Accurate and Fast Algorithms for Dense Motion Estimation Based on Global Energy Minimisation

Andrés Bruhn

Mathematical Image Analysis Group
Saarland University
Saarbrücken, Germany

joint work with

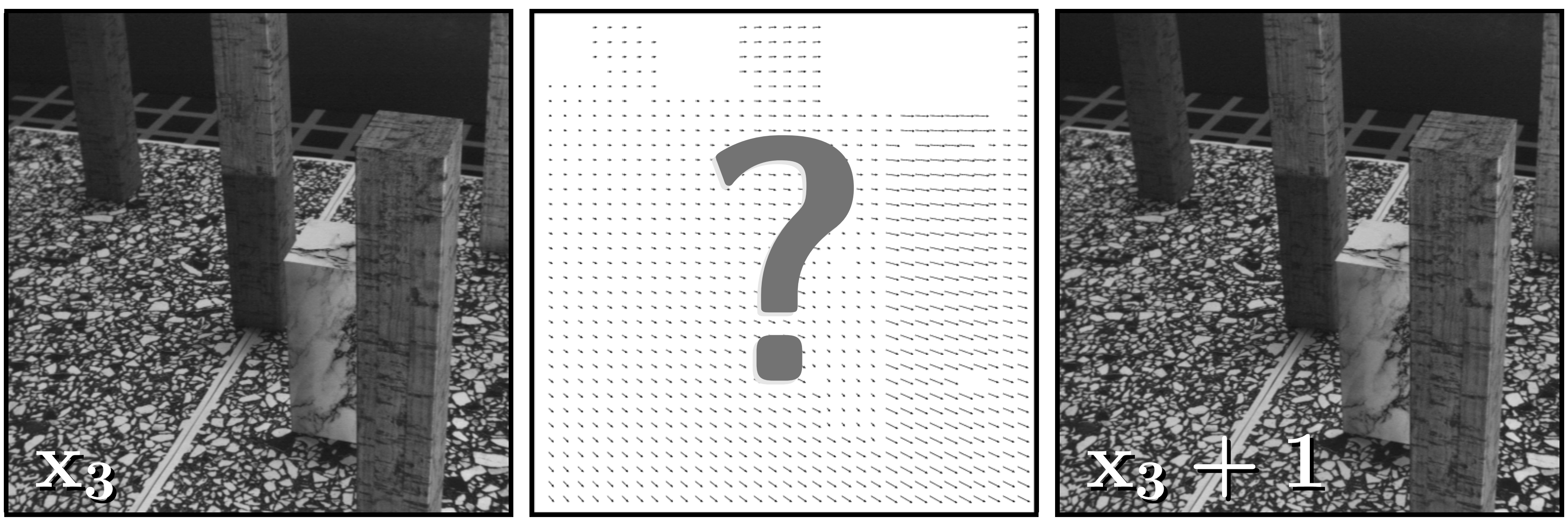
Joachim Weickert	–	Saarland University, Germany
Levi Valgaerts	–	Saarland University, Germany
Henning Zimmer	–	Saarland University, Germany
Thomas Brox	–	UC Berkeley, United States
Christoph Schnörr	–	University of Heidelberg, Germany

What is the Goal?

- **Given**
  - image sequence $f(x_1, x_2, x_3)$

location	(x_1, x_2)	$\in$	Ω
time	x_3	$\in$	$[0, T]$

- **Wanted**
  - interframe displacement field $\mathbf{u} = (u_1, u_2, 1) \rightarrow$ **optic flow**



What is the Goal?

- **Given**
  - image sequence $f(x_1, x_2, x_3)$

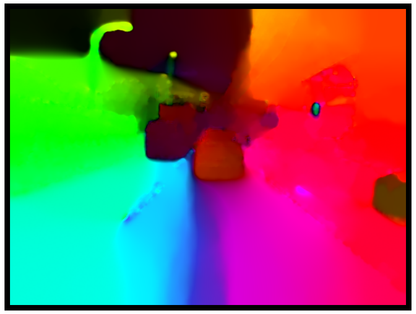
location	(x_1, x_2)	$\in$	Ω
time	x_3	$\in$	$[0, T]$

- **Wanted**
  - interframe displacement field $\mathbf{u} = (u_1, u_2, 1) \rightarrow$ **optic flow**

What is Optic Flow Good for?

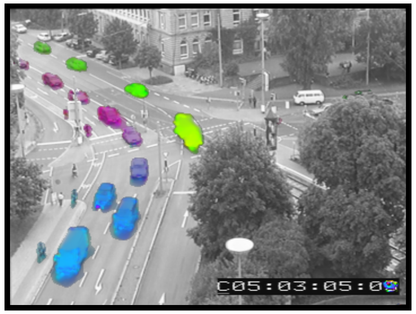
Extraction of Motion Information

- navigation
- obstacle detection
- tracking



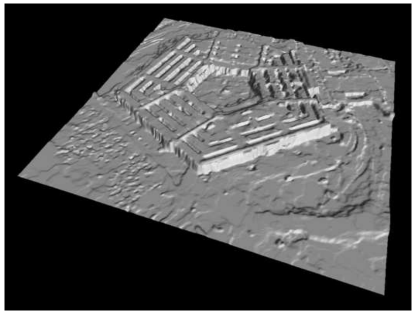
Processing of Image Sequences

- compact coding (compression $\longrightarrow$ MPEG)
- restoration and editing
- motion compensation



Related Correspondence Problems

- stereo reconstruction
- structure-from-motion
- medical image registration



Why Variational Methods?

- **Many Advantages**
  - transparent modelling
  - well-posedness and simple minimisation
  - highest accuracy in the literature
  - dense flow fields

- **Main Drawback**
  - large linear/nonlinear systems of equations
    ($\longrightarrow$ very slow with basic numerical solvers)

Goals of this Talk

- **Quality:** introduction to the design of high accuracy methods
- **Efficiency:** discussion of specifically adapted real-time algorithms
- **Practical Relevance:** presentation of current applications

Outline

- **Modelling**
  - general structure
  - data and smoothness term
  - qualitative benchmarks
- **Numerics**
  - minimisation and discretisation
  - efficient multigrid algorithms
  - performance benchmarks
  - real-time live demo
- **Applications**
- **Summary**

PART I

Modelling

Motion Estimation as Optimisation Problem

- **Optic Flow $\mathbf{u}$ as Minimiser of the Energy Functional**
  (*Horn/Schunck 1981)*

$$E(\mathbf{u}) = \int \Big(\underbrace{D(\mathbf{u})}_{\text{Data Term}} + \alpha \underbrace{G(\nabla \mathbf{u})}_{\text{Smoothness Term}} \Big) \, d\mathbf{x}$$

- **data term** penalises deviations from constancy assumptions on image features
- **smoothness term** penalises deviations from smoothness of solution
- regularisation parameter $\alpha > 0$ determines smoothness
- **global method:** integration over single image or full video

Standard Data Term

- **Constancy Assumption on the Image Brightness**
  (*e.g. Horn/Schunck 1981, Lucas/Kanade 1981*)
  - implicit formulation

$$0 = f(x_1 + u_1, x_2 + u_2, x_3 + 1) - f(x_1, x_2, x_3)$$

  - Taylor linearisation

$$0 = f_{x_1} u_1 + f_{x_2} u_2 + f_{x_3} 1 = \mathbf{u}^\top \nabla_3 f$$

  - quadratic penalisation

$$(\mathbf{u}^\top \nabla_3 f)^2$$

- **Drawback**
  - image brightness not invariant under **varying illumination**

Higher Order Constancy Assumptions

◆ Constancy Assumptions on Image Derivatives

(*Uras* et al. *1988, Schnörr 1993, Papenberg/Bruhn/Brox/Didas/Weickert IJCV 2006*)

Constancy	Data Term	Motion Type
gradient	$\sum_{i=1}^{2} (\mathbf{u}^\top \nabla_3 f_{x_i})^2$	translational divergent slow rotational
Hessian	$\sum_{i=1}^{2} \sum_{j=1}^{2} (\mathbf{u}^\top \nabla_3 f_{x_i x_j})^2$	translational divergent slow rotational
gradient magnitude	$(\mathbf{u}^\top \nabla_3 \lvert \nabla_2 f \rvert)^2$	**any**
Hessian trace	$(\mathbf{u}^\top \nabla_3 (\Delta_2 f))^2$	**any**
Hessian determinant	$(\mathbf{u}^\top \nabla_3 \det(\mathcal{H}_2 f))^2$	**any**

◆ Drawback

- images derivatives only invariant under **global additive** illumination changes

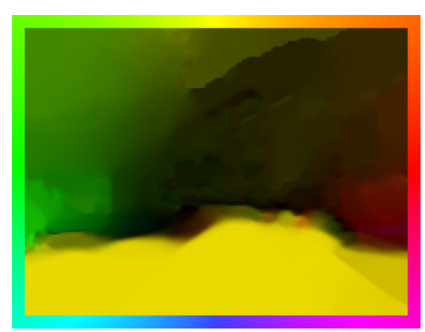
Colour Constancy Assumptions

- **Constancy Assumptions on Photometric Invariants**
  (*Mileva/Bruhn/Weickert DAGM 2007*)
  - colour images offer three measurements per pixel (R,G,B)
  - exploit redundancy by computing differences and ratios
  - transformations of the colour space / normalisation of RGB channels

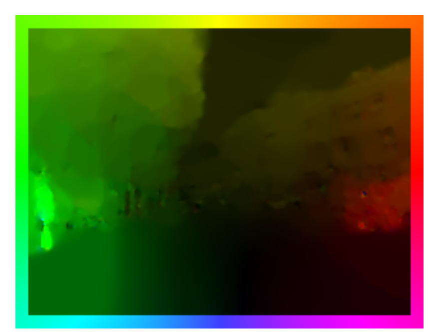
$$(R,G,B)^\top \mapsto \left(\frac{R}{N},\frac{G}{N},\frac{B}{N}\right)^\top, \qquad N = \frac{R+G+B}{3}$$

  - invariant under more realistic **local multiplicative** illumination changes

DIPLODOC Road Sequence



RGB Constancy



Invariant Constancy

Generic Framework for the Data Term

- **Motion Tensor Formalism**
  (*Bigün/Granlund/Wiklund 1991, Bruhn/Weickert/Kohlberger/Schnörr IJCV 2006*)
  - compact representation for combined data term
  - consider n constancy assumptions on $p_1,\ldots,p_n$ with weights $\lambda_1,\ldots,\lambda_n$

$$\sum_{i=1}^{n} \lambda_i \, (\mathbf{u}^\top \nabla_3 p_i)^2 = \sum_{i=1}^{n} \lambda_i \, \left(\mathbf{u}^\top \nabla_3 p_i \, \nabla_3 p_i^\top \, \mathbf{u}\right) = \mathbf{u}^\top \left(\sum_{i=1}^{n} \lambda_i \, \nabla_3 p_i \, \nabla_3 p_i^\top \right) \mathbf{u} = \mathbf{u}^\top J \, \mathbf{u}$$

  - single quadratic form with positive semi-definite 3×3 motion tensor J

- **Advantages**
  - framework for all presented data terms
  - rank analysis specifies degrees of freedom

Generic Framework for the Data Term

- **Overview of Motion Tensors**
  (*Bruhn/Weickert/Kohlberger/Schnörr IJCV 2006*)

Constancy	Motion Tensor J
brightness	$\nabla_3 f \; \nabla_3 f^\top$
gradient	$\sum_{i=1}^{2} (\nabla_3 f_{x_i})(\nabla_3 f_{x_i})^\top$
Hessian	$\sum_{i=1}^{2} \sum_{j=1}^{2} (\nabla_3 f_{x_i x_j})(\nabla_3 f_{x_i x_j})^\top$
gradient norm	$\frac{(f_{x_1} \nabla_3 f_{x_1} + f_{x_2} \nabla_3 f_{x_2})(f_{x_1} \nabla_3 f_{x_1} + f_{x_2} \nabla_3 f_{x_2})^\top}{f_{x_1}^2 + f_{x_2}^2}$
Hessian trace	$(\nabla_3 \sum_{i=1}^{2} f_{x_i x_i})(\nabla_3 \sum_{i=1}^{2} f_{x_i x_i})^\top$
Hessian determinant	$(f_{x_2 x_2} \nabla_3 f_{x_1 x_1} + f_{x_1 x_1} \nabla_3 f_{x_2 x_2} - 2 f_{x_1 x_2} \nabla_3 f_{x_1 x_2}) \cdot$ $(f_{x_2 x_2} \nabla_3 f_{x_1 x_1} + f_{x_1 x_1} \nabla_3 f_{x_2 x_2} - 2 f_{x_1 x_2} \nabla_3 f_{x_1 x_2})^\top$

Robustification against Noise and Outliers

- **Local Least Squares Fit**
  (*Lucas/Kanade 1981, Bruhn/Weickert/Schnörr IJCV 2005*)
  - integration over a neighbourhood of fixed size

$$K_\rho * \left(\mathbf{u}^\top J\, \mathbf{u}\right) = \mathbf{u}^\top \left(K_\rho * J\right) \mathbf{u} = \mathbf{u}^\top J_\rho\, \mathbf{u}$$

- **Robust Statistics – Single Assumption**
  (*Black/Anandan 1991, Mémin/Pérez 1998*)
  - subquadratic penalisation with increasing function $\Psi(s^2)$

$$\Psi\left(\mathbf{u}^\top J\, \mathbf{u}\right)$$

  - reduce influence of outliers, e.g. by replacing L_2 with L_1 norm

$$\Psi(s^2) := s^2 \quad \longrightarrow \quad \Psi(s^2) := \sqrt{\varepsilon^2 + s^2} - \varepsilon$$

Robustification of Multiple Assumptions

- **Robust Statistics – Correlated Assumptions**
  (*Brox/Bruhn/Papenberg/Weickert ECCV 2004*)
  - **joint** robustification, e.g. in the case of RGB colour images

$$\Psi\left(\sum_{i=1}^{n} \lambda_i \, \mathbf{u}^\top J_i \, \mathbf{u}\right)$$

- **Robust Statistics – Independent Assumptions**
  (*Bruhn/Weickert ICCV 2005*)
  - **separate** robustification, e.g. in the case of HSV colour images

$$\sum_{i=1}^{n} \lambda_i \, \Psi\left(\mathbf{u}^\top J_i \, \mathbf{u}\right)$$

Large Displacements

- **Theoretically Justified Warping**
  (*Nagel 1983, Brox/Bruhn/Papenberg/Weickert ECCV 2004*)

  - original constancy assumption

$$0 = f(\mathbf{x} + \mathbf{u}) - f(\mathbf{x})$$

  - incremental computation

$$\mathbf{u}^{k+1} = \mathbf{u}^{k} + \Delta\mathbf{u}^{k}$$

  - linearisation only by increment

$$0 = \Delta\mathbf{u}^{k\top} \nabla_3 f(\mathbf{x} + \mathbf{u}^{k})$$

Large Displacements

◆ **Multiscale Strategy**
(*Bergen/Anandan/Hanna/Hingorani 1992*)

- large displacements become **small** displacements

fine scale (displacements up to 10 pixels)

coarse scale (displacements up to 1 pixel)

◆ **Modified Motion Tensors for Large Displacements**
(*Bruhn/Weickert/Kohlberger/Schnörr IJCV 2006*)

- motion tensor notation still applicable (in the incremental computation)
- specific multiscale representation

Spatial vs. Spatiotemporal Regularisation

- **Spatial Regularisation**
  (*Horn/Schunck 1981*)
  - penalises deviations from smoothness in the spatial domain

$$\sum_{i=1}^{2} |\nabla_2 u_i|^2$$

- **Spatiotemporal Regularisation**
  (*Nagel 1990, Weickert/Schnörr 1999*)
  - extending spatial smoothness to the temporal domain

$$\sum_{i=1}^{2} |\nabla_3 u_i|^2$$

  - computationally hardly more expensive than spatial approach
  - better results but delayed computation (stack of frames required)

