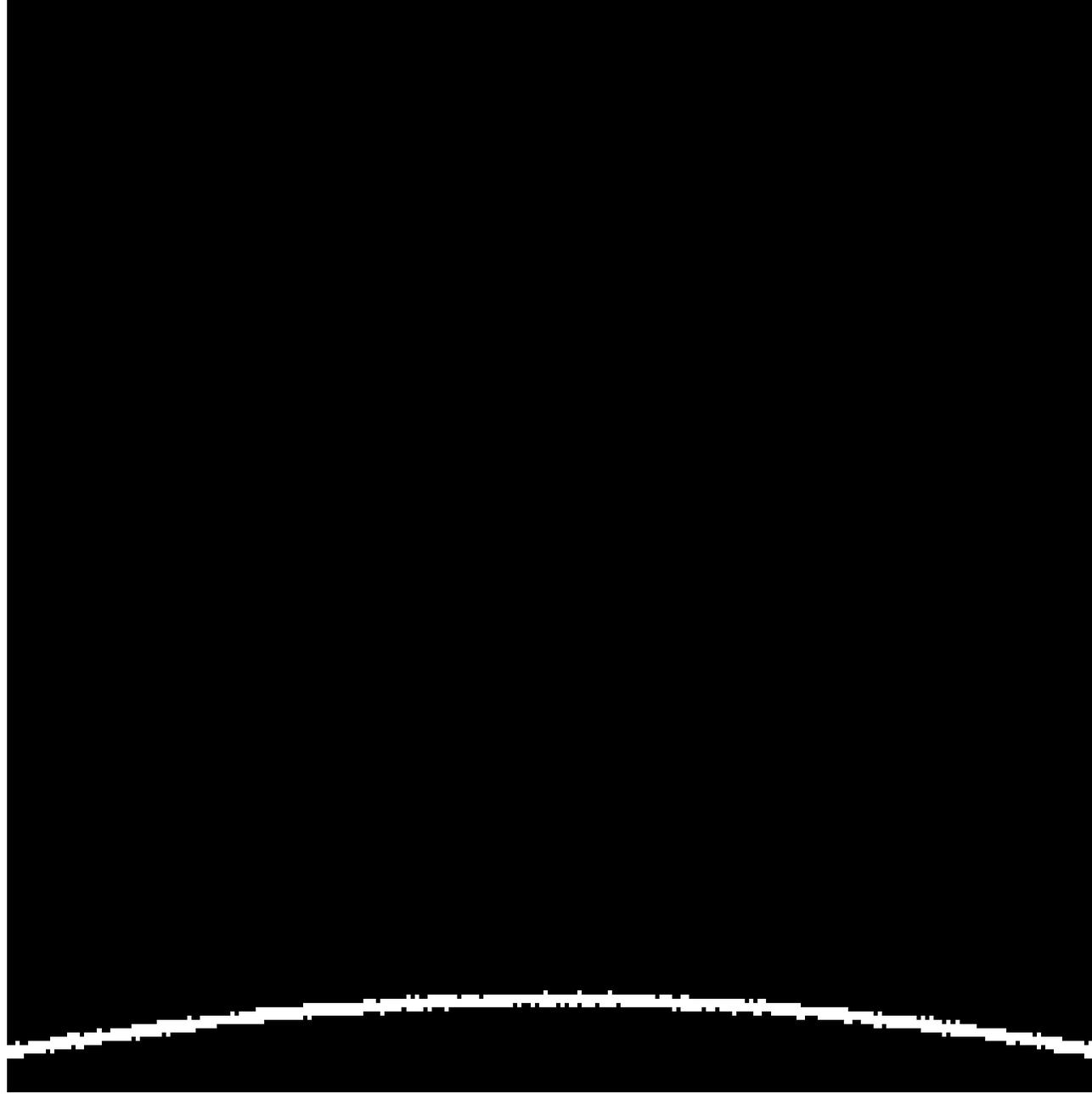
Impulse response



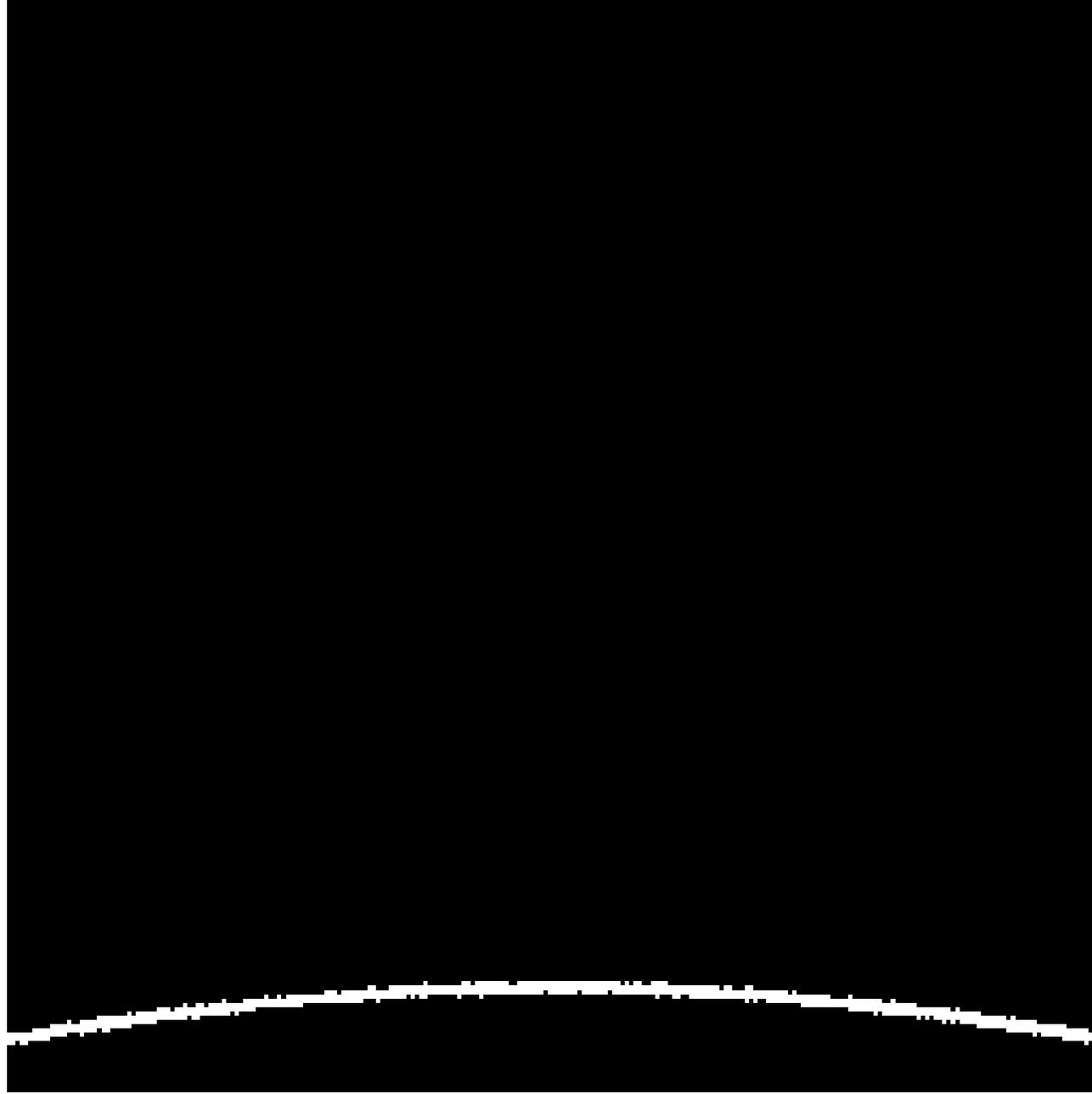
Impulse response



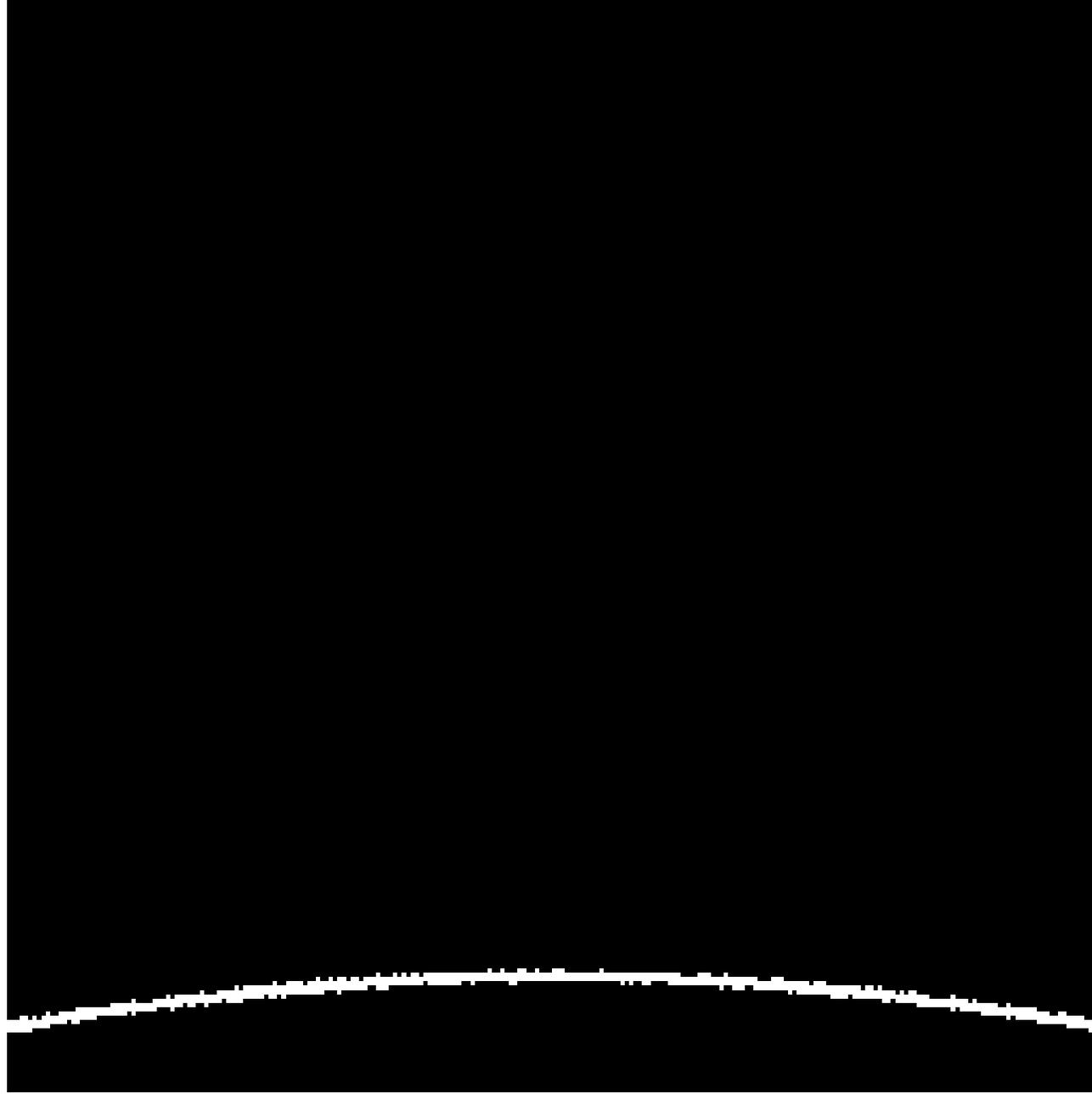
Impulse response



Impulse response



Impulse response



Impulse response



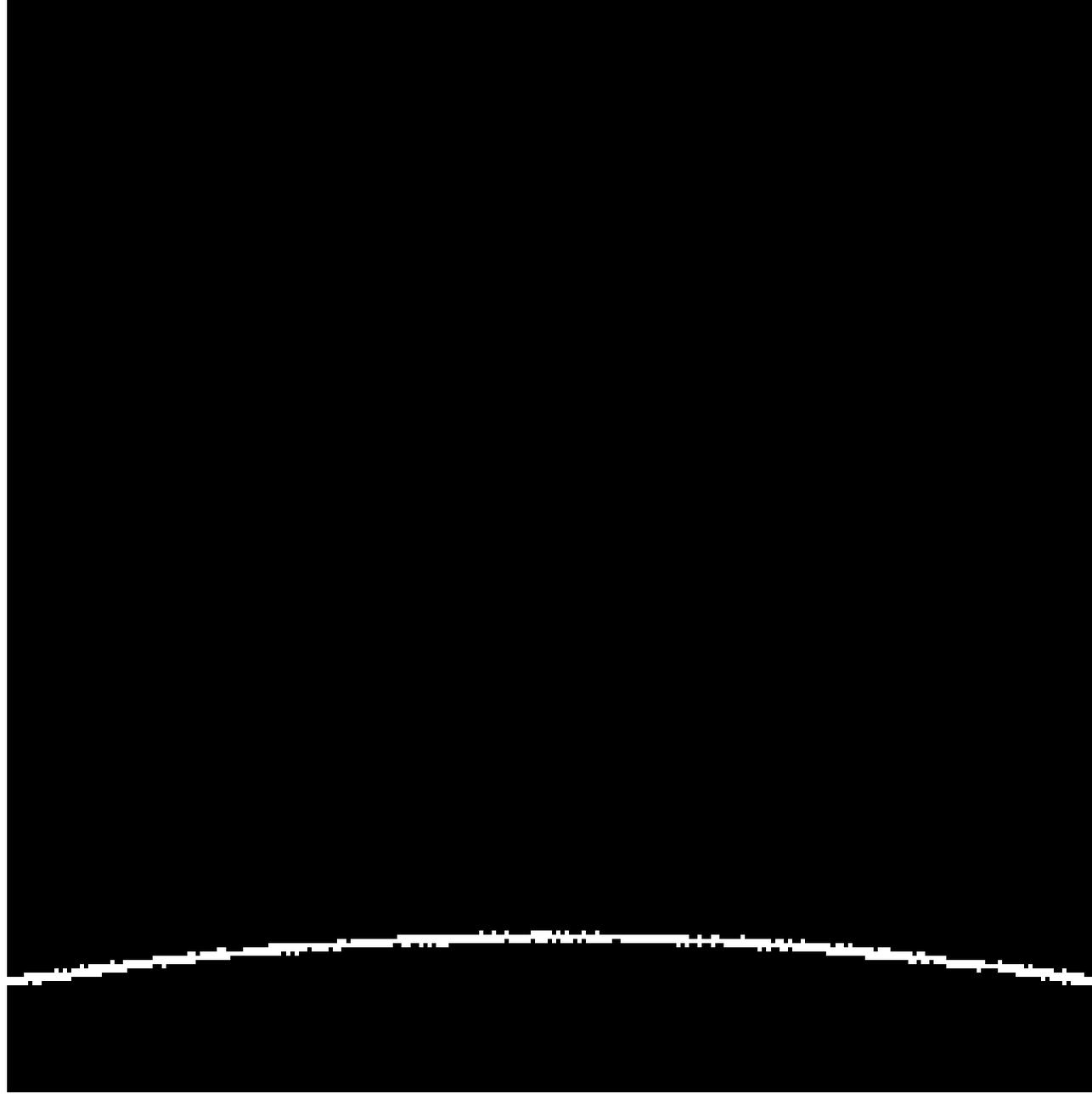
Impulse response



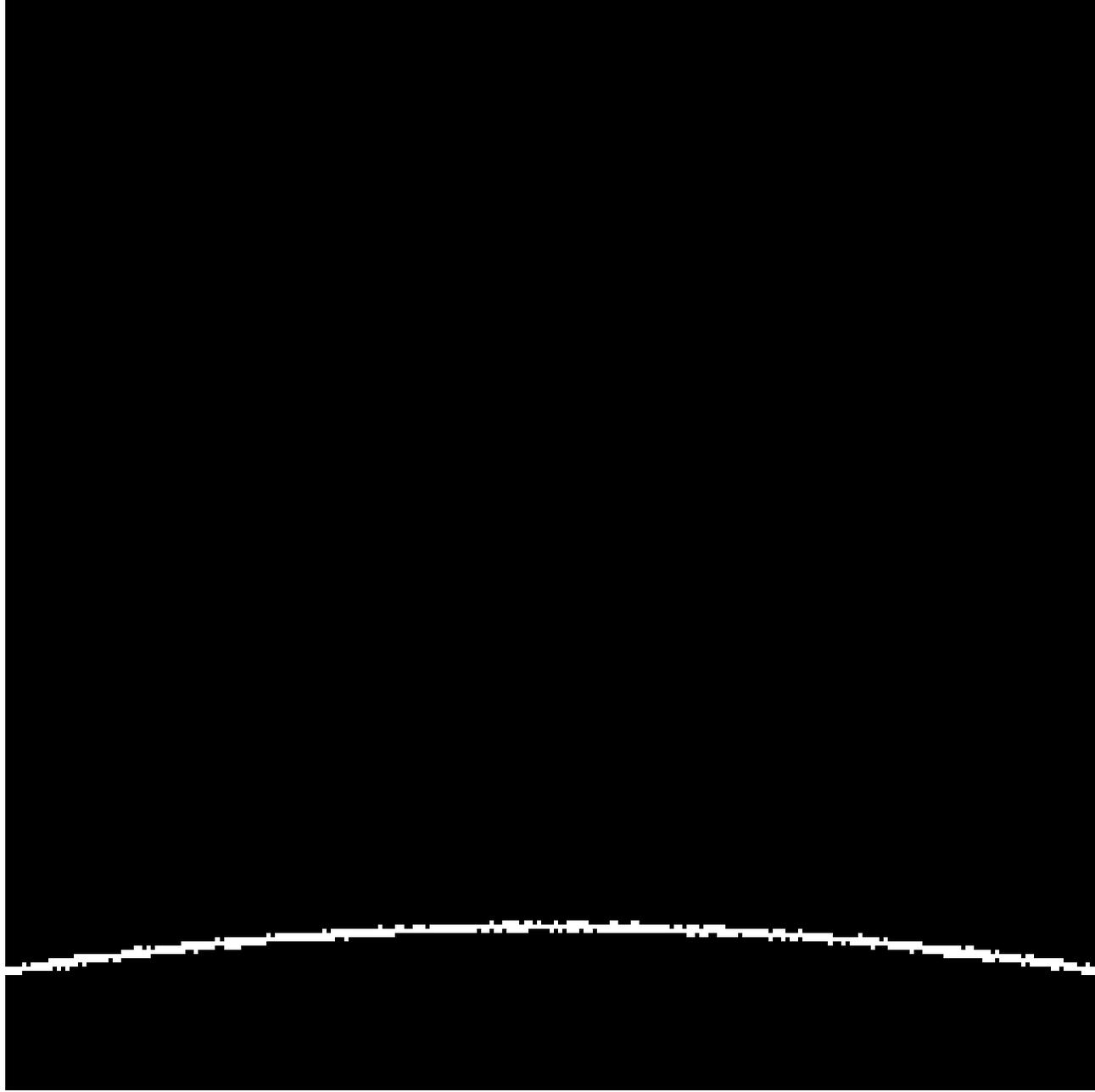
Impulse response



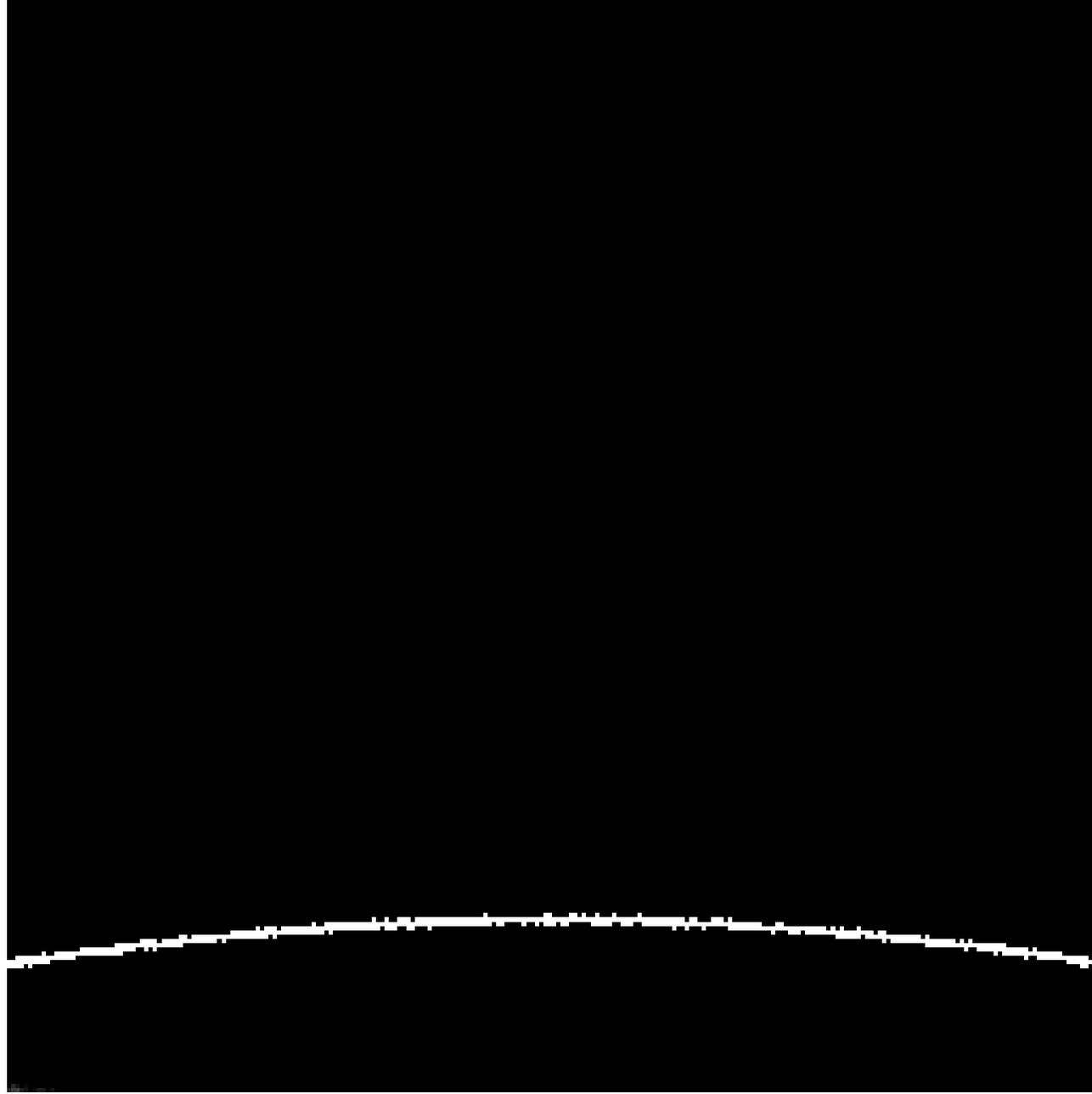
Impulse response



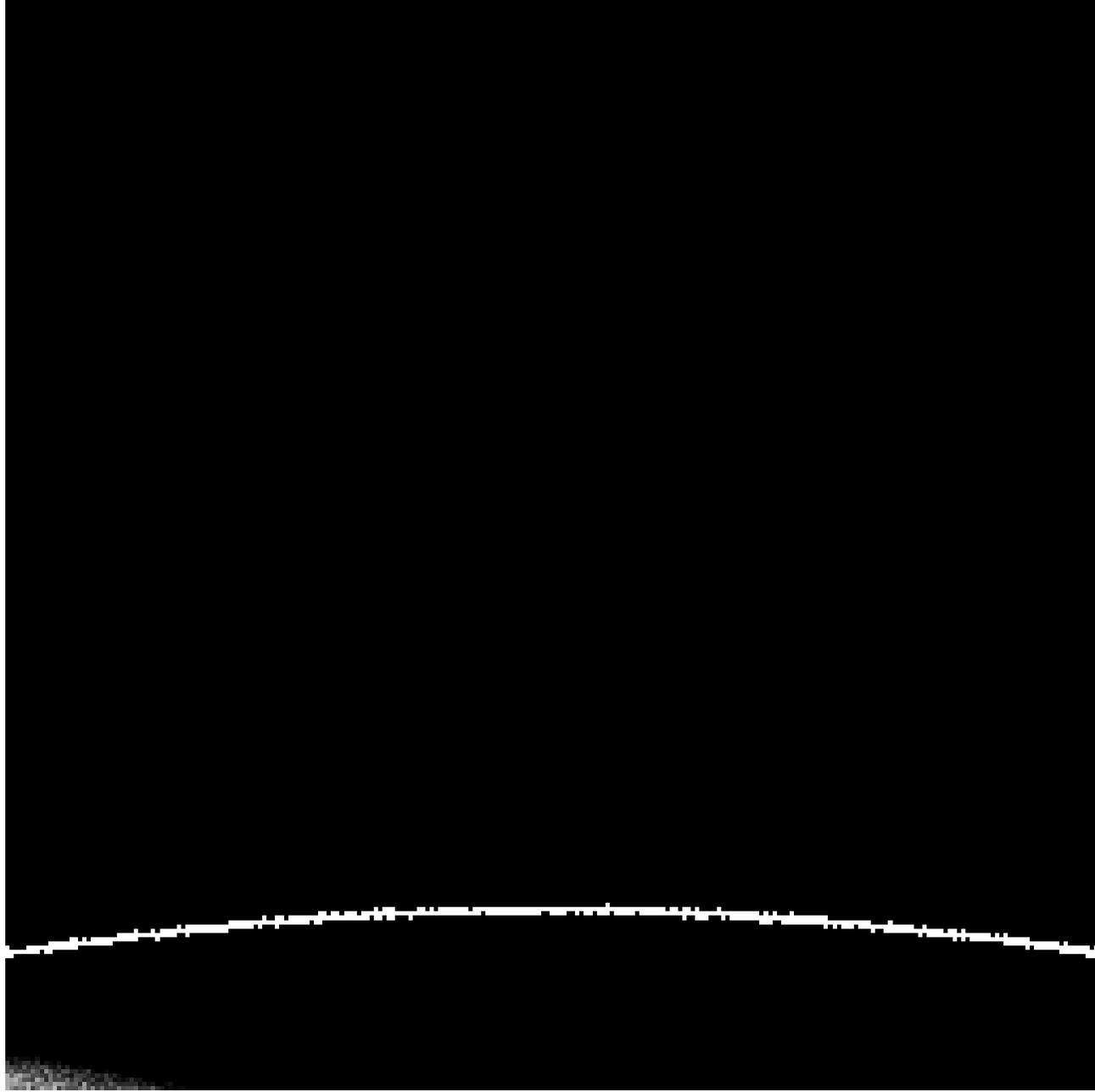
Impulse response



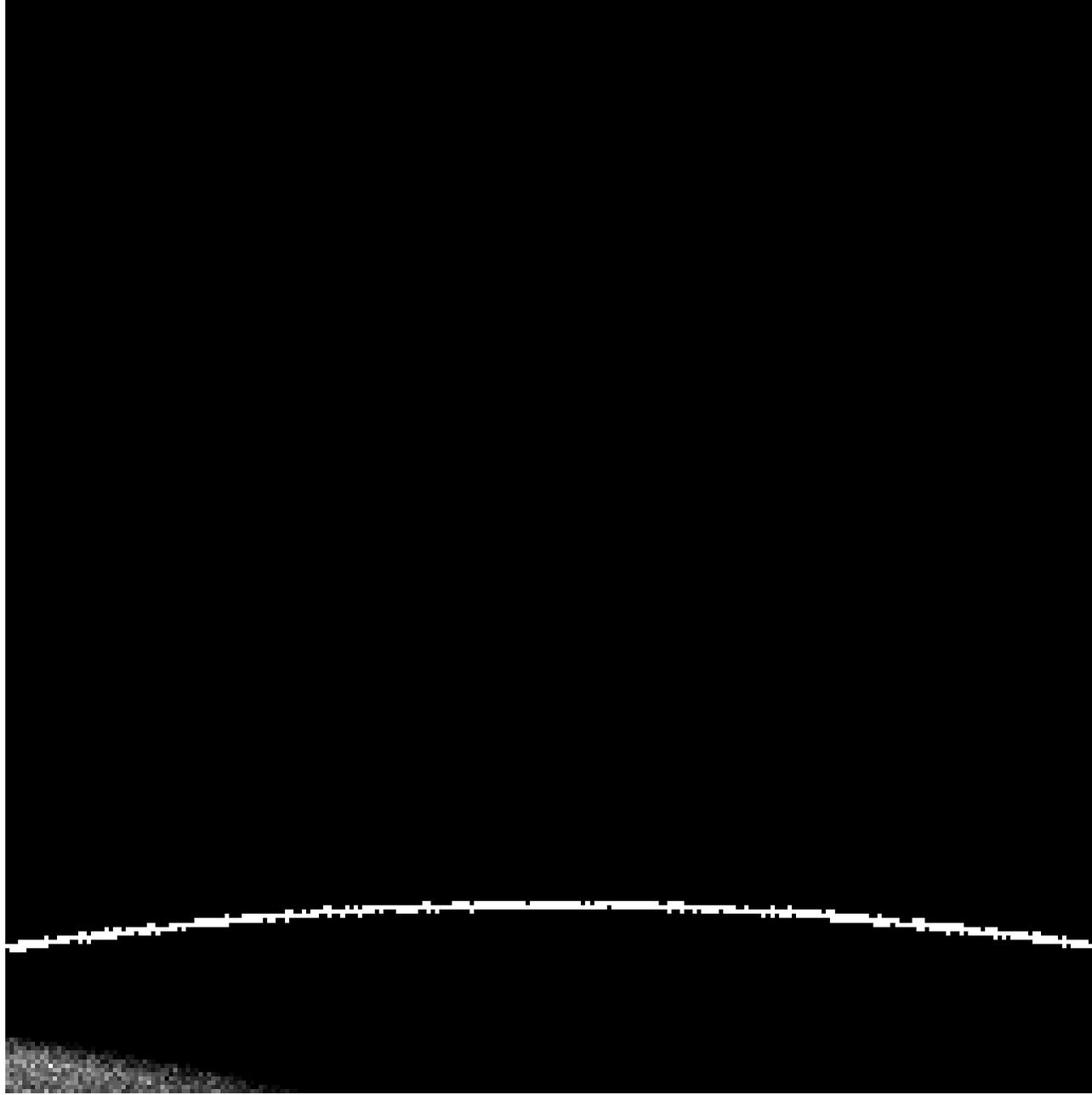
Impulse response



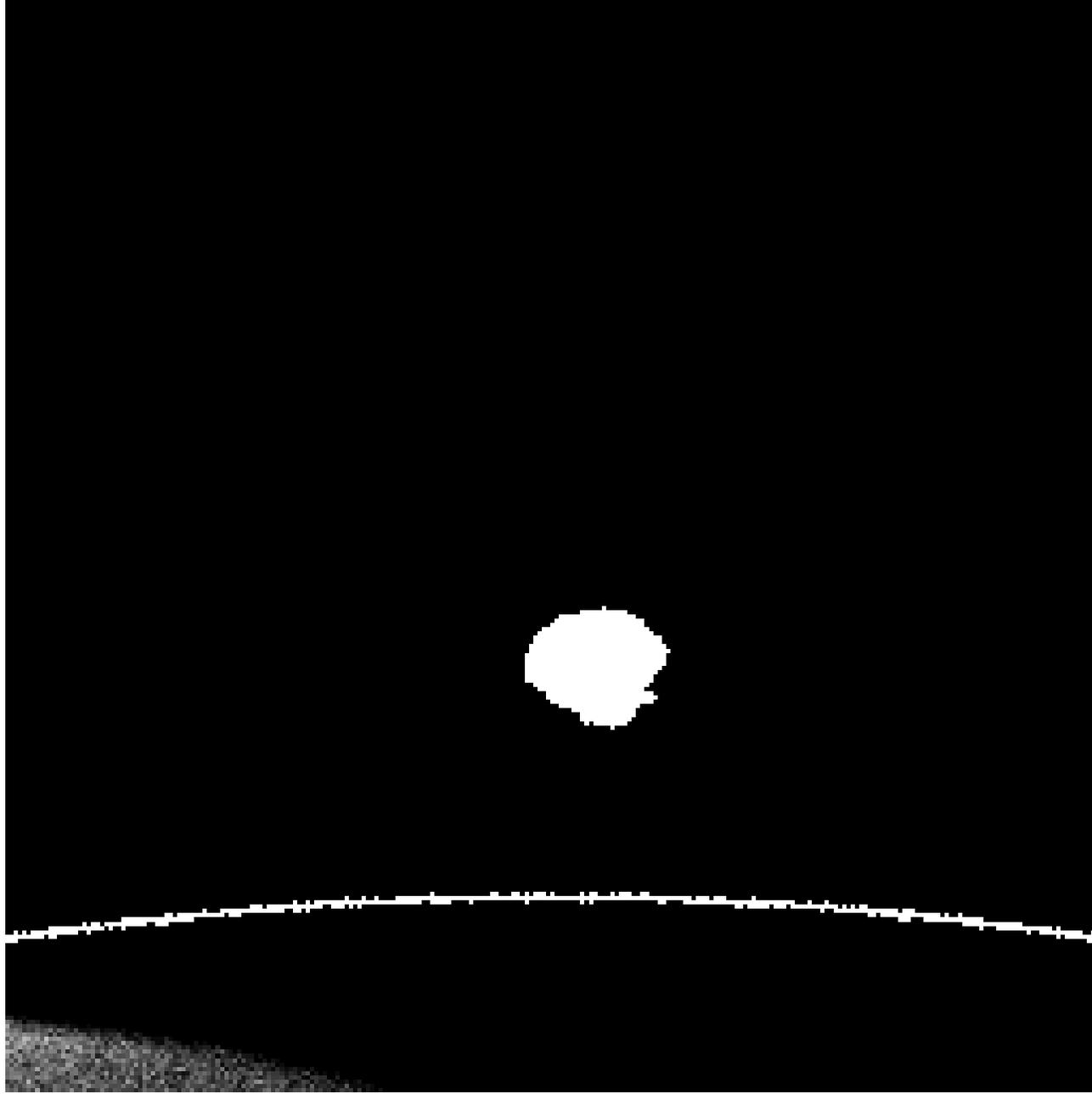
Impulse response



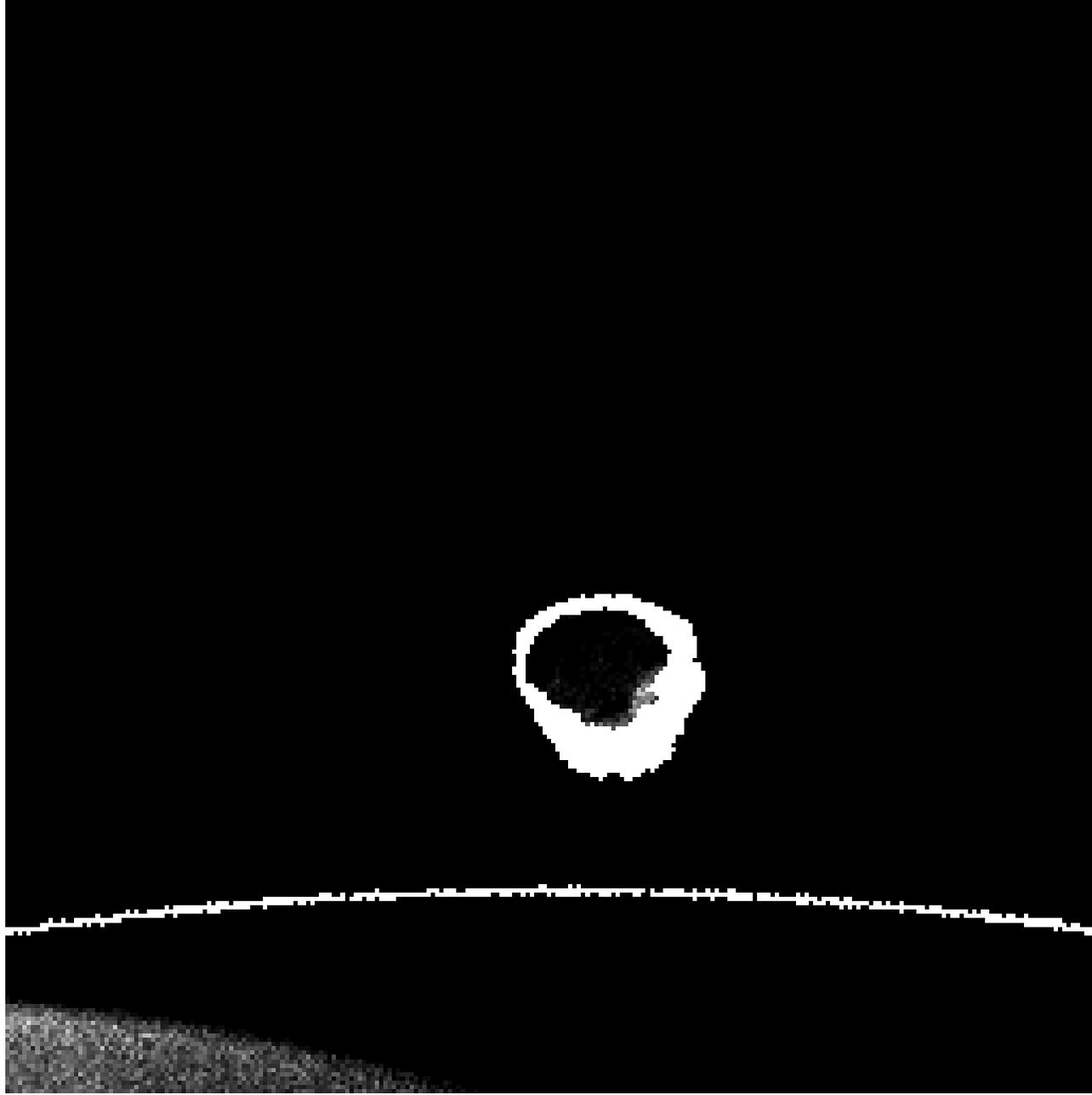
Impulse response



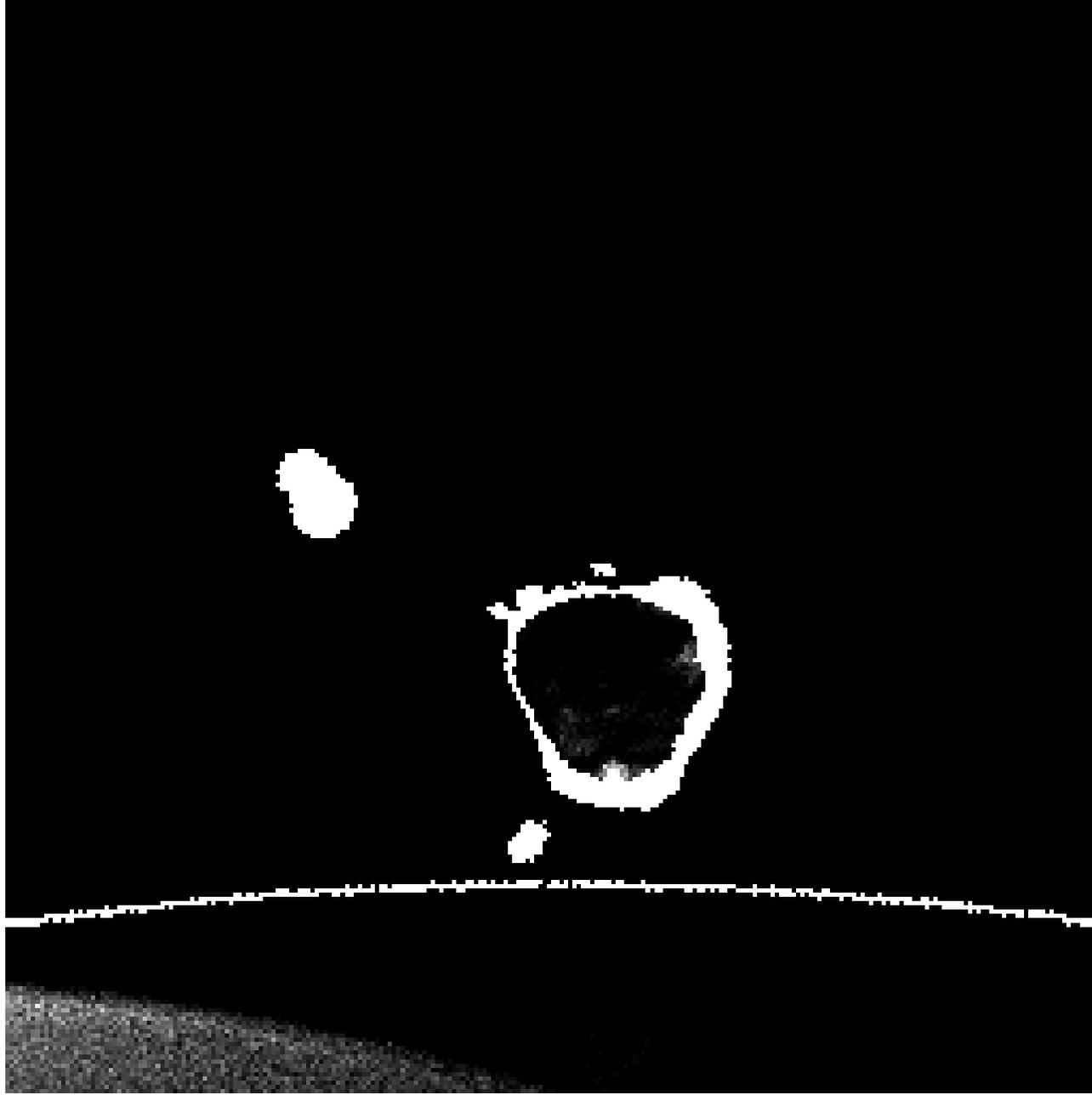
Impulse response



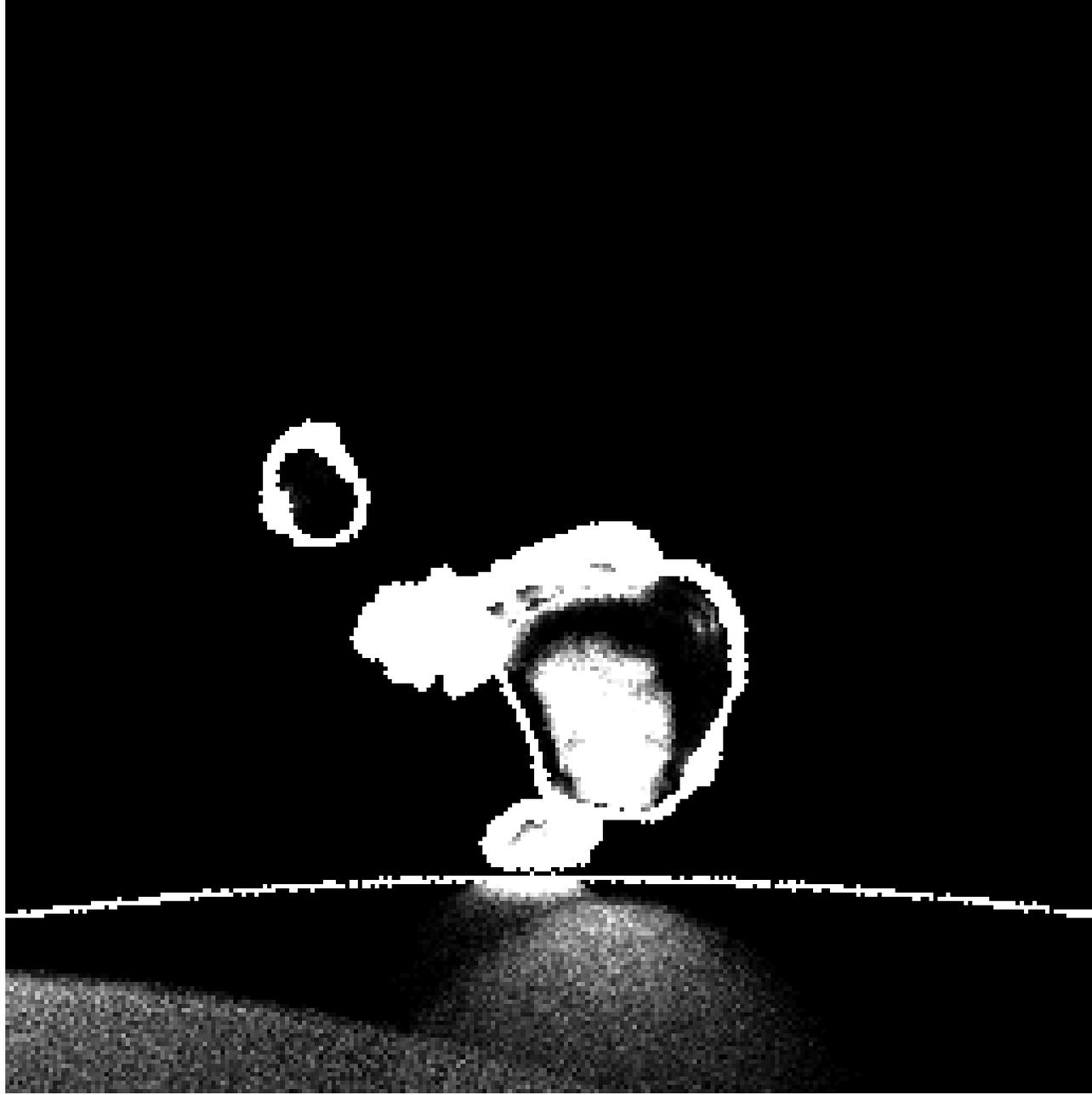
Impulse response



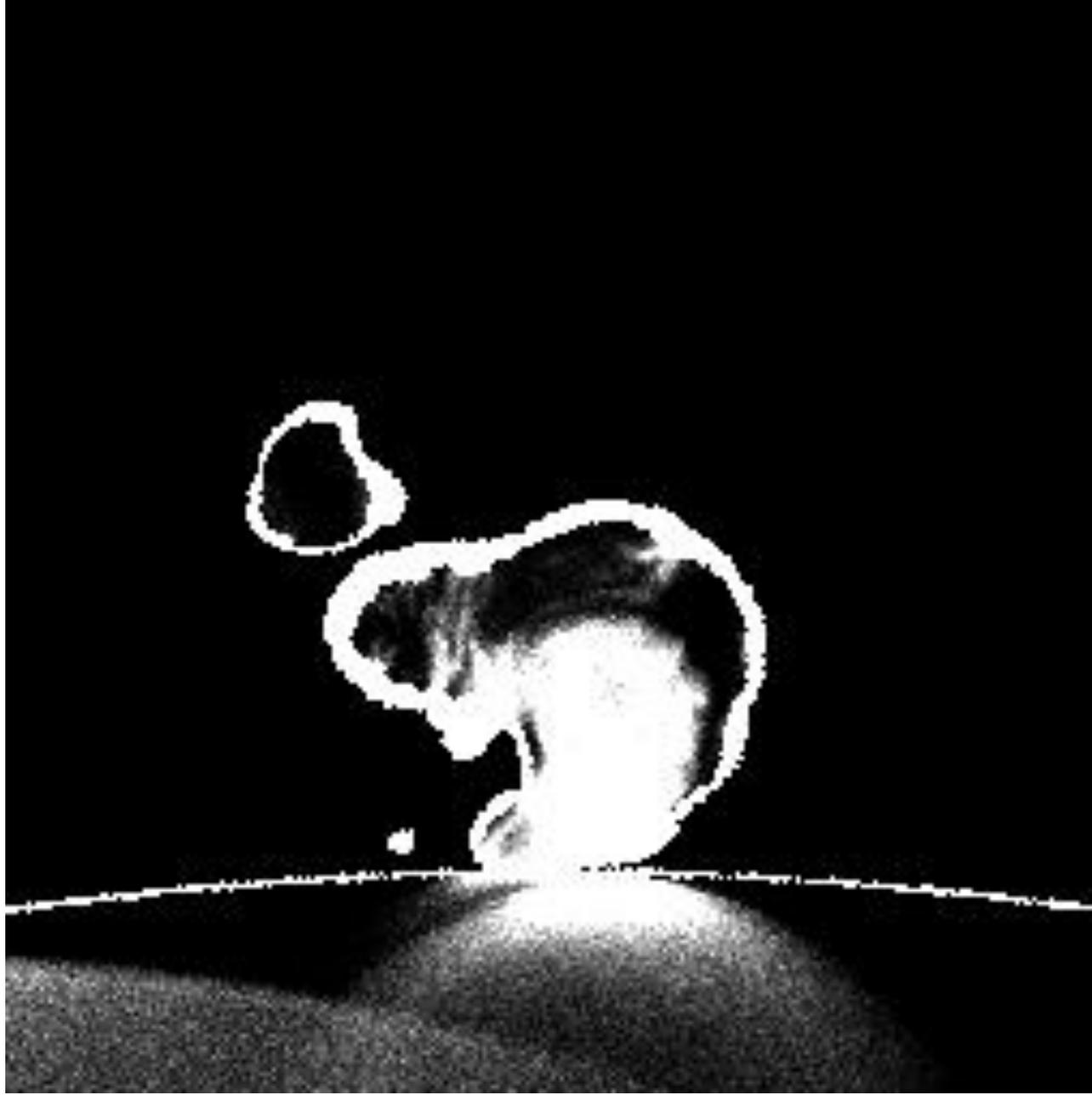
Impulse response



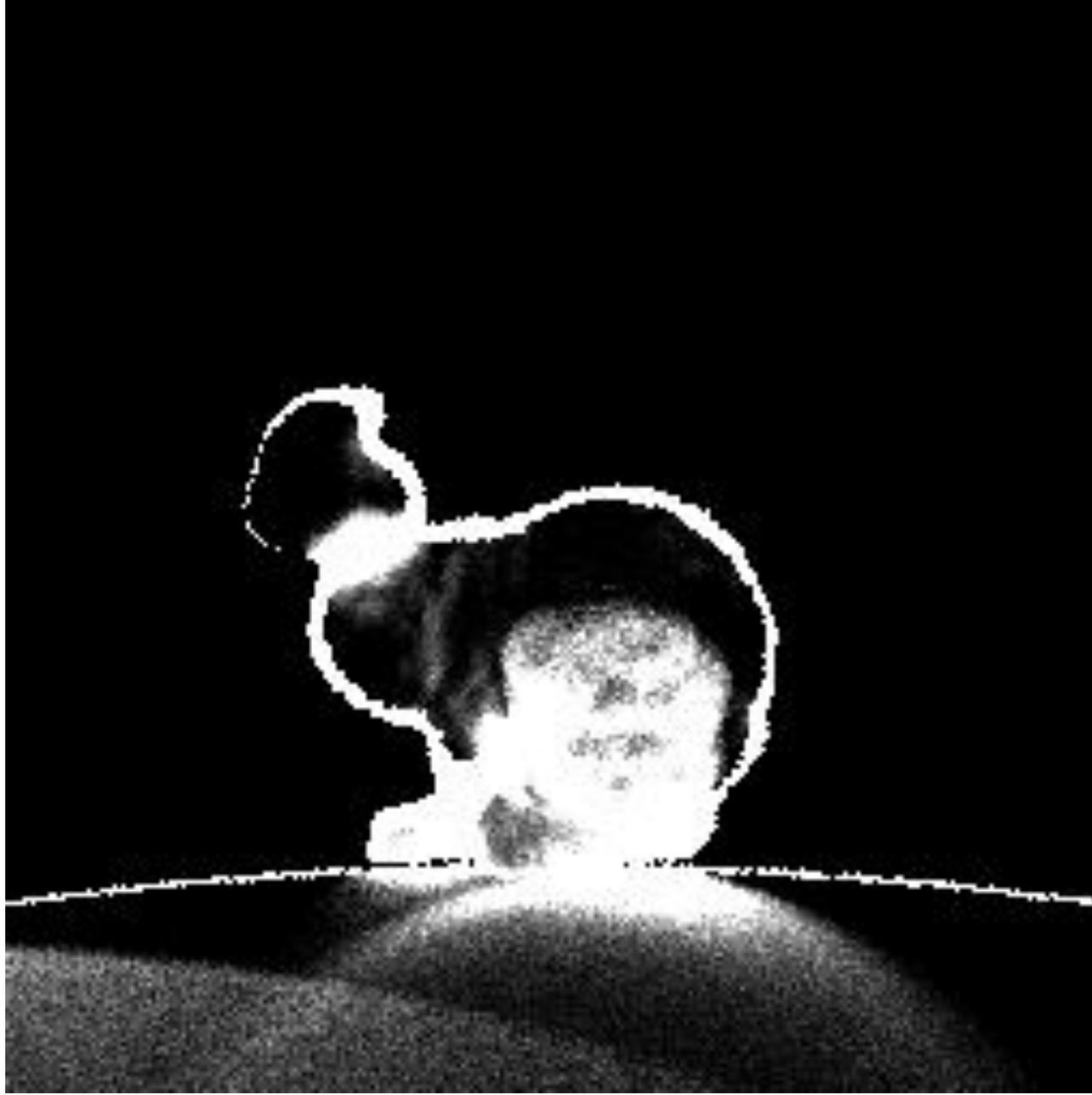
Impulse response



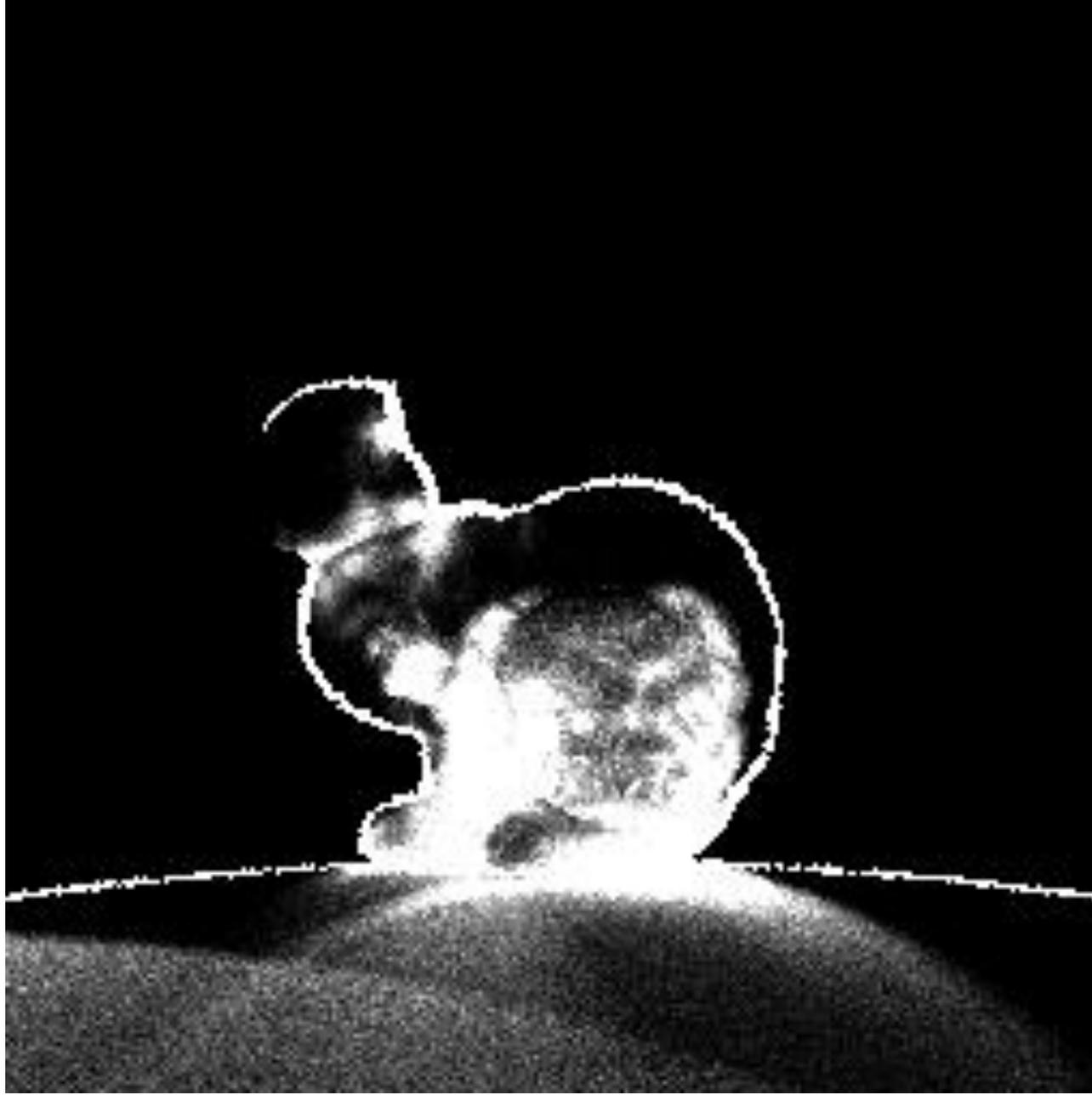