Adaptive Smoothness Terms

- **Diffusion-Inspired Regularisers (yield Diffusion Tensors)**
  (*Weickert/Schnörr 2001, Zimmer/Bruhn at al. EMMCVPR 2009*)

Strategy	Smoothness Term
homogeneous (*Horn/Schunck 1981*)	$\sum_i \lvert\nabla \boldsymbol{u}_i\rvert^2$
image-driven, isotropic/anistropic (*Alvarez et al. 1999, Nagel 1983*)	$g(\lvert\nabla \boldsymbol{f}\rvert^2) \sum_i \lvert\nabla \boldsymbol{u}_i\rvert^2 \,\Big/\, \sum_i \nabla \boldsymbol{u}_i^\top \boldsymbol{D}(\nabla \boldsymbol{f})\, \nabla \boldsymbol{u}_i$
flow-driven, isotropic/anisotropic (*Schnörr 1994, Weickert et al. 2001*)	$\boldsymbol{\Psi}\left(\sum_i \lvert\nabla \boldsymbol{u}_i\rvert^2\right) \,\Big/\, \mathrm{tr}\left(\boldsymbol{\Psi}\left(\sum_i \nabla \boldsymbol{u}_i \nabla \boldsymbol{u}_i^\top\right)\right)$
combined, anisotropic (*Sun et al. 2008*)	$\sum_r \boldsymbol{\Psi}_r\left(\sum_i \nabla \boldsymbol{u}_i^\top \boldsymbol{D}_r(\nabla \boldsymbol{f})\, \nabla \boldsymbol{u}_i\right)$
complementary, anisotropic (*Zimmer/Bruhn et al. 2009*)	not yet available

Modelling - The Smoothness Term (3)

Adaptive Smoothness Terms

- **Comparison of Different Strategies**
  (*Zimmer/Bruhn/Weickert/Valgaerts/Salgado/Rosenhahn/Seidel EMMCVPR 2009*)

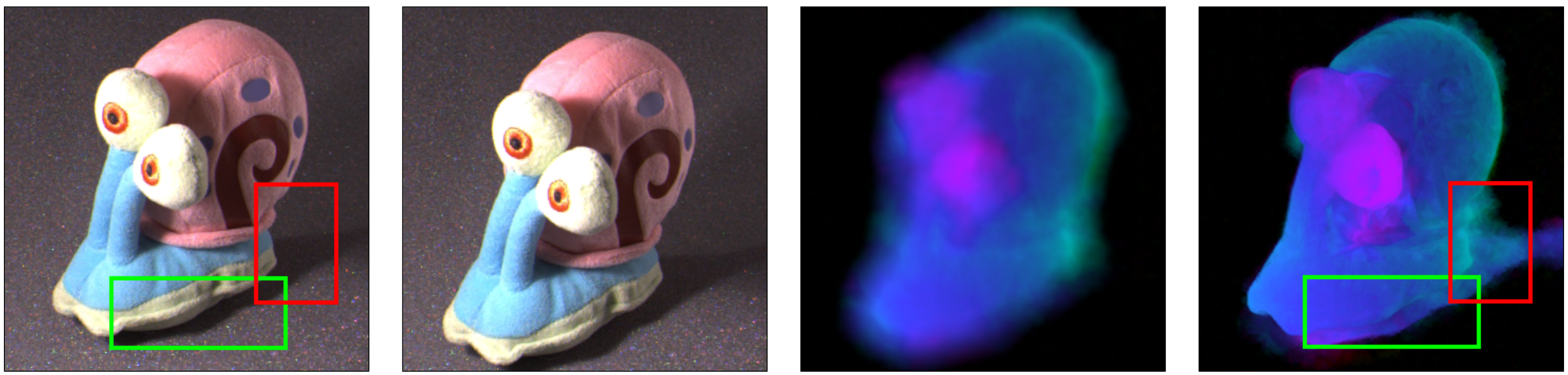
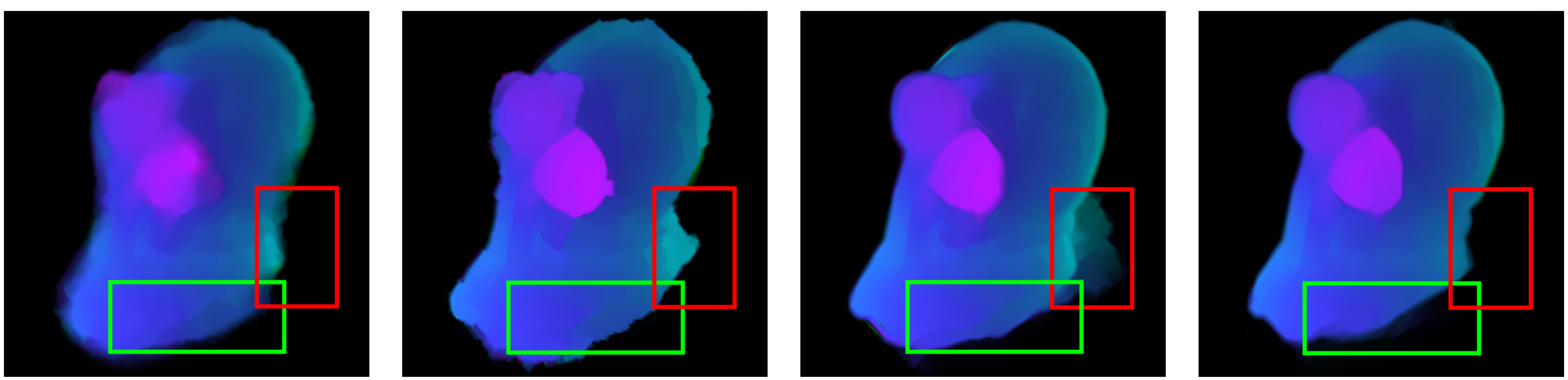


Image 1 | Image 2 | homogeneous | image-driven

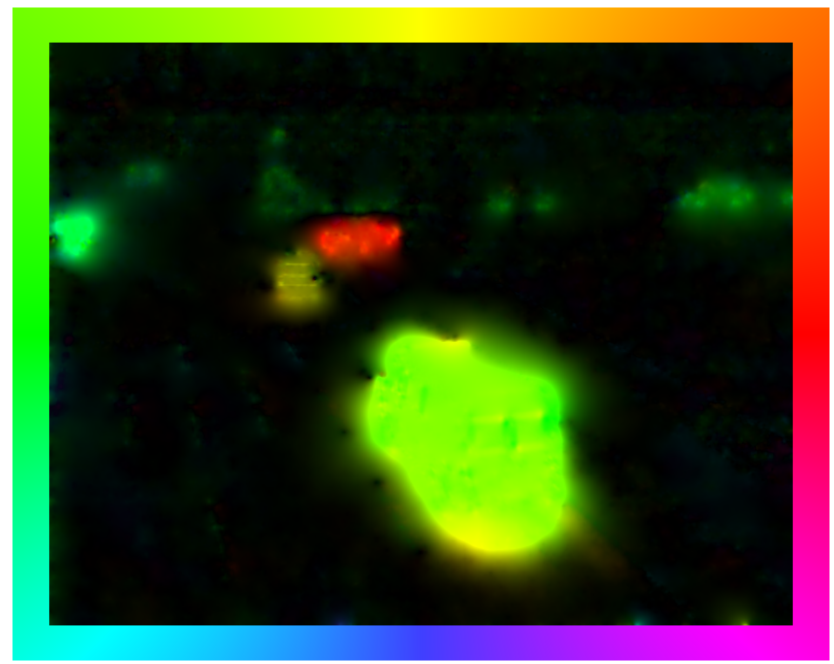


flow-driven | combined | complementary RGB | complementary HSV

Real-World Sequences - Qualitative Evaluation

- **Rheinhafen Sequence**
  (*Nagel, Size 688 × 565 × 1000*)

Frame 1130

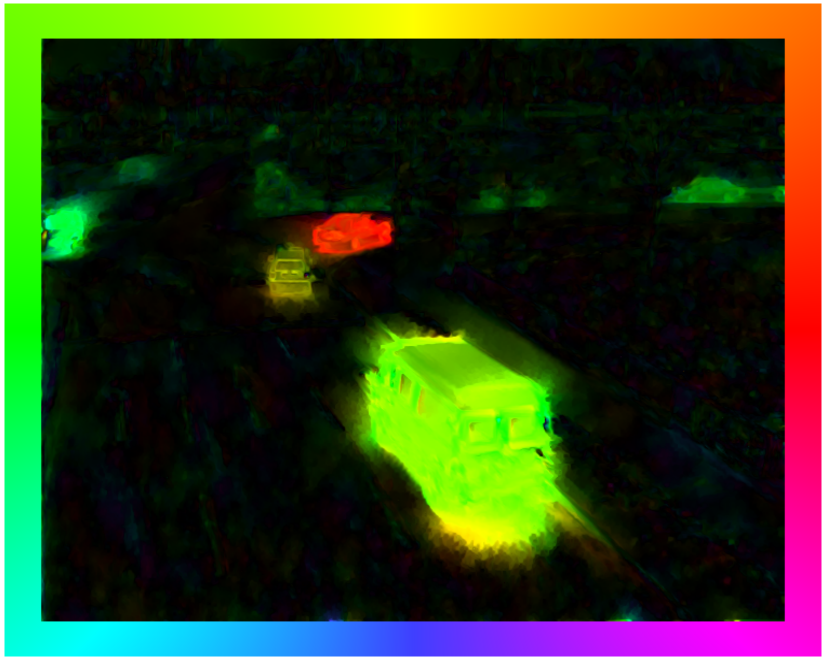


Brightness Constancy
Homogeneous Regulariser

Real-World Sequences - Qualitative Evaluation

- **Rheinhafen Sequence**
  (*Nagel, Size 688* $\times$ *565* $\times$ *1000*)

Frame 1130

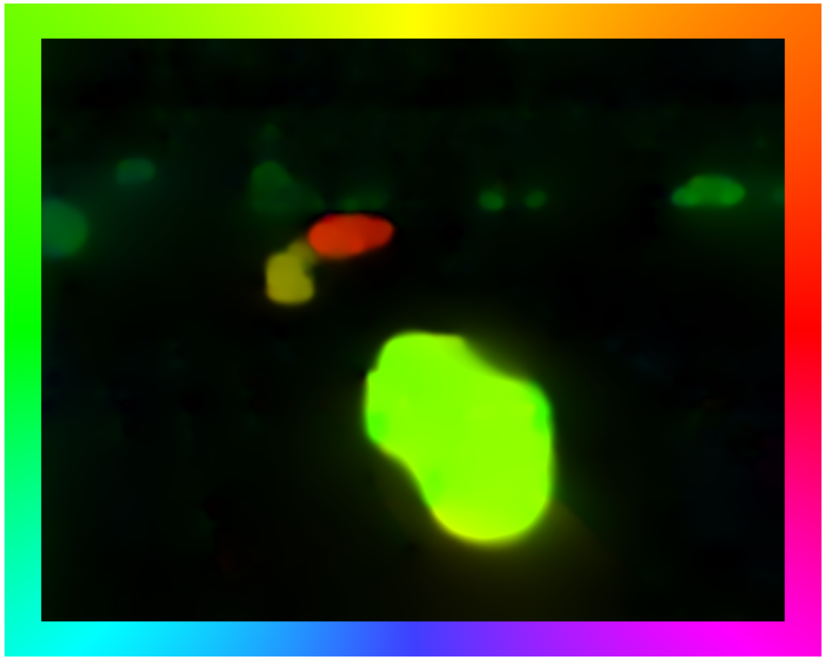


Brightness Constancy
Image-Driven Anisotropic Regulariser

Real-World Sequences - Qualitative Evaluation

- **Rheinhafen Sequence**
  (*Nagel, Size 688 $\times$ 565 $\times$ 1000*)

Frame 1130

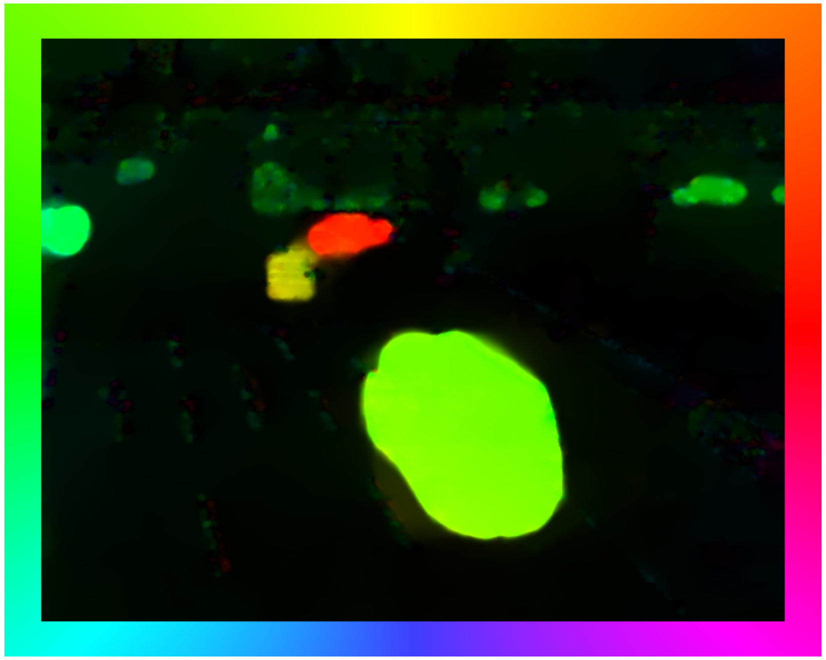


Robust Brightness Constancy
Flow-Driven Isotropic Regulariser

Real-World Sequences - Qualitative Evaluation

- **Rheinhafen Sequence**
  (*Nagel, Size 688* $\times$ *565* $\times$ *1000*)

Frame 1130



Robust Brightness Constancy
Robust Gradient Constancy
without Linearisation
Flow-Driven Isotropic Regulariser

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35

Real-World Sequences - Qualitative Evaluation

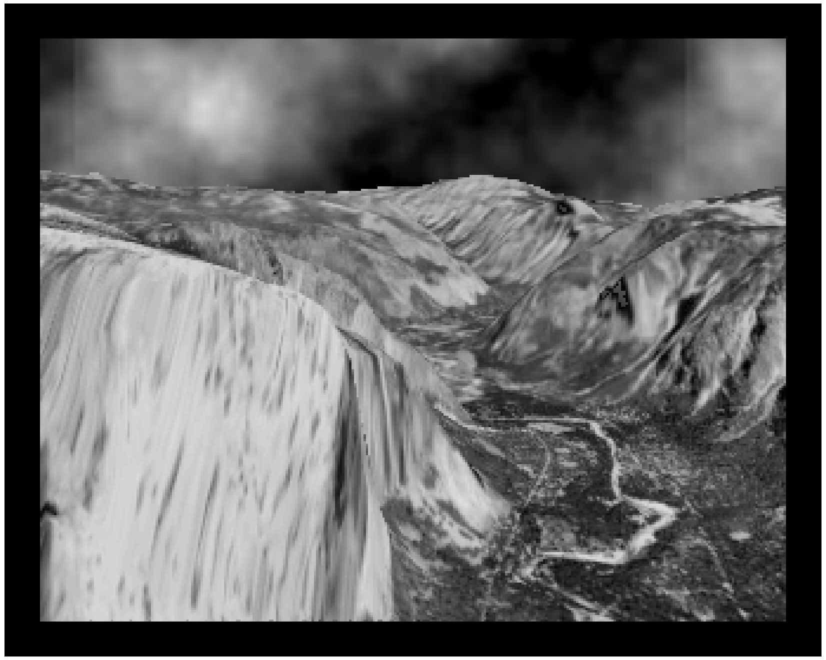
- **Karl Wilhelm Street and Ettlinger Tor Sequence**
  (*Nagel, Size 351* $\times$ *283* $\times$ *1034 and Size 512* $\times$ *512* $\times$ *50*)