Impulse response



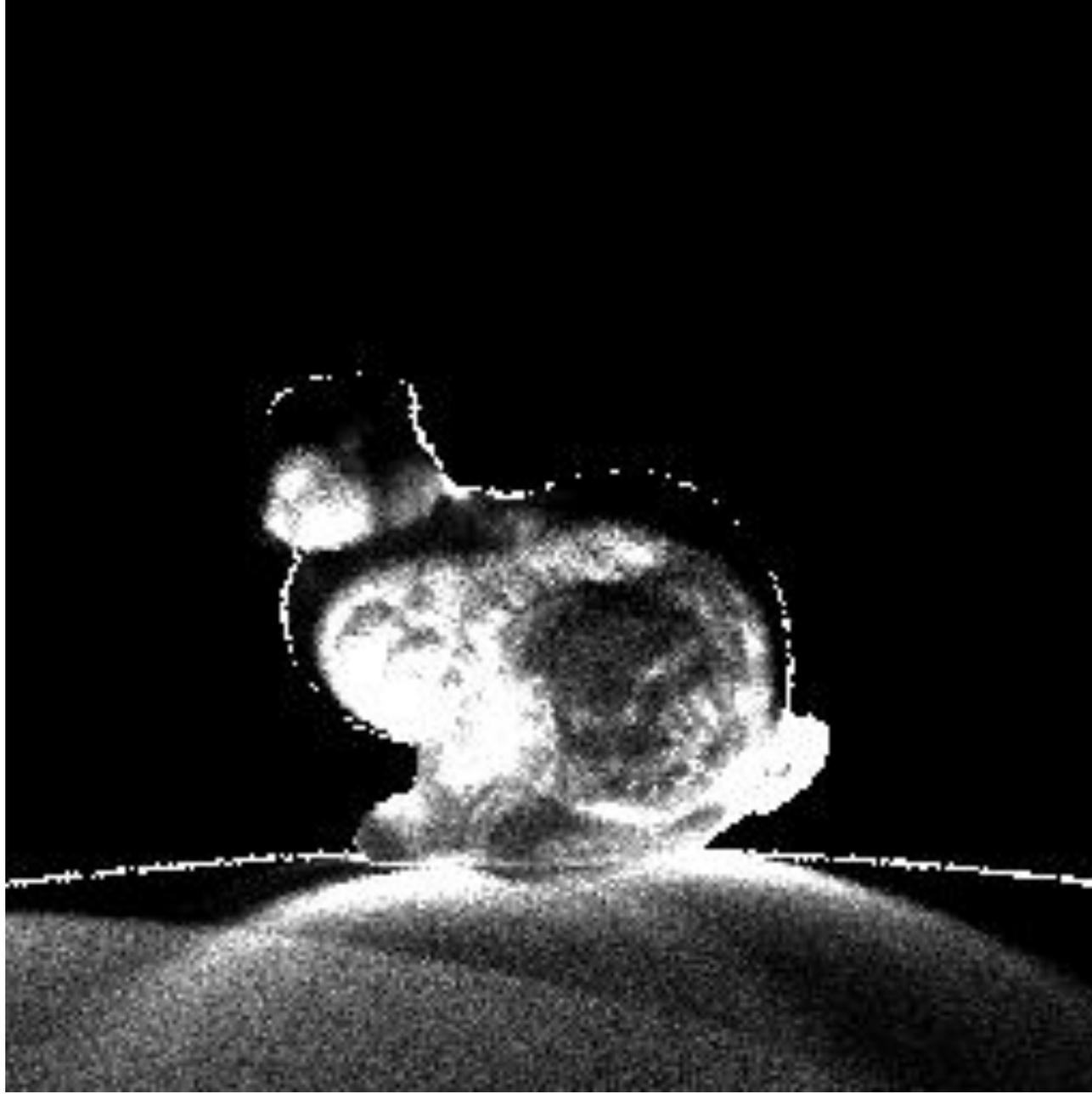
Impulse response



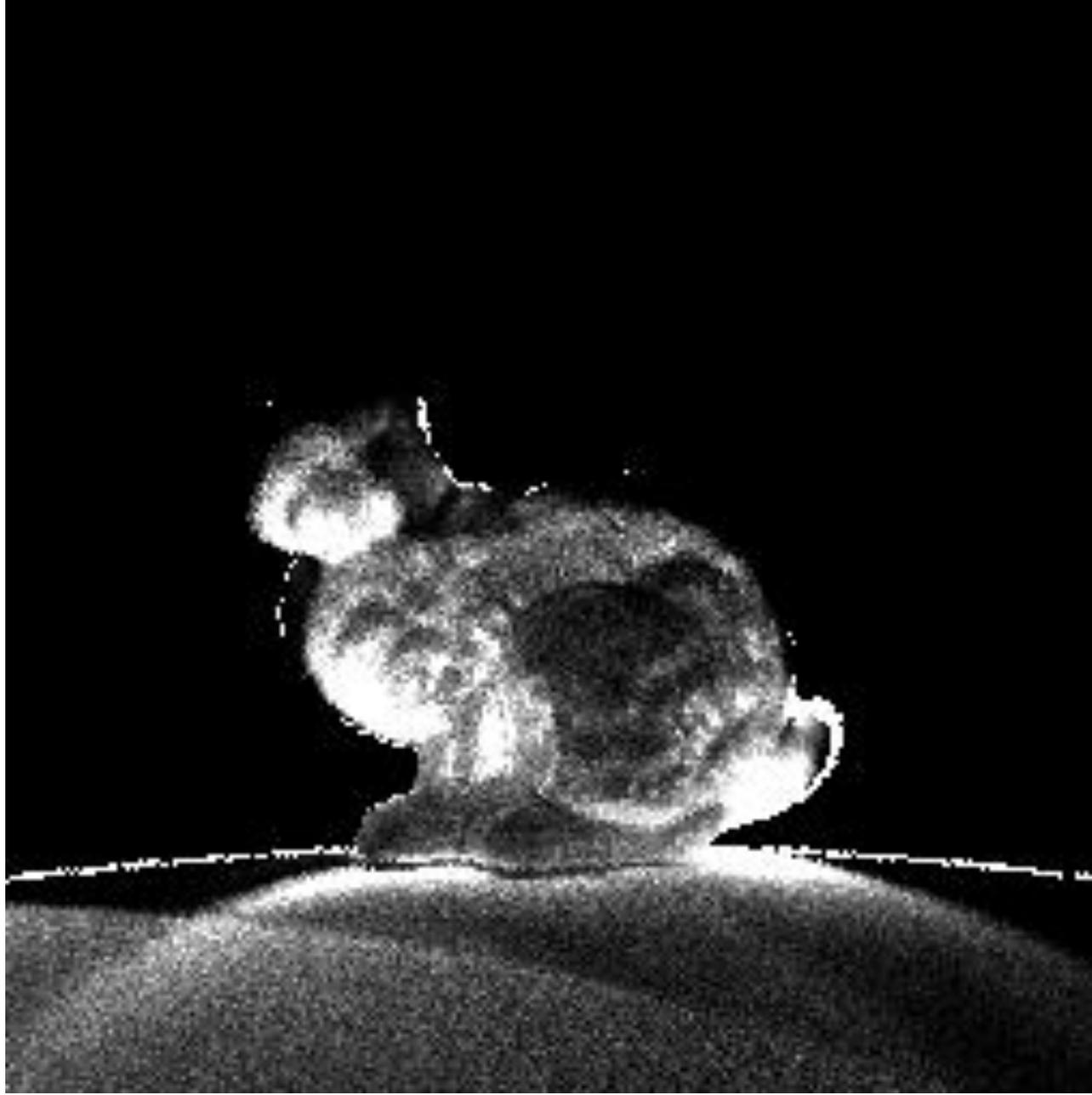
Impulse response



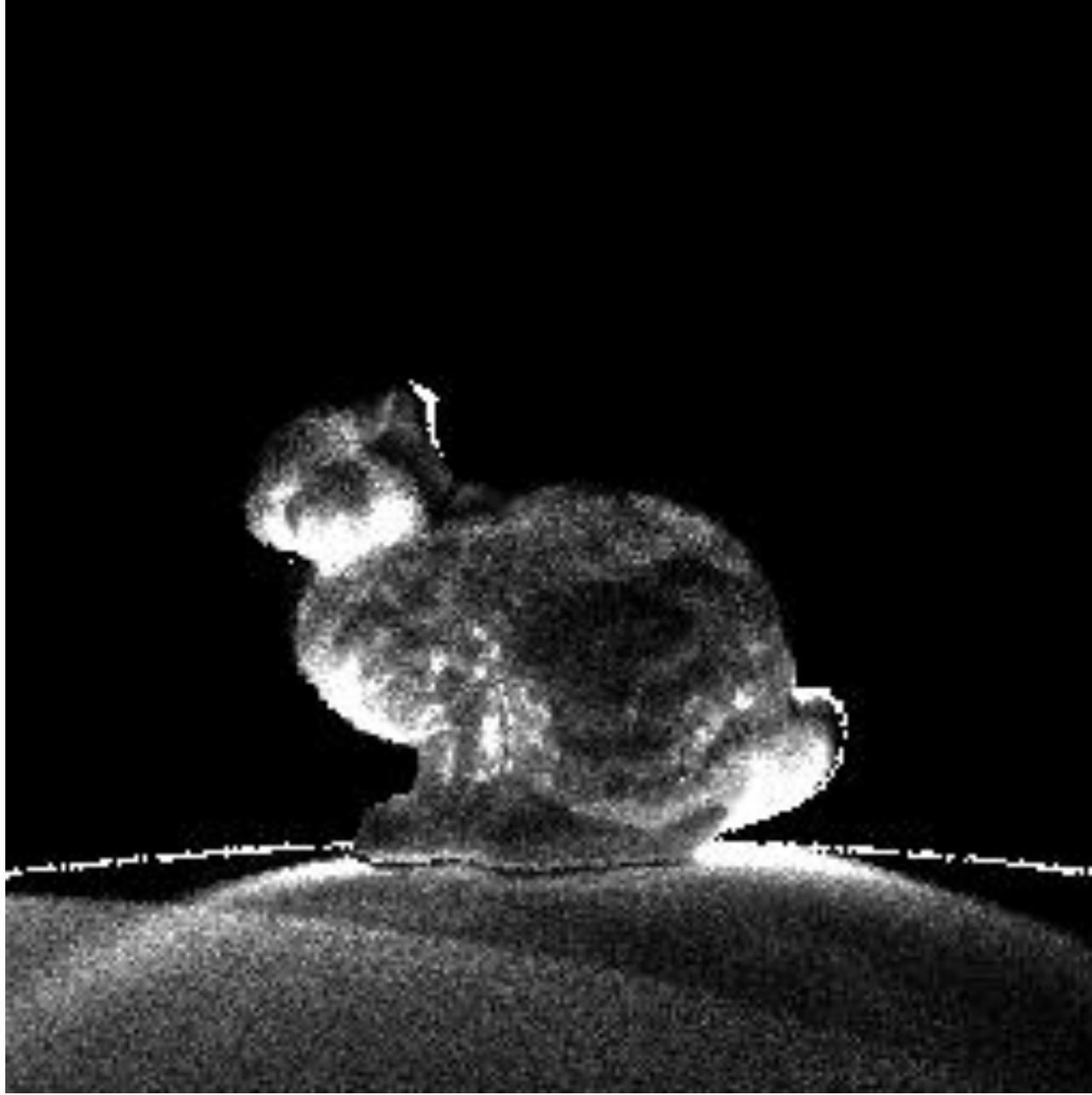
Impulse response



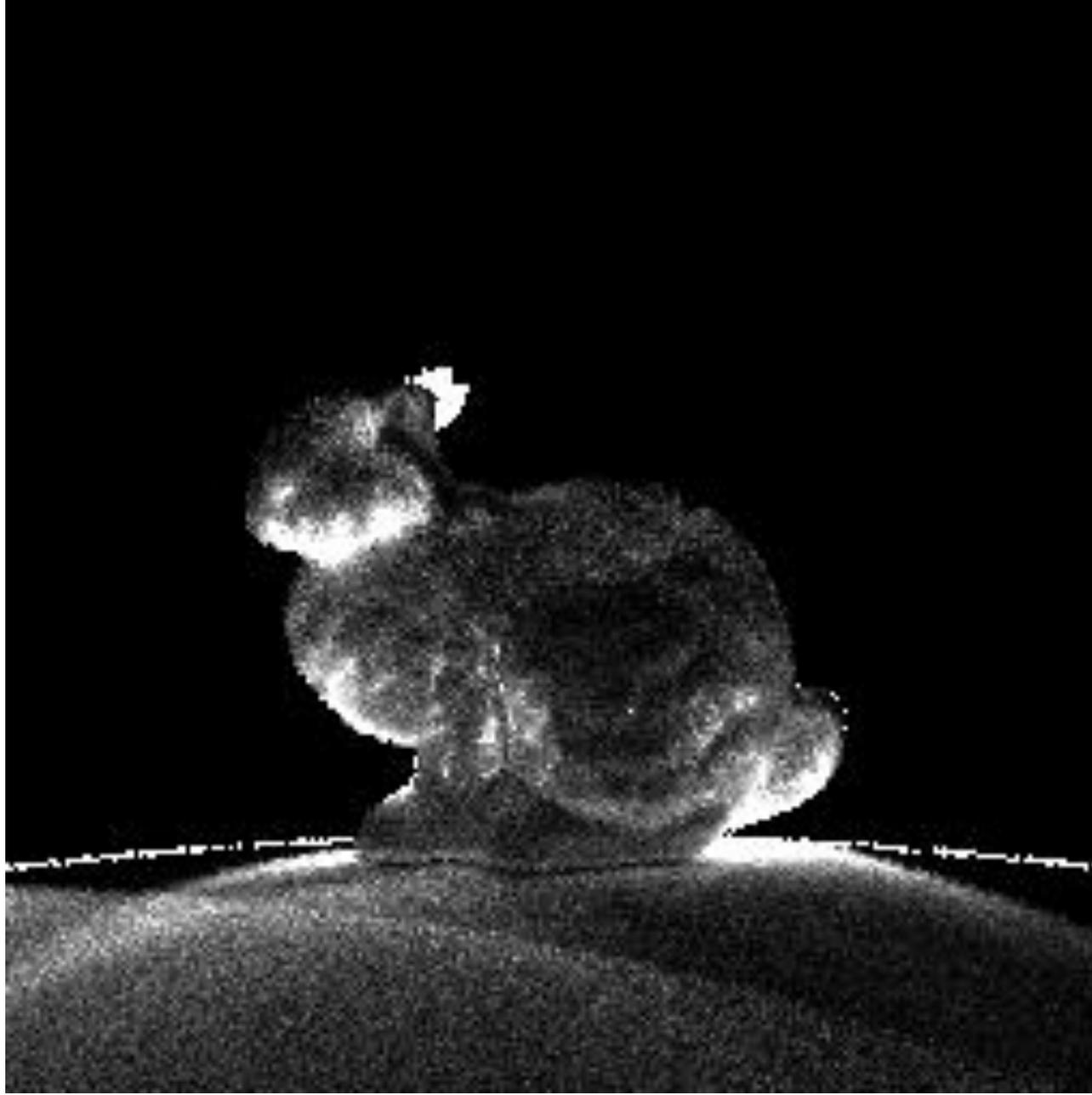
Impulse response



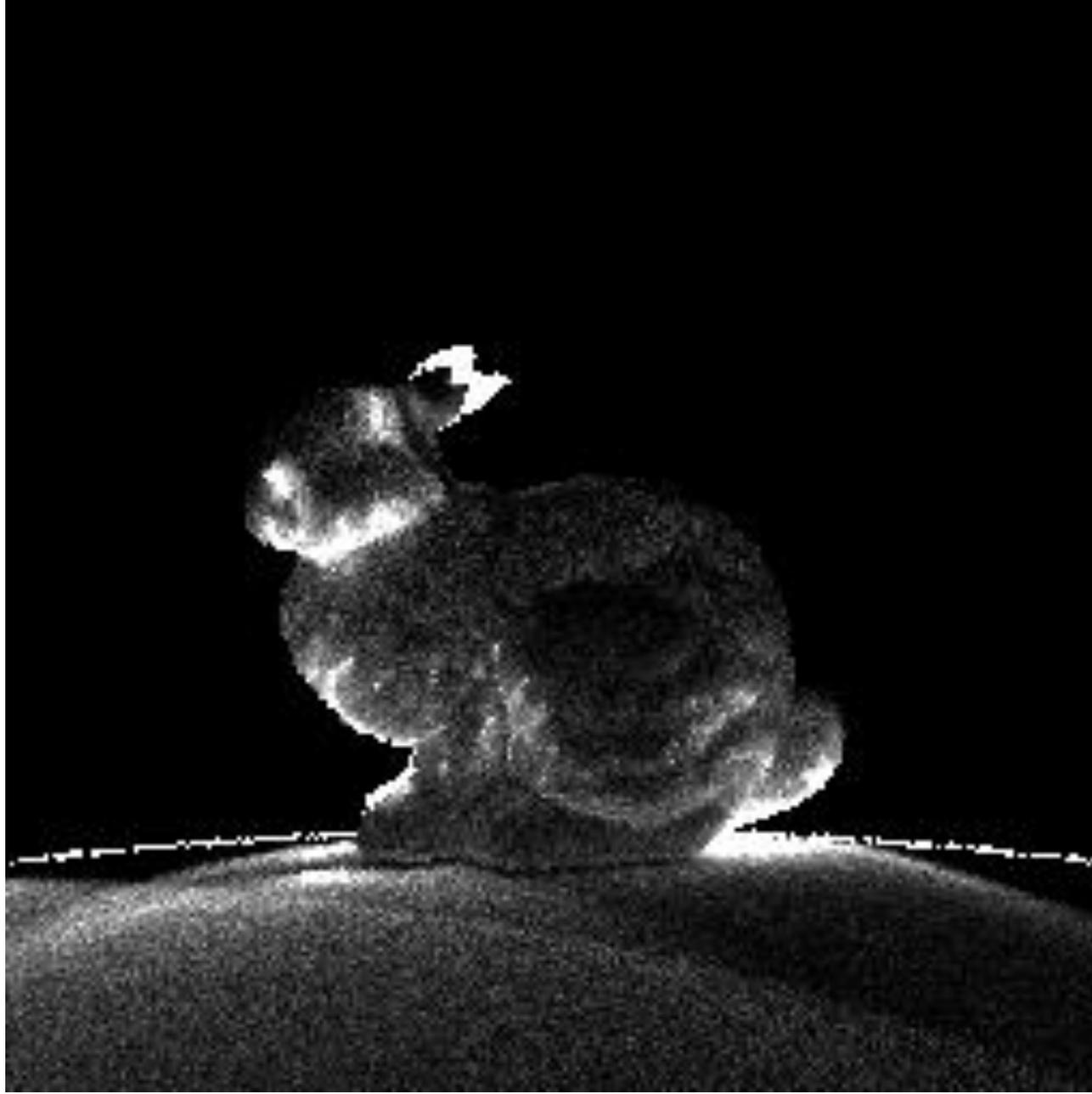
Impulse response



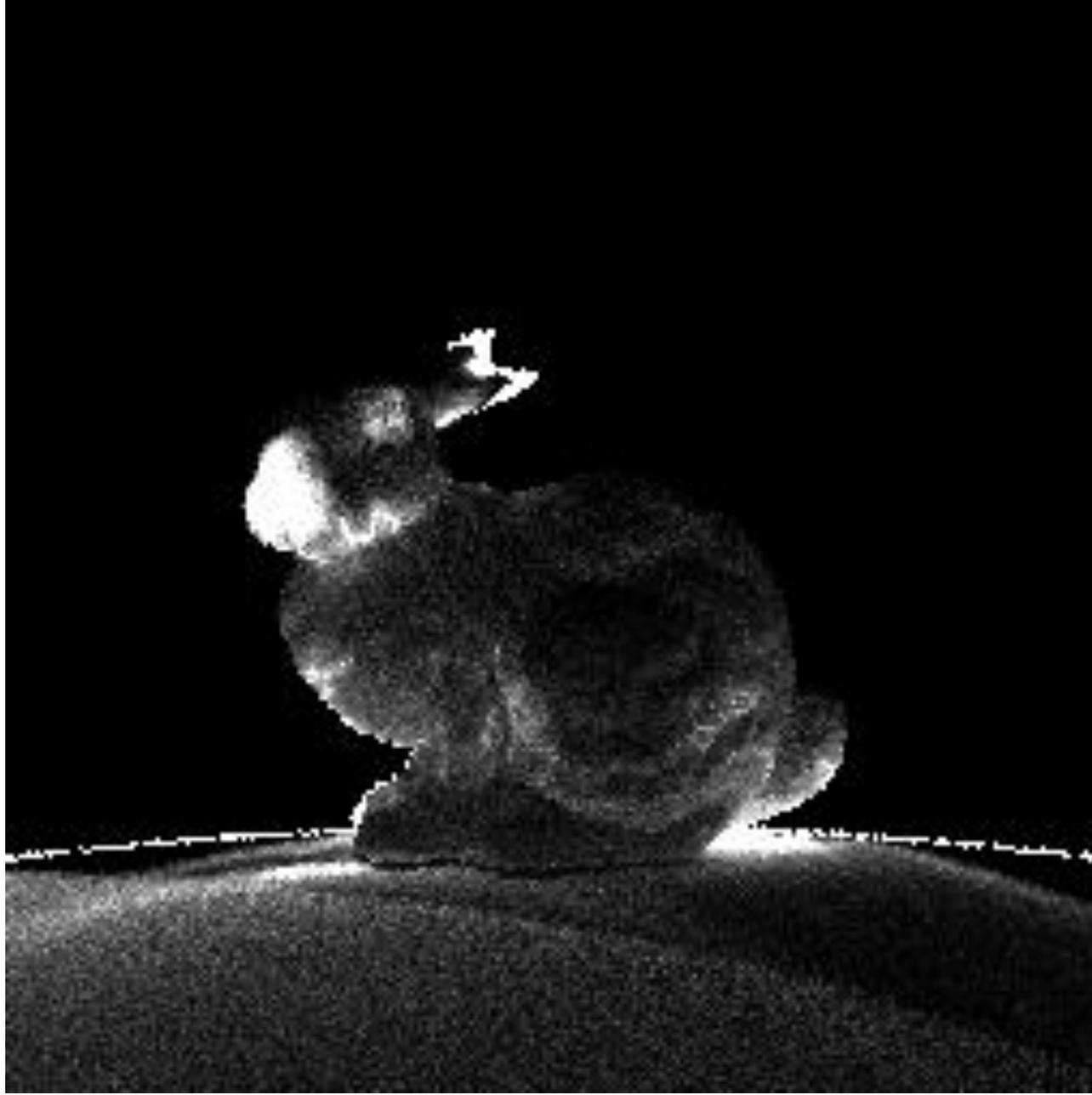
Impulse response



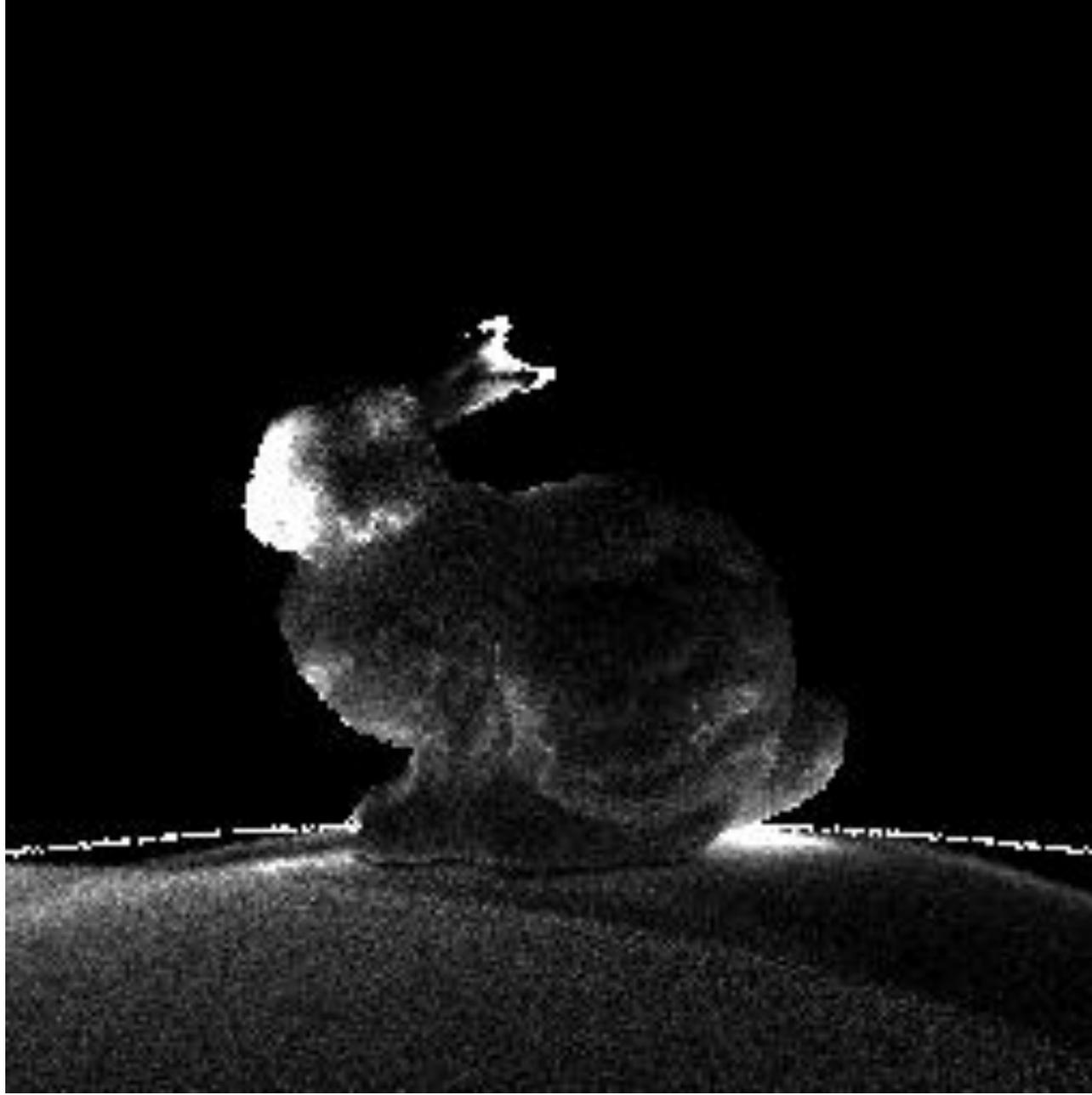
Impulse response



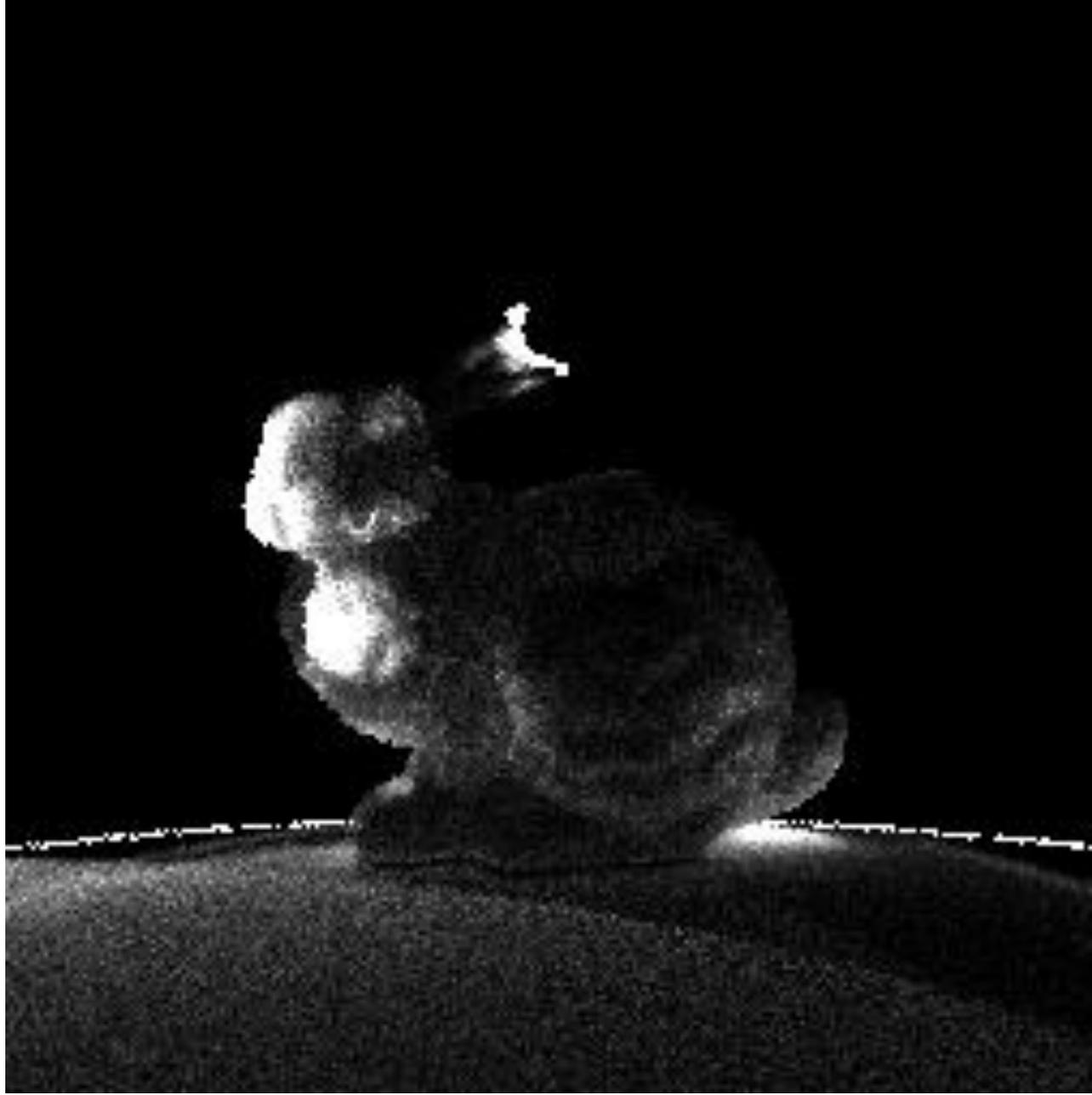
Impulse response



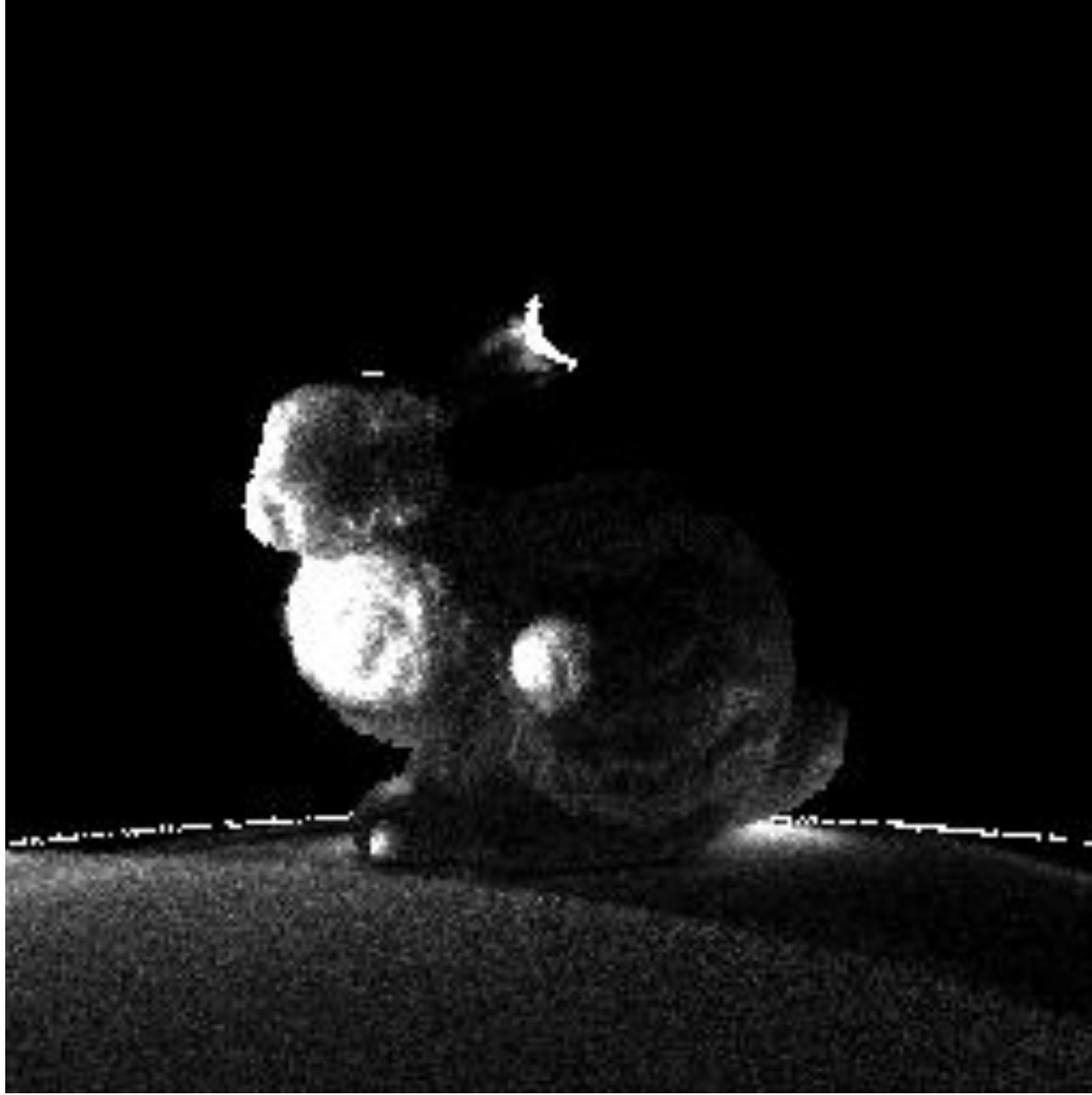
Impulse response



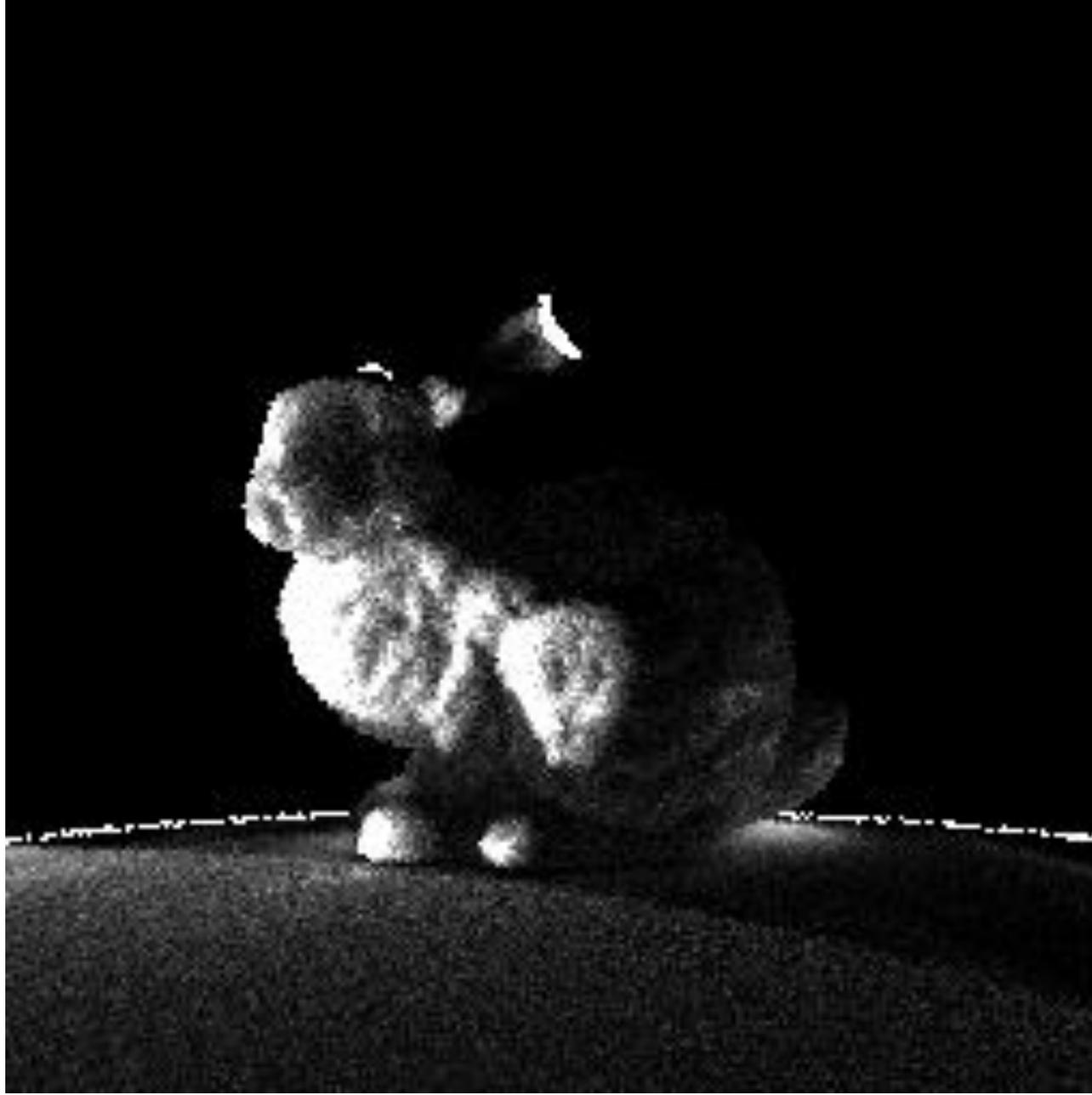
Impulse response



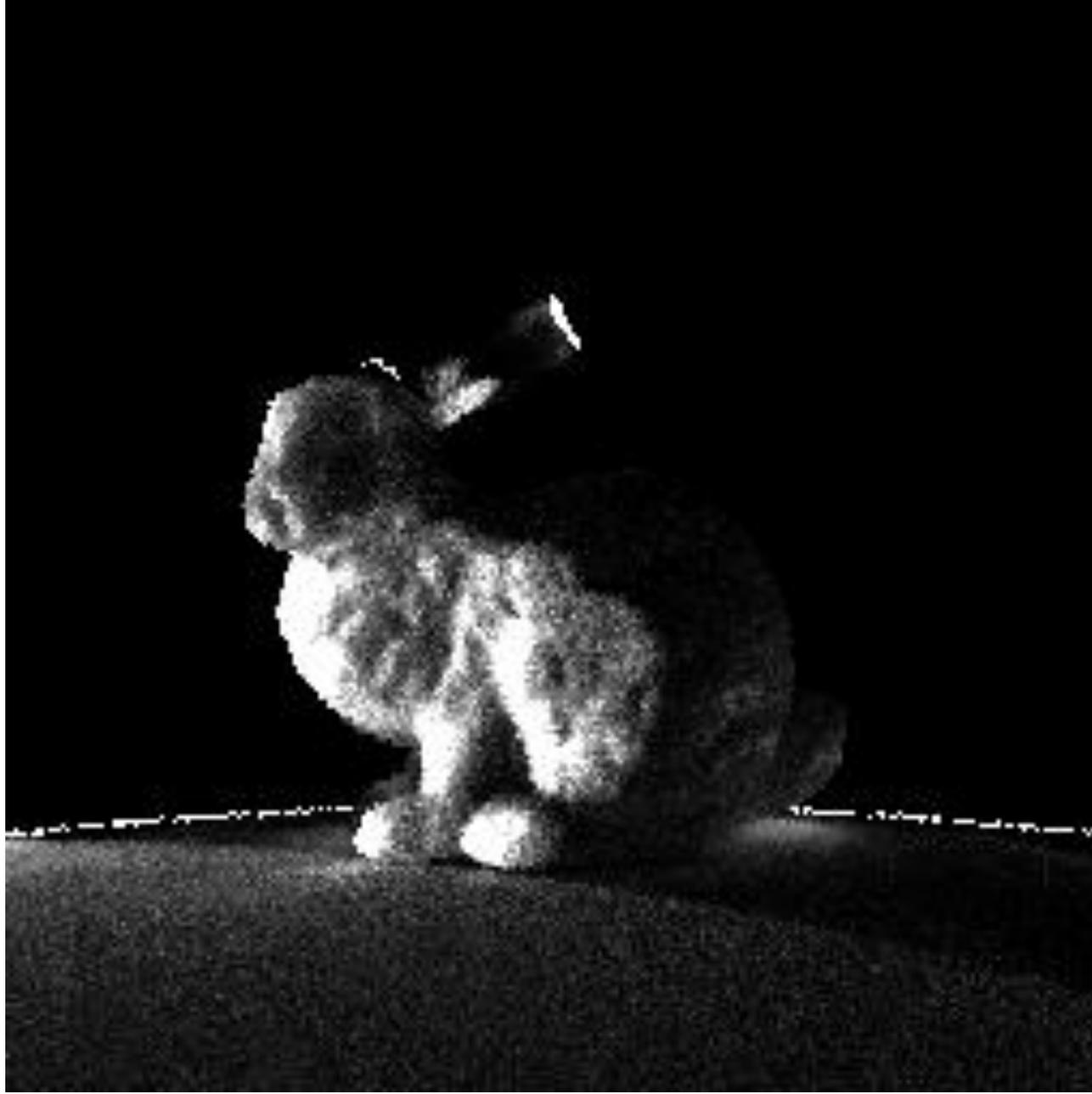
Impulse response



Impulse response



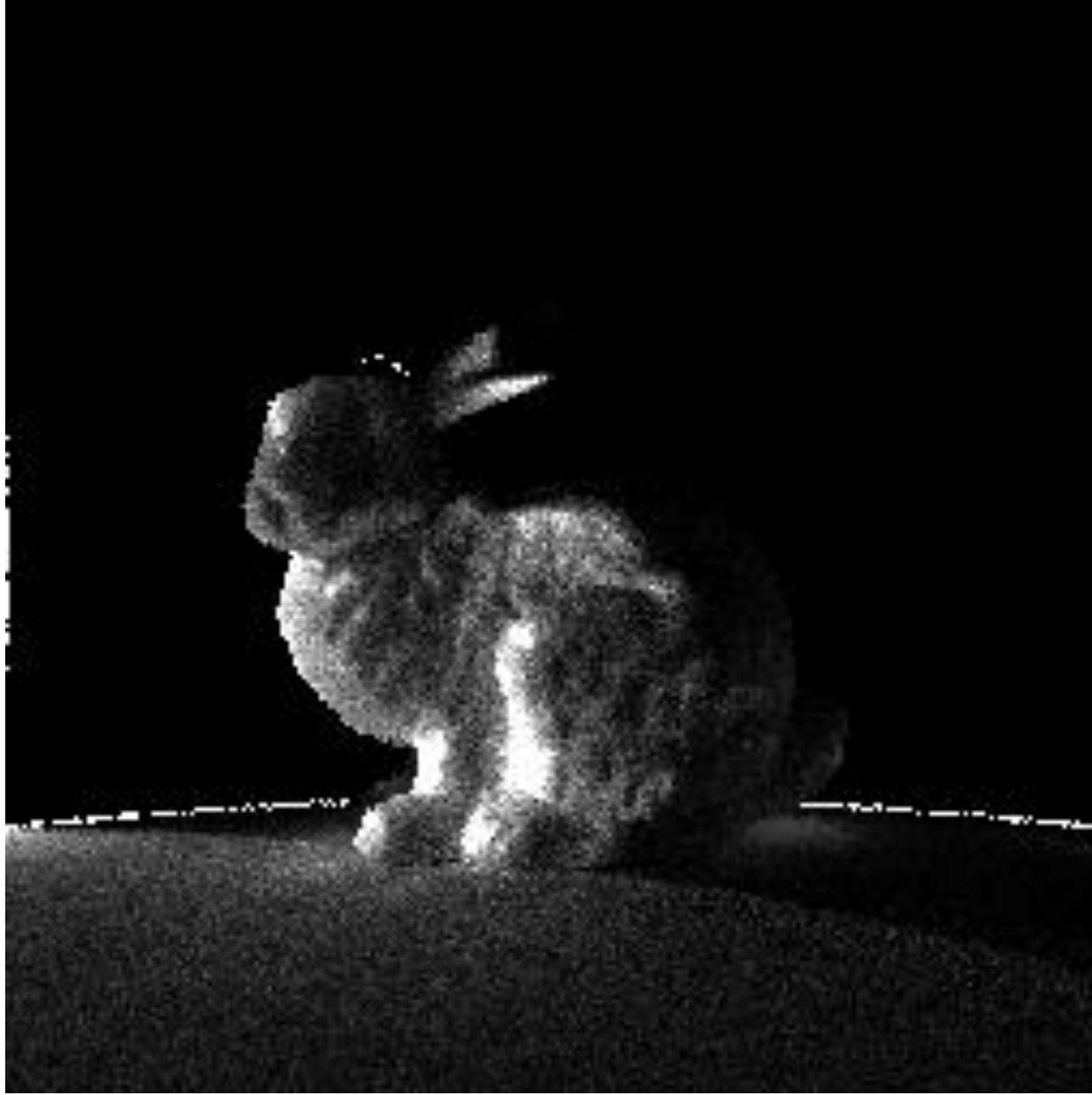
Impulse response



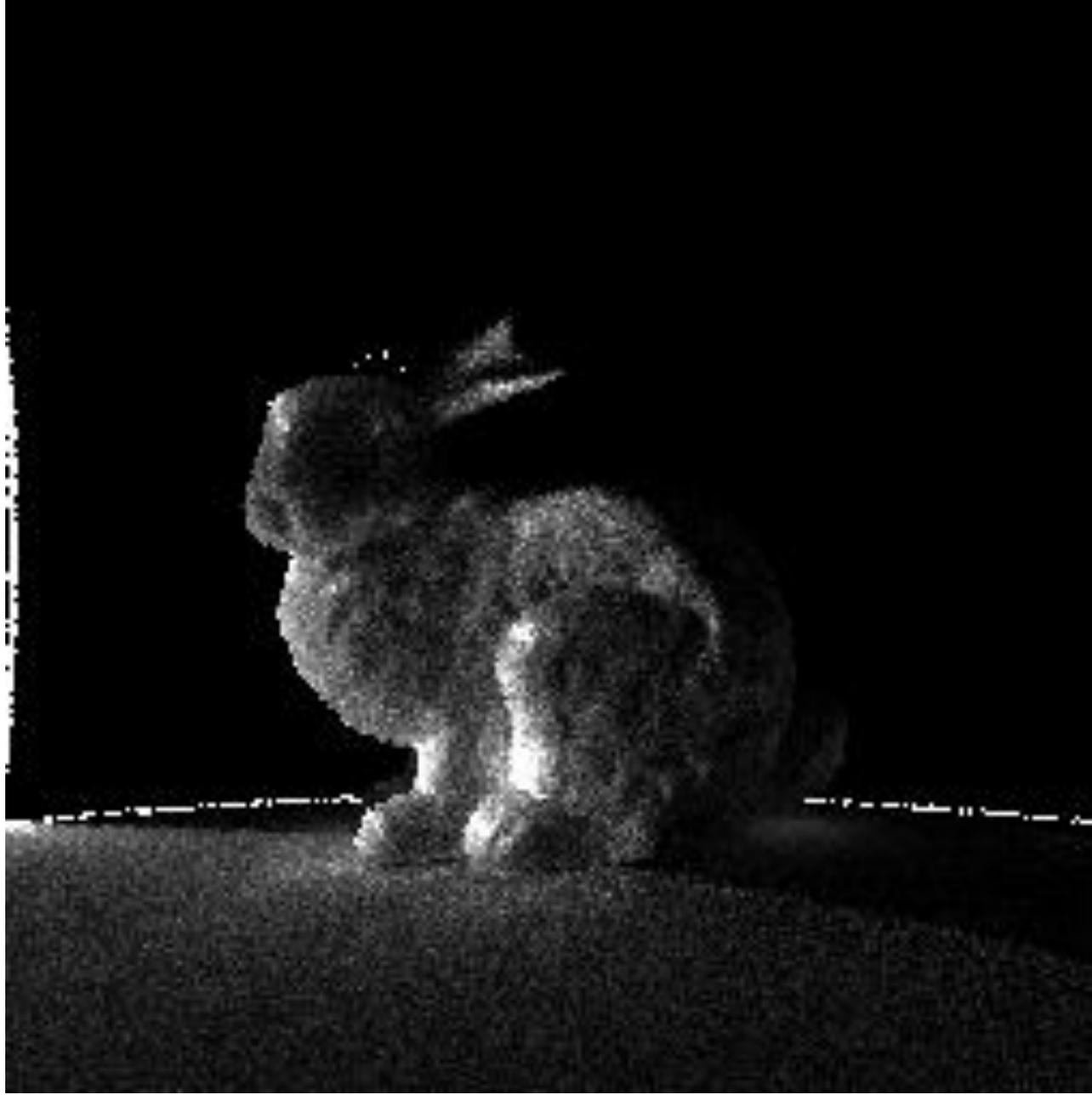
Impulse response



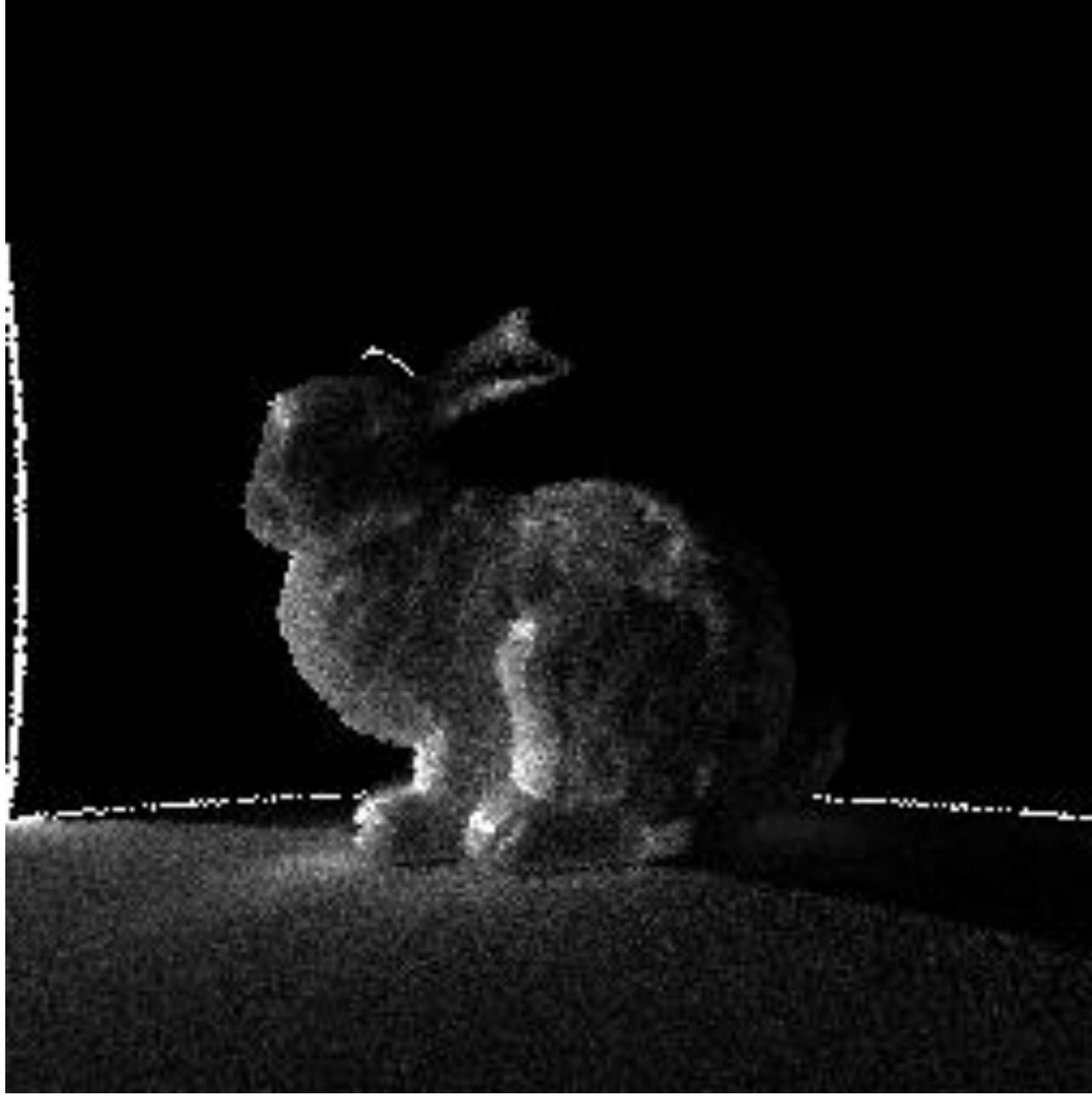
Impulse response



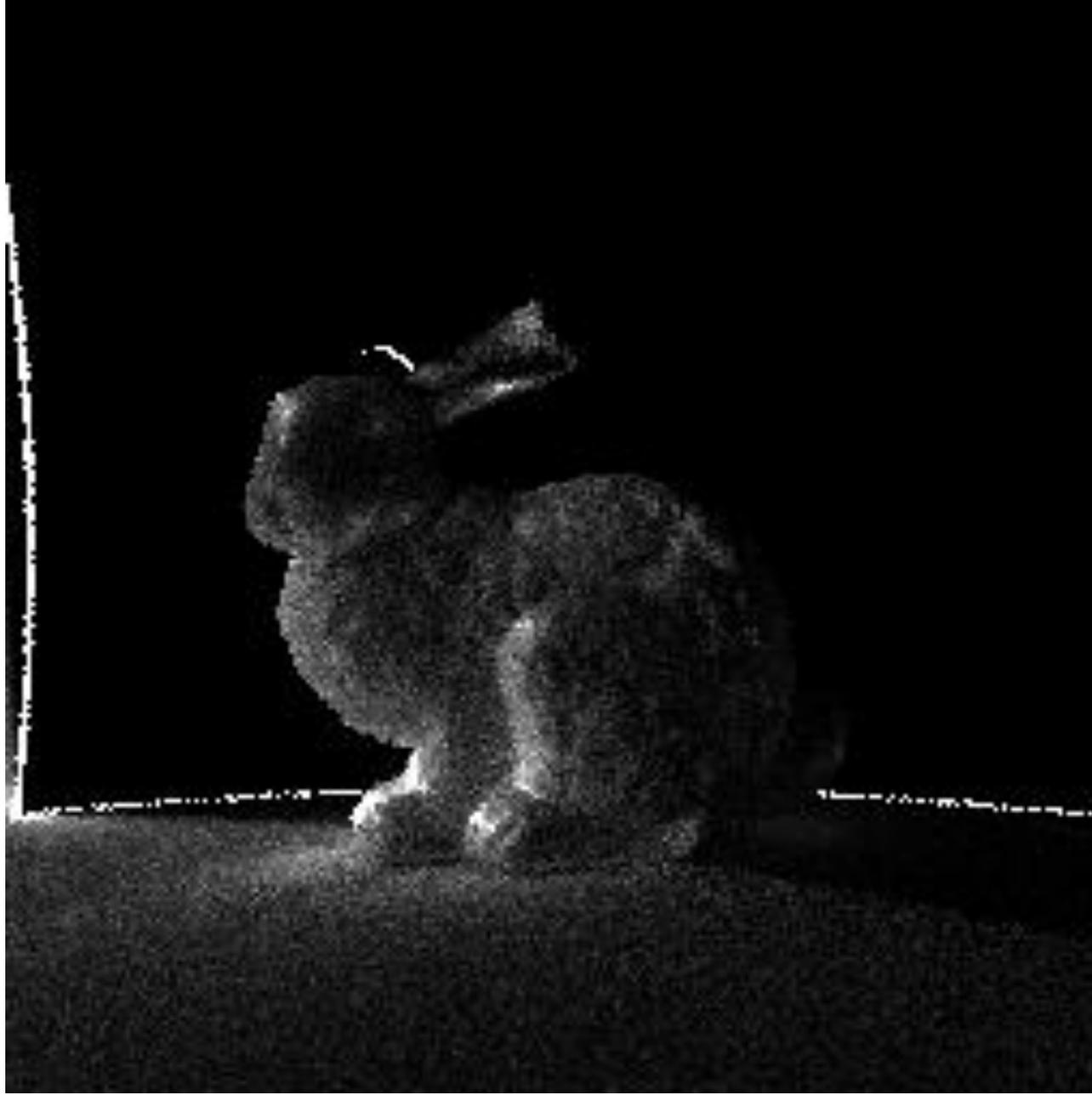
Impulse response



Impulse response



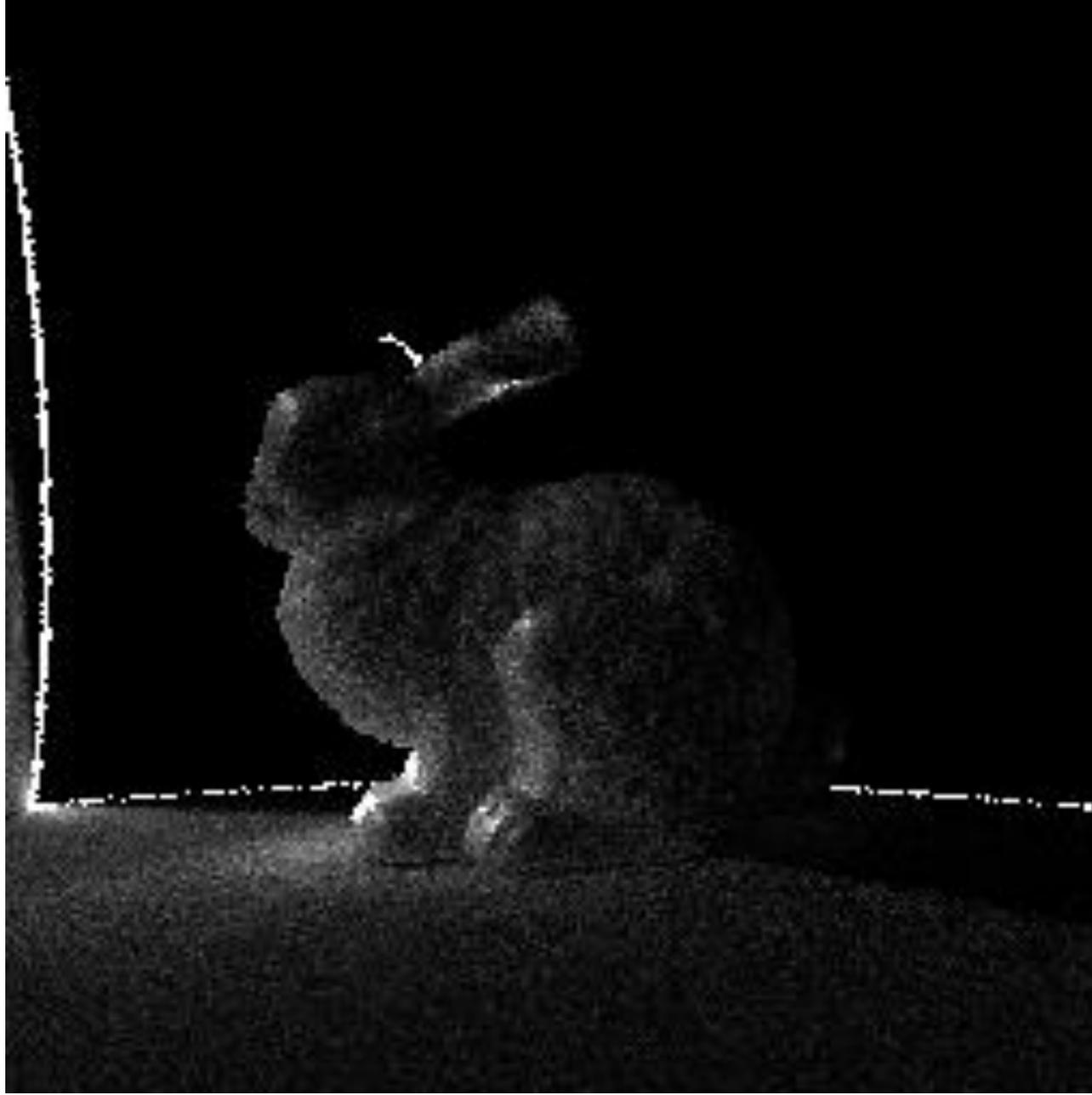
Impulse response



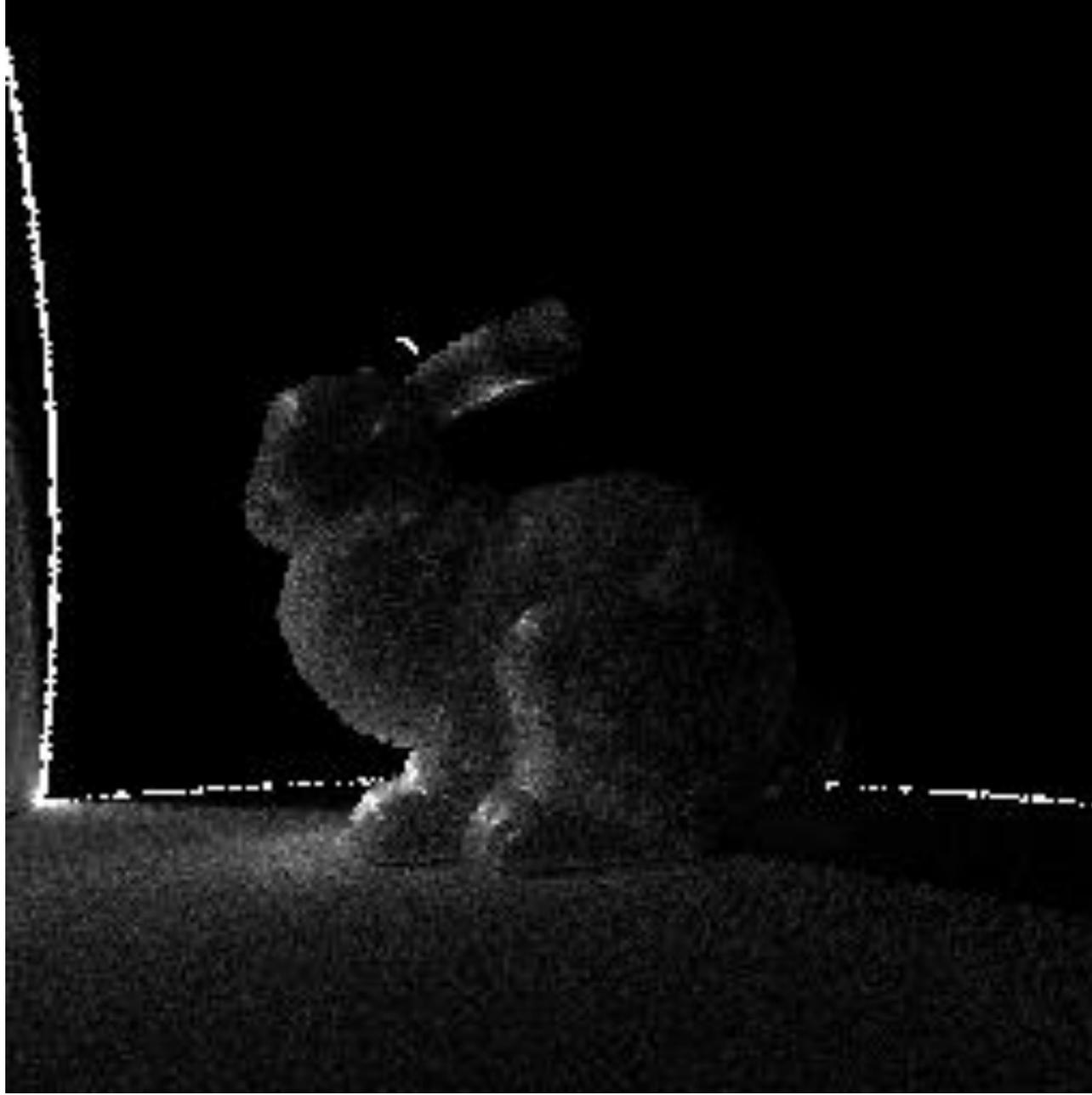
Impulse response



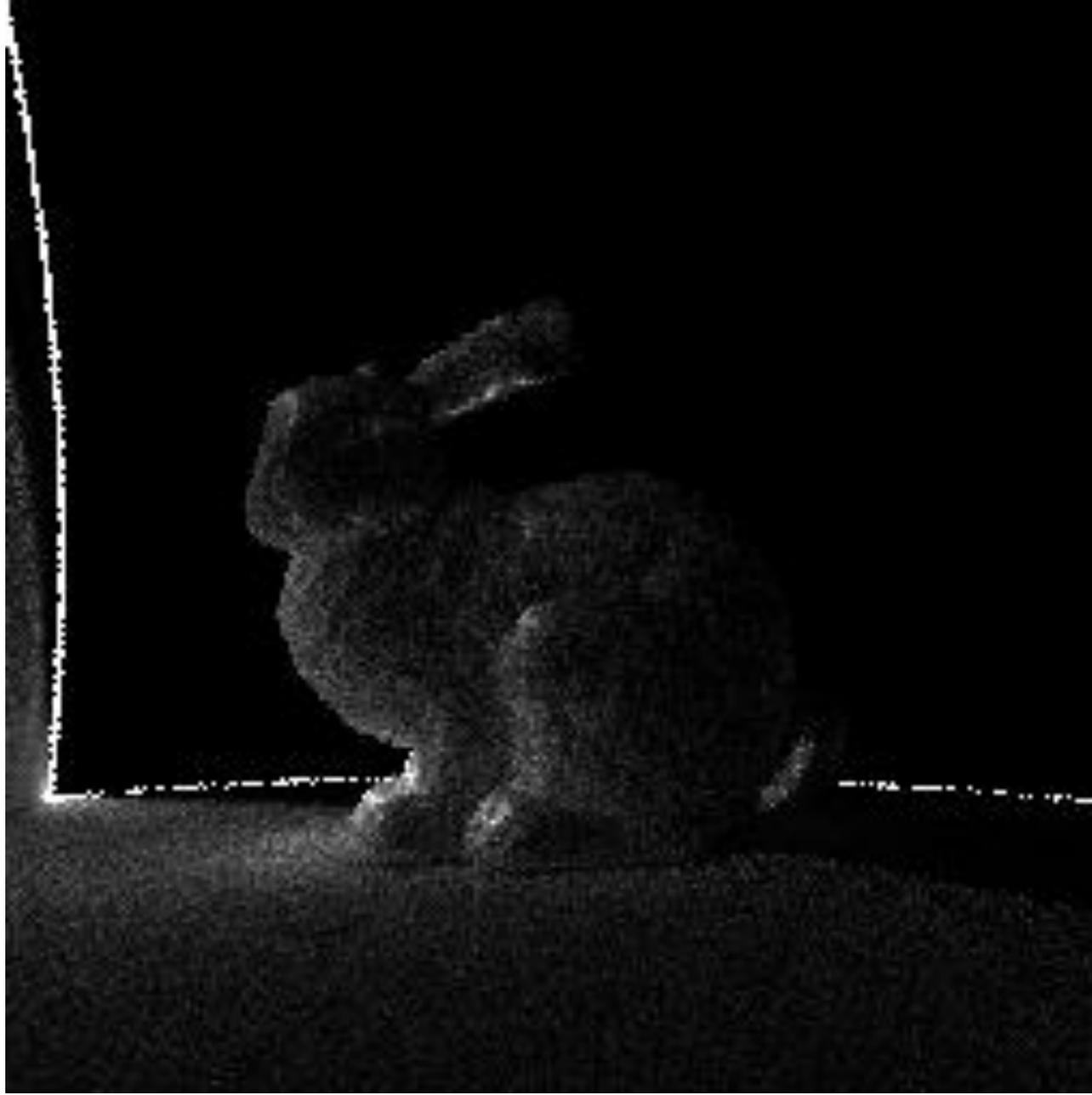
Impulse response



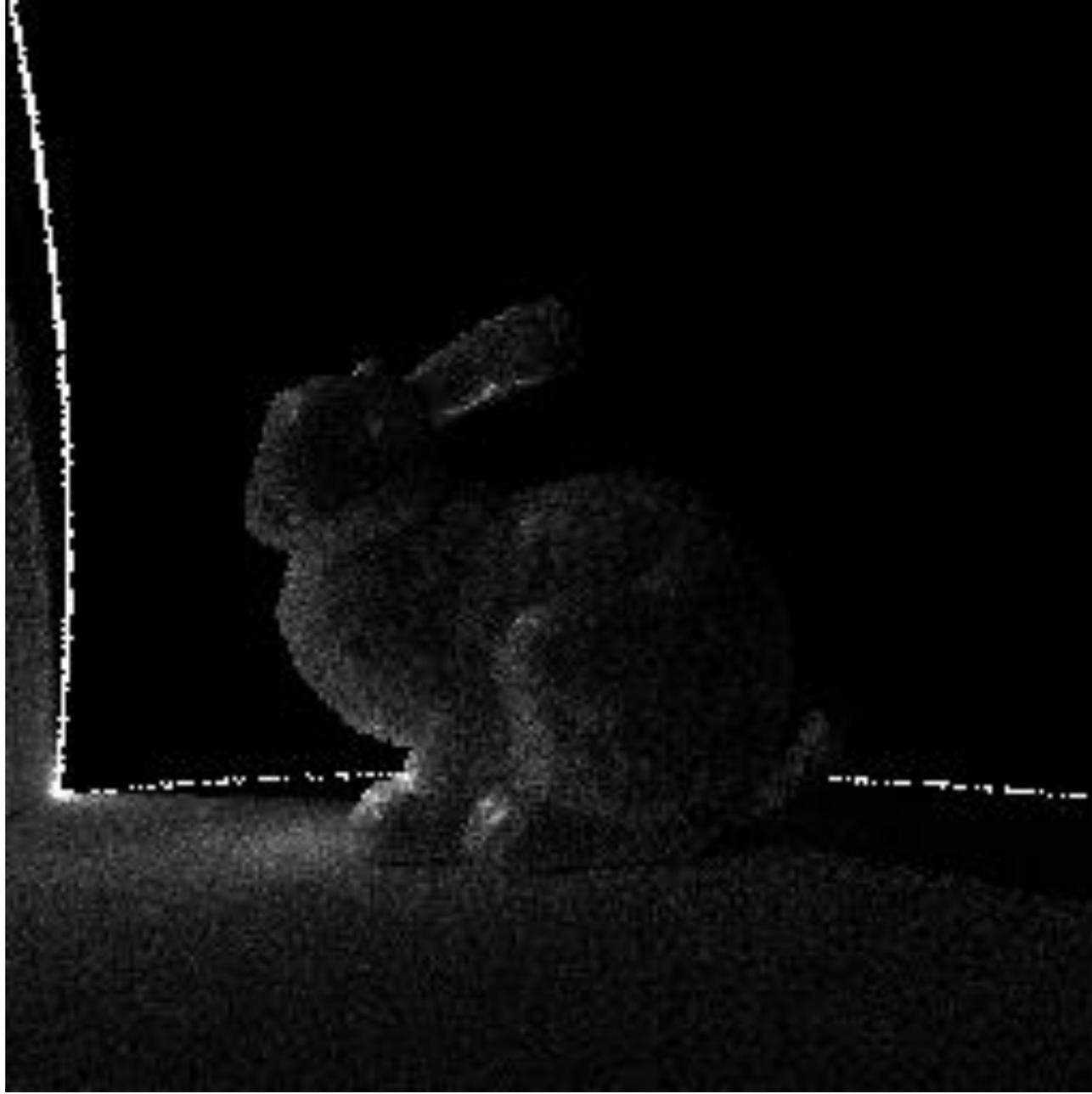
Impulse response



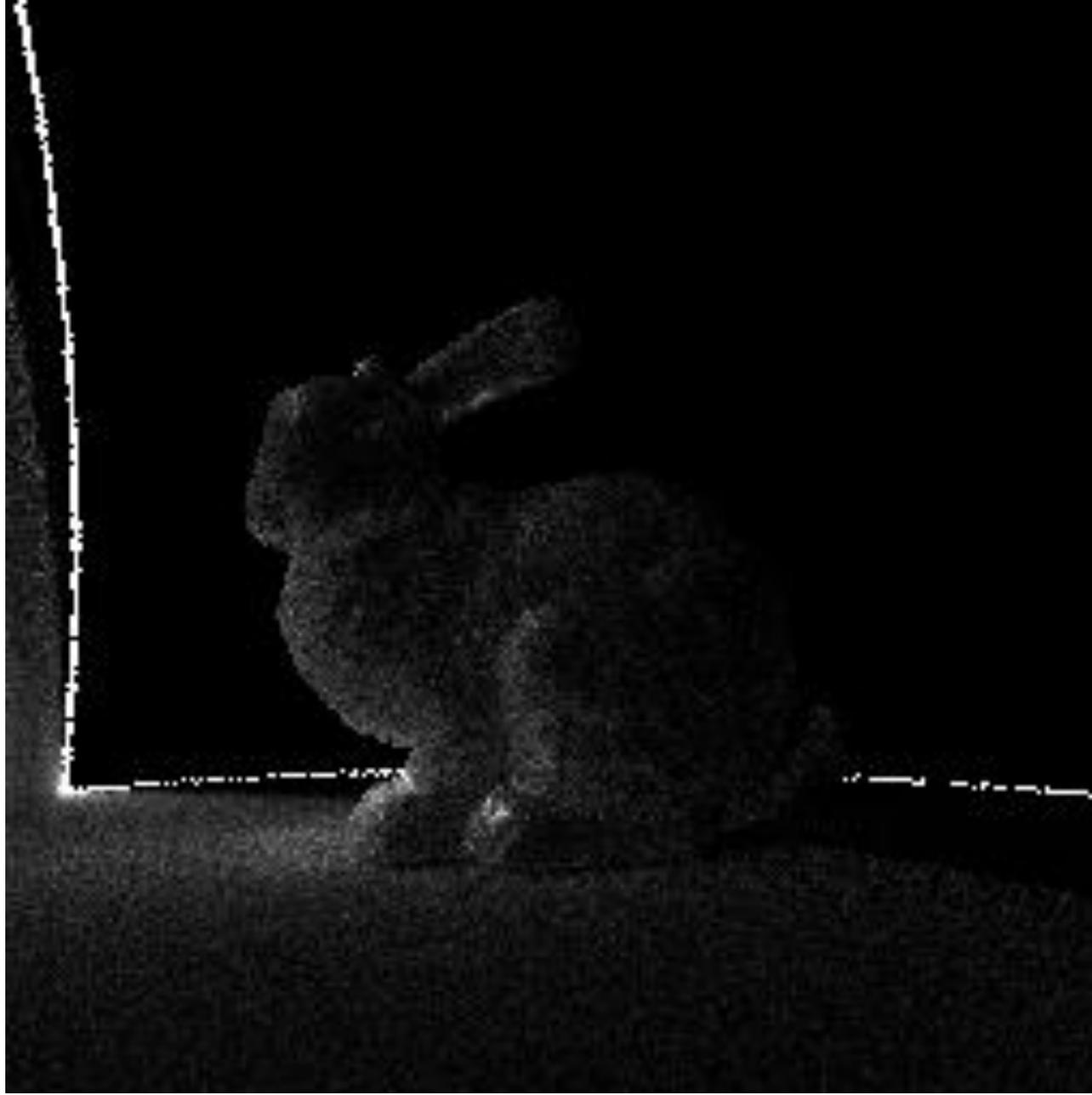
Impulse response



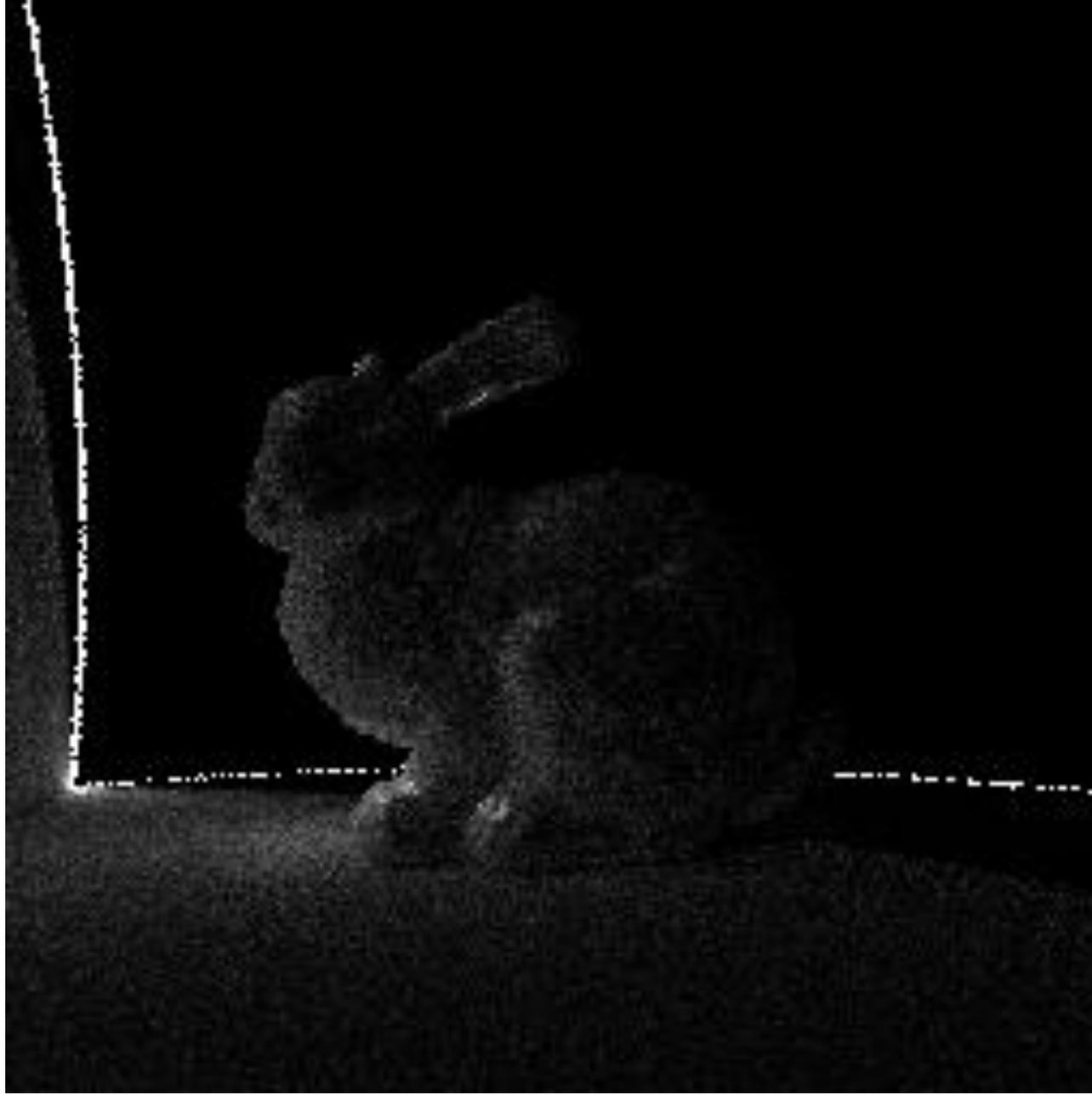
Impulse response



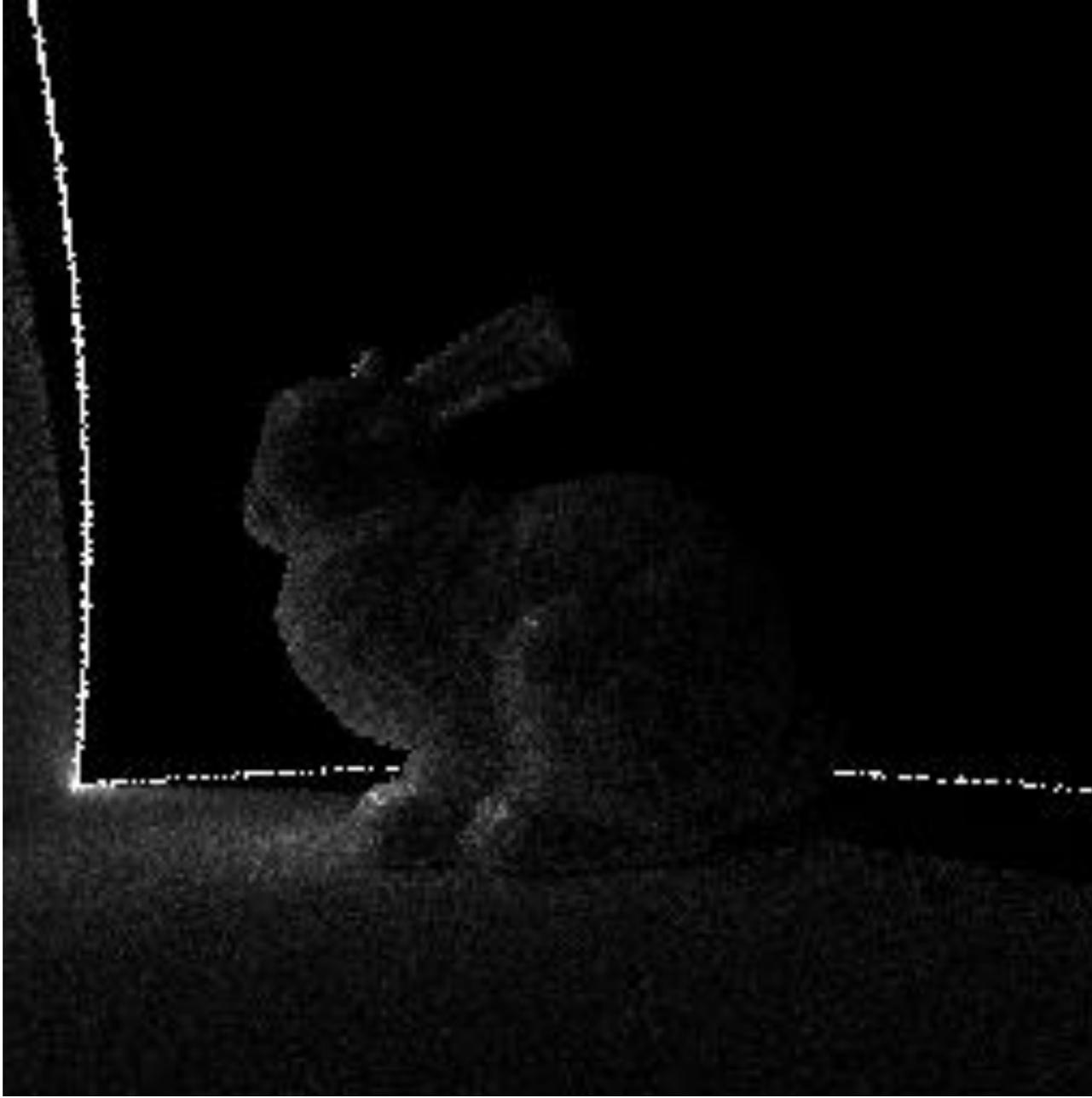
Impulse response



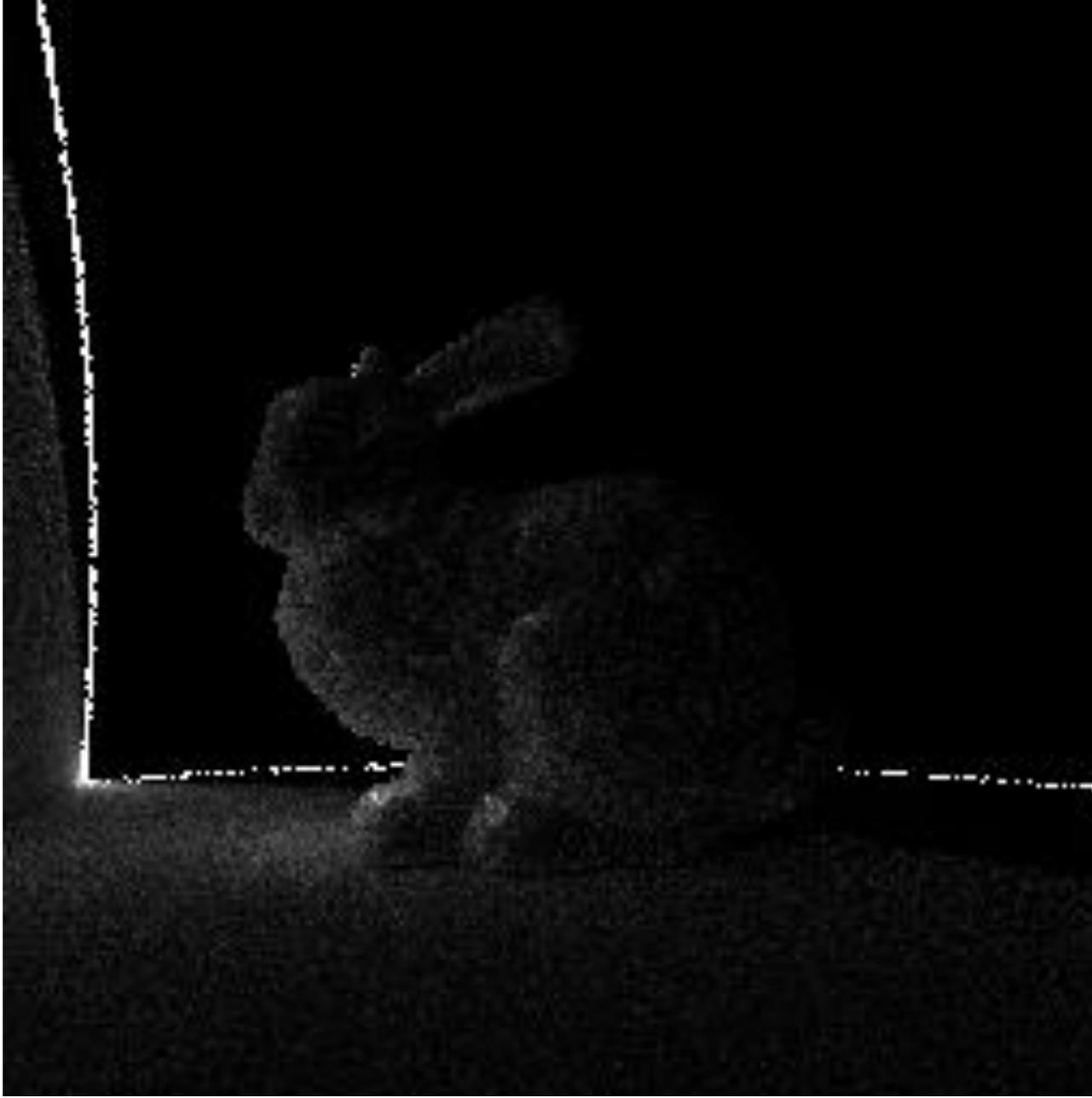
Impulse response



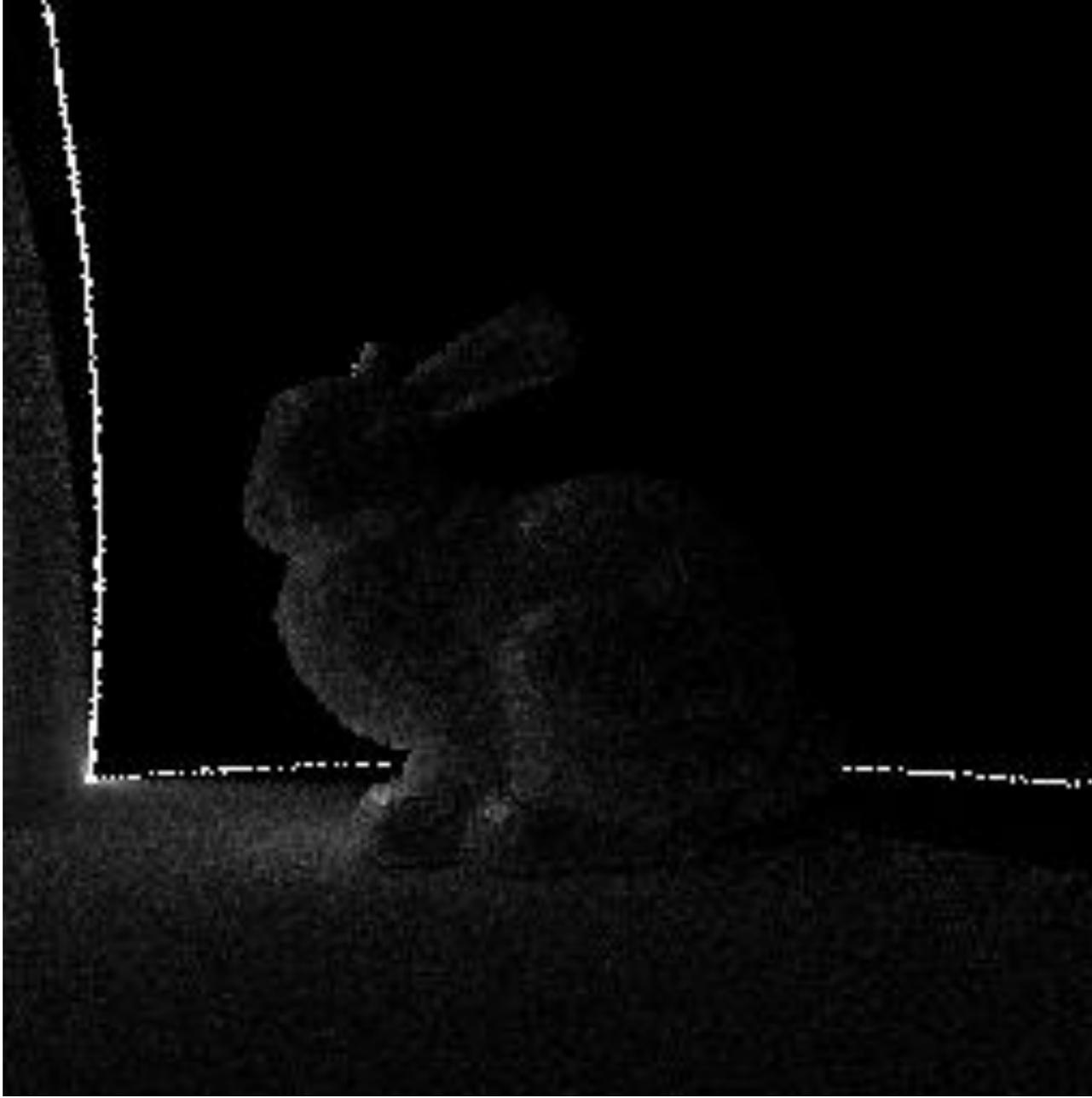
Impulse response



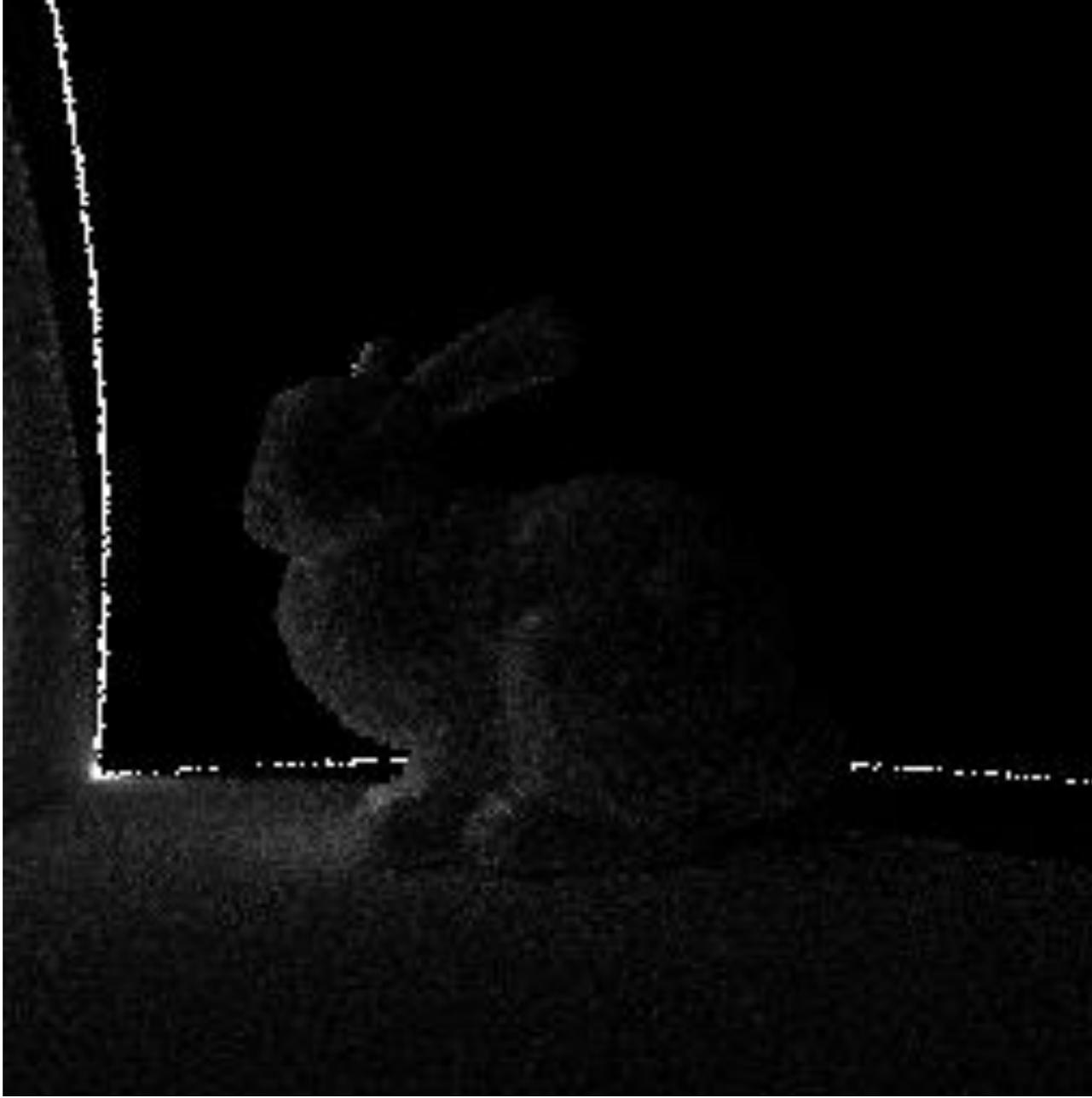
Impulse response



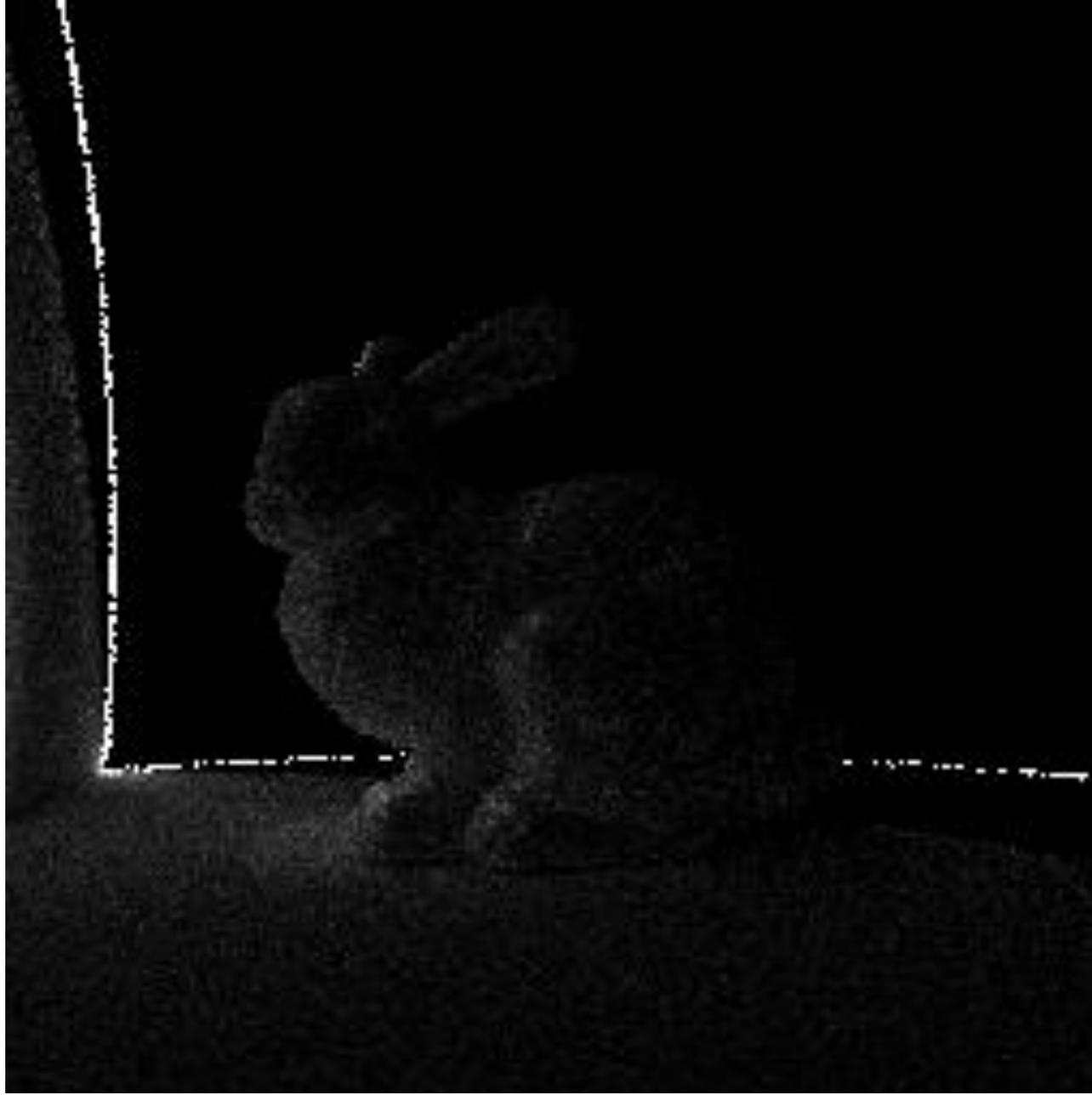
Impulse response



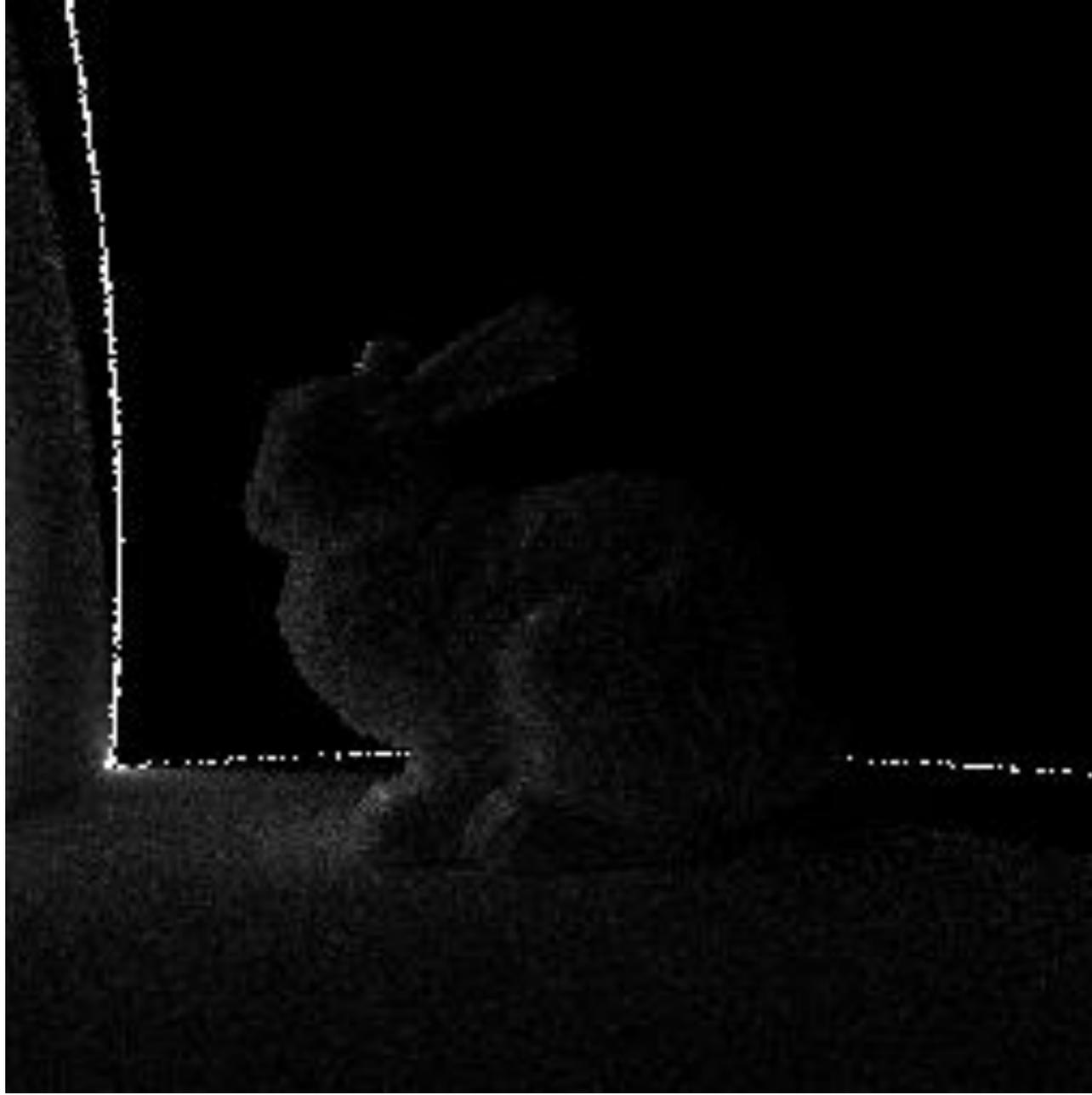
Impulse response



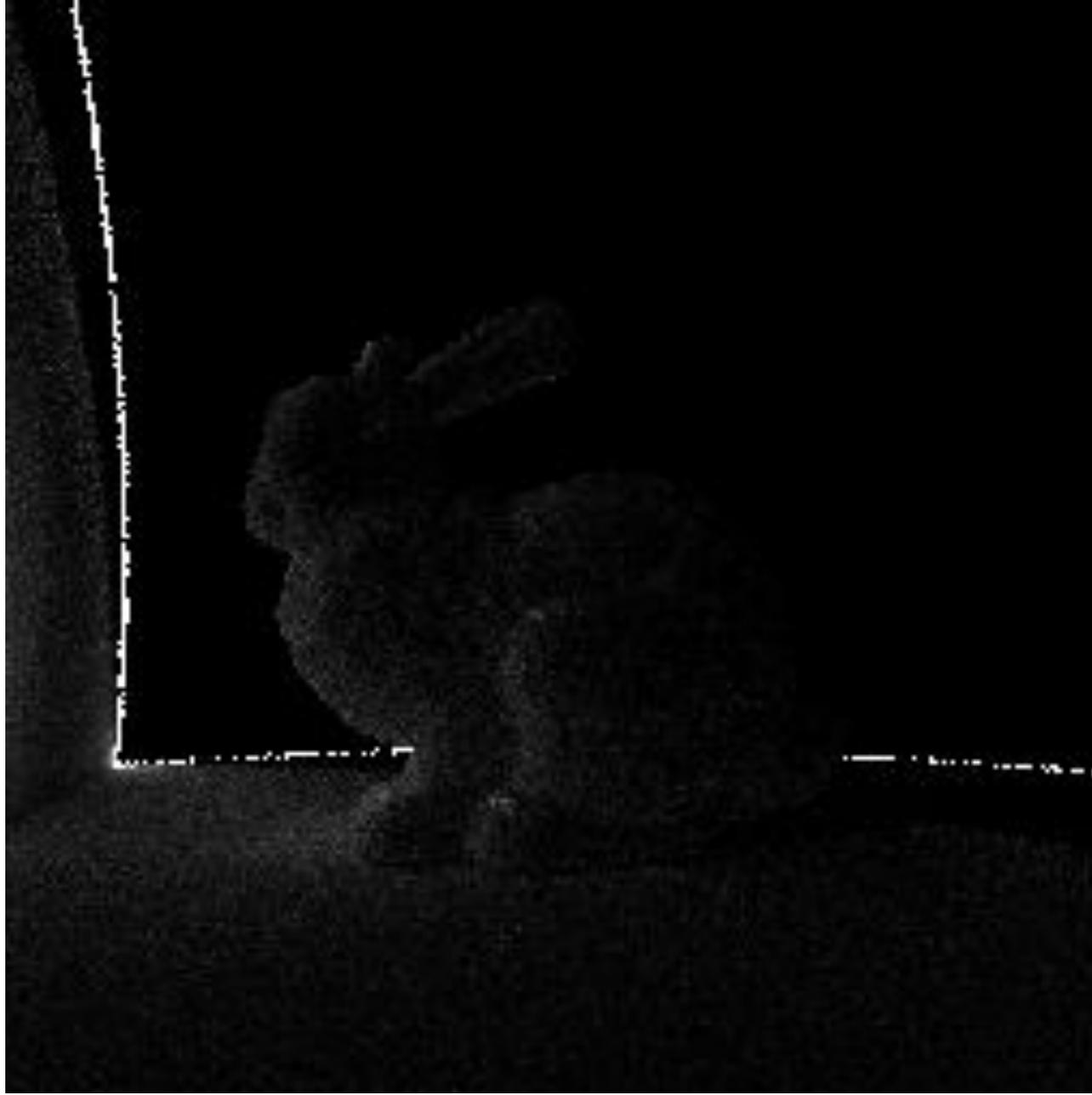
Impulse response



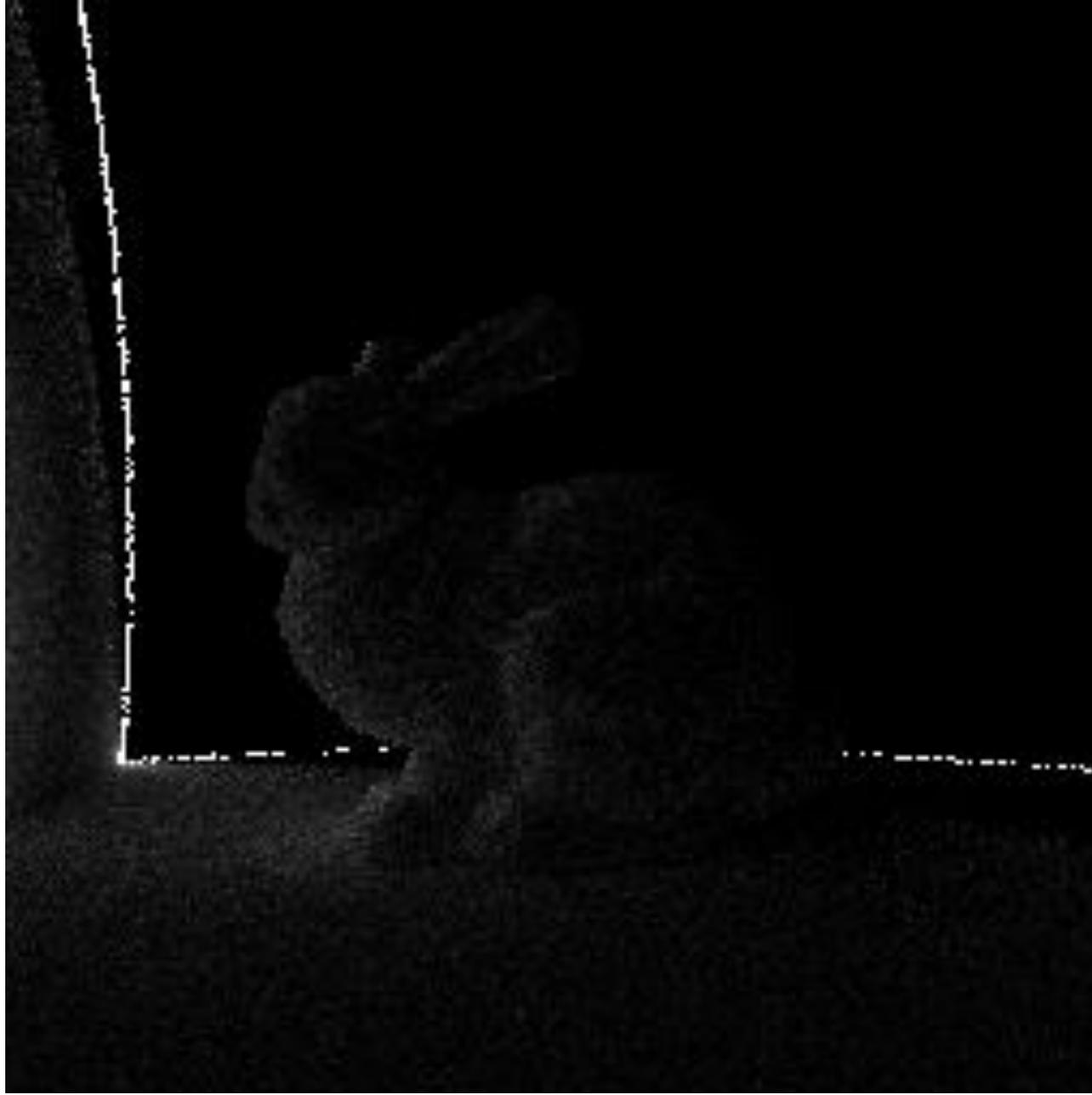
Impulse response



Impulse response



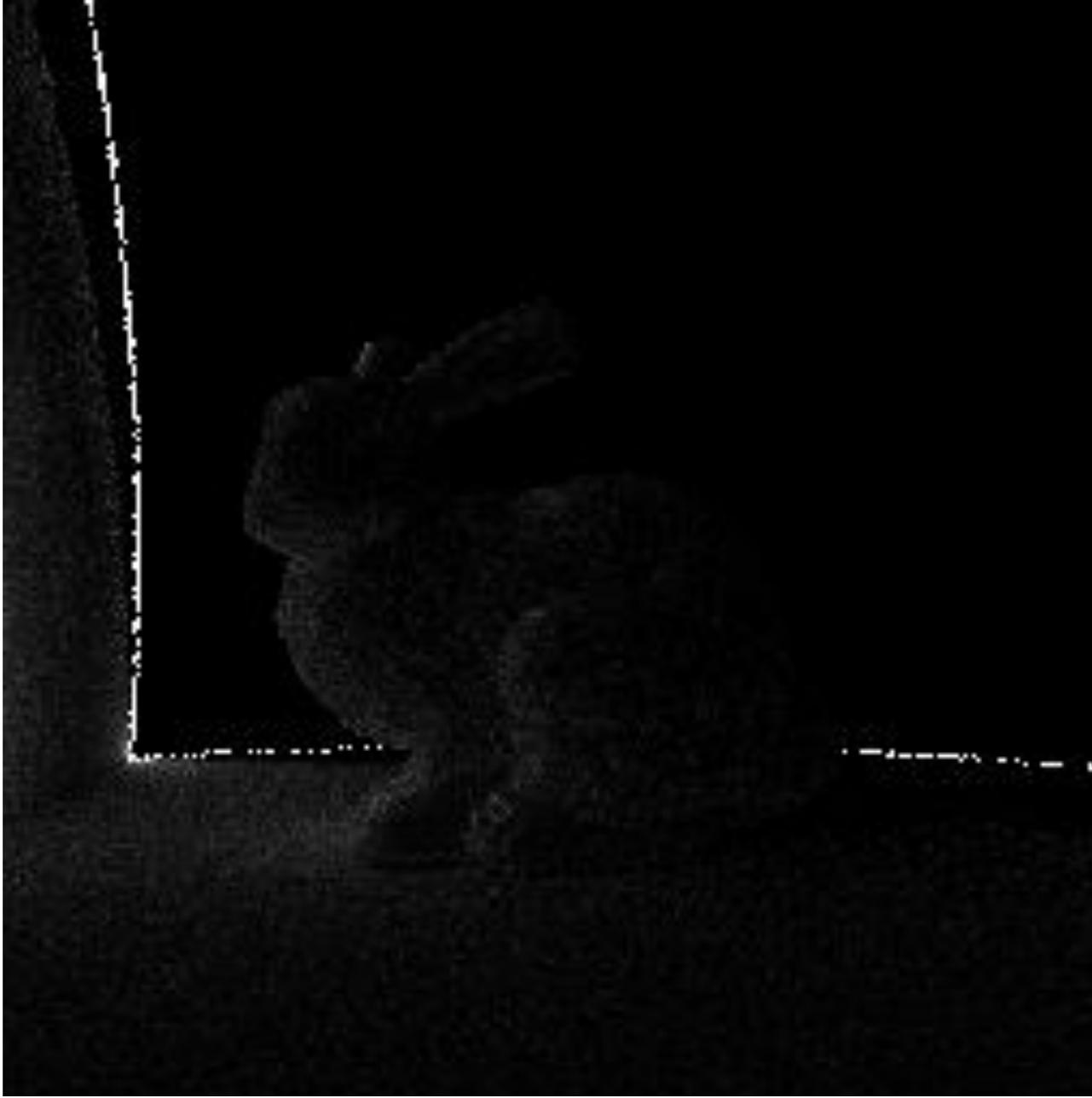
Impulse response



Impulse response