Karl Wilhelm Street

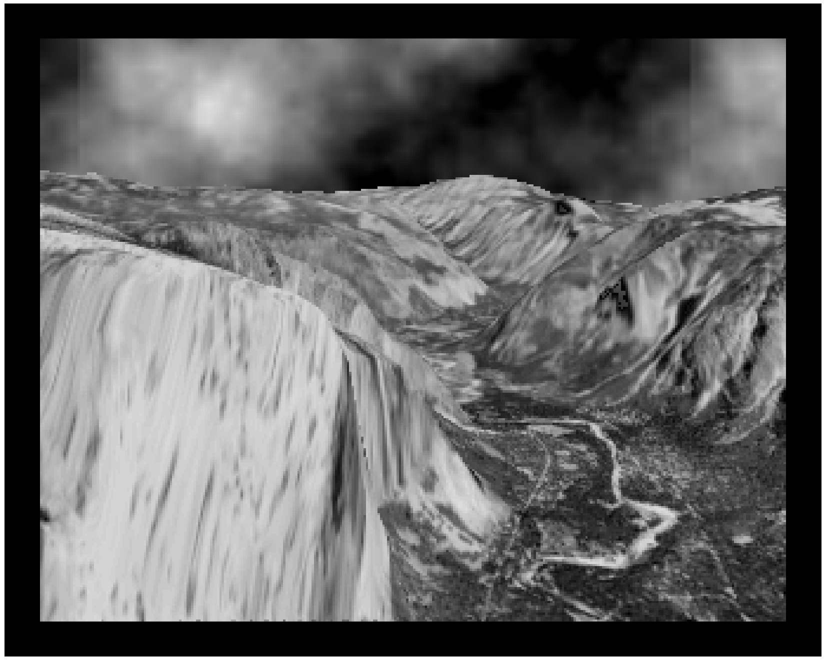
Ettlinger Tor

Synthetic Sequences - Qualitative Evaluation

- **Yosemite Sequence with Clouds**
  (*Quam 1984, Size 316 $\times$ 252 $\times$ 15*)



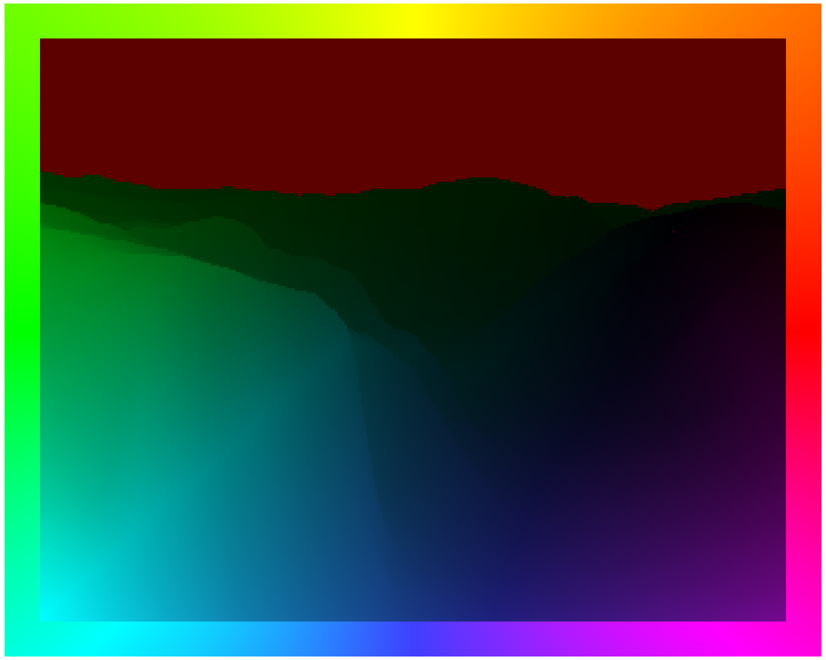
Frame 8



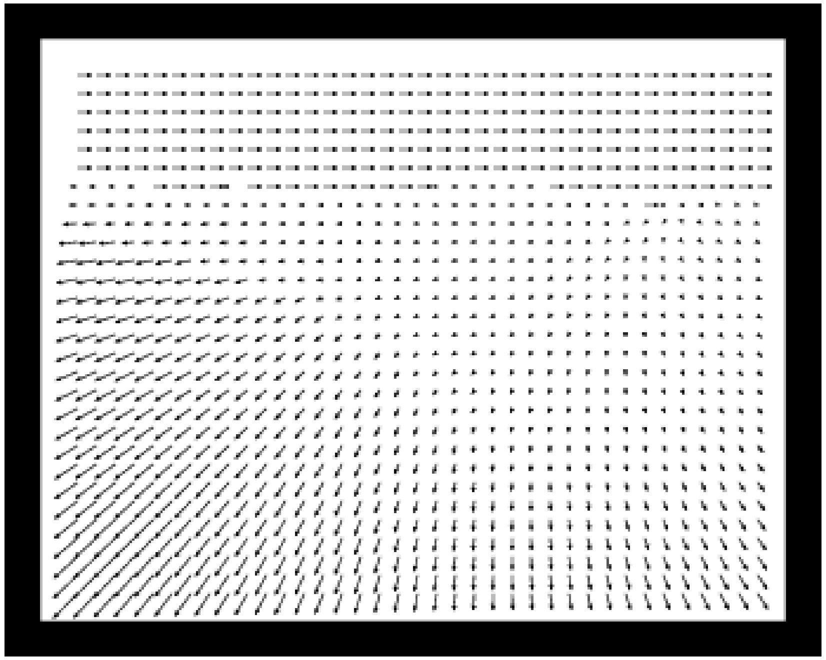
Frame 9

Synthetic Sequences - Qualitative Evaluation

- **Yosemite Sequence with Clouds**
  (*Quam 1984, Size 316 × 252 × 15*)



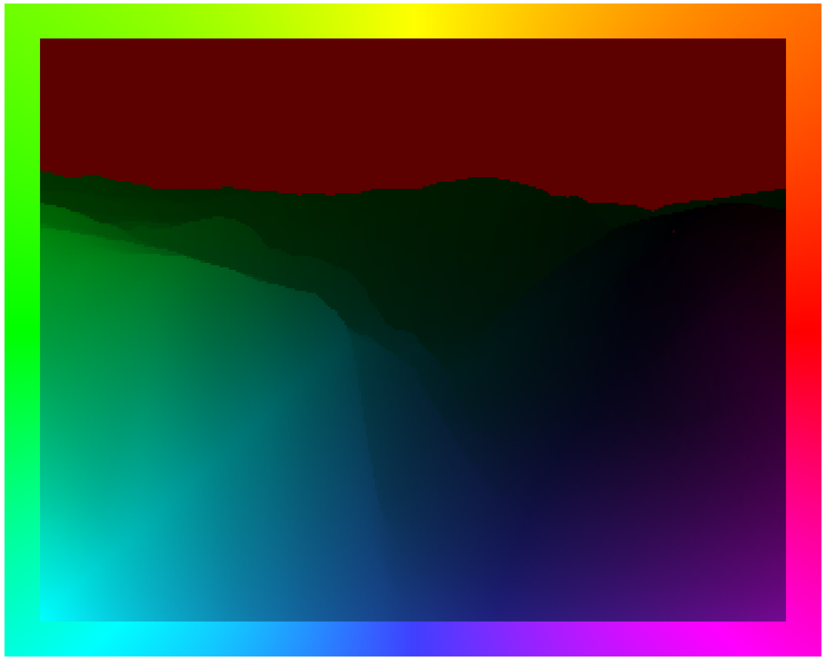
Ground Truth (Colour Plot)



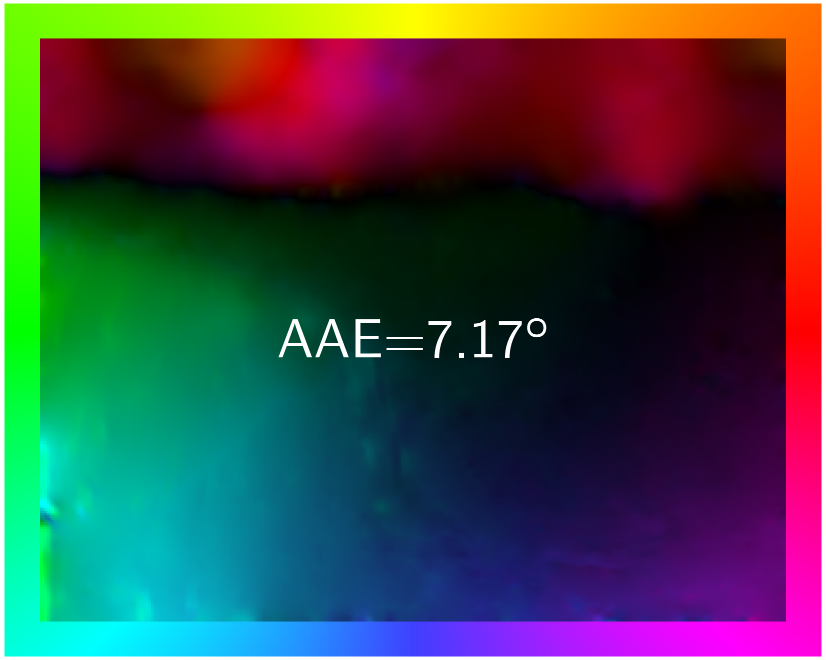
Ground Truth (Vector Plot)

Synthetic Sequences - Qualitative Evaluation

- **Yosemite Sequence with Clouds**
  (*Quam 1984, Size 316 × 252 × 15*)



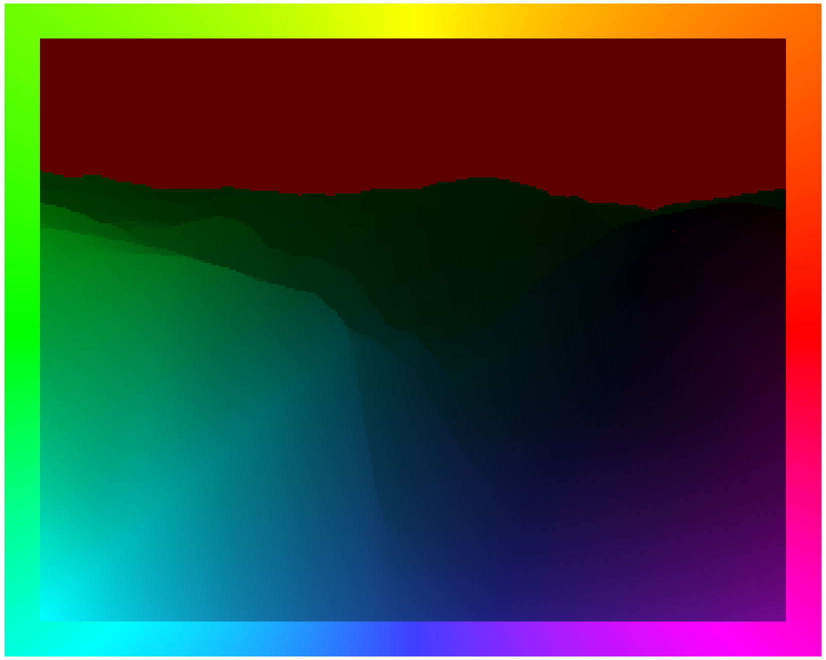
Ground Truth (Colour Plot)



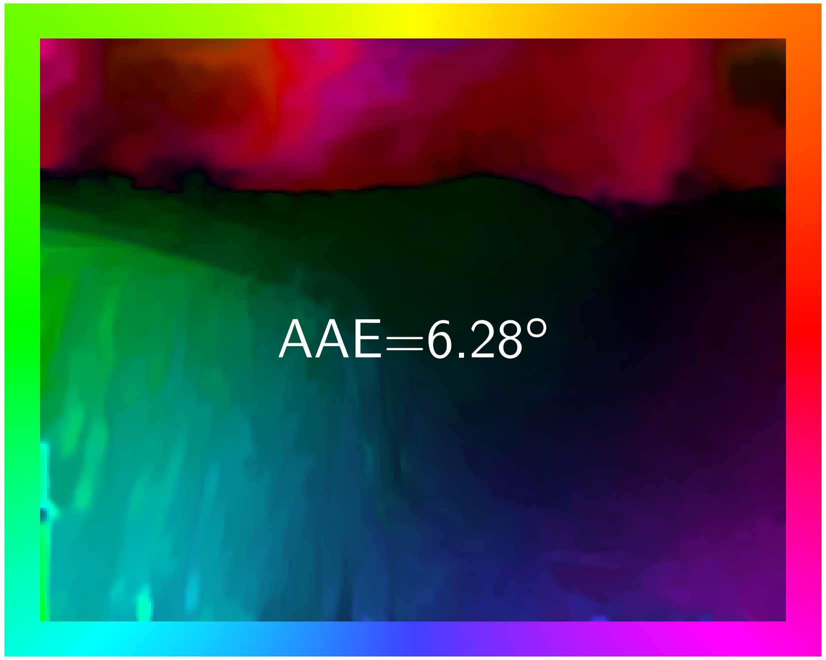
Brightness Constancy
Homogeneous Regulariser

Synthetic Sequences - Qualitative Evaluation

- **Yosemite Sequence with Clouds**
  (*Quam 1984, Size 316 × 252 × 15*)



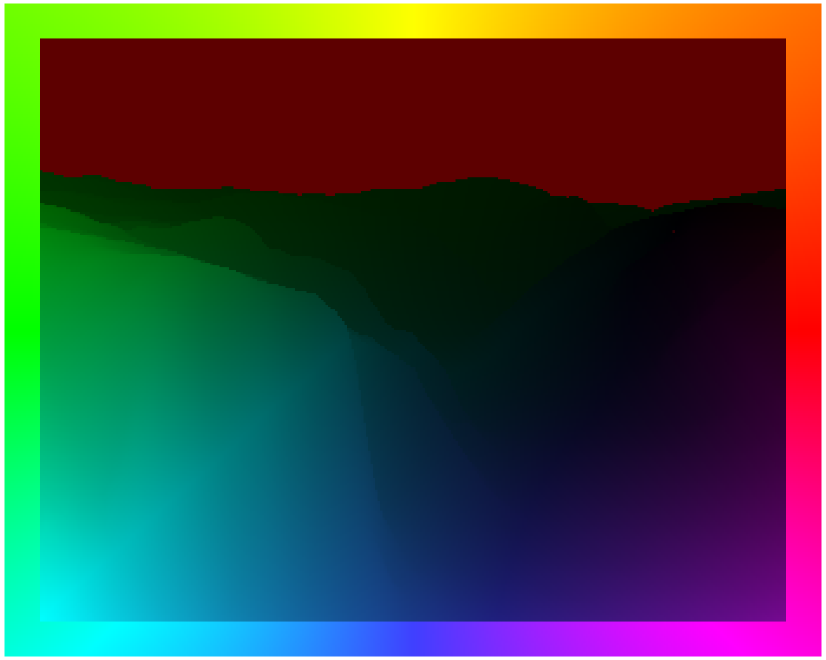
Ground Truth (Colour Plot)