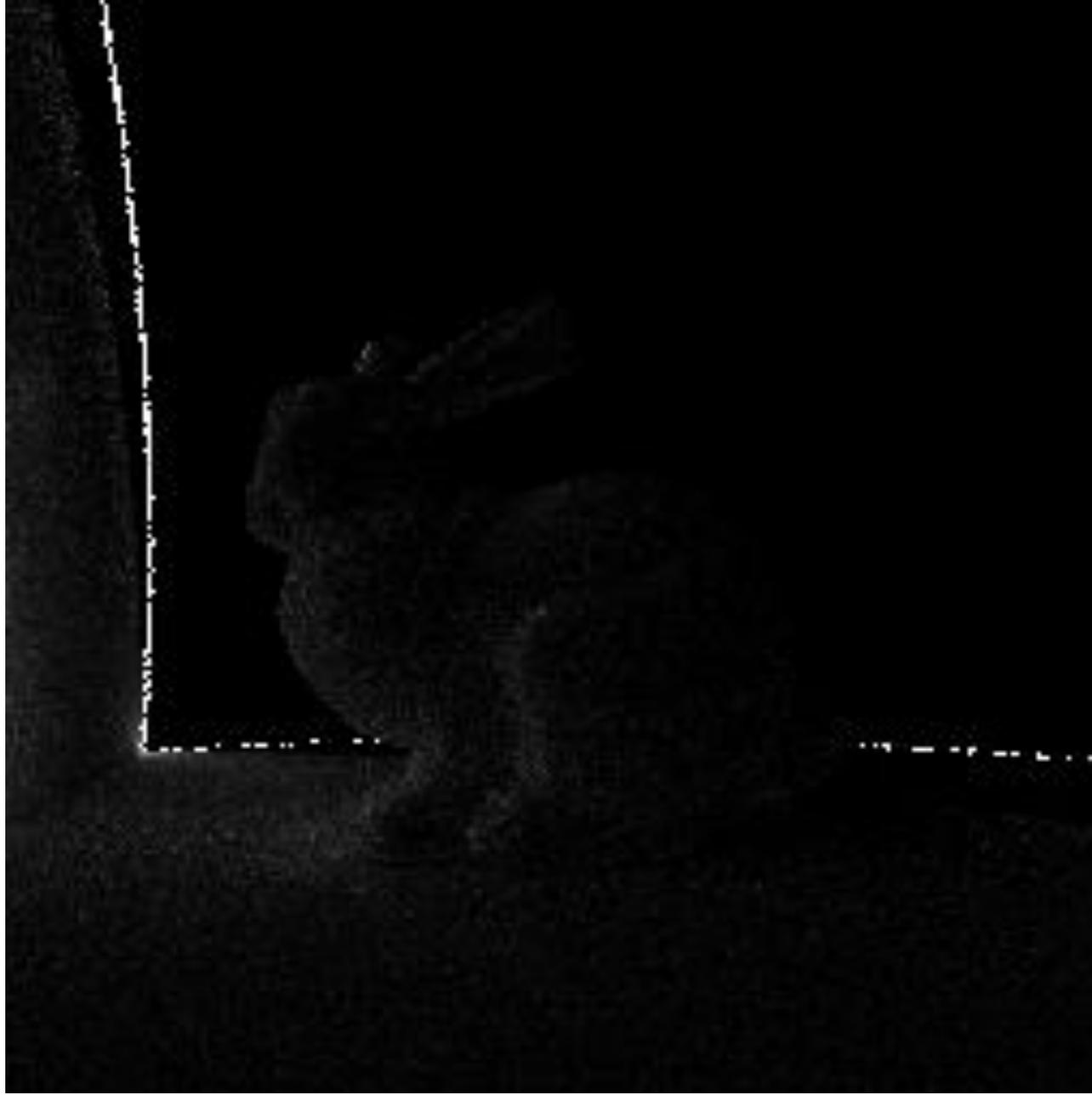
Impulse response



Impulse response



Impulse response



Impulse response



Impulse response



Impulse response



Impulse response



Impulse response



Impulse response



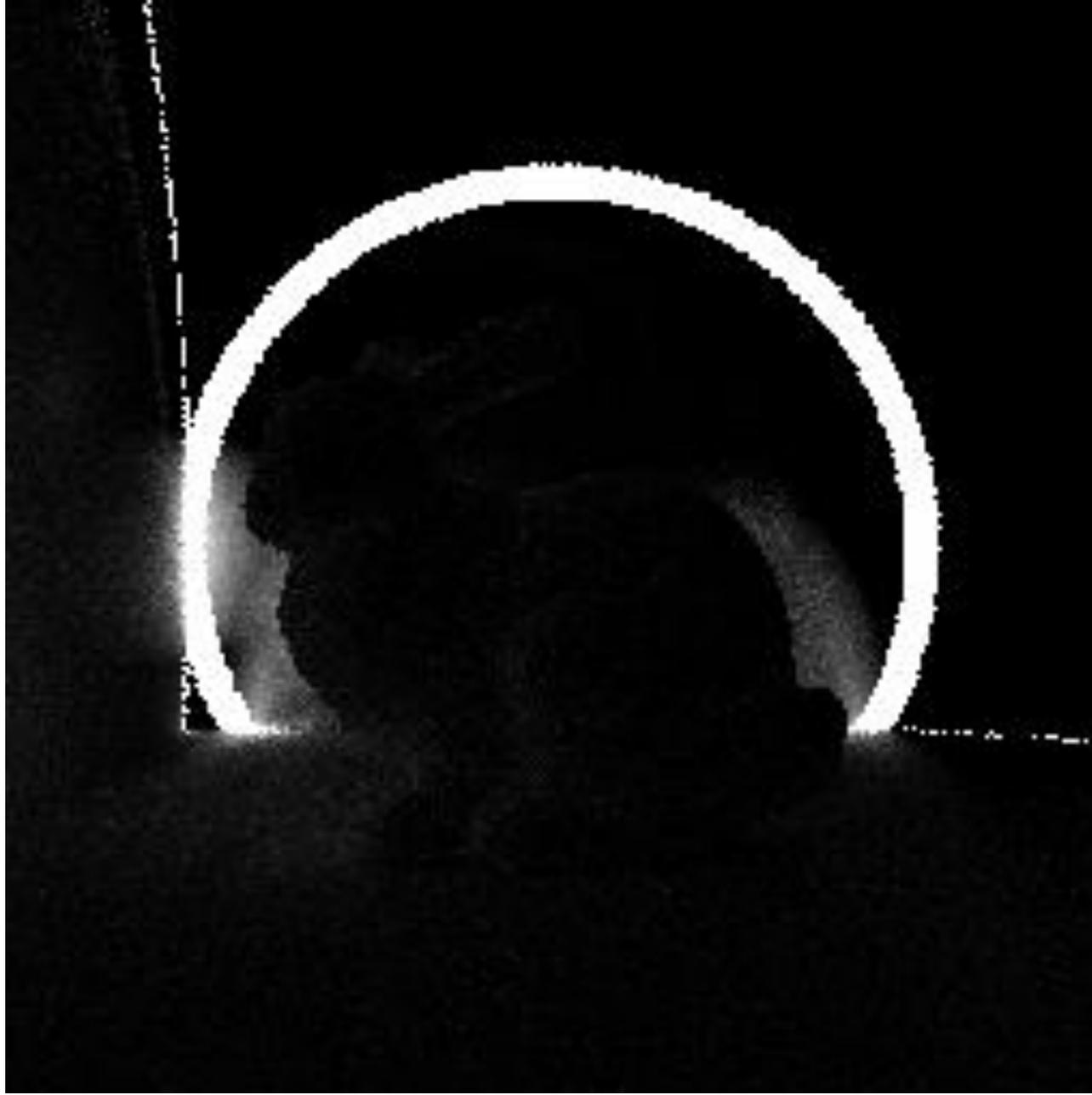
Impulse response



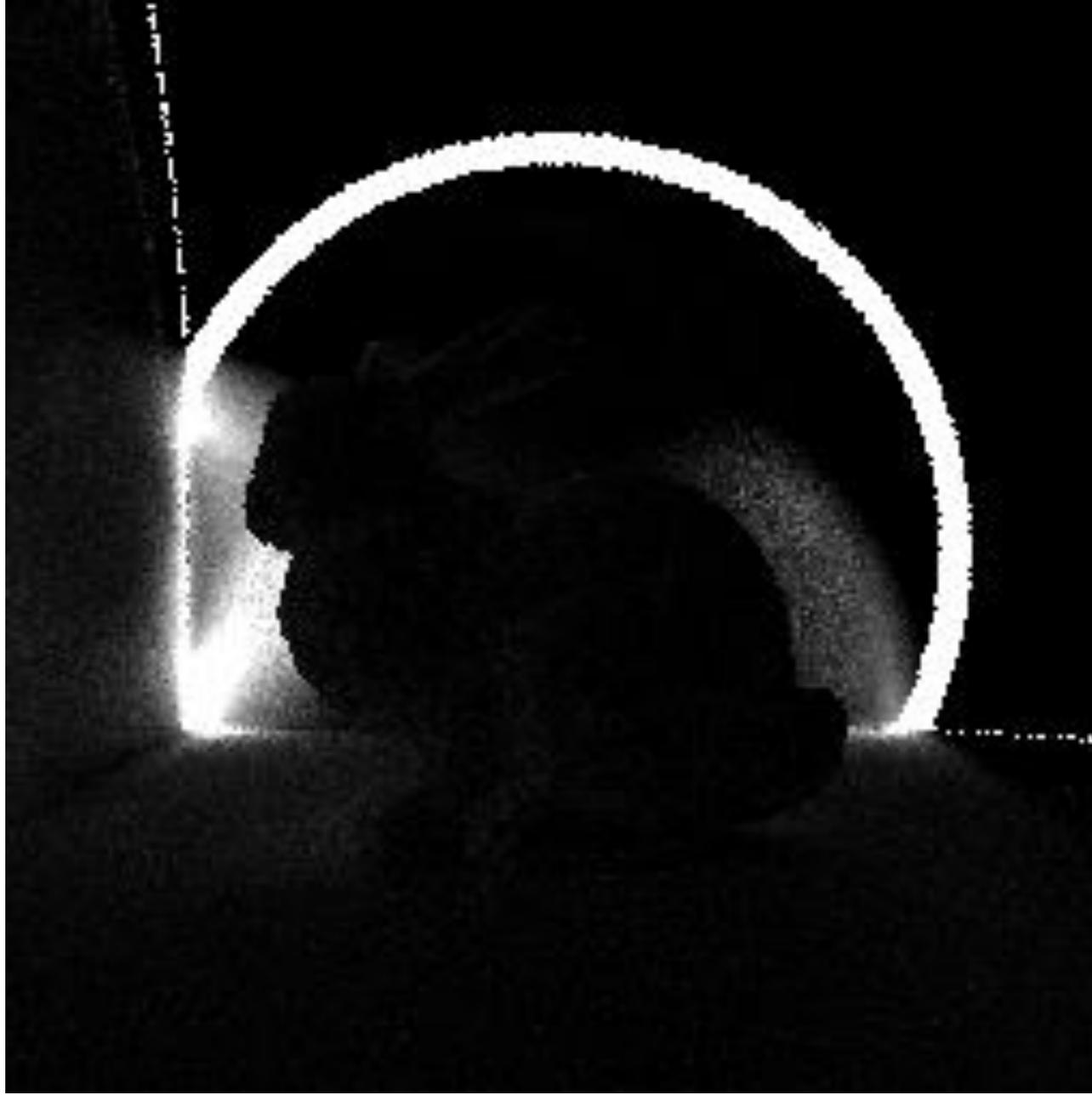
Impulse response



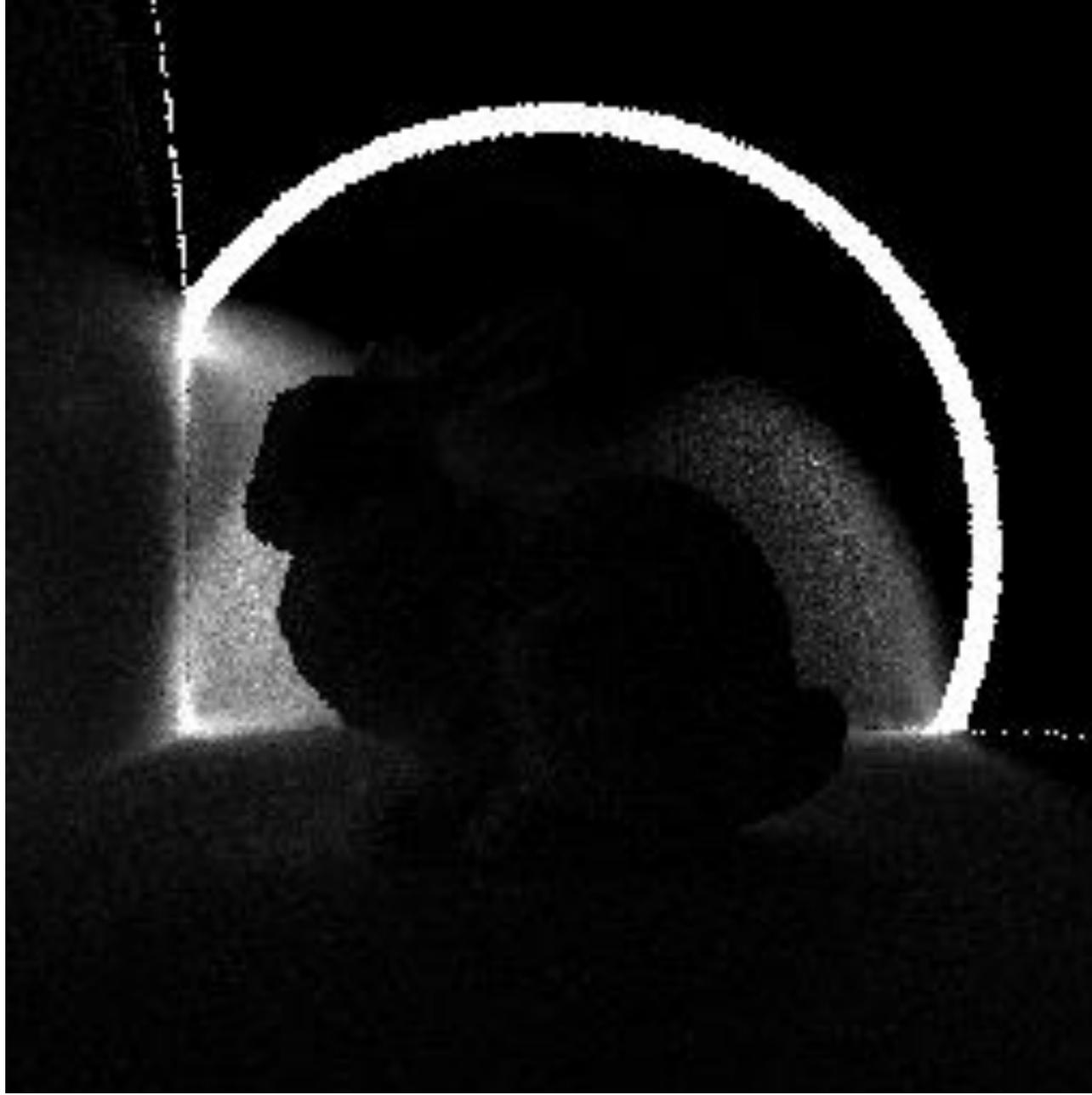
Impulse response



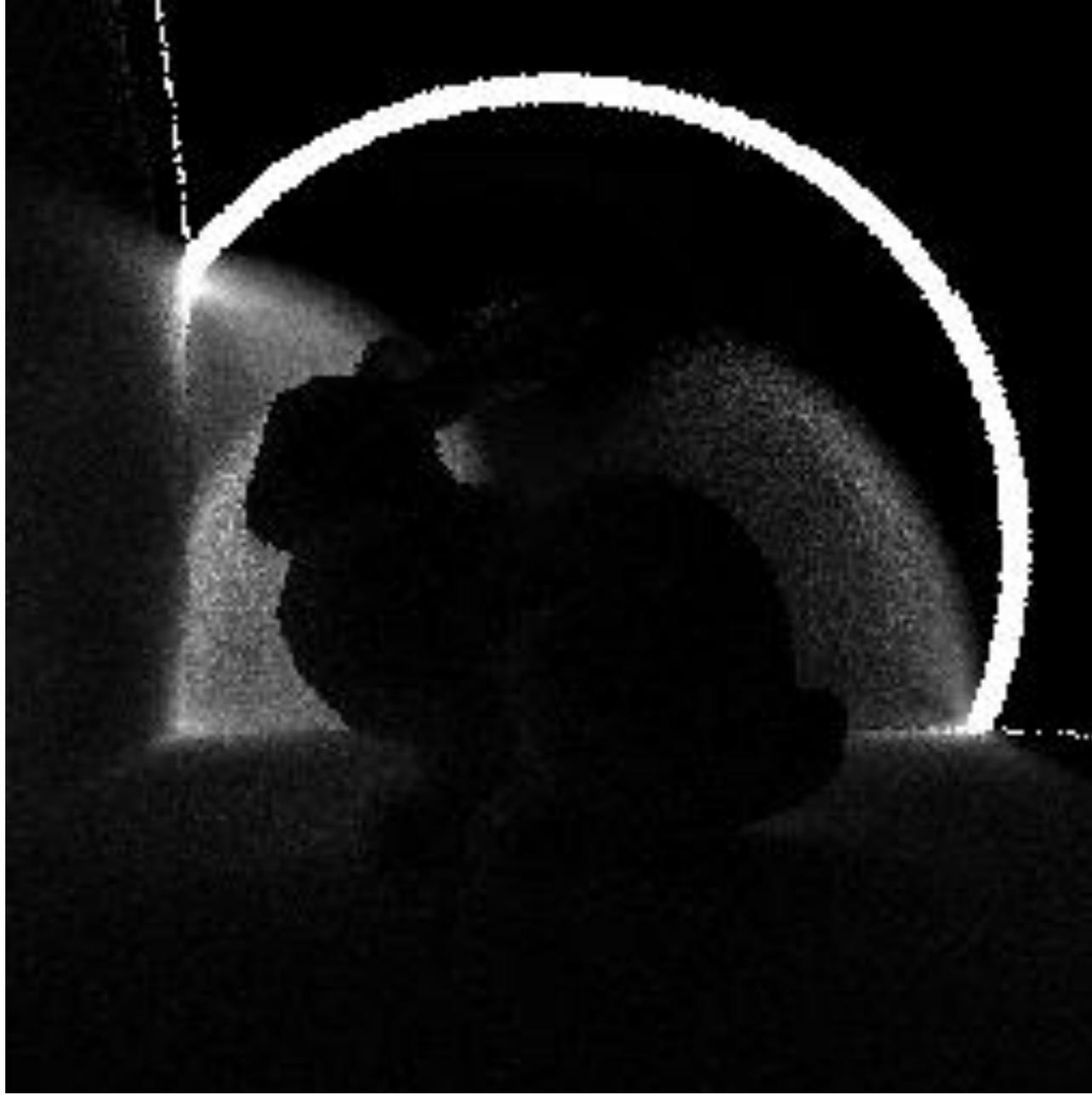
Impulse response



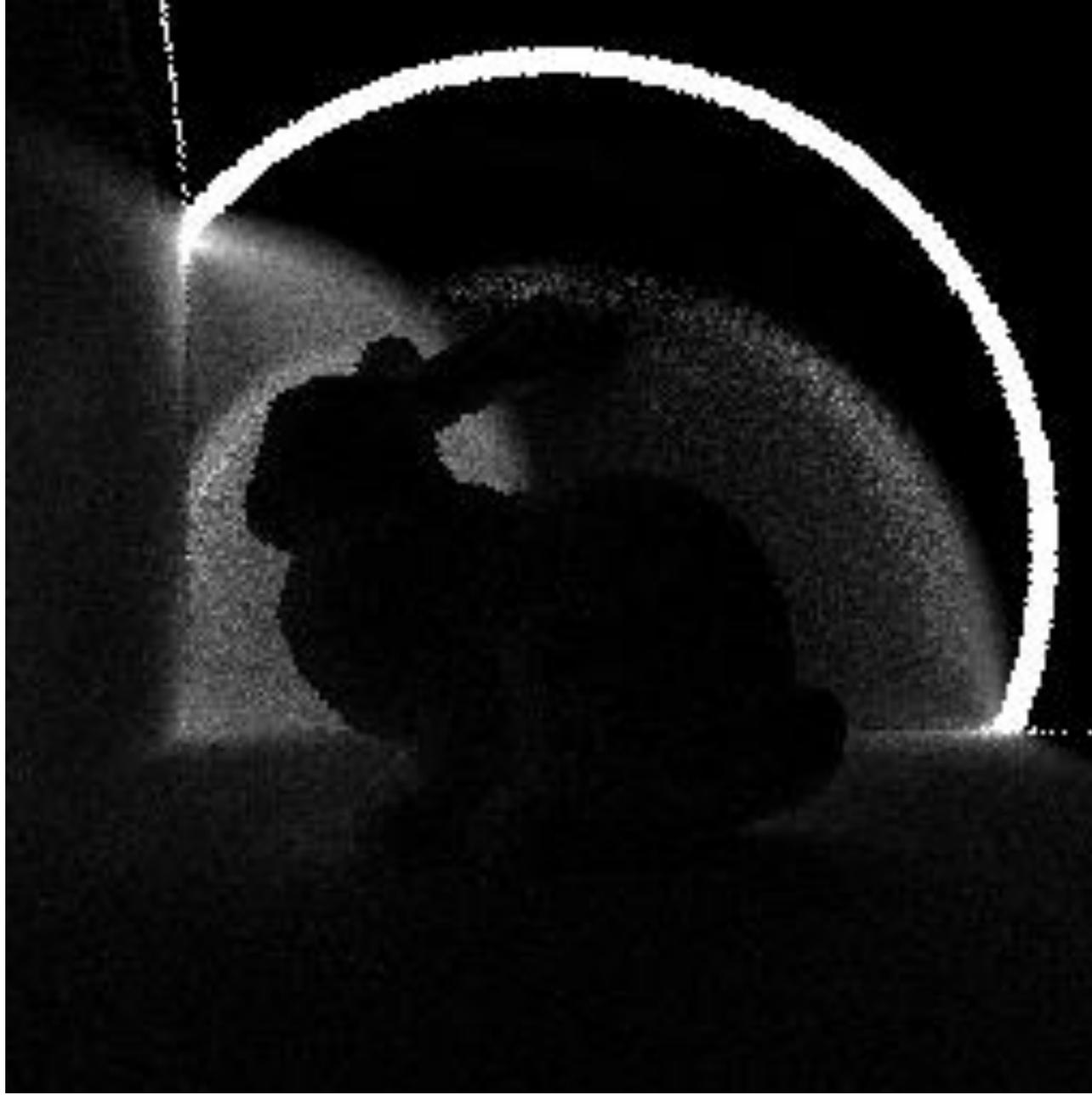
Impulse response



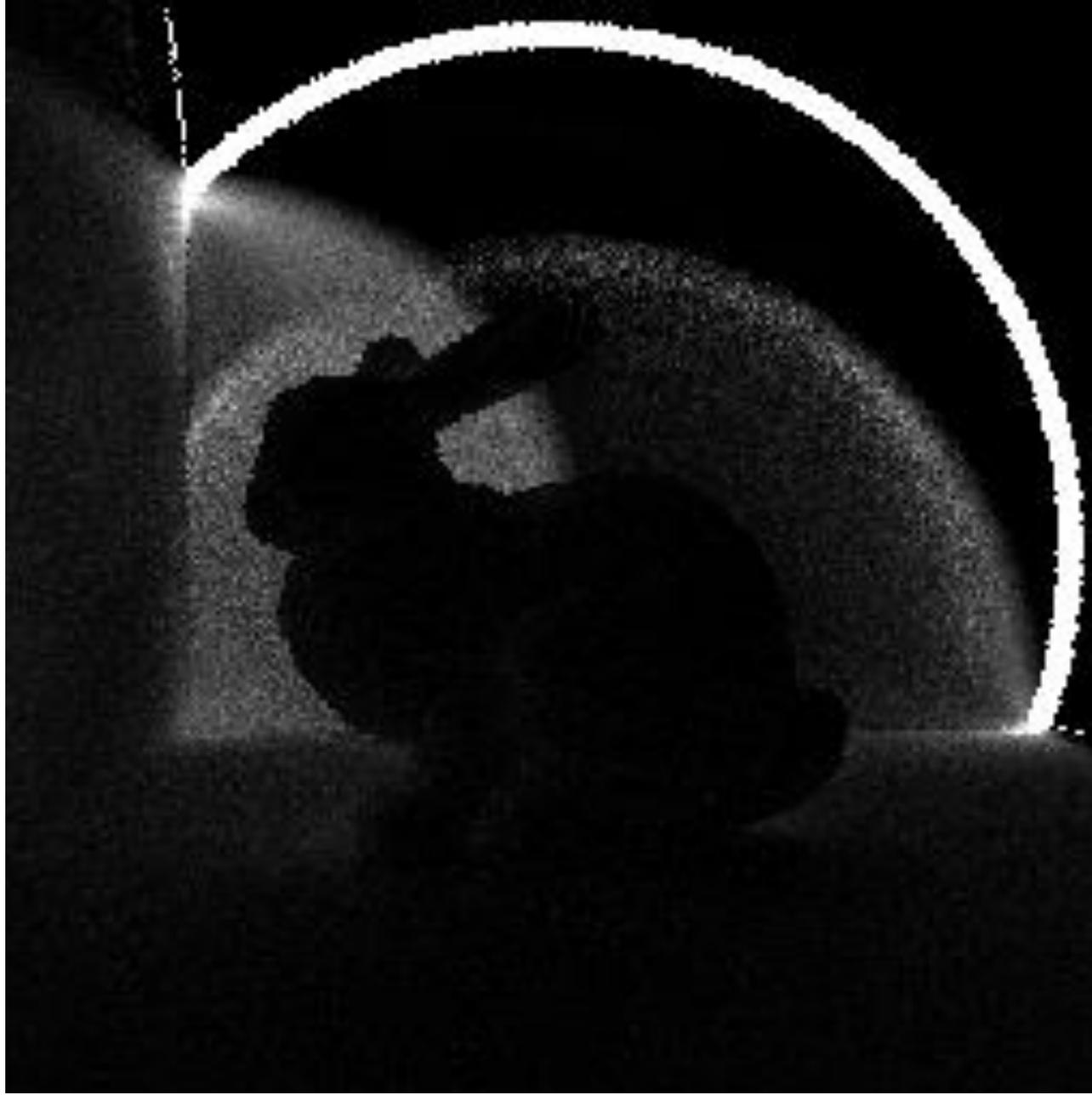
Impulse response



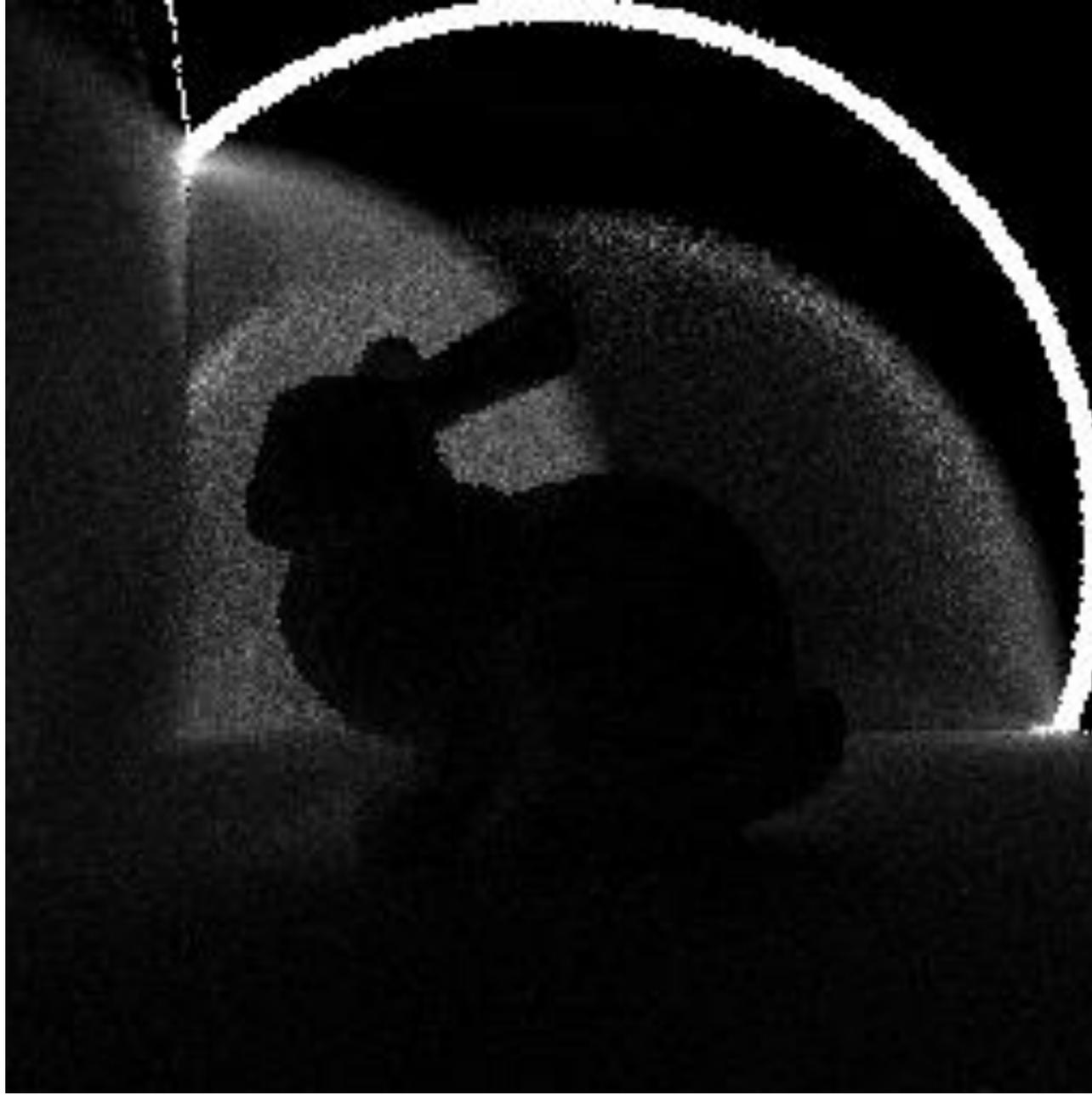
Impulse response



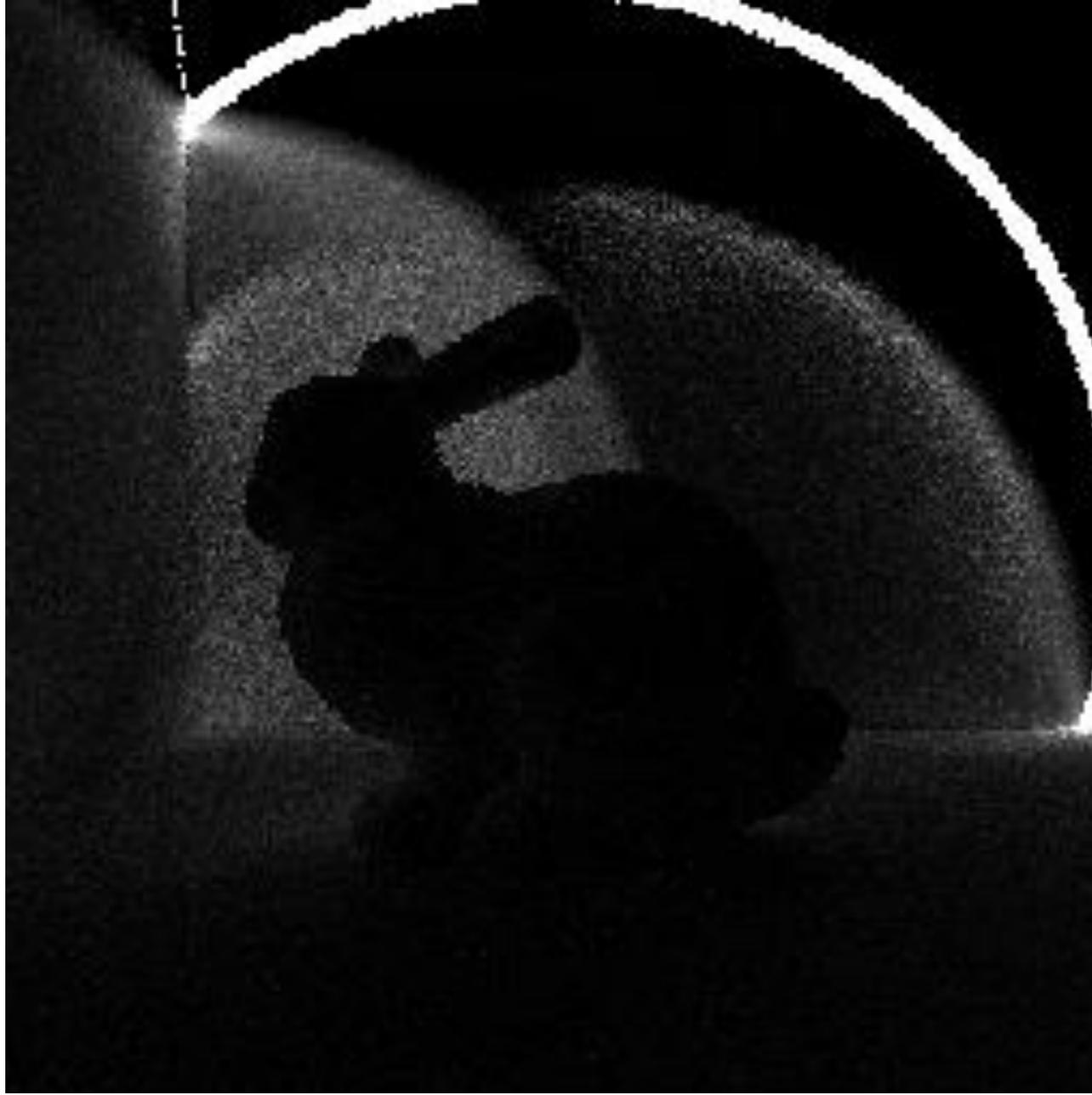
Impulse response



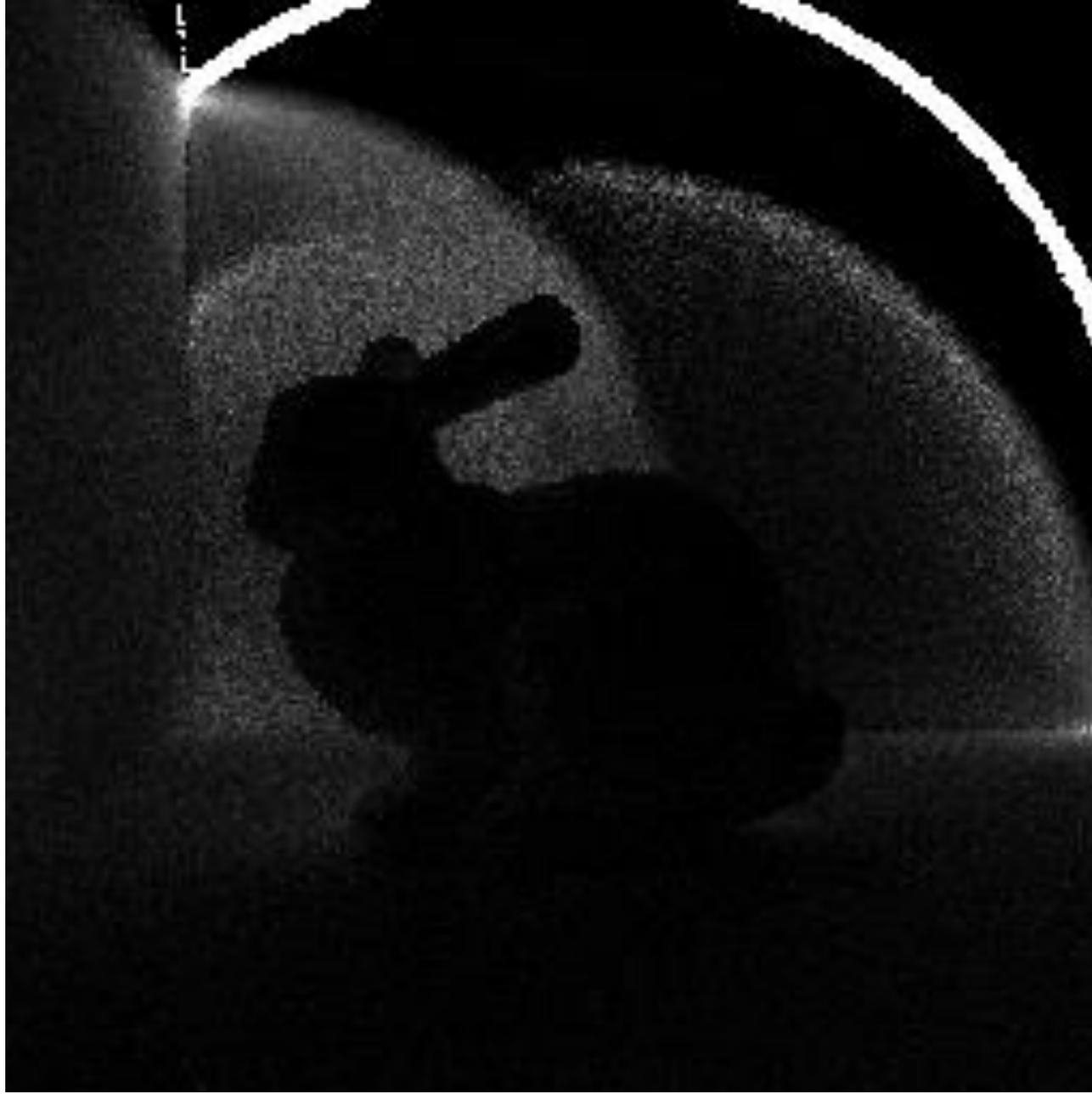
Impulse response



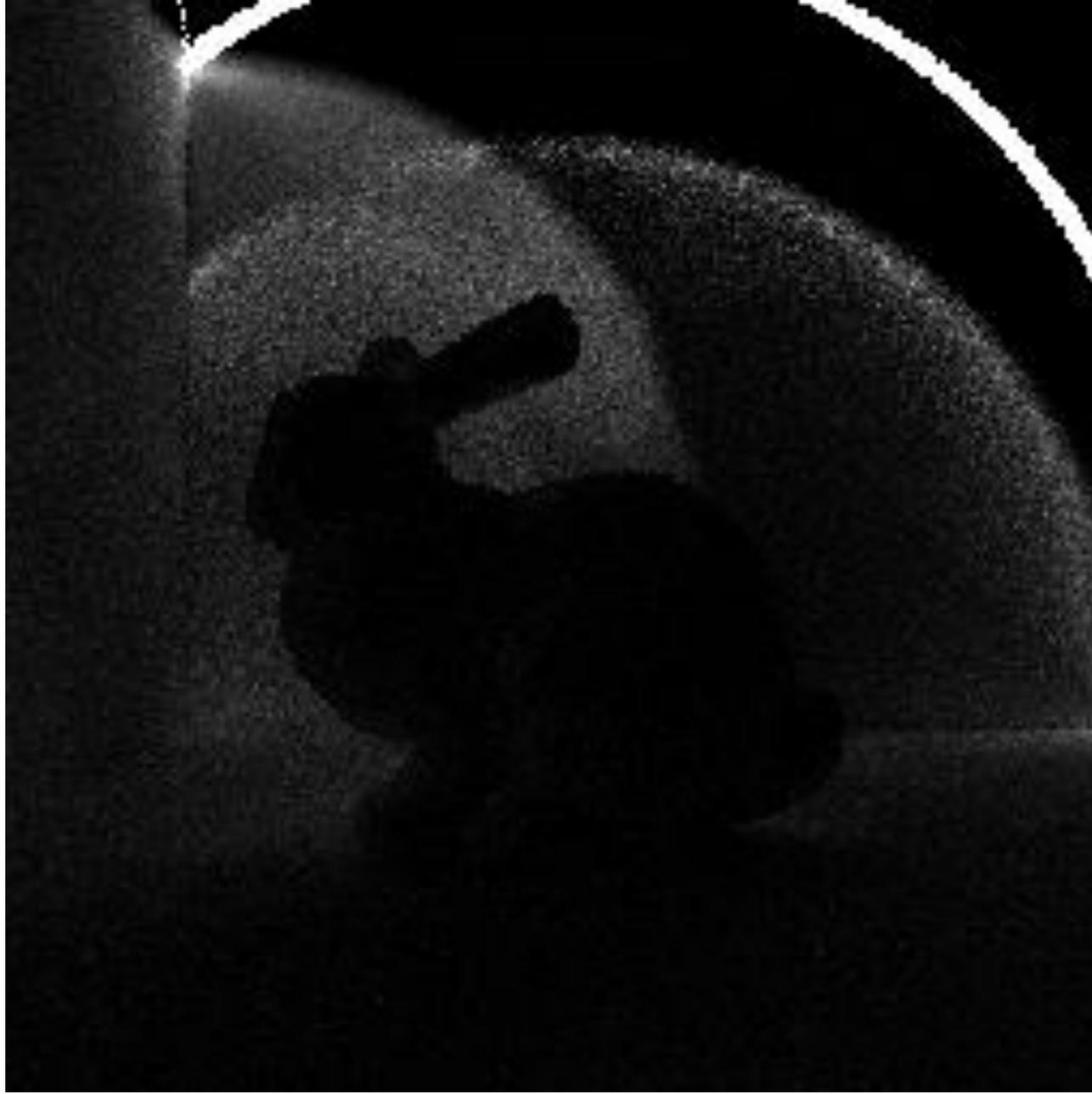
Impulse response



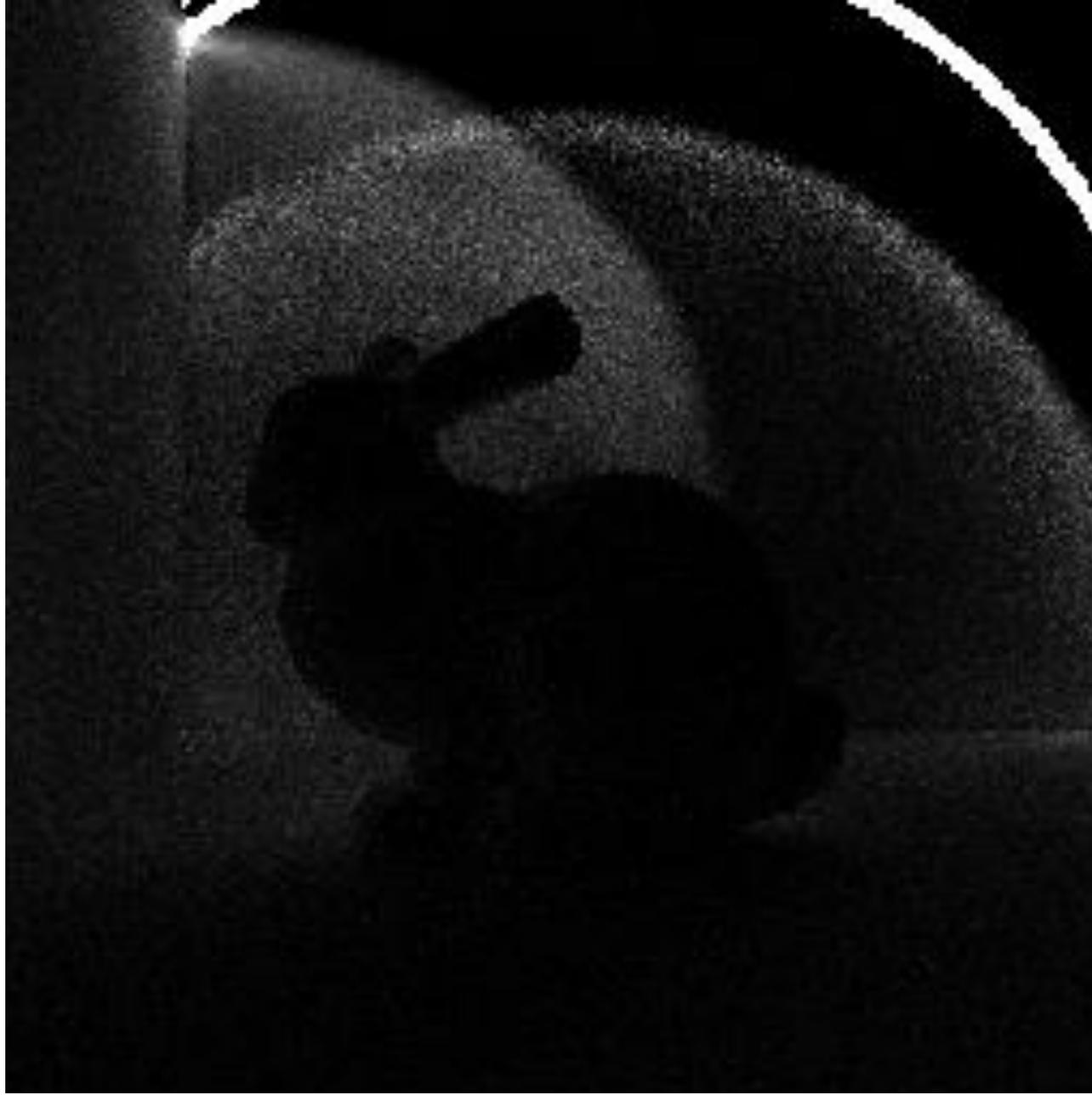
Impulse response



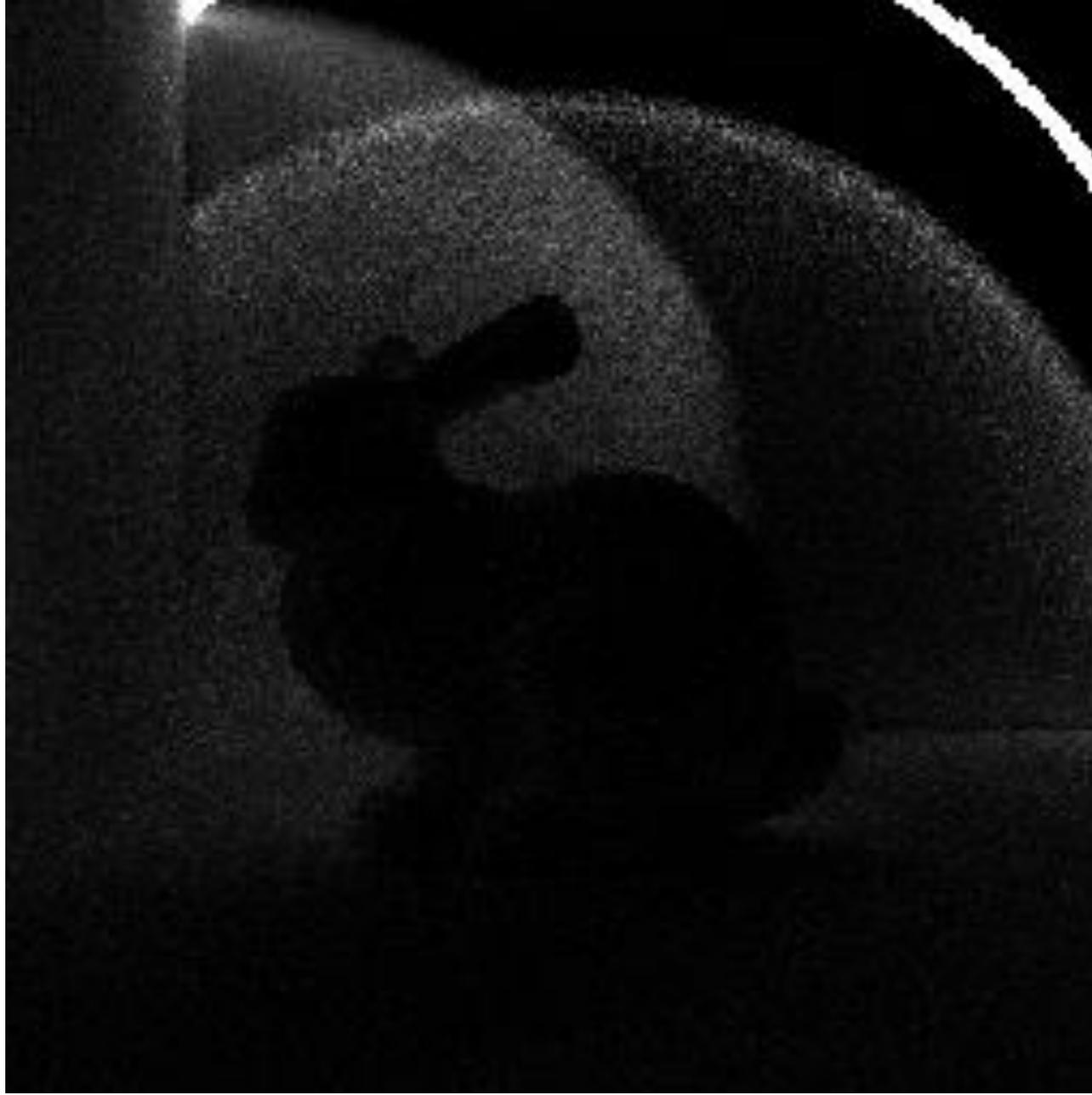
Impulse response



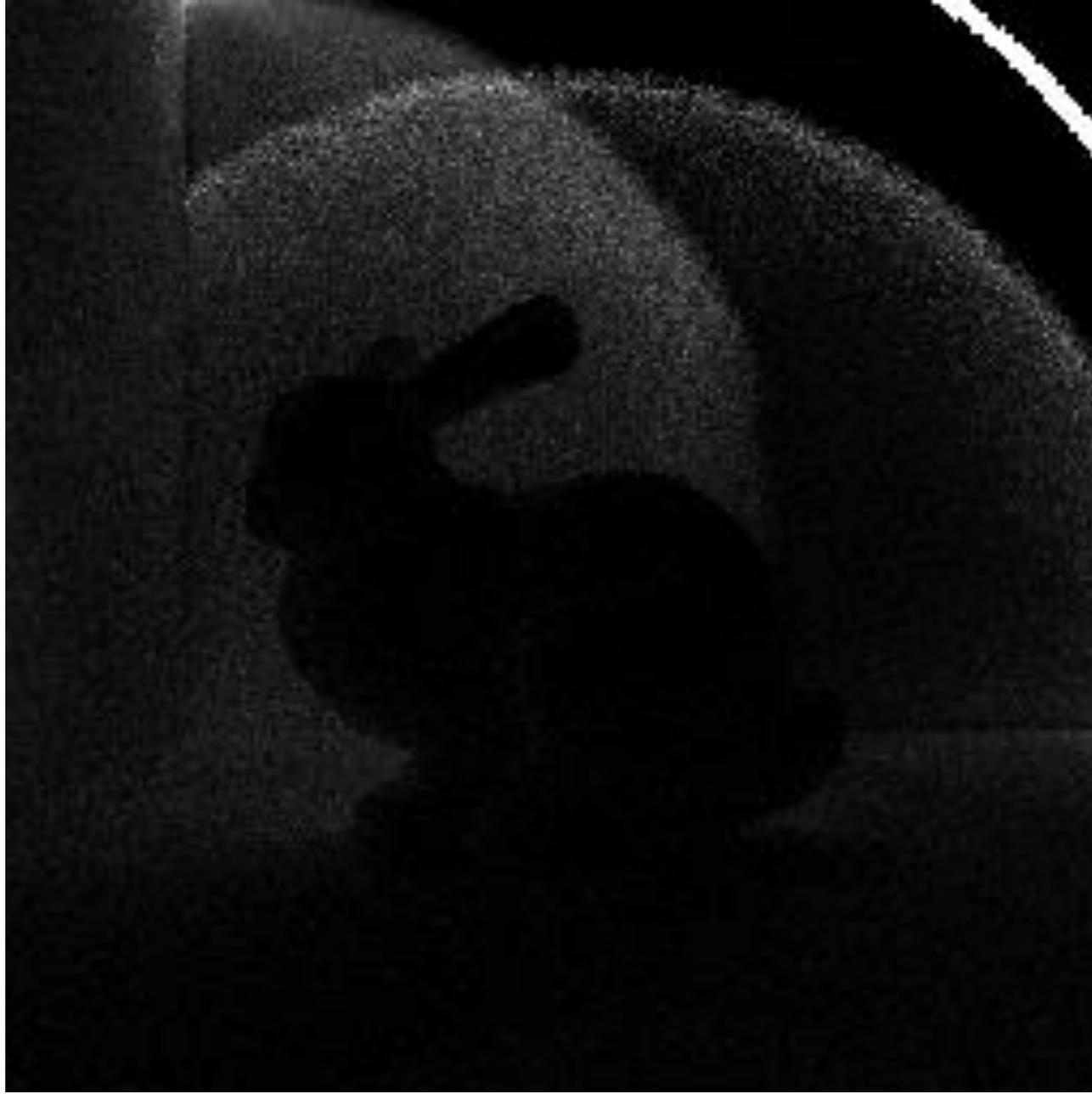
Impulse response



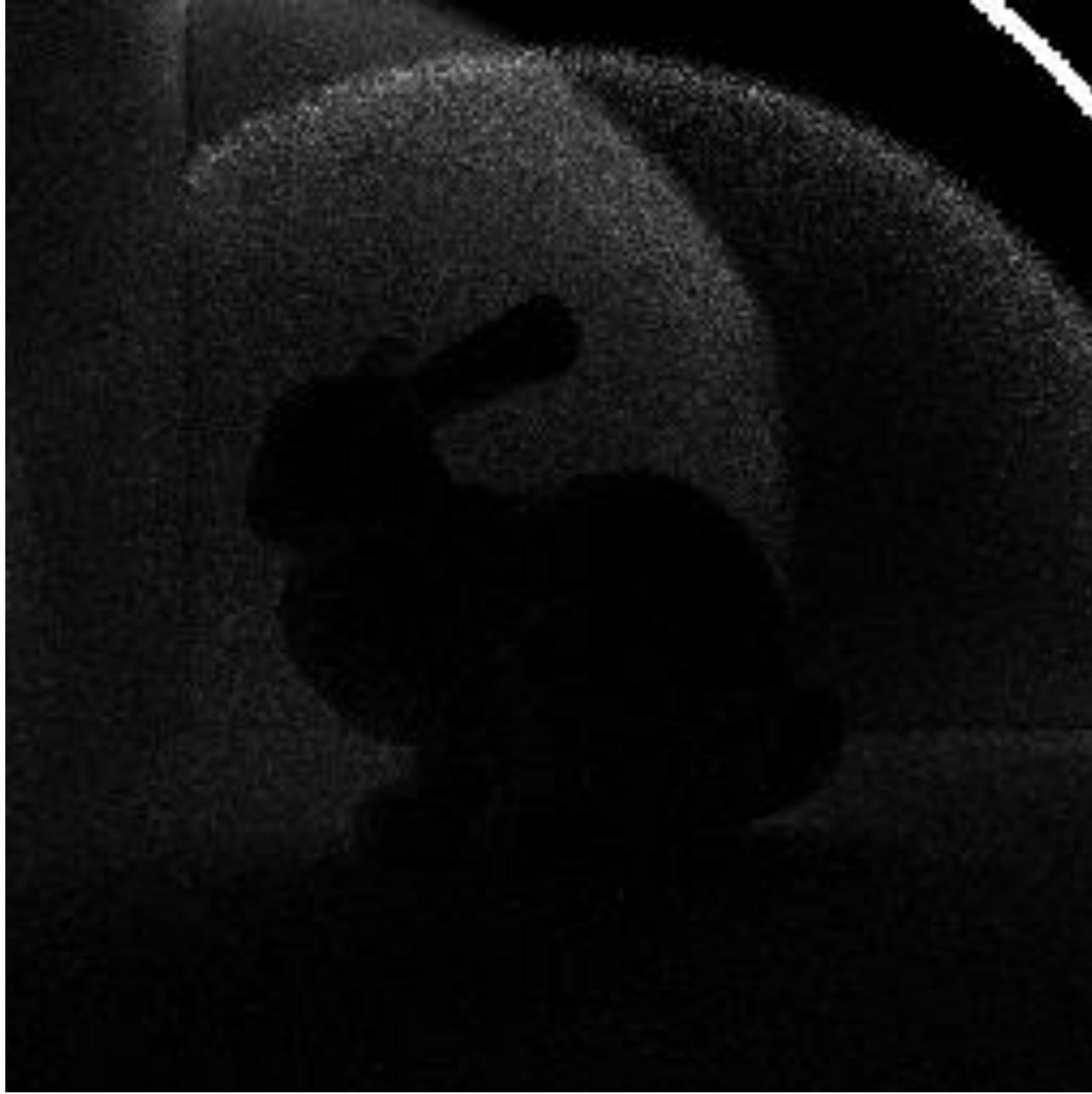
Impulse response



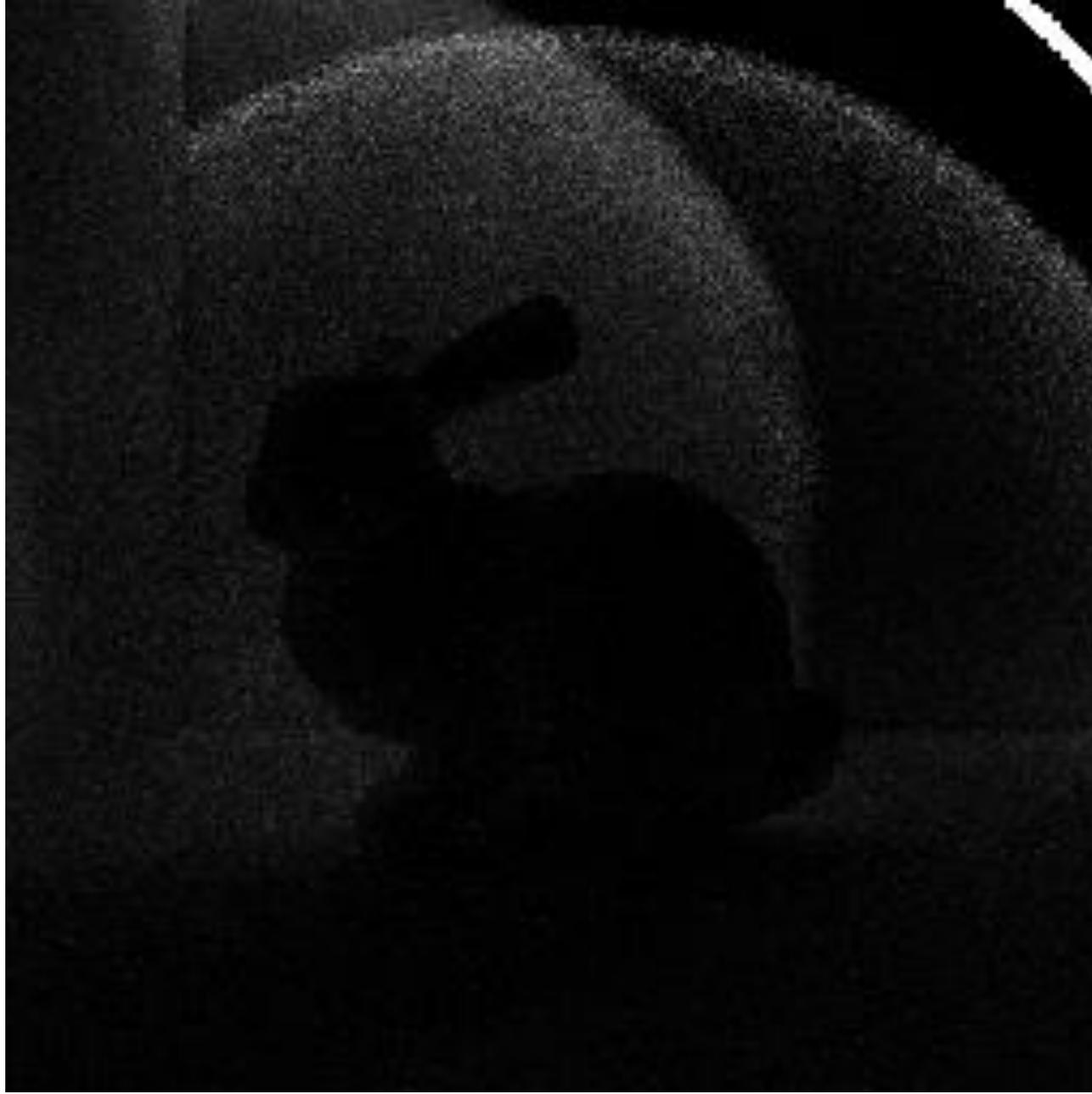
Impulse response



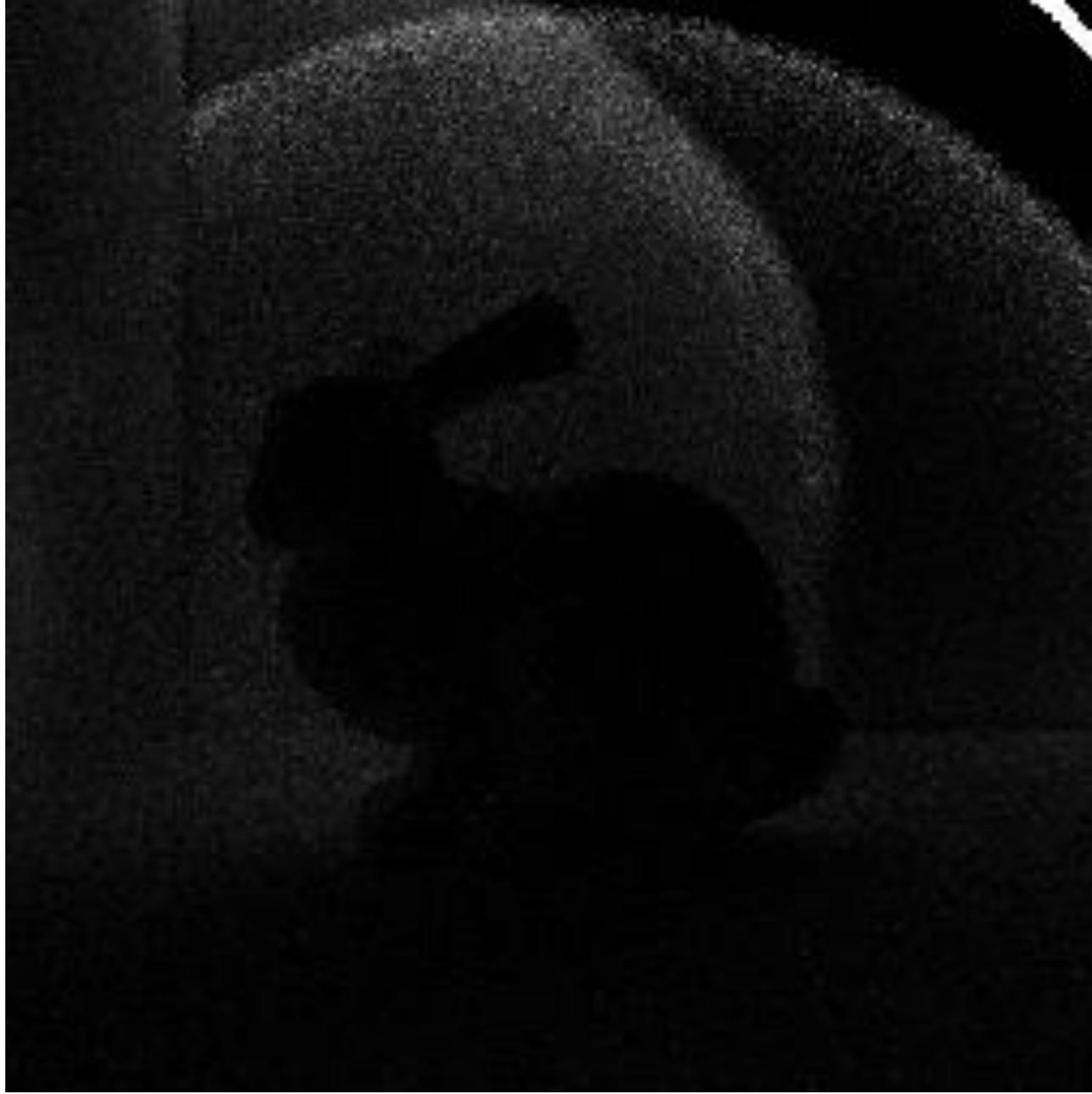
Impulse response



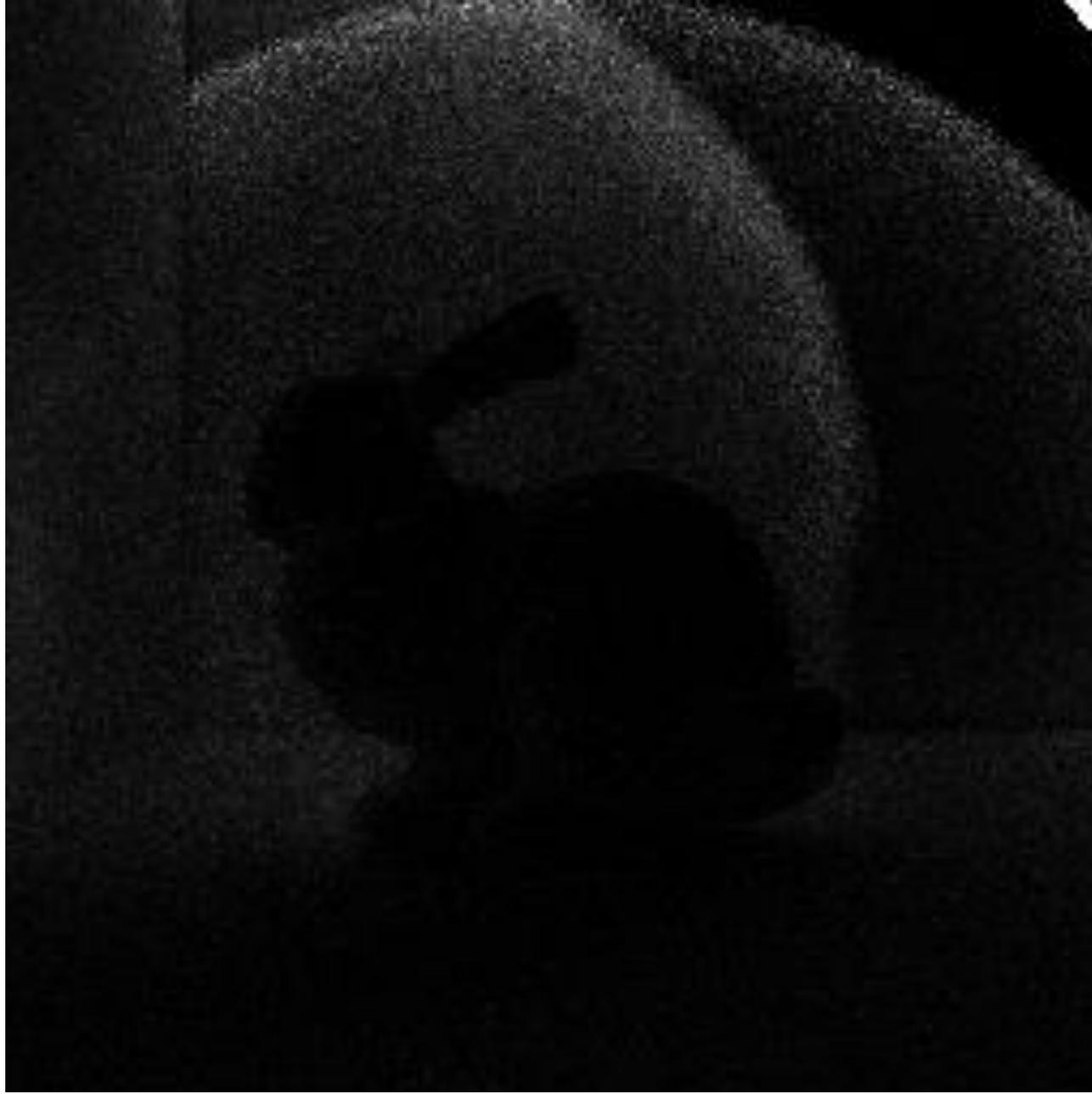
Impulse response



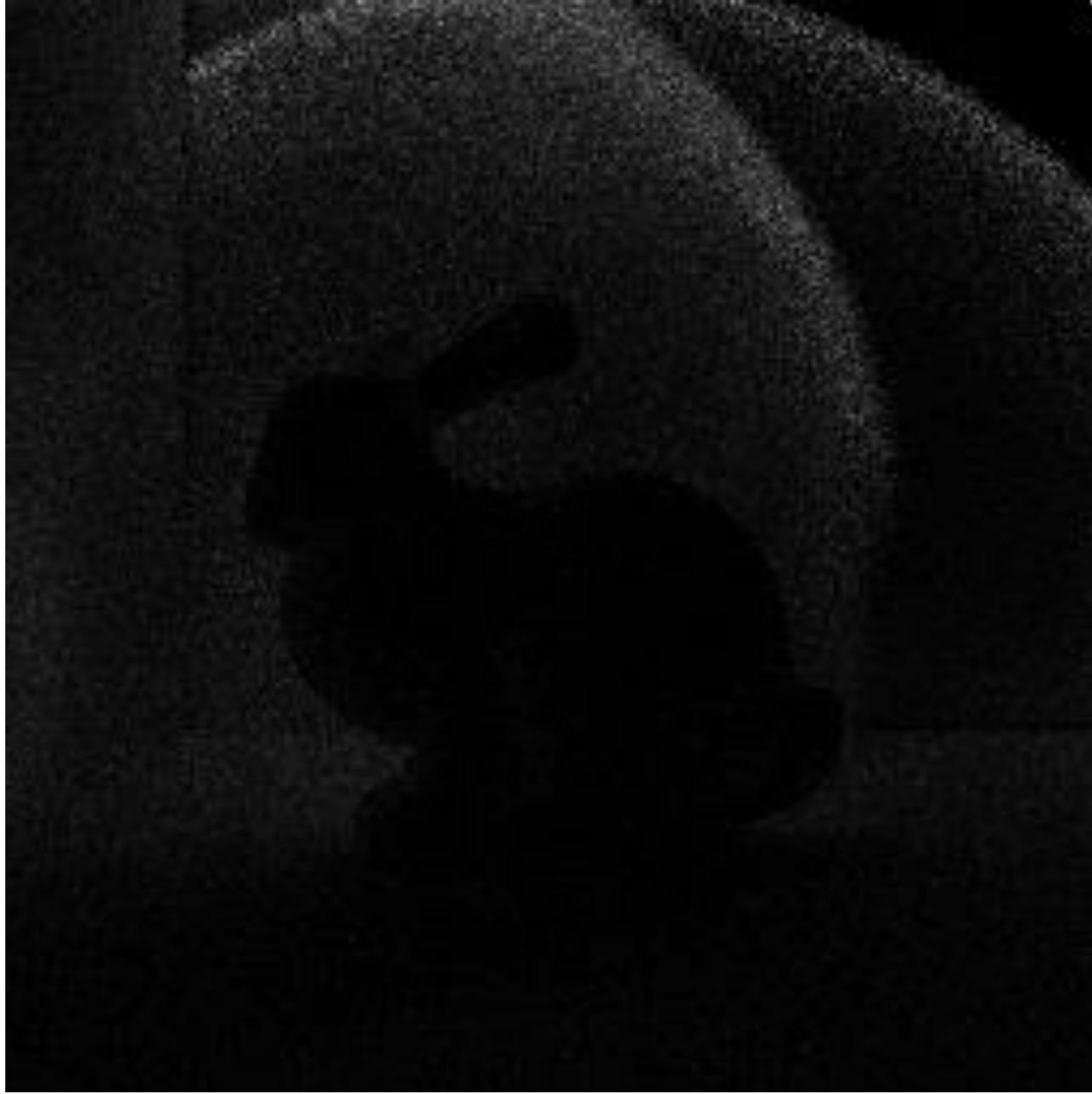
Impulse response



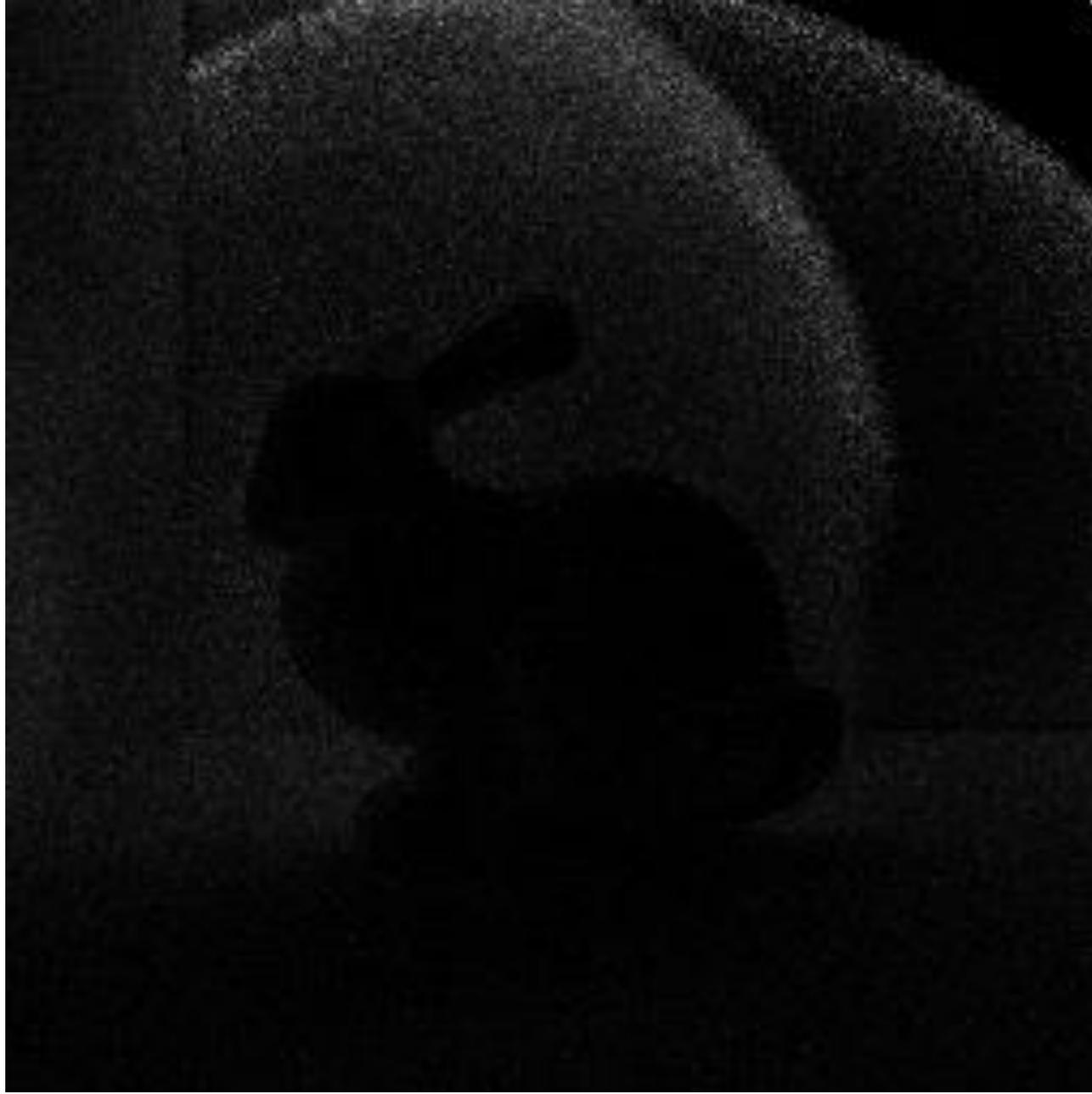
Impulse response



Impulse response



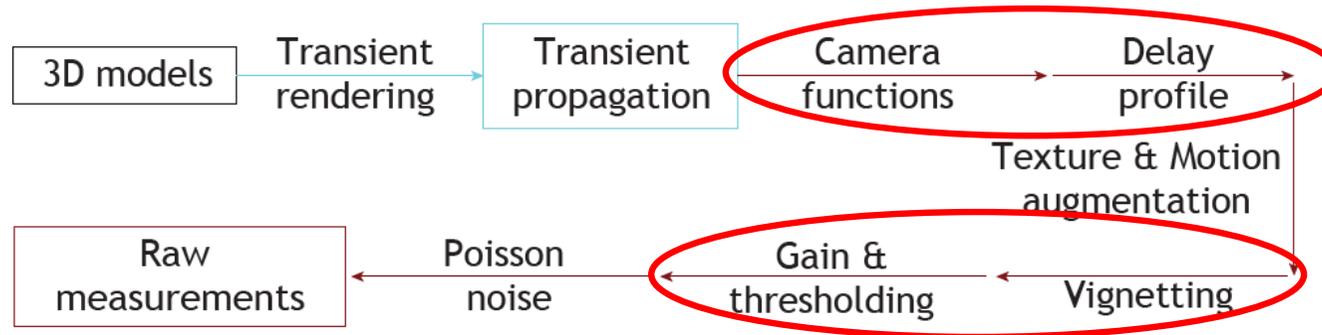
Impulse response



THE FLAT DATASET

Flexible, Large, Augmentable, ToF (FLAT)

Different cameras (beyond Kinect 2) can be simulated, after calibration.



Flexible: Separately models **scenes** and **cameras**

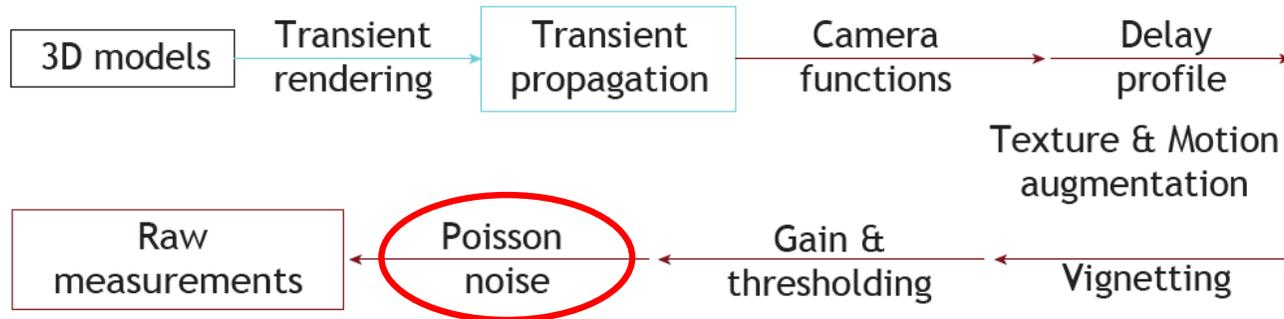
Large: 2061 static and 29 dynamic scenes for now, numbers are increasing

Augmentable: approximate textures and motions can be added on the fly

THE FLAT DATASET

Flexible, Large, Augmentable, ToF (FLAT)

Noise can be added...