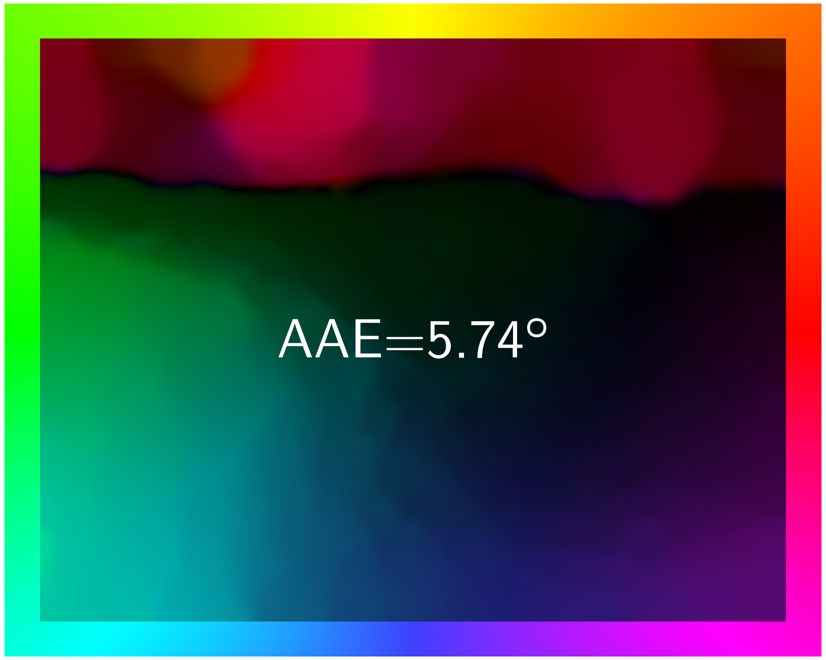
Brightness Constancy
Image-Driven, Anisotropic Regulariser

Synthetic Sequences - Qualitative Evaluation

- **Yosemite Sequence with Clouds**
  (*Quam 1984, Size 316 × 252 × 15*)



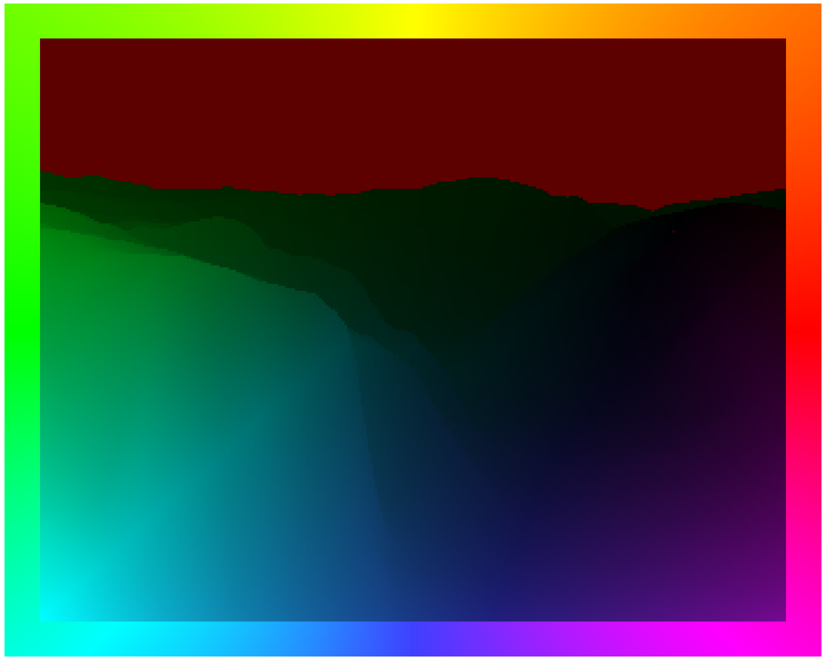
Ground Truth (Colour Plot)



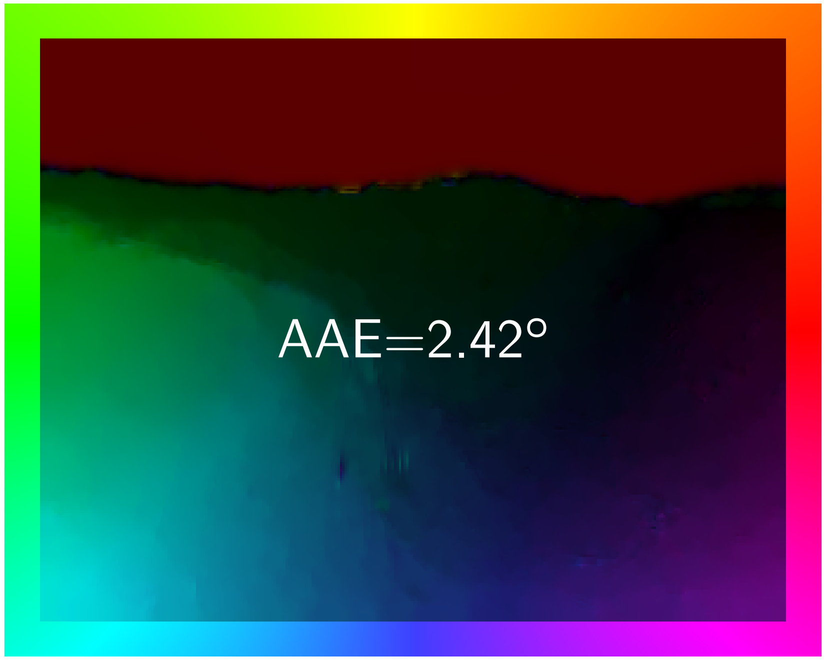
Robust Brightness Constancy
Flow-Driven, Isotropic Regulariser

Synthetic Sequences - Qualitative Evaluation

- **Yosemite Sequence with Clouds**
  (*Quam 1984, Size 316 × 252 × 15*)



Ground Truth (Colour Plot)



Robust Brightness Constancy
Robust Gradient Constancy
without Linearisation
Flow-Driven, Isotropic Regulariser

Synthetic Sequences - Comparison to Literature

- **Yosemite Sequence with Clouds**
  (*Quam 1984, Size 316 × 252 × 15*)

Technique	AAE
Horn/Schunck, orig.	31.69°
Singh, step 1	15.28°
Anandan	13.36°
Singh, step 2	10.44°
Nagel	10.22°
Horn/Schunck, mod.	9.78°
Uras *et al.*	8.94°
Prototype A	**7.17°**
Liu *et al.*	6.85°
Prototype B	**6.44°**
Prototype E	**6.42°**
Prototype D	**6.32°**
Prototype C	**6.28°**
Prototype F (2-D, SD)	**5.74°**
Alvarez *et al.*	5.53°

Technique	AAE
Mémin/Pérez	5.38°
Prototype F (3-D, SD)	**5.18°**
Farnebäck	4.84°
Mémin/Pérez	4.69°
Prototype F (3-D, LD)	**4.17°**
Wu *et al.*	3.54°
Prototype G (2-D, SD)	**3.50°**
Prototype G (3-D, SD)	**2.78°**
Teng *et al.*	2.70°
Prototype H (2-D, LD)	**2.42°**
Amiaz/Kiryati	2.04°
Prototype G (3-D, LD)	**1.78°**
Amiaz/Kiryati	1.73°
Prototype H (3-D, LD)	**1.72°**
Brox/Bruhn/Weickert	**1.22°**

Synthetic Sequences - Comparison to Literature

- **Middlebury Benchmark**
  (*Baker/Scharstein/Lewis/Roth/Black/Szeliski 2007, 8 Sequences with Different Sizes*)

Average angle error	avg. rank	Army (Hidden texture) GT im0 im1			Mequon (Hidden texture) GT im0 im1			Schefflera (Hidden texture) GT im0 im1			Wooden (Hidden texture) GT im0 im1			Grove (Synthetic) GT im0 im1			Urban (Synthetic) GT im0 im1			Yosemite (Synthetic) GT im0 im1			Teddy (Stereo) GT im0 im1		
		all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext
ComplOF [27]	4.2	4.44 8	11.2 7	4.04 8	2.51 2	9.77 3	**1.74** 1	3.93 4	10.6 4	**2.04** 1	3.87 7	18.8 4	2.19 6	**3.17** 1	**4.00** 1	2.92 2	4.64 5	13.8 2	3.64 4	2.17 6	3.36 3	2.51 12	3.08 2	7.04 2	3.65 5
Adaptive [26]	4.3	**3.29** 1	**9.43** 1	**2.28** 1	3.10 6	11.4 7	2.46 7	6.58 9	15.7 8	2.52 5	**3.14** 1	**15.6** 1	**1.56** 1	3.67 6	4.46 5	3.48 6	**3.32** 1	**13.0** 1	**2.38** 1	2.76 11	4.39 10	1.93 7	3.58 3	8.18 3	2.88 2
Spatially variant [22]	5.9	3.73 3	10.2 5	3.33 4	3.02 5	11.0 6	2.67 8	5.36 5	13.8 6	2.35 2	3.67 2	19.3 6	1.84 3	3.81 8	4.81 13	3.69 9	4.48 4	16.0 7	3.90 5	2.11 4	3.26 2	2.12 9	4.66 8	9.41 8	4.35 10
TV-L1-improved [20]	6.8	3.36 2	9.63 2	2.62 2	2.82 4	10.7 5	2.23 3	6.50 8	15.8 9	2.73 6	3.80 5	21.3 11	1.76 2	3.34 2	4.38 4	**2.39** 1	5.97 7	18.1 13	5.67 11	3.57 15	4.92 16	3.43 17	4.01 7	9.84 9	3.44 3
Multicue MRF [24]	7.9	4.50 9	10.1 3	4.18 12	2.52 3	**7.07** 1	2.36 6	**3.09** 1	**7.41** 1	2.36 3	4.46 10	20.8 10	2.73 9	3.51 4	4.11 2	4.06 14	6.08 9	15.6 6	5.40 9	5.25 22	5.36 18	9.02 22	3.63 4	8.39 4	4.15 8
F-TV-L1 [18]	8.5	5.44 12	12.5 11	5.69 16	5.46 13	15.0 13	4.03 13	7.48 13	16.3 11	3.42 11	5.08 12	23.3 15	2.81 10	3.42 3	4.34 3	3.03 3	4.05 3	15.1 4	3.18 2	2.43 8	3.92 8	1.87 6	3.90 6	9.35 7	**2.61** 1
DPOF [21]	10.2	5.63 13	10.9 6	4.16 11	4.05 10	12.1 8	3.31 9	3.87 3	8.82 2	3.17 9	4.34 8	16.2 2	3.13 12	3.95 12	4.78 12	4.17 16	6.69 15	15.2 5	6.27 14	5.62 23	6.89 24	6.60 21	**2.44** 1	**4.83** 1	3.74 7
Brox et al. [8]	10.2	4.80 11	14.4 15	4.29 13	4.05 10	13.5 10	3.71 11	6.63 10	16.0 10	7.26 13	5.22 13	22.7 14	3.22 13	4.56 17	6.09 23	3.40 4	3.97 2	17.9 11	3.41 3	2.07 3	3.76 6	1.18 2	5.14 10	11.9 12	4.28 9
Fusion [9]	10.7	4.43 7	13.7 13	4.08 9	**2.47** 1	8.91 2	2.24 4	3.70 2	9.68 3	3.12 8	3.68 3	19.8 7	2.54 8	4.26 15	5.16 14	4.31 18	6.32 11	16.8 9	6.15 13	4.55 19	5.78 20	3.10 16	7.12 18	13.6 18	7.86 19
SegOF [13]	11.2	5.85 14	13.5 12	3.98 7	7.40 15	14.9 12	8.13 19	8.55 15	17.3 15	9.01 14	6.50 16	18.1 3	5.14 16	3.90 11	4.53 6	4.81 21	6.57 14	21.7 19	6.81 17	**1.65** 1	3.49 5	**1.08** 1	3.71 5	9.23 6	3.63 4
Dynamic MRF [10]	11.2	4.58 10	12.4 10	4.14 10	3.25 8	13.9 11	2.27 5	6.02 7	16.8 12	2.36 3	4.39 9	22.6 13	2.51 7	3.61 5	4.55 7	3.46 5	6.81 16	22.2 21	6.78 16	2.41 7	3.48 4	3.69 18	9.26 22	17.8 22	10.2 22
CBF [15]	12.1	3.95 5	10.1 3	3.44 6	3.70 9	10.6 4	3.85 12	5.64 6	13.5 5	3.34 10	3.71 4	21.5 12	1.99 4	4.36 16	5.50 16	3.55 7	11.3 23	19.1 15	9.05 22	6.79 25	7.37 26	11.6 25	5.50 11	11.8 11	5.66 13
GraphCuts [17]	13.2	6.25 15	14.3 14	5.53 15	8.60 17	20.1 19	6.61 15	7.91 14	15.4 7	10.9 15	4.88 11	19.0 5	3.05 11	3.78 7	4.71 10	3.94 12	8.74 19	16.4 8	5.39 8	4.04 18	4.87 14	4.85 20	6.35 14	12.2 13	6.05 16
Learning Flow [14]	13.2	4.23 6	11.7 9	3.41 5	4.16 12	15.3 14	3.42 10	6.78 11	16.9 13	3.83 12	6.41 15	25.3 17	4.25 14	4.66 20	6.01 22	4.00 13	6.33 13	20.7 17	5.30 6	3.09 13	4.84 13	2.91 14	7.08 17	15.0 20	5.27 12
Second-order prior [11]	13.5	3.84 4	11.2 7	3.11 3	3.12 7	12.9 9	2.17 2	6.96 12	17.2 14	2.83 7	3.84 6	20.5 9	2.09 5	4.83 22	5.83 20	3.90 11	14.0 25	21.8 20	8.28 19	7.74 26	6.88 23	11.7 26	6.74 16	13.4 17	5.80 14

PART II

Numerics

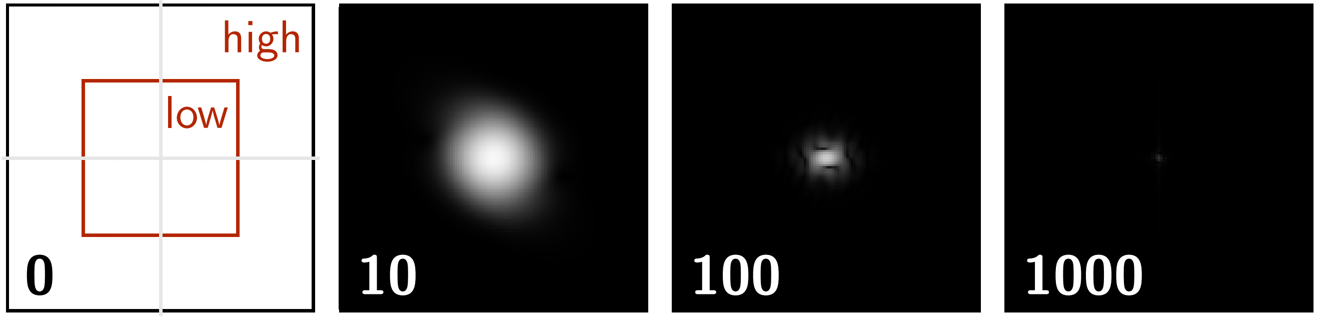
Minimisation Strategy

Euler–Lagrange Equations

- necessary conditions for a minimiser
- coupled system of partial differential equations
- discretisation yields large **linear** or **nonlinear** system of equations
- typically solved by iterative methods (Jacobi, Gauß-Seidel)

Drawback of Iterative Methods

- slow convergence after a few iterations
- logarithmic error spectrum reveals slow decrease of lower frequency parts ($\longrightarrow$ only efficient damping of **higher error frequency parts**)