Flexible: Separately models **scenes** and **cameras**

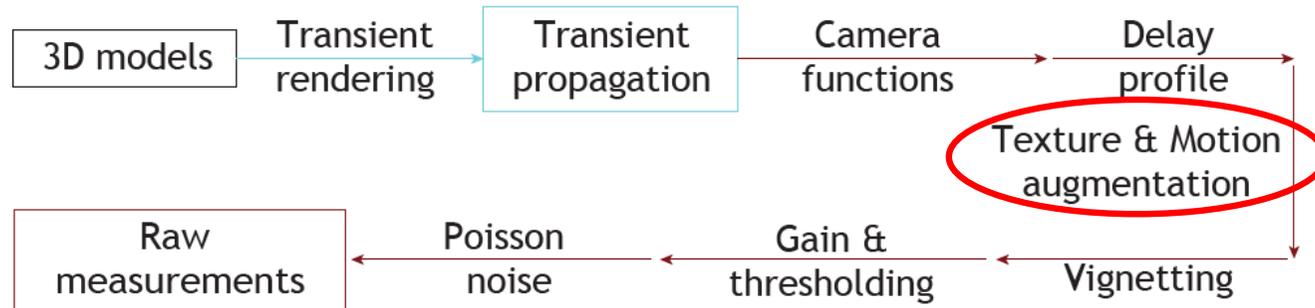
Large: 2061 static and 29 dynamic scenes for now, numbers are increasing

Augmentable: approximate textures and motions can be added on the fly

THE FLAT DATASET

Flexible, Large, Augmentable, ToF (FLAT)

... As well as motion (approximate model) and texture...



Flexible: Separately models **scenes** and **cameras**

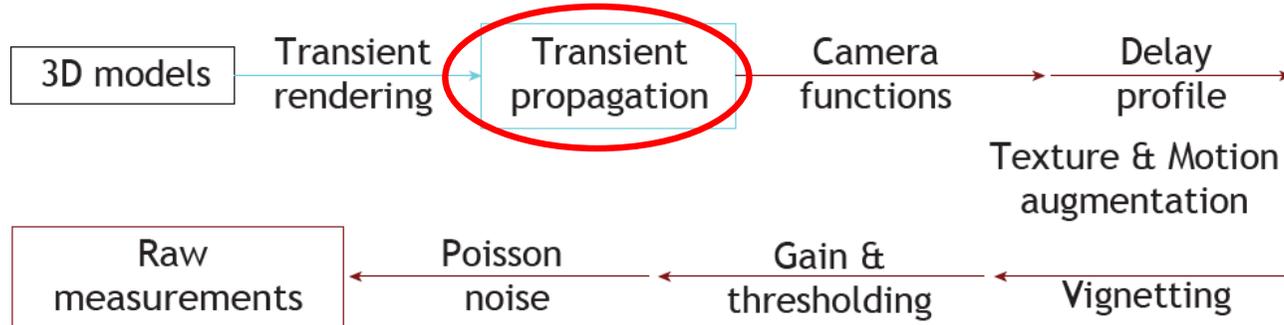
Large: 2061 static and 29 dynamic scenes for now, numbers are increasing

Augmentable: approximate textures and motions can be added on the fly

THE FLAT DATASET

Flexible, Large, Augmentable, ToF (FLAT)

... and multiple reflections.



Flexible: Separately models **scenes** and **cameras**

Large: 2061 static and 29 dynamic scenes for now, numbers are increasing

Augmentable: approximate textures and motions can be added on the fly

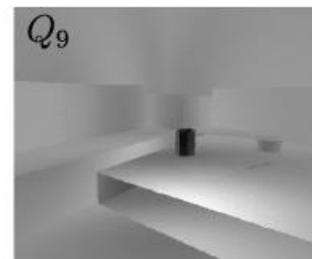