Multigrid Methods

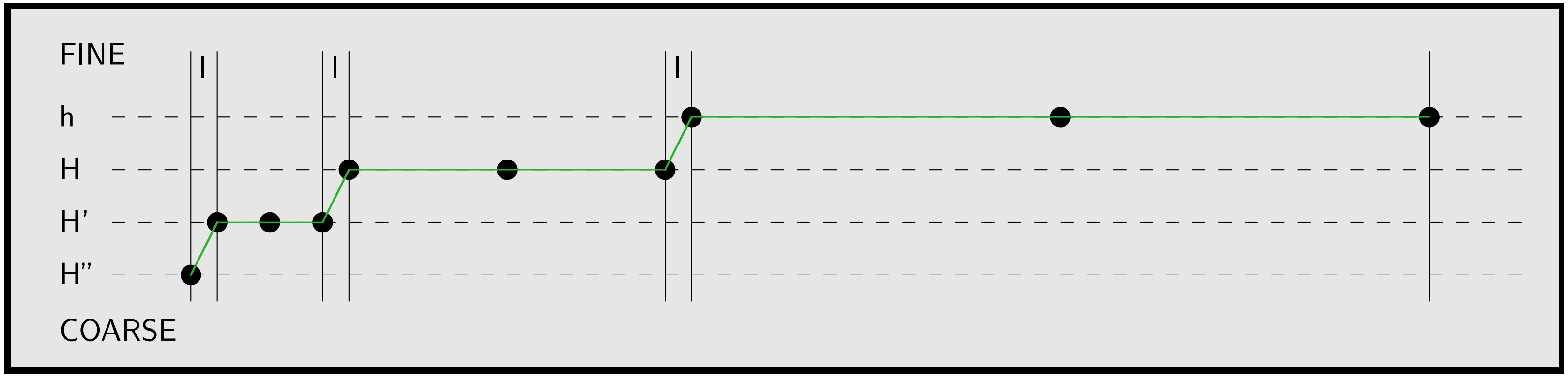
Basic Idea

(*Brandt 1977, Hackbusch 1985*)

- transfer and compute **error** on coarser grids
- low frequencies reappear as higher frequencies
  (⟶ also efficient damping of **lower error frequency parts**)

Recursive Strategies, Linear Complexity

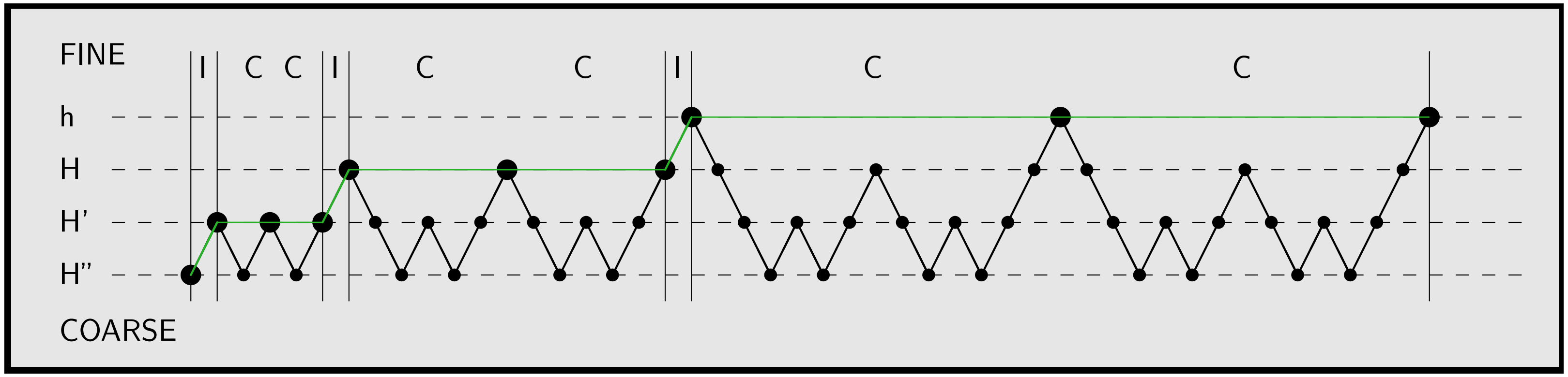
- hierarchical application (⟶ V–cycle, W–cycle)
- additional usage of hierarchical initialisation (⟶ Full Multigrid)



Multigrid Methods

- **Basic Idea**
  (*Brandt 1977, Hackbusch 1985*)
  - transfer and compute **error** on coarser grids
  - low frequencies reappear as higher frequencies
    (⟶ also efficient damping of **lower error frequency parts**)

- **Recursive Strategies, Linear Complexity**
  - hierarchical application (⟶ V–cycle, W–cycle)
  - additional usage of hierarchical initialisation (⟶ Full Multigrid)



Multigrid for Optic Flow

Specific Adaptations

- improved coarse grid correction scheme (**tensor based**)
  (*Bruhn/Weickert/Kohlberger/Schnörr IJCV 2006*)
- improved intergrid transfer operators (**non-dyadic**) and solvers (**coupled**)
  (*Bruhn/Feddern/Weickert/Kohlberger/Schnörr IEEE TIP 2005*)
- extended to large displacements (**combination with warping**)
  (*Bruhn/Weickert ICCV 2005*)

Overview of Multigrid Implementations

Type	MG Solver	Cycles	Basic Solver	Pre/Post
A - Homogeneous	FMG-W	1	GS-CPR	1-1
B - Image-Driven Isotropic	FMG-W	2	GS-CPR	2-2
C - Image-Driven Anisotropic	FMG-W	4	GS-ALR	1-1
D - Flow-Driven Isotropic	FAS-FMG-W	2	GS-CPR	2-2
E - Flow-Driven Anisotropic	FAS-FMG-W	4	GS-ALR	1-1
F - Bruhn *et al.* 2-D, SD	FAS-FMG-W	2	GS-CPR	2-2
G - Papenberg *et al.* 3-D, SD	FAS-FMG-W	2	GS-CPR	2-2
H - Bruhn/Weickert 2-D, LD	WARP-FAS-FMG-W	2	GS-CPR	3-3

Comparison of Numerical Solvers (Image Size 160 × 120)

Overview for Different Numerical Prototypes

- C implementation on standard desktop PC (3.06 GHz Pentium4)
- stopping criterion: norm of error less than 1% of norm of solution

Method with Image-Driven Anisotropic Regularisation: Linear Multigrid

Solver	Iterations	Time [s]	FPS [s^{-1}]	Speedup
Mod. Explicit Scheme	36433	47.08	0.02	1
Gauß-Seidel (ALR)	607	3.60	0.27	13
Full Multigrid	**1**	**0.17**	**5.88**	**275**

Method with Flow-Driven Isotropic Regularisation: Nonlinear Multigrid

Solver	Iterations	Time [s]	FPS [s^{-1}]	Speedup
Mod. Explicit Scheme	10633	30.492	0.033	1
Gauß-Seidel (ALR)	2679	6.911	0.145	4
FAS Full Multigrid	**1**	**0.082**	**12.172**	**372**

Multigrid Speedups (Image Size 160 × 120)

- **Overview For Different Model Prototypes**
  (*Bruhn/Weickert/Kohlberger/Schnörr IJCV 2006*)

  - **two to three orders of magnitude** for different regularisation strategies

Type	Solver	FPS	Speedup
A - Homogeneous	Full Multigrid	62.7	220
B - Image-Driven Isotropic	Full Multigrid	20.8	251
C - Image-Driven Anisotropic	**Full Multigrid**	**5.8**	**275**
D - Flow-Driven Isotropic	**FAS Full Multigrid**	**12.1**	**372**
E - Flow-Driven Anisotropic	FAS Full Multigrid	2.0	120

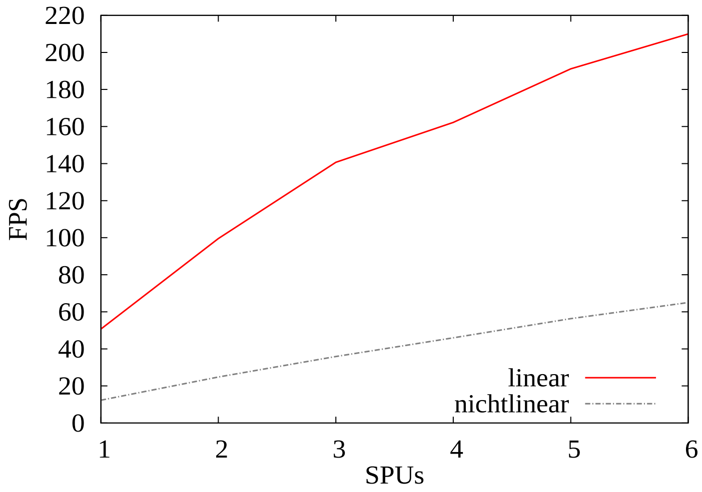
  - **three to four orders of magnitude** for high accuracy methods

Type	Solver	FPS	Speedup
F - Bruhn *et al.* 2-D, SD	FAS Full Multigrid	11.5	**2836**
G - Papenberg *et al.* 3-D, SD	FAS Full Multigrid	9.9	**10588**
H - Bruhn/Weickert 2-D, LD	Warp FAS Full Multigrid	2.9	**5454**

Speedup by Parallel Hardware (Image Size 316 × 252)

- **Cell Processor - Sony Playstation 3**
  (*Gwosdek/Bruhn/Weickert VMV 2008, Gwosdek/Bruhn/Weickert JRTIP 2009 submitted*)

  - 6 SPUs with ringbus memory interface
  - Speedup of 6.5 compared to a 3.2 GHz desktop PC
  - **linear case:** up to 210 dense flow fields per second (13.6 million pixels)
  - **nonlinear case:** up to 65 dense flow fields per second (4.2 million pixels)



Real-Time Live Demo

- **Live Computation with Webcam (160×120)**

Flow fields are computed with a 1.7 GHz PentiumM CPU

Start
Stop

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35

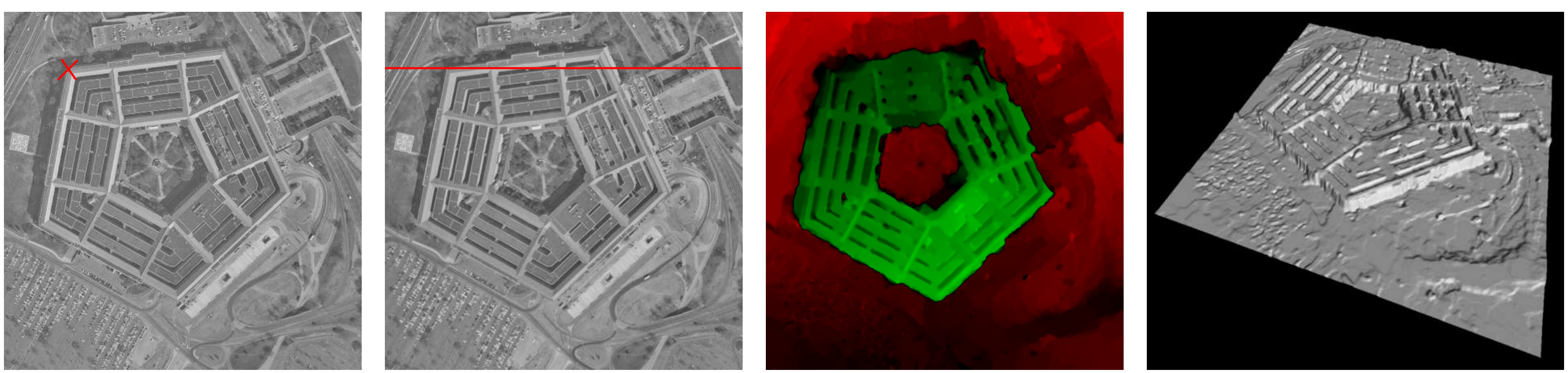
PART III

Applications

Stereo - Calibrated Case

- **Integration of Prior Knowledge on the Stereo Geometry**
  (*Slesareva/Bruhn/Weickert DAGM 2005*)
  - restriction of search space to given stereo geometry
  - depth can be directly computed from displacements (disparity)

- **Example: Reconstruction of the Pentagon from Aerial Views**
  (*CMU Stereo Database, Size 512 $\times$ 512*)



CMU Pentagon Image Pair with Stereo Geometry | Displacement Field | Reconstruction with Illumination

Stereo - Uncalibrated Case

- **Joint Estimation of Displacements and Stereo Geometry**
  (*Valgaerts/Bruhn/Weickert DAGM 2008)*
  - more precise and more robust estimation of correspondences
  - more exact estimation of camera poses (essential matrix)

- **Example: Face Reconstruction from Uncalibrated Images**
  (*Pascal Gwosdek, Size 280* $\times$ *430*)

Image Pair without Geometry

Reconstruction with Texture

Deinterlacing

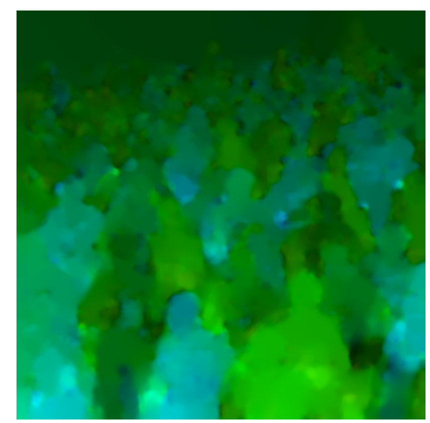
- **Conversion from Interlaced to Progressive Format**
  (*Ghodstinat/Bruhn/Weickert SGAVMA 2009*)
  - alternatingly only even and odd lines given (z.B. PAL)
  - inpainting of missing information respecting motion trajectories

- **Example: Motion Compensation in Broadcasts of Sports Events**
  (*European Broadcasting Union, Zoom-In, Size 300* $\times$ *300*)

Progressive Image

Interlaced Image

Deinterlaced Result



Displacement Field

Hairstyle Simulation

- **Automatic Registration of Reference Hairstyles onto Customer Faces**
  (*Demetz/Weickert/Bruhn/Welk SSVM 2007*)
  - adaptation of reference hairstyle according to deformation field
  - preregistration of eyes, masking of hairstyle, flow computation

- **Example: Registration of a Short Hairstyle**
  (*Style Concept, Size 900* $\times$ *900*)

Customer Face

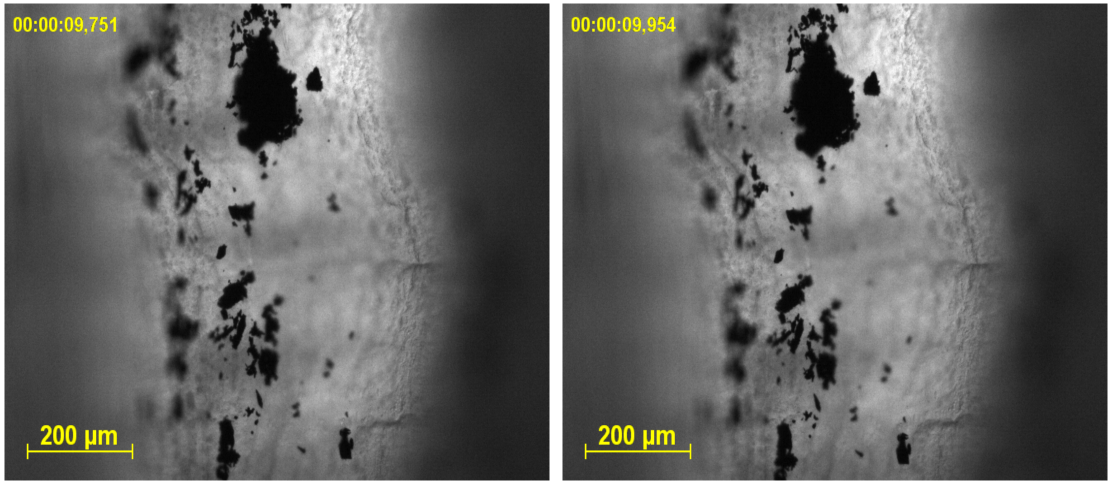
Reference Face with Hairstyle

Hairstyle Simulation

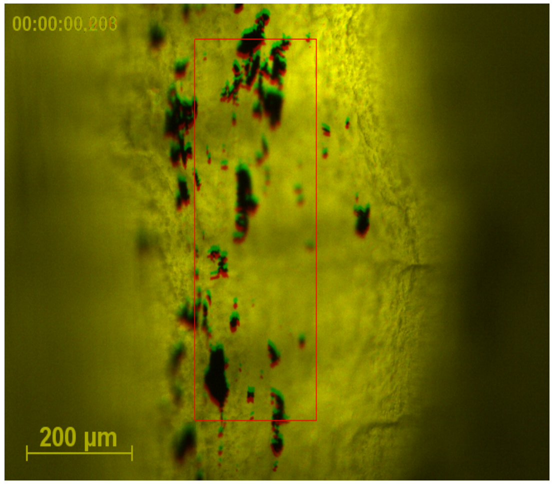
Motion Analysis

- **Motion Estimation of Nanoparticles in Tracheal Tissue**
  (*Kariger/Bruhn/Henning/Weickert/Lehr - Cooperation with Dept. of Pharmaceutical Technology*)
  - investigation of mucociliary clearance (protection of respiratory system)
  - smoothness constraints from the field of particle image velocimetry (PIV)