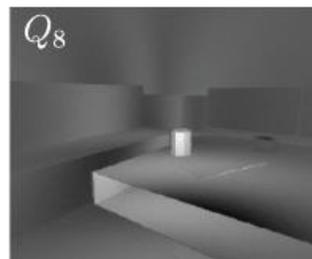
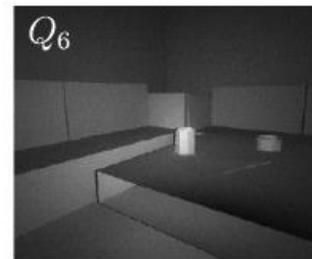
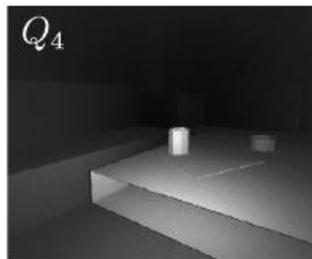
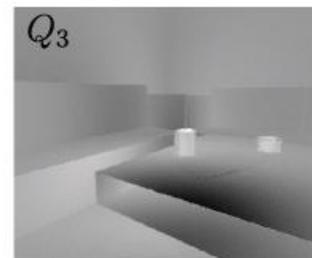
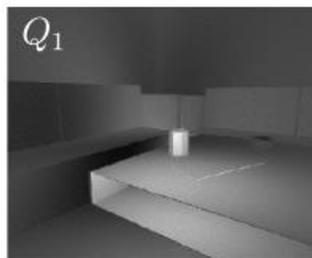
THE FLAT DATASET

Samples

Sample 3D Models



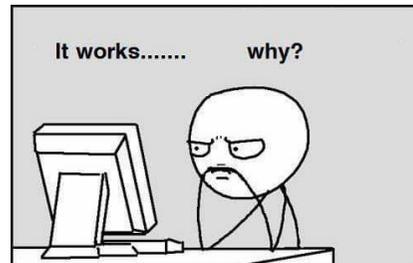
Sample raw measurements (Simulating Kinect 2)



TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*
- Back to physics
- DNN architecture
- Results
- Conclusion



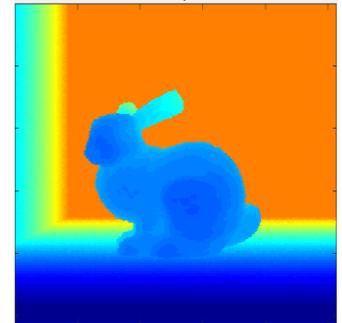
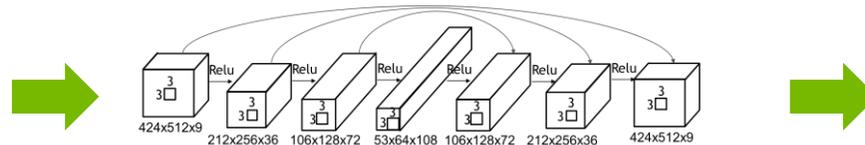
* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

NAÏVE MACHINE LEARNING (ML) FOR TOF RECONSTRUCTION

Supervised learning

Take it easy: supervised learning, from raw data to 3D map.

Training input/output pairs from the FLAT dataset.



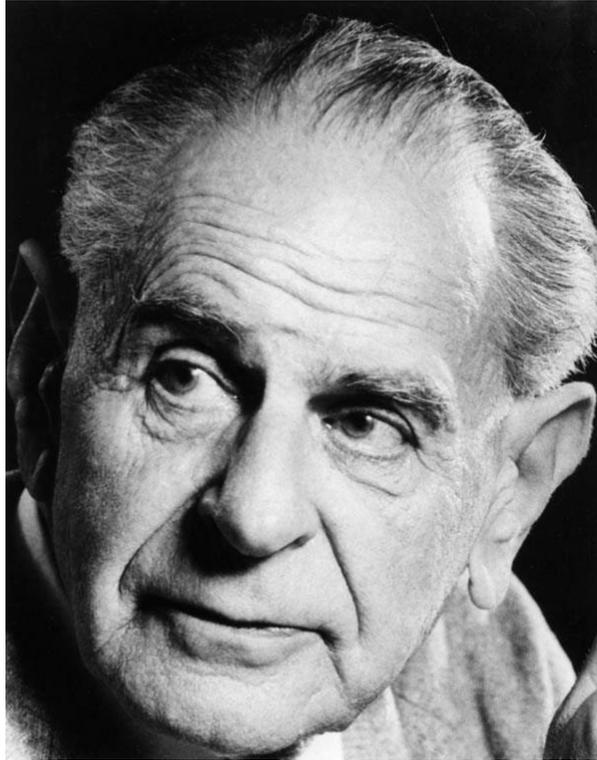
NAÏVE MACHINE LEARNING (ML) FOR TOF RECONSTRUCTION

Supervised learning



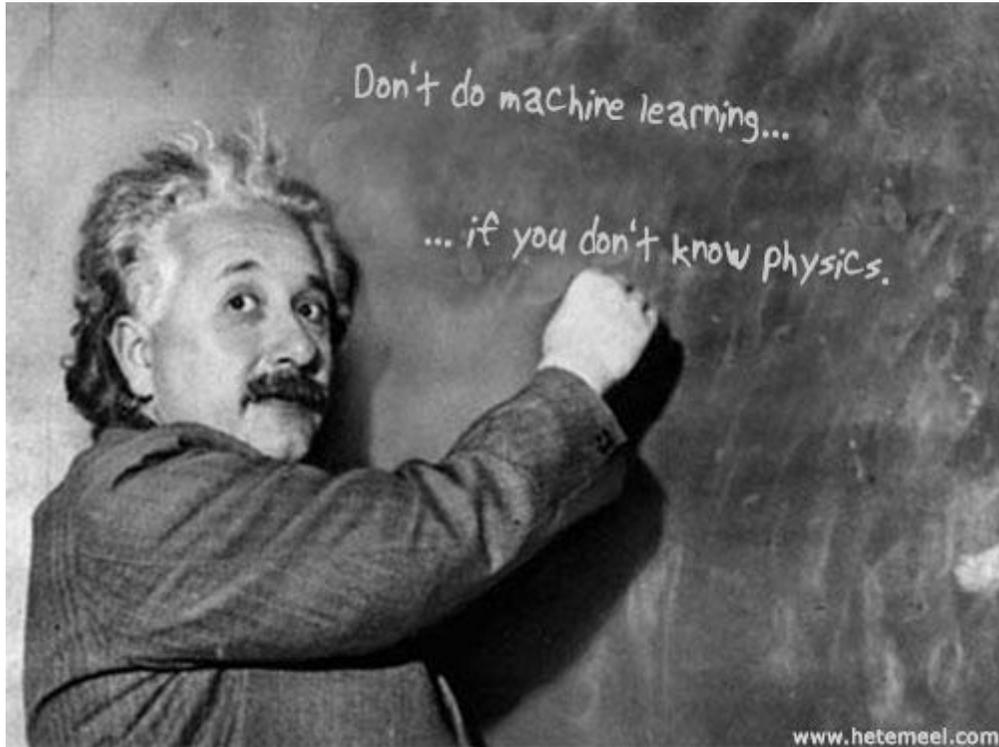
THE LESSON WE LEARNED*...