- **Example: Transport of Coal Particles of Size 1-100 μm**
  (*Andreas Henning, Size 660 $\times$ 492*)



Images with Time Interval 200 ms



Registered Images
Velocity 2.55 mm/min

Summary

High Accuracy Models

- generic framework for the design of novel methods
- highest precision in the literature
- robust under noise and illumination changes

Real-Time Algorithms

- speedups of two to four orders of magnitude
- additional acceleration using parallel hardware

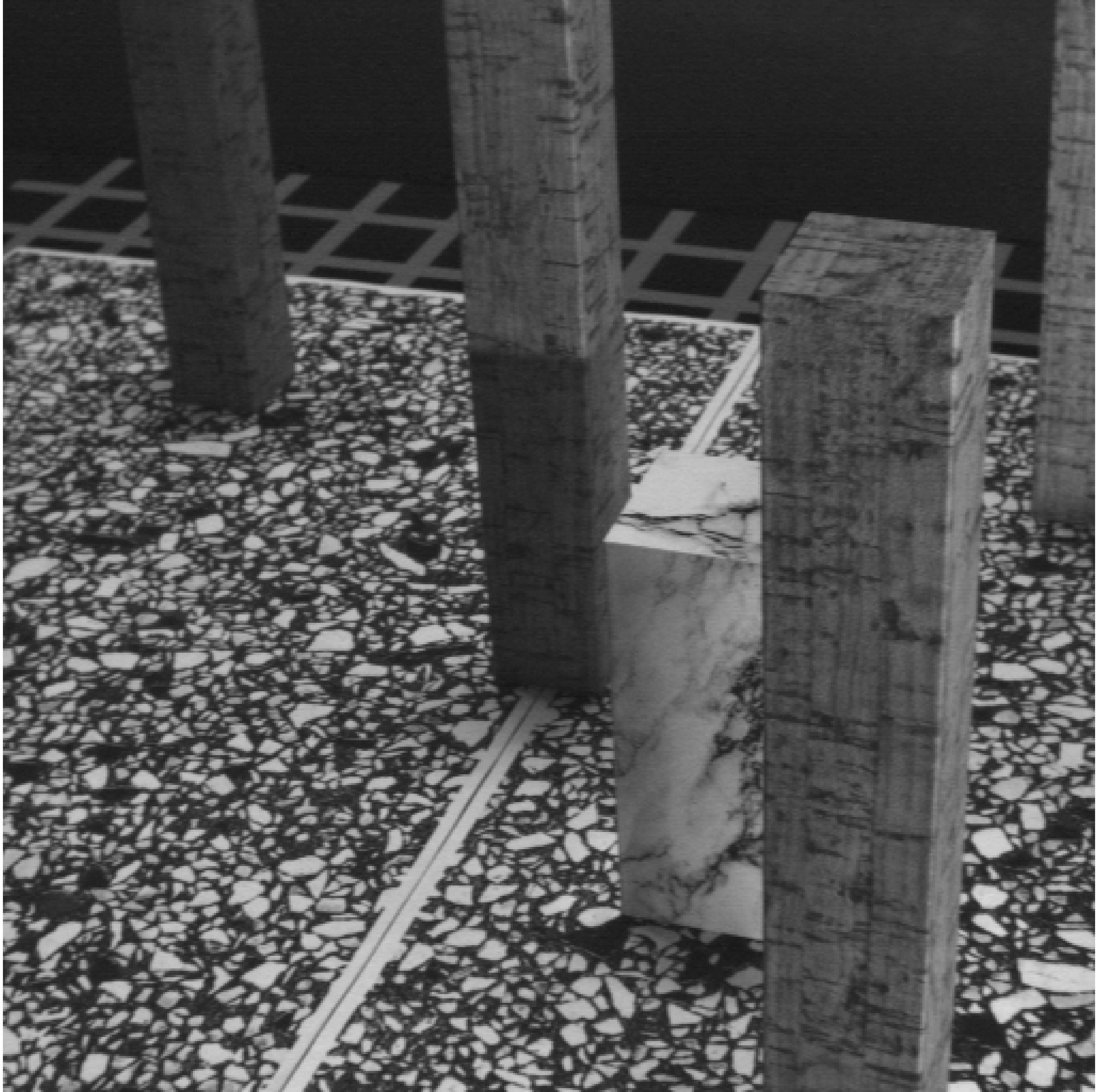
Numerous Applications

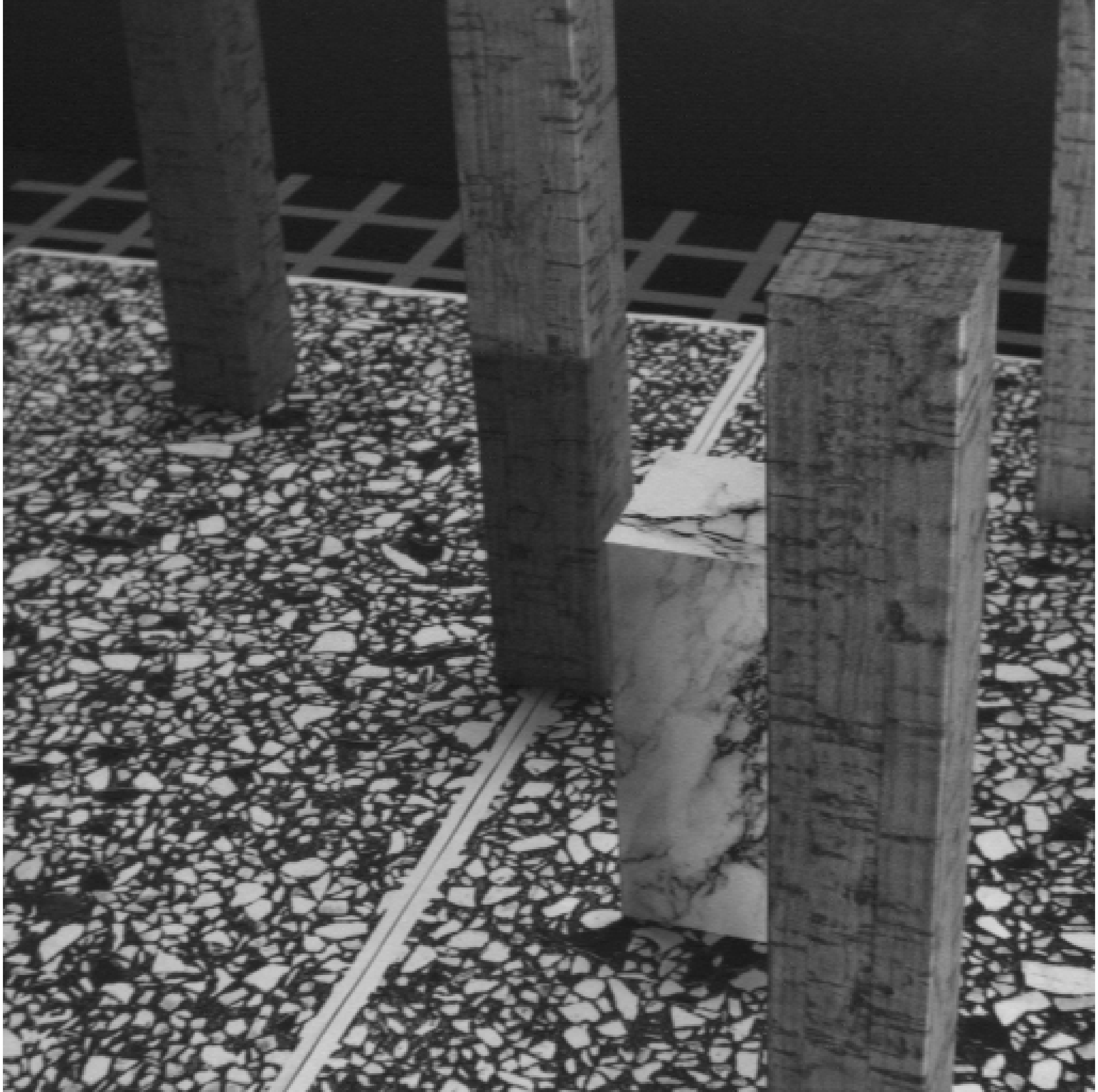
- stereo: calibrated and uncalibrated case
- video processing: deinterlacing, re-timing
- image registration: hairstyle simulation, particle matching

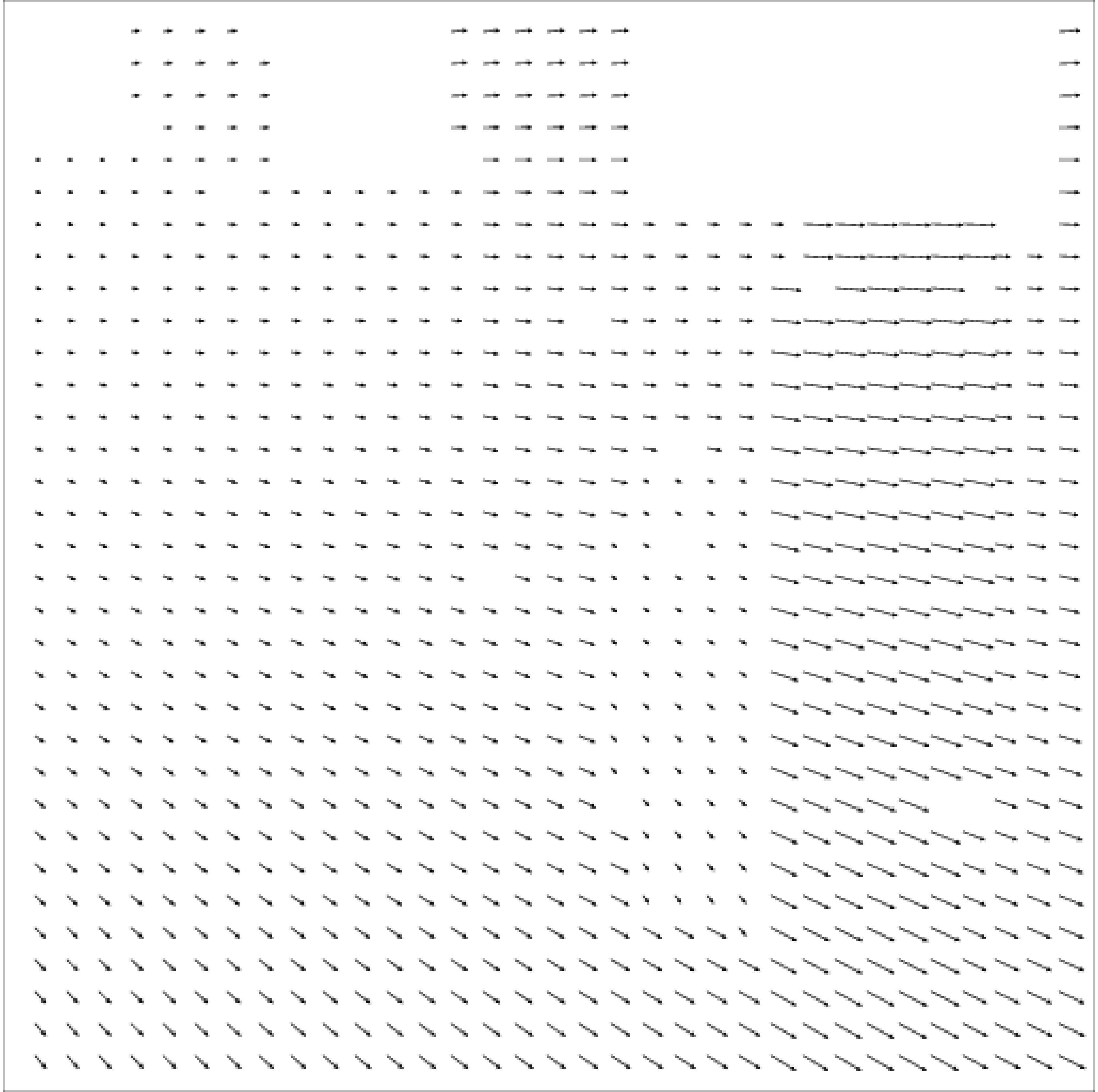
Thank you very much!

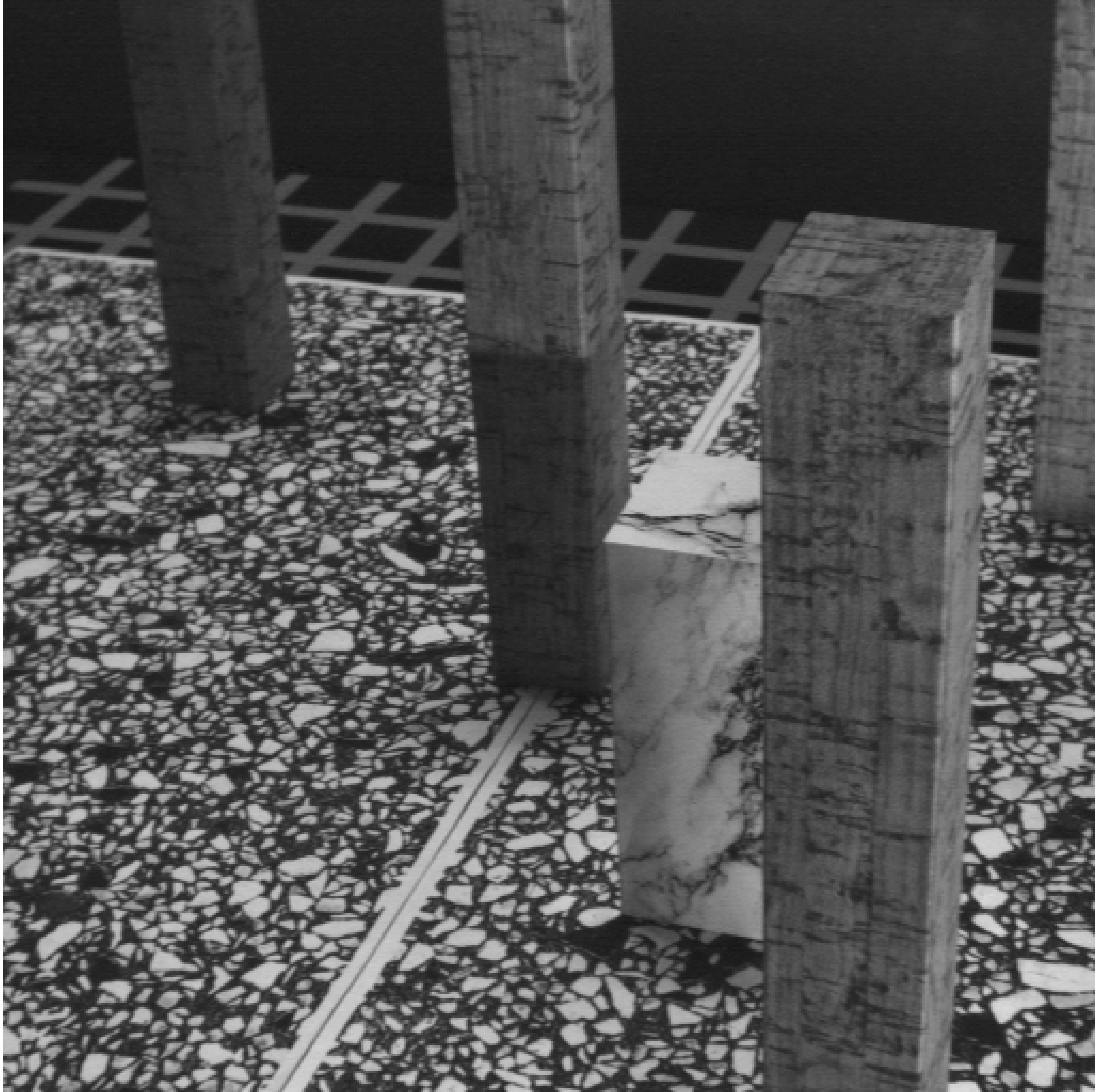
more information:

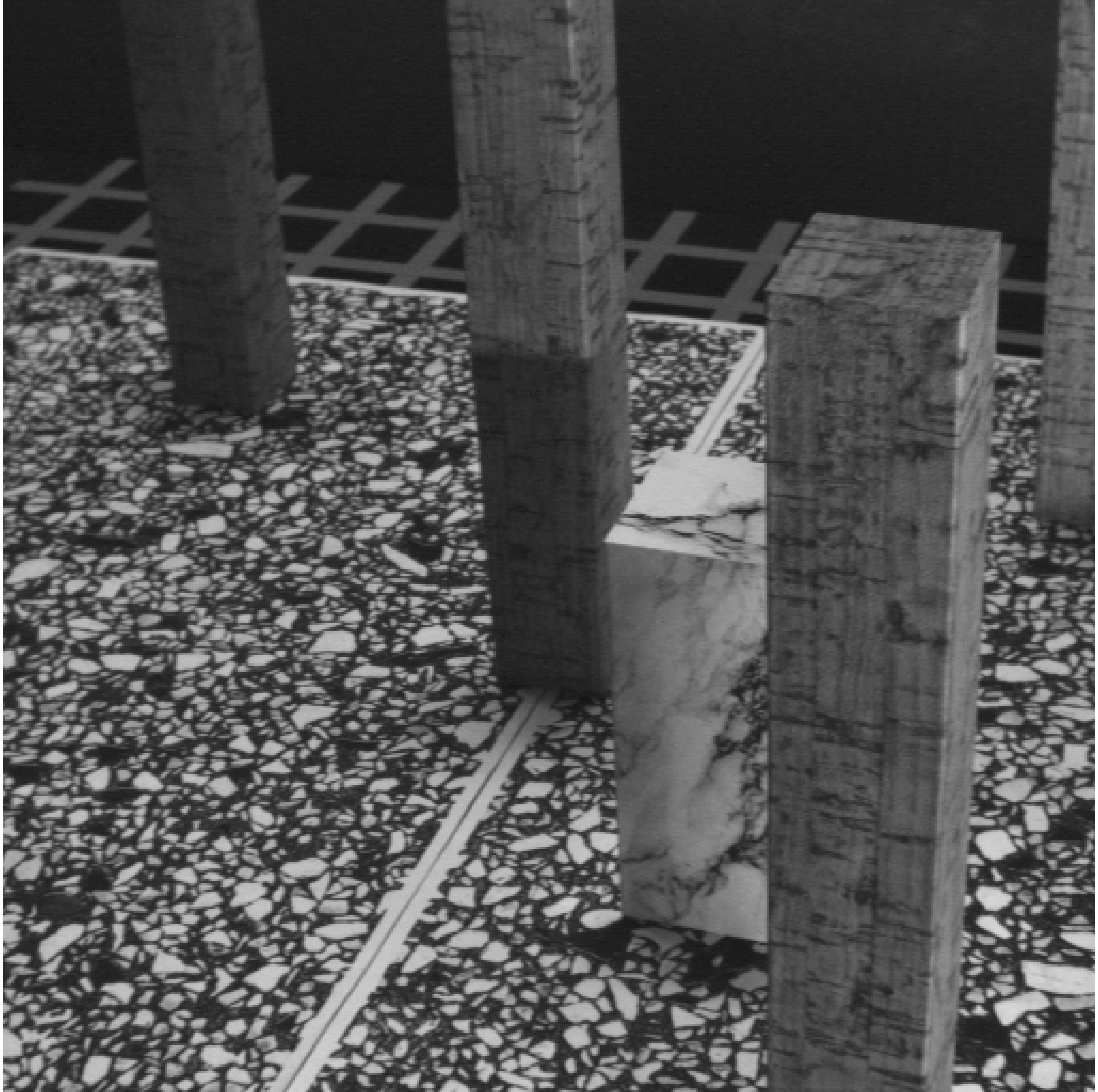
www.mia.uni-saarland.de

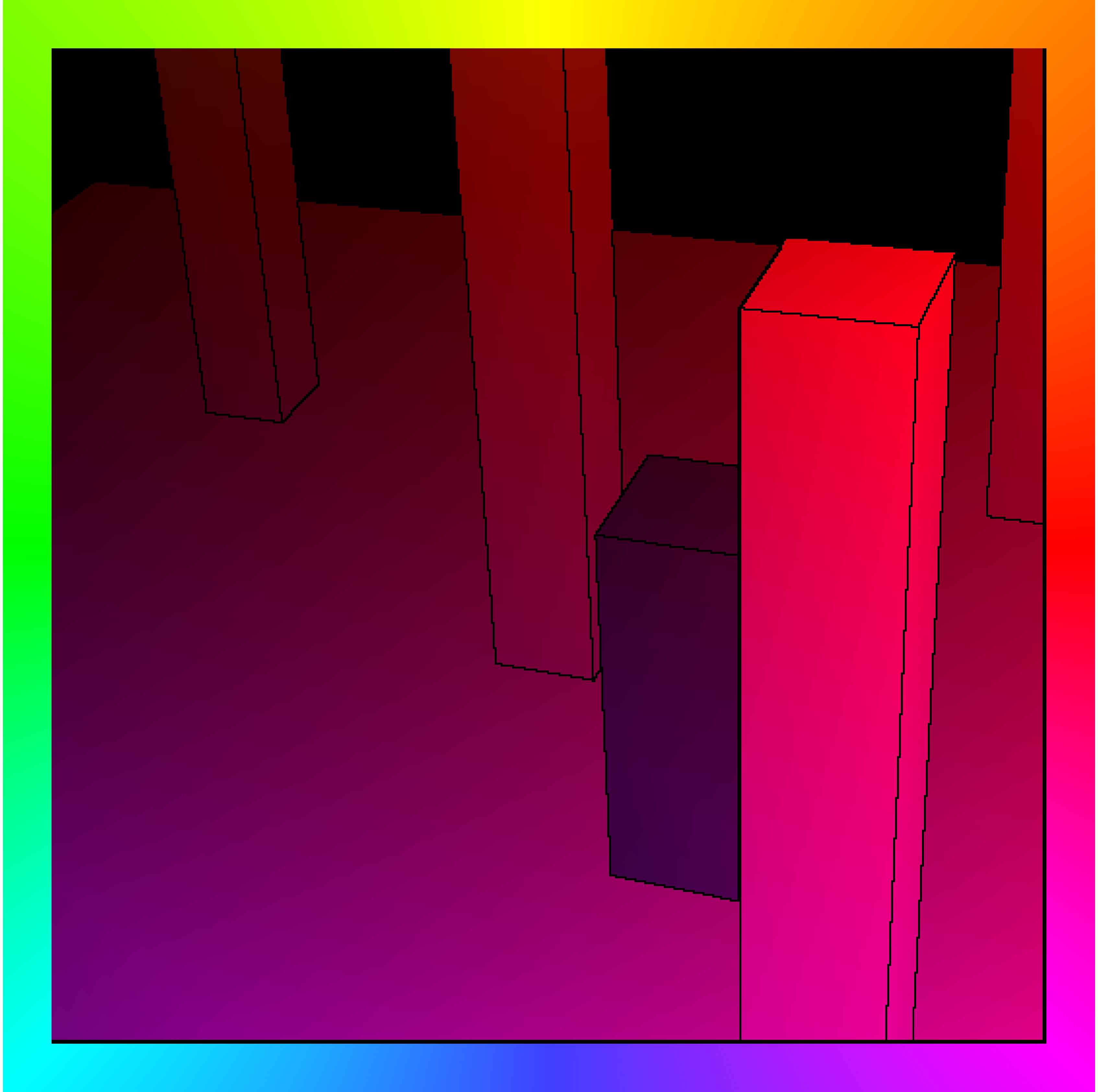


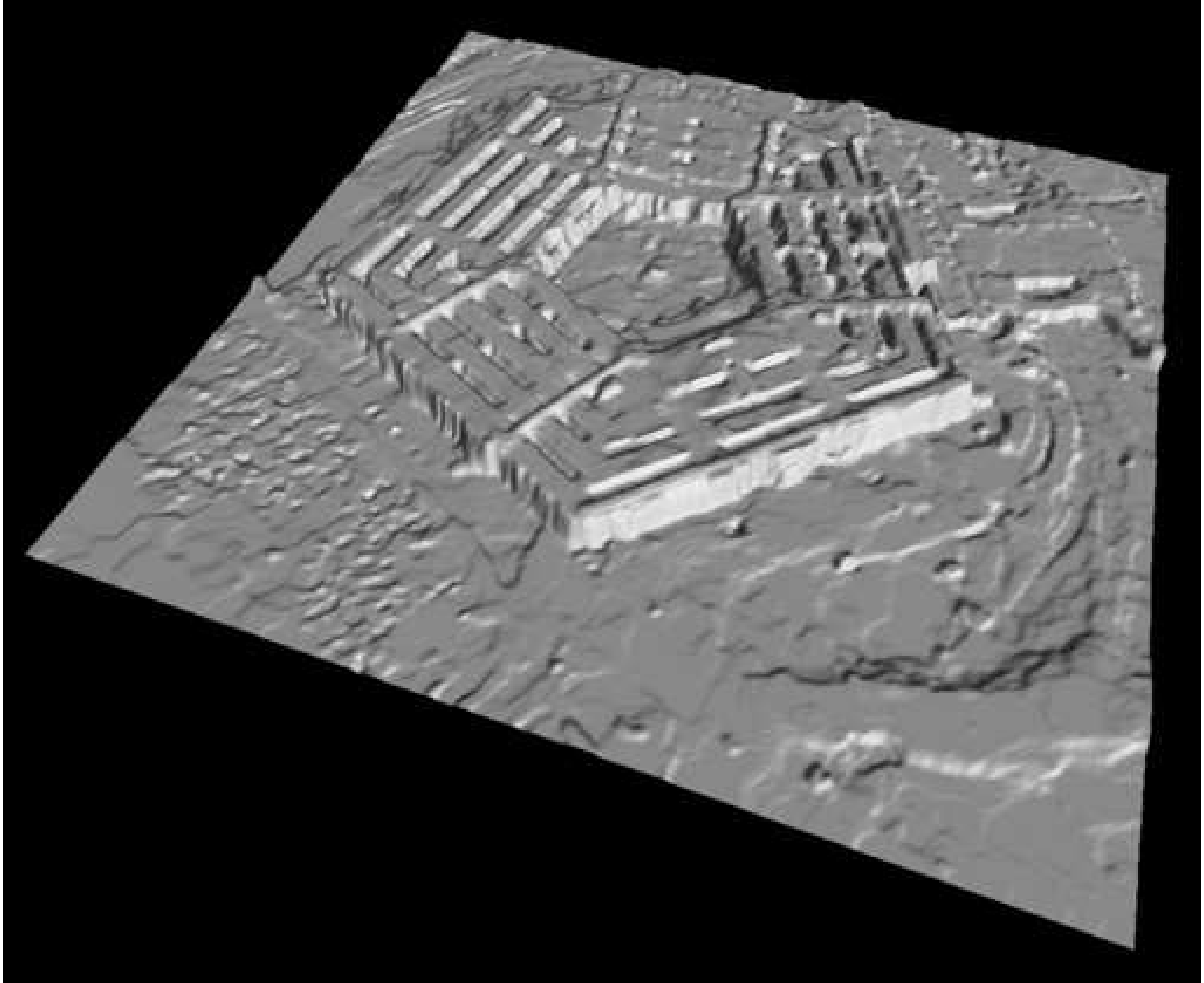




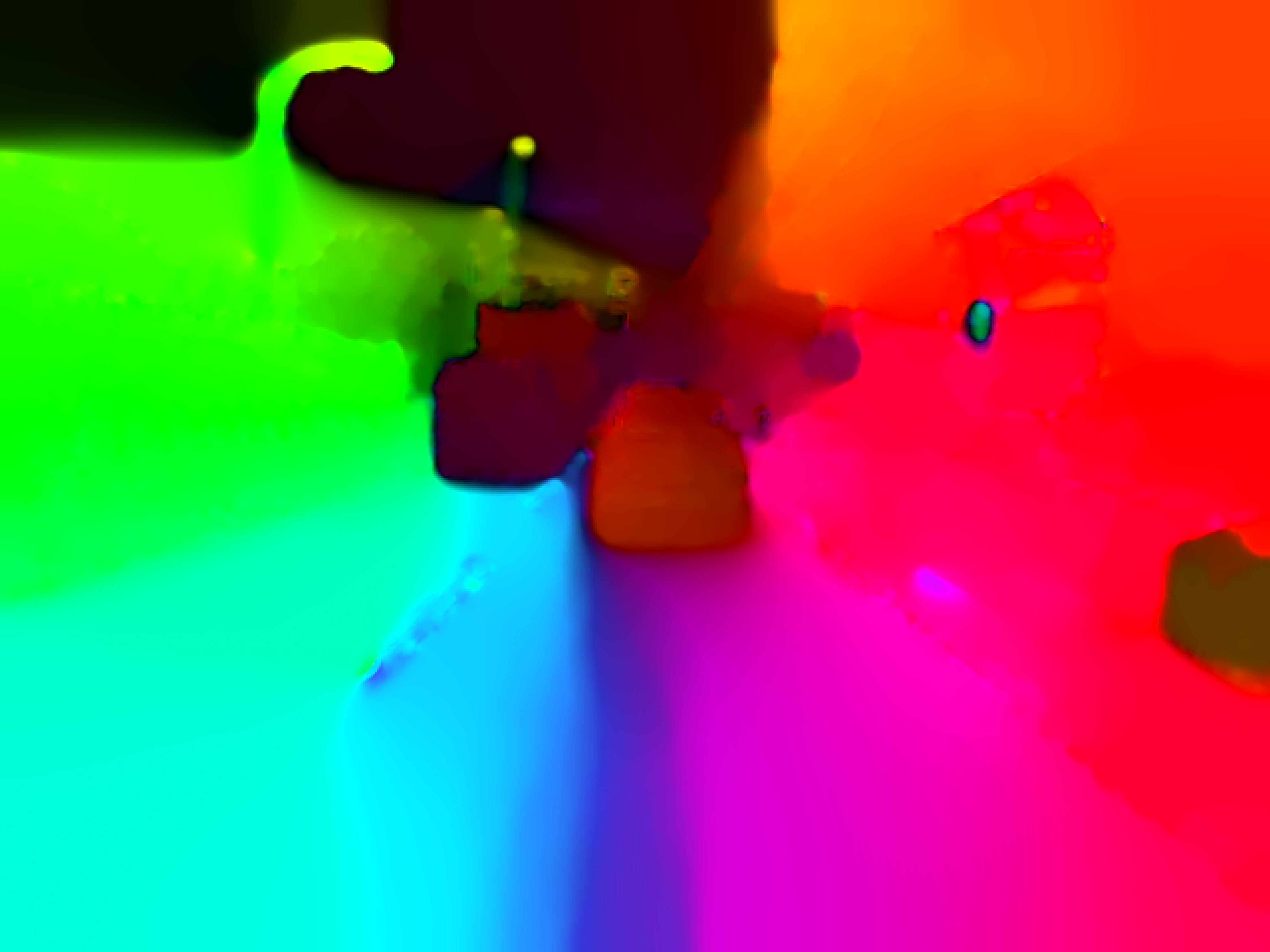