* To advance science.



That's a nice
negative
result!

... AND HOW WE IMPROVED



*1

*2

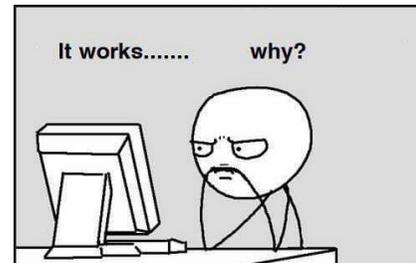
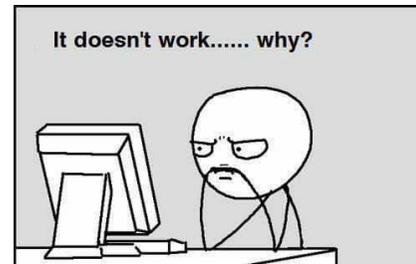
*1 Yes, it's a fake picture...

*2 ... But the message is correct.

TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

- Time Of Flight (TOF) cameras & artifacts
 - Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
 - Spoiler: our non-Naïve ML solution works*
-
- Back to physics
 - DNN architecture
 - Results
 - Conclusion



* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

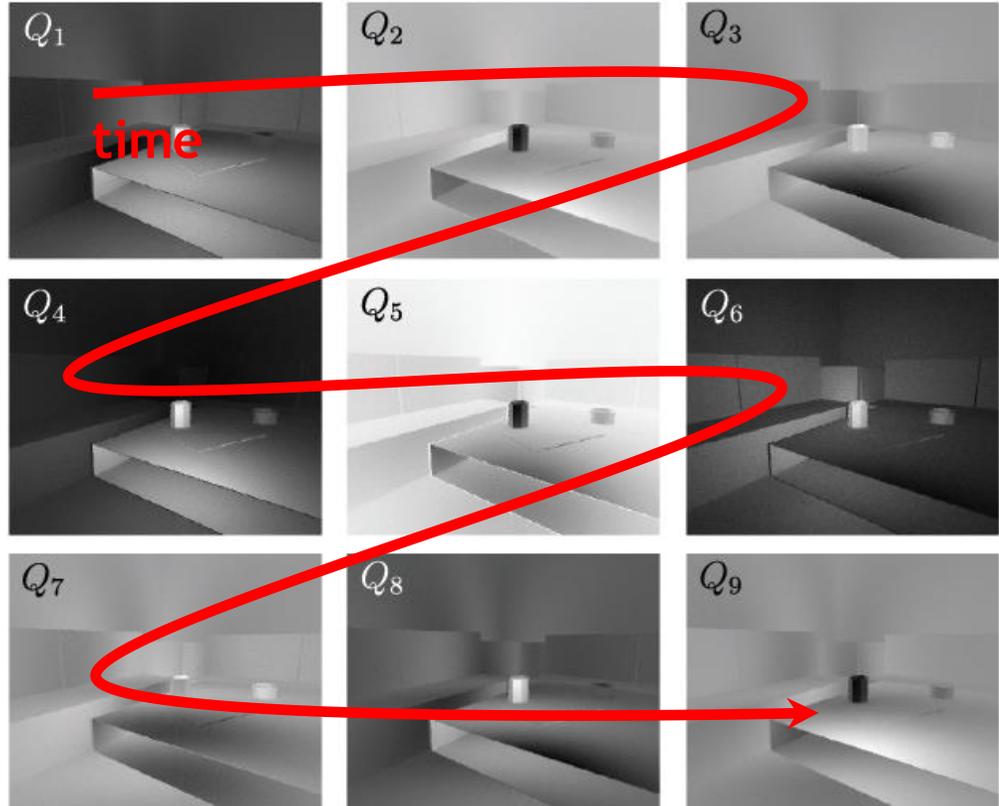
BACK TO PHYSICS

And, more generally speaking, any a-priori knowledge.

Cause: Sequential measurements

Effect: Misaligned moving object

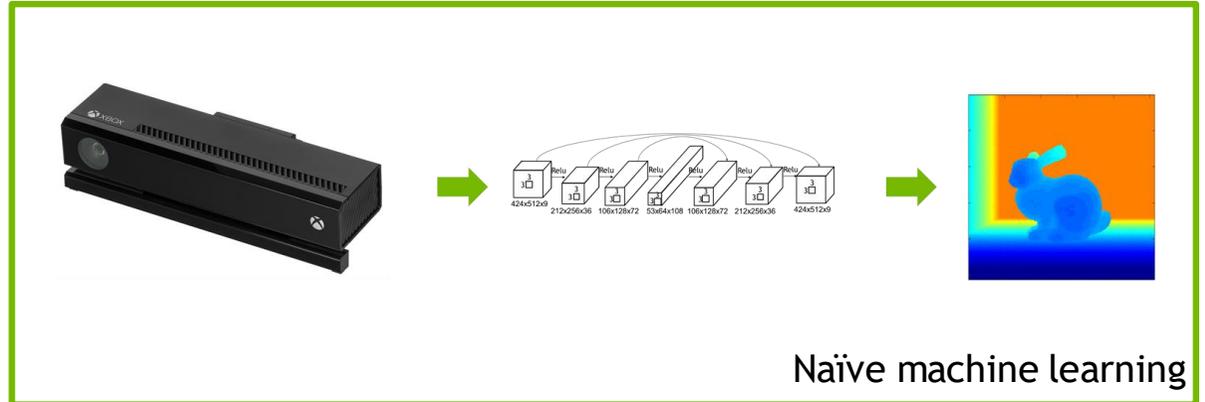
Solution:Warping



BACK TO PHYSICS

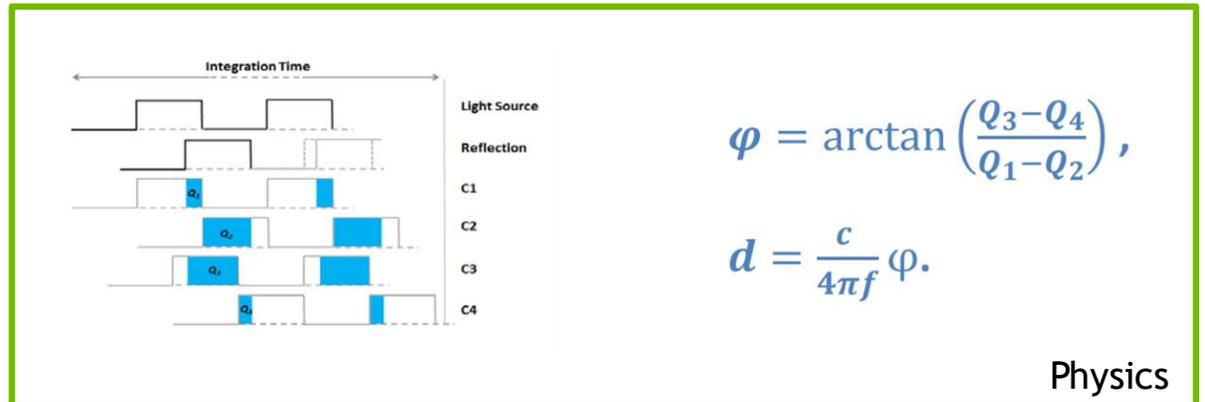
And, more generally speaking, any a-priori knowledge.

Cause: DNN architecture and learning mocks physics



Effect: Sub-optimal results

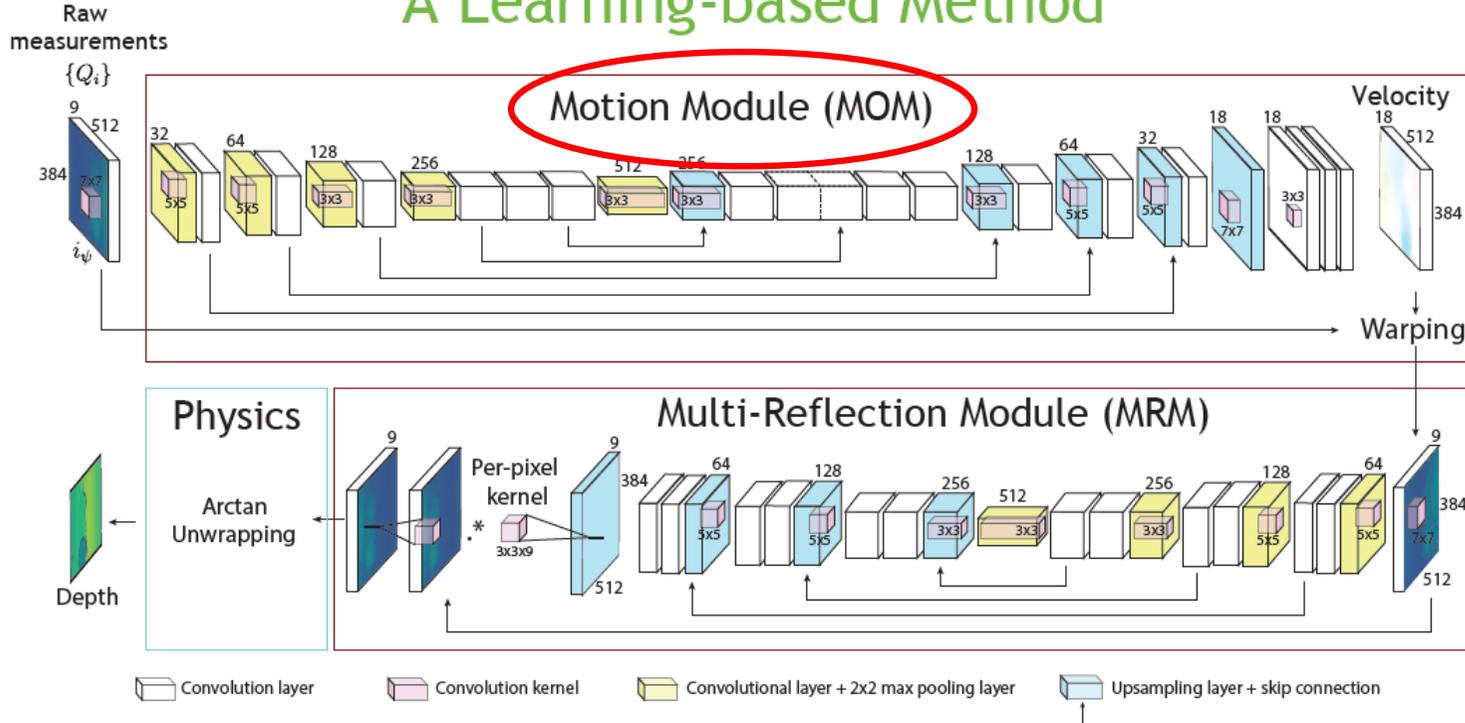
Solution: Include physics in the DNN architecture / reconstruction pipeline.



DNN ARCHITECTURE

#1: Motion Correction Module

A Learning-based Method

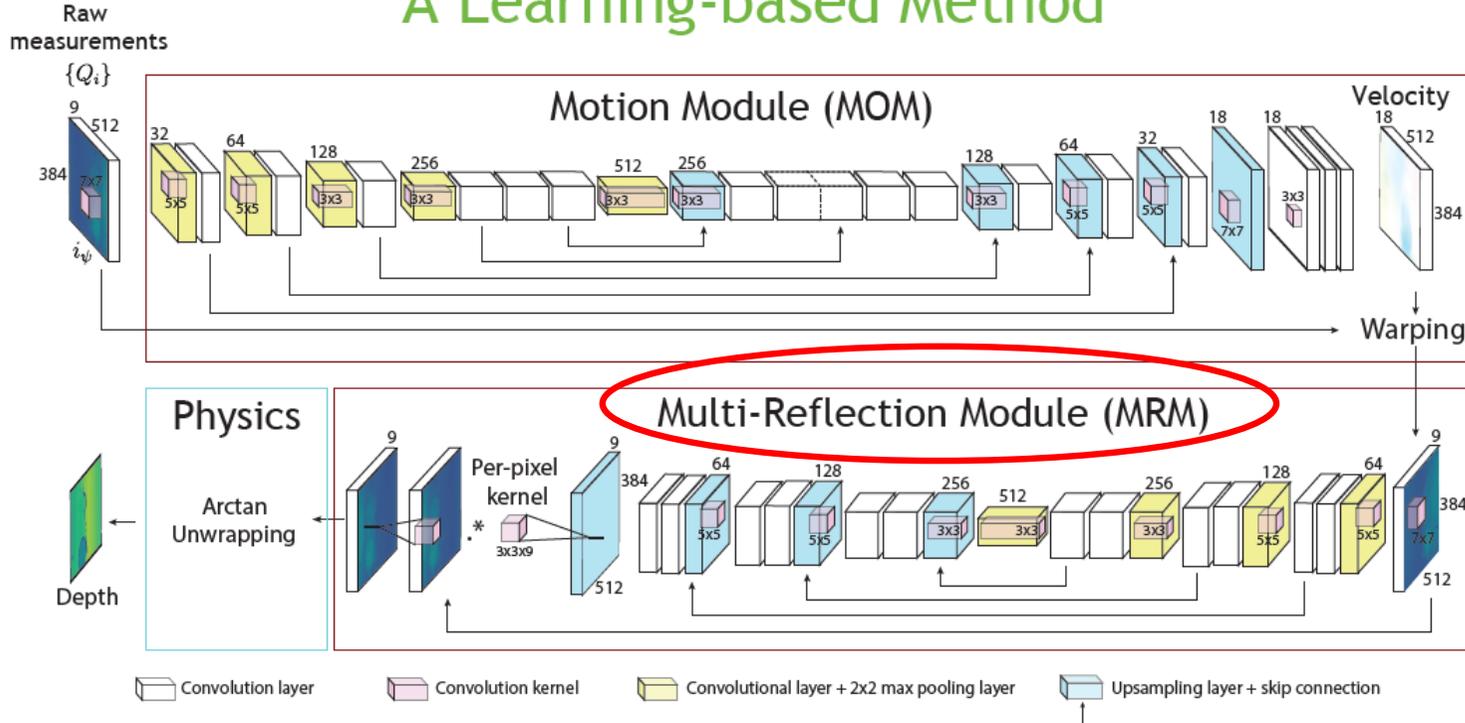


Trained to warp images to the central one

DNN ARCHITECTURE

#2: Motion Reflection Module

A Learning-based Method

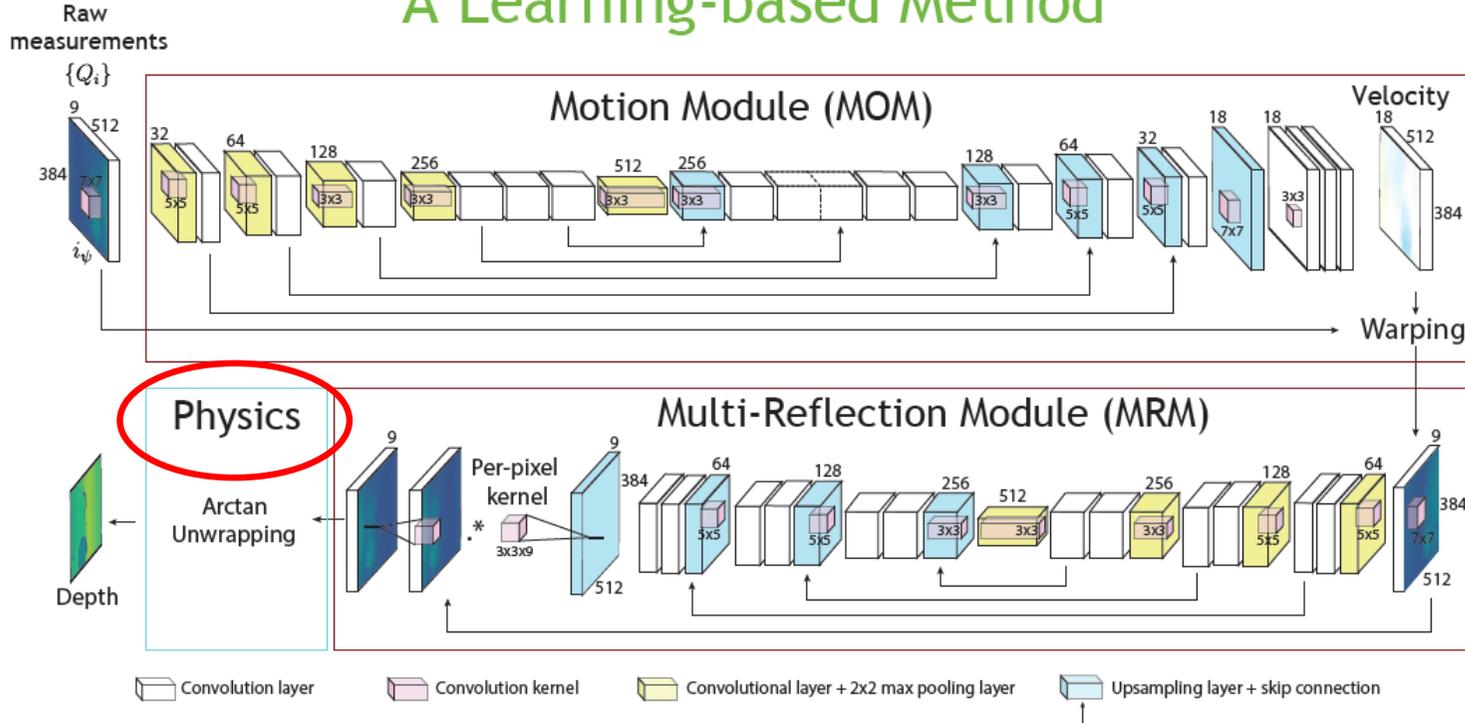


Trained to reduce multiple reflection after re-alignment.

DNN ARCHITECTURE

#3: Differential reconstruction pipeline

A Learning-based Method

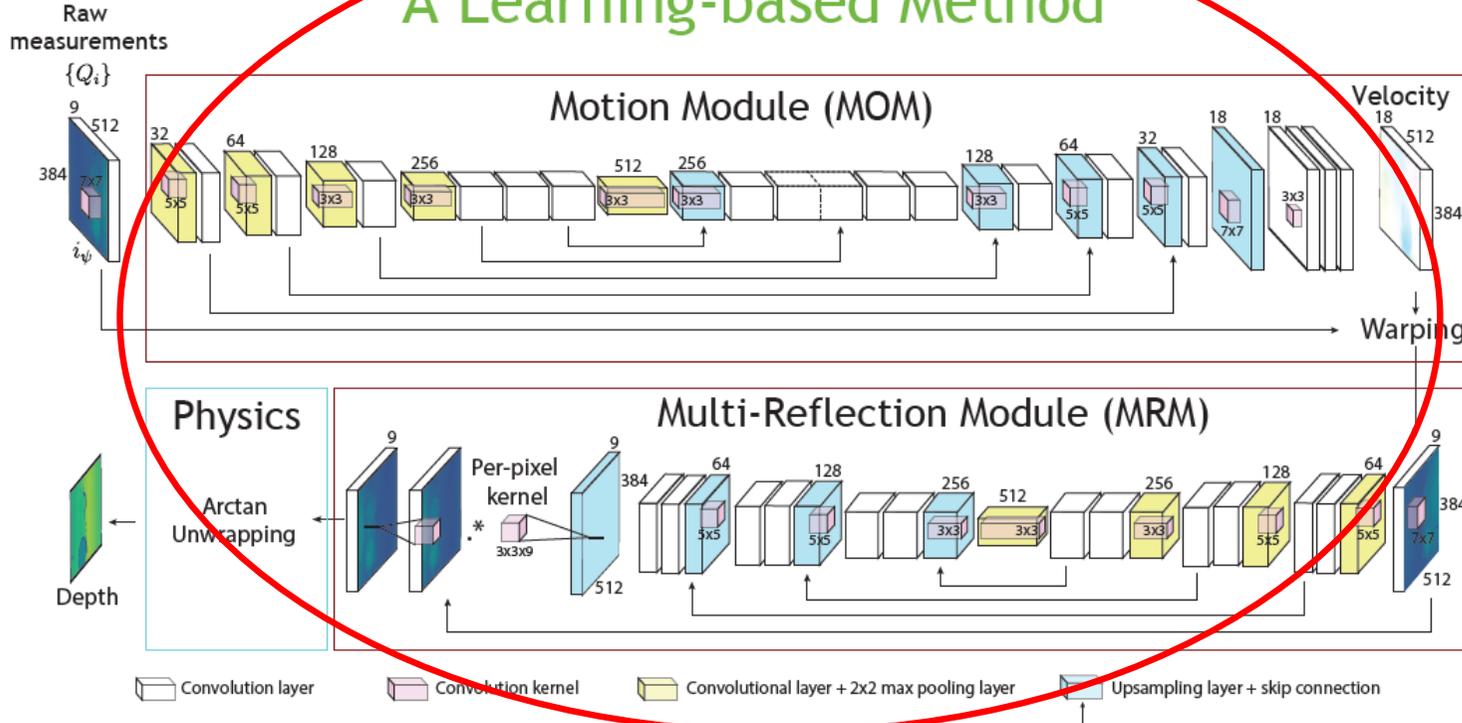


Non-trainable, but differentiable, physics-based reconstruction pipeline.

DNN ARCHITECTURE

#3: Differential reconstruction pipeline

A Learning-based Method



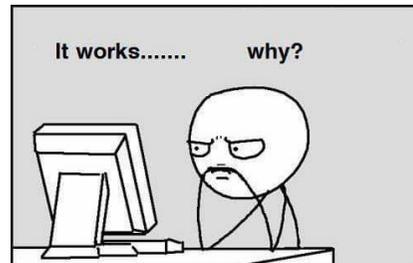
Can be refined with end-to-end training.

TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*

- Back to physics
- DNN architecture
- Results
- Conclusion



* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

RESULTS

Competitors & ablation study

Ablation study

MOM [motion only]

MRM [multiple reflection and noise only]

MOM-MRM [motion, multiple reflection and noise]

Compare against

LF2 [1] [Kinect, non DL]

DToF [3] [DL, Raw to 3D, no motion]

Phasor [2] [High frequencies reduce MPI]

[1] Xiang, et al. libfreenect2: Release 0.2

[2] Marco, et al. DeepToF: Off-the-shelf real-time correction of multipath interference in time-of-flight imaging. In: ACM Transactions on Graphics (SIGGRAPH ASIA).

[3] Gupta, et al. Phasor imaging: A generalization of correlation-based time-of-flight imaging. ACM Transactions on Graphics.

RESULTS

Ablation study: none, MRM, MOM+MRM [simulation]

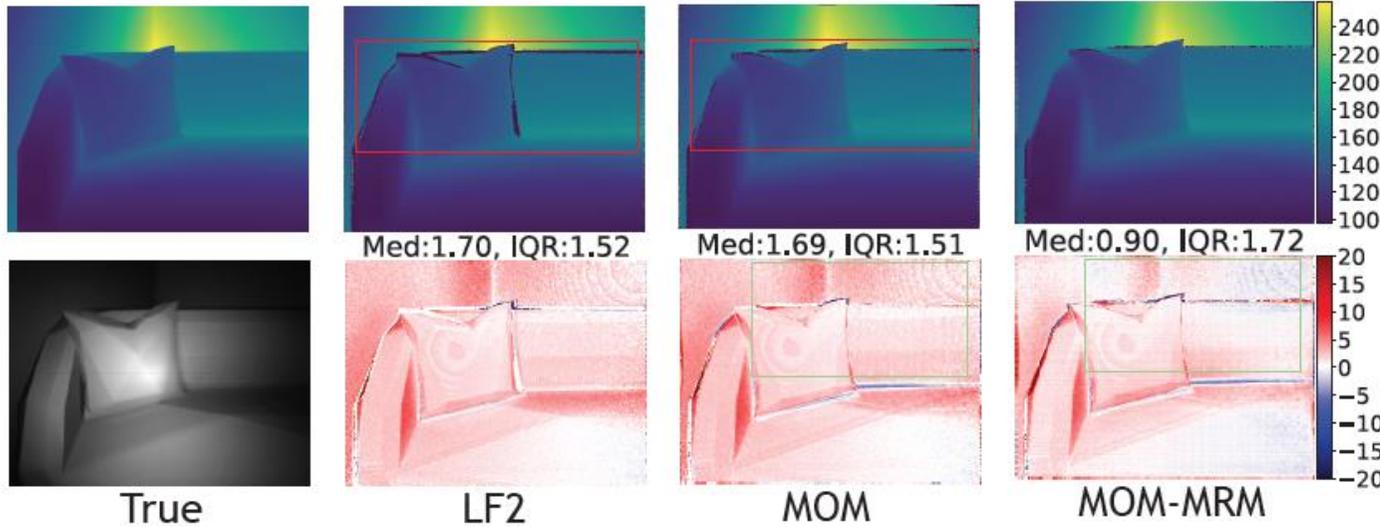
Motion, multi-reflection and shot noise, synthetic



Median [Med] and Inter Quartile Range [IQR] of the error decreased by MRM / MOM-MRM, in cm.

RESULTS

Ablation study: none, MRM, MOM+MRM [simulation]

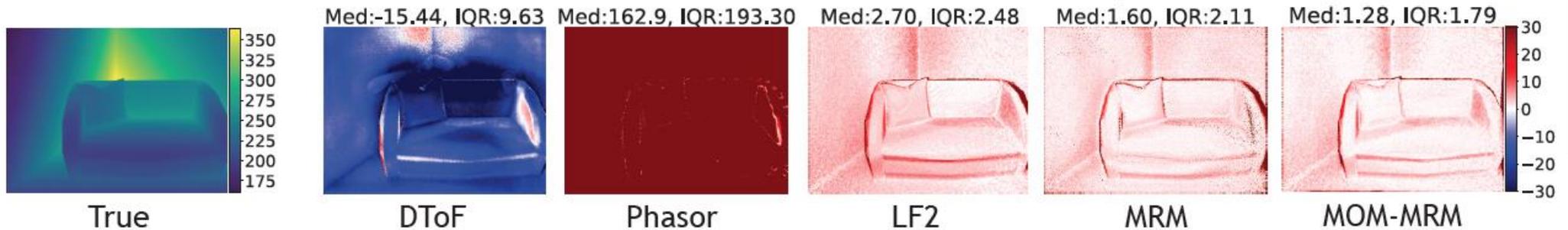


MOM aligns object boundaries and allows a more dense reconstruction (red boxes). MRM mostly corrects multi-reflection artifacts in the smooth areas (green boxes).

RESULTS

Compare against: DToF, Phasor imaging [simulation]

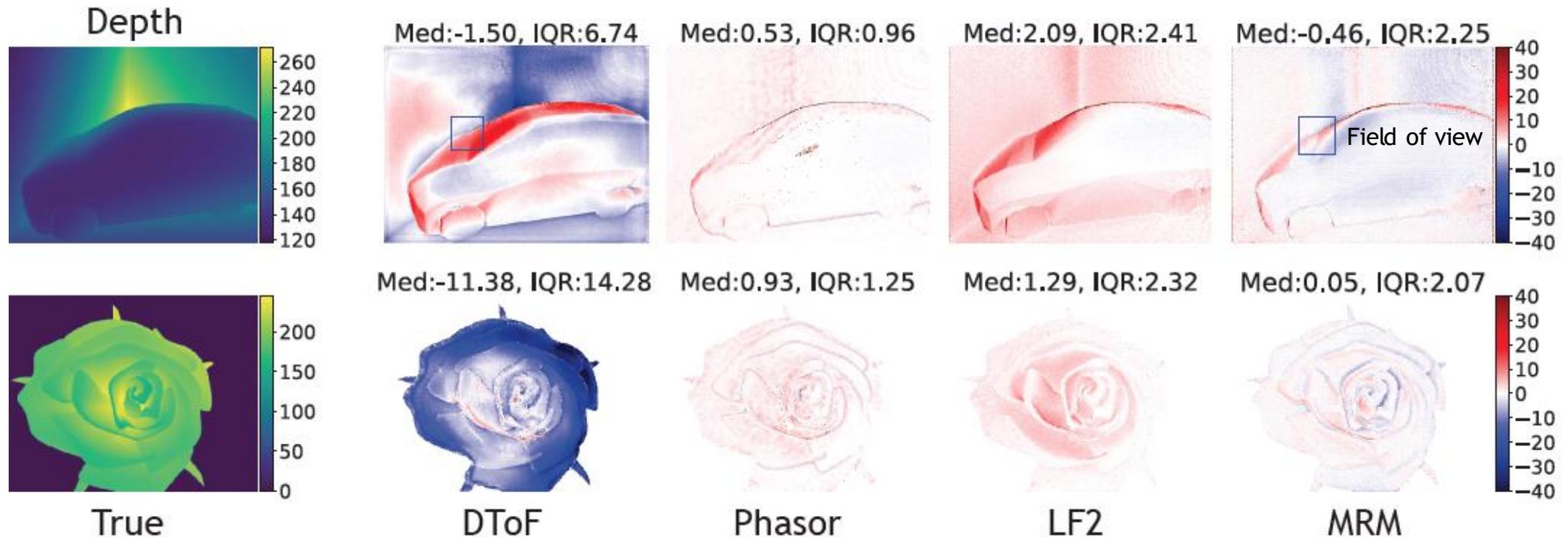
Motion, multi-reflection and shot noise, synthetic



Smaller error when compared to DToF or Phasor.

RESULTS

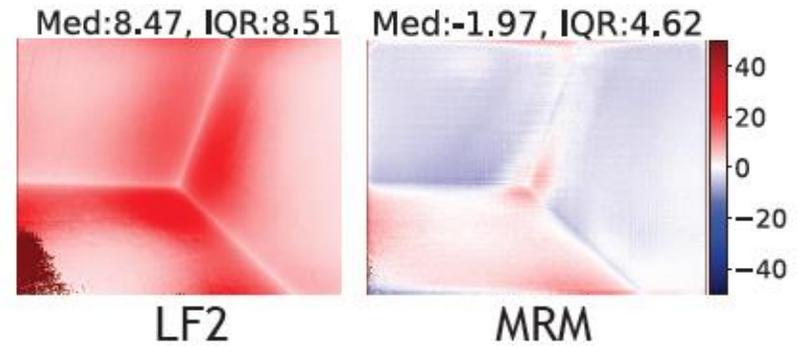
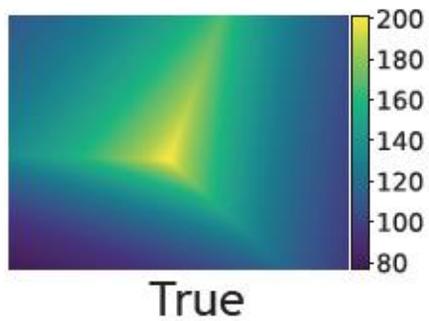
Compare against: DTOF, Phasor imaging on multi-reflection and shot noise [simulation]



Multiple reflection removed through local reflection / a-priori information, no bias.

RESULTS

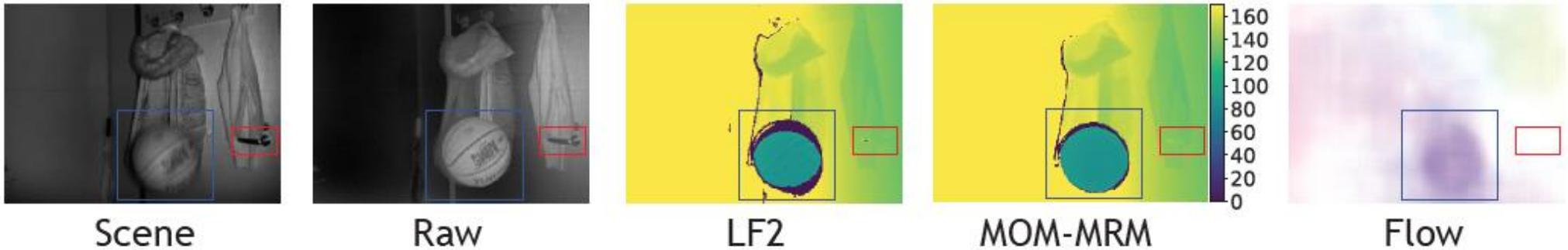
Compare against: LF2, on multi-reflection and shot noise [real data]



Multiple reflection removed through local reflection / coherence / a-priori information, no bias.

RESULTS

Compare against: LF2, on movement [real data]

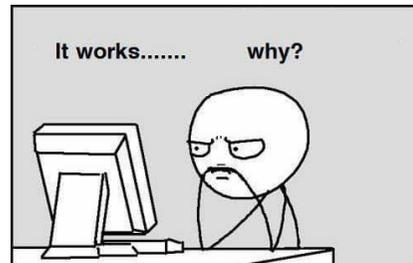


Realignment of raw data reduce motion artifacts, specular reflections (red box) generate errors.

TACKLING 3D TOF ARTIFACTS THROUGH LEARNING AND THE FLAT DATASET

Agenda

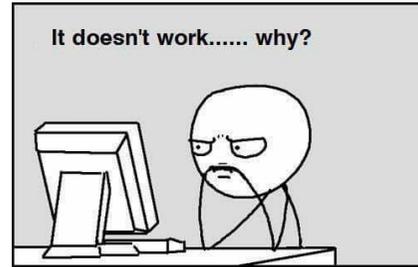
- Time Of Flight (TOF) cameras & artifacts
- Naïve Machine Learning (ML) for TOF reconstruction
 - TOF cameras: working principles
 - Camera calibration
 - The FLAT dataset
- Spoiler: our non-Naïve ML solution works*
- Back to physics
- DNN architecture
- Results
- Conclusion



* See Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz, *Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset*, ECCV 2018, Munich (Germany), Sept. 2018.

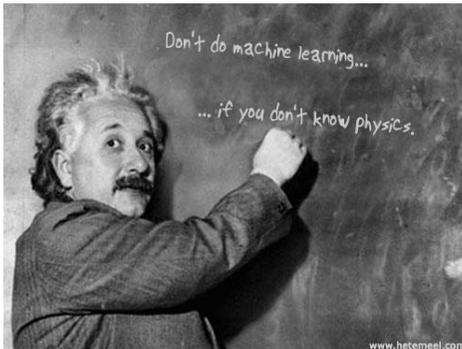
CONCLUSION

1. Naïve ML does not always work...



CONCLUSION

1. Naïve ML does not always work...



2.But going back to a priori knowledge may help.

CONCLUSION

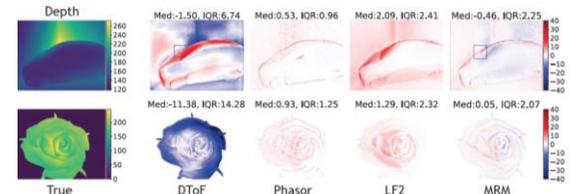
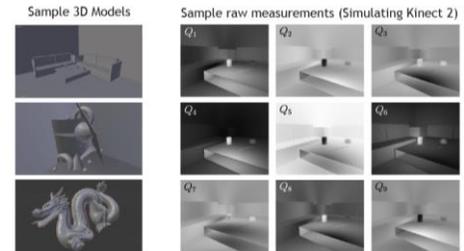
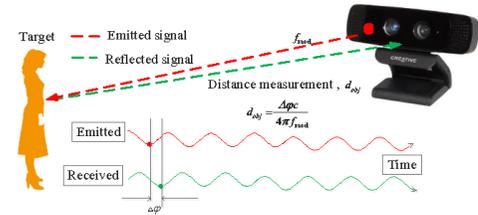
The physics of ToF cameras: acquisition, reconstruction, artifacts

Photon shot noise, motion artifacts, multiple reflection

A large dataset of simulated data

Design the DNN architecture accordingly to a-priori knowledge

Effective reduction of reconstruction artifacts



RESOURCES

Online



Tackling 3D ToF Artifacts Through Learning and the FLAT Dataset, Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, Jan Kautz; The European Conference on Computer Vision (ECCV), 2018, pp. 368-383,
http://openaccess.thecvf.com/content_ECCV_2018/html/Qi_Guo_Tackling_3D_ToF_ECCV_2018_paper.html

The FLAT dataset (code and data): <https://github.com/NVlabs/FLAT>

Contact: {ifrosio, ogallo}@nvidia.com, qiguo@g.harvard.edu