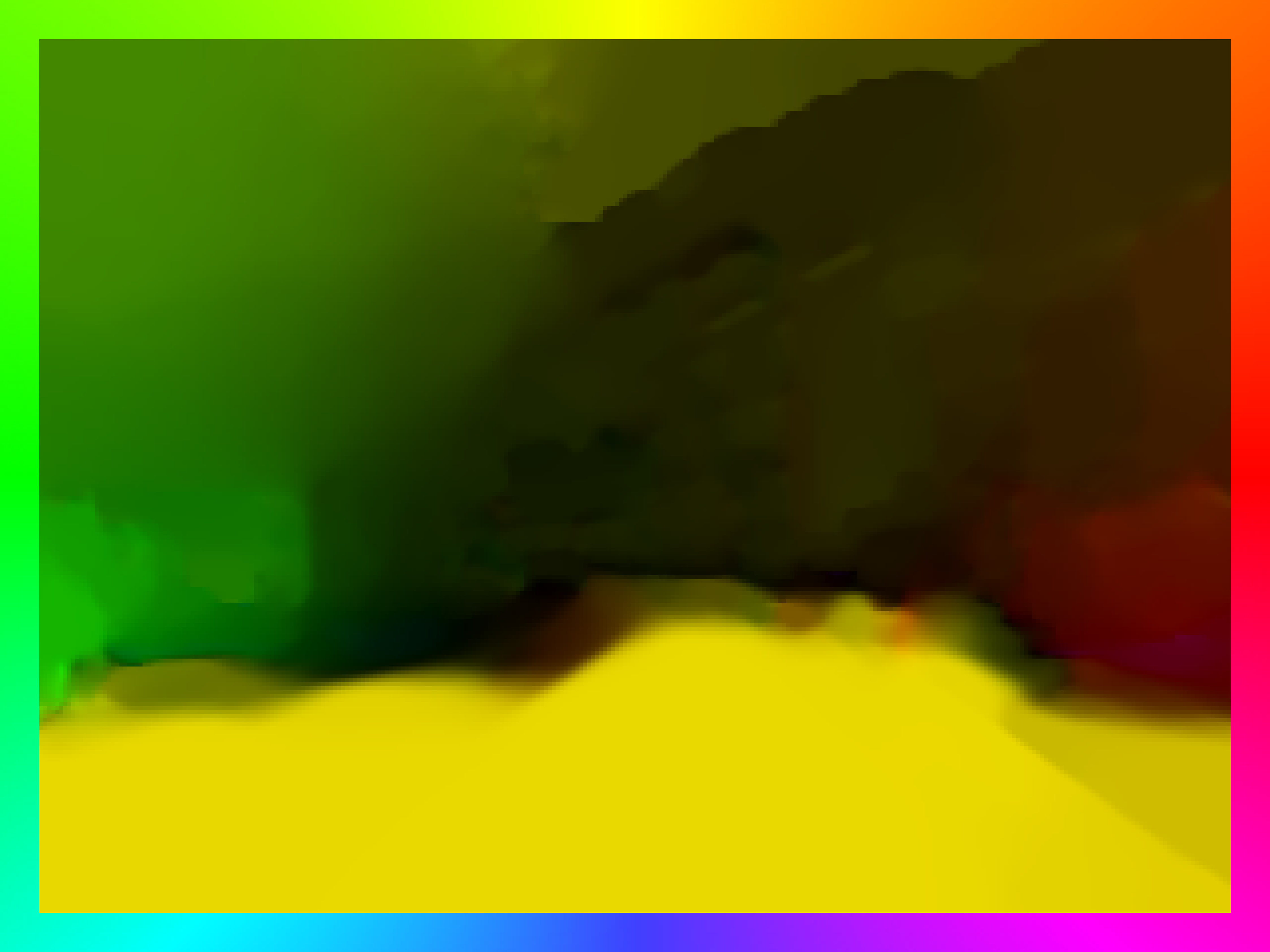


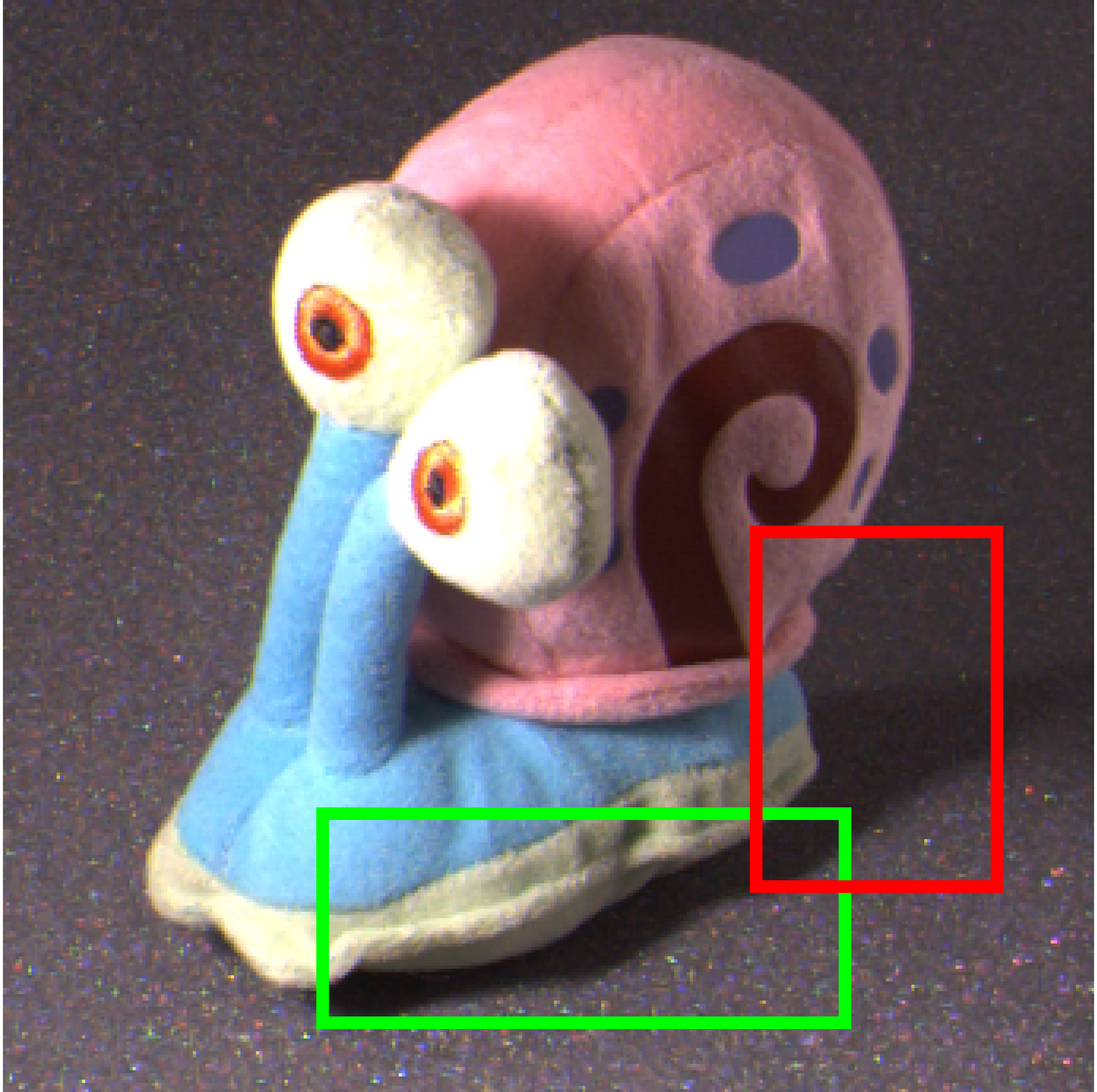


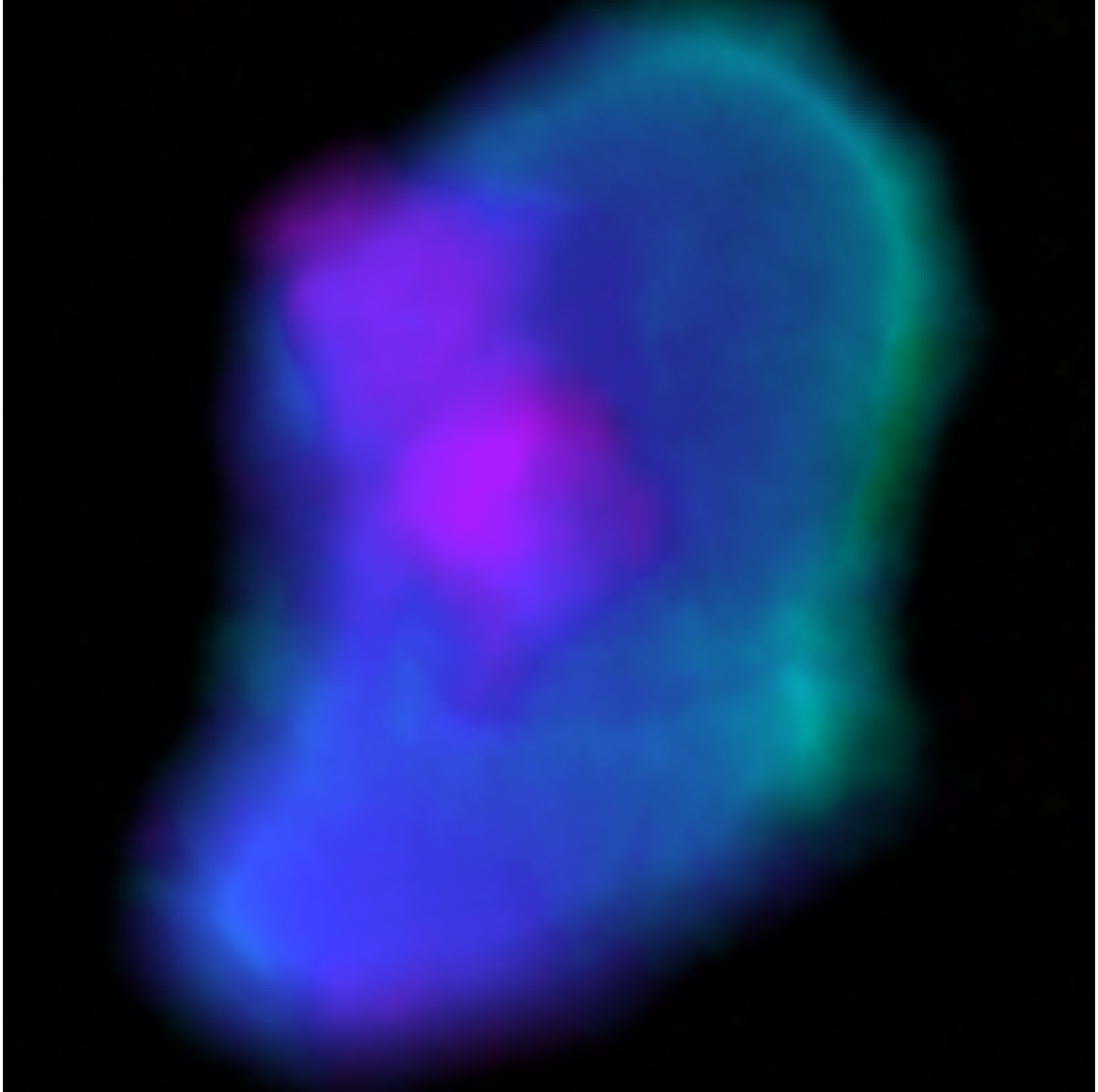


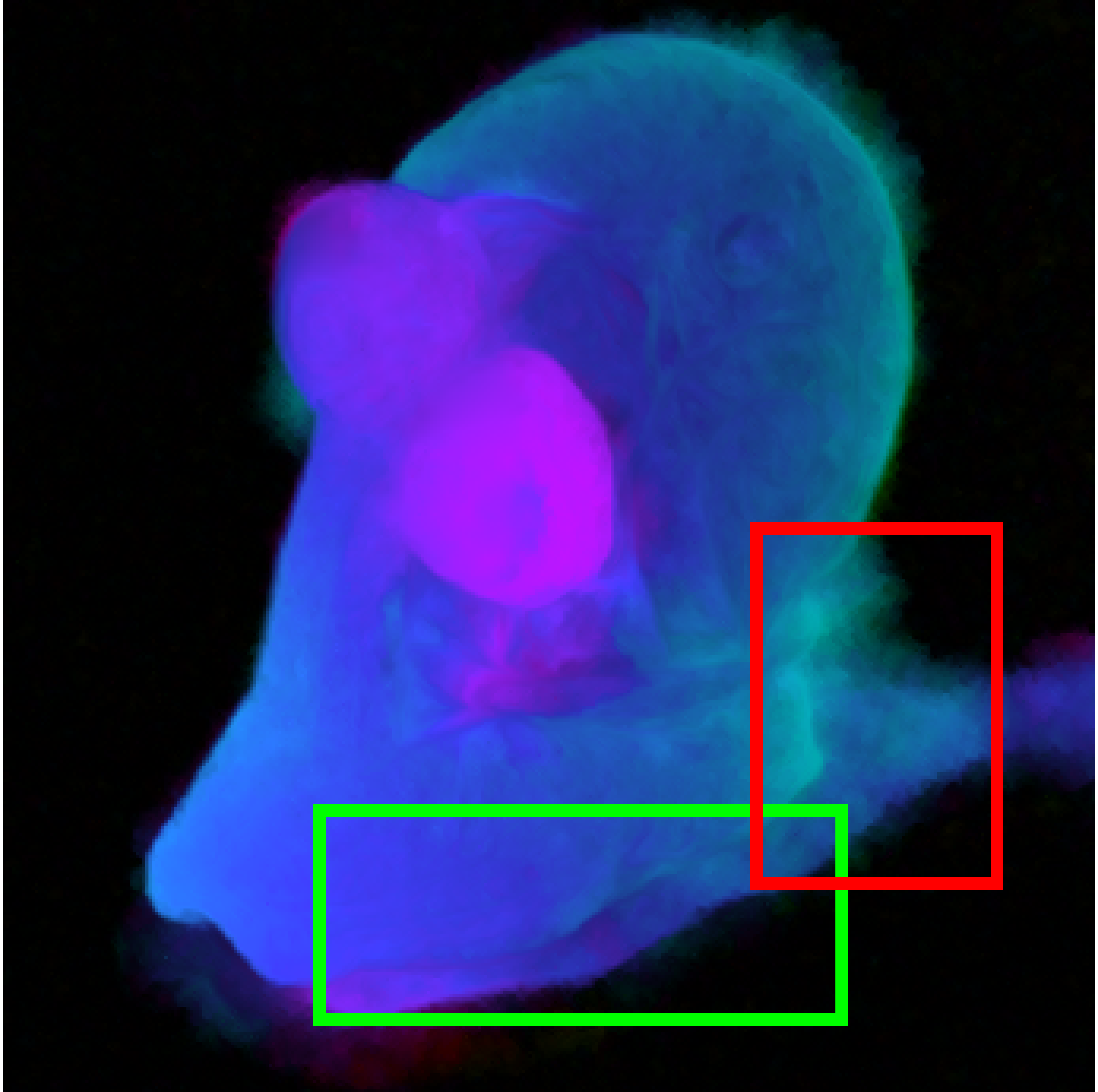

C05:03:05:09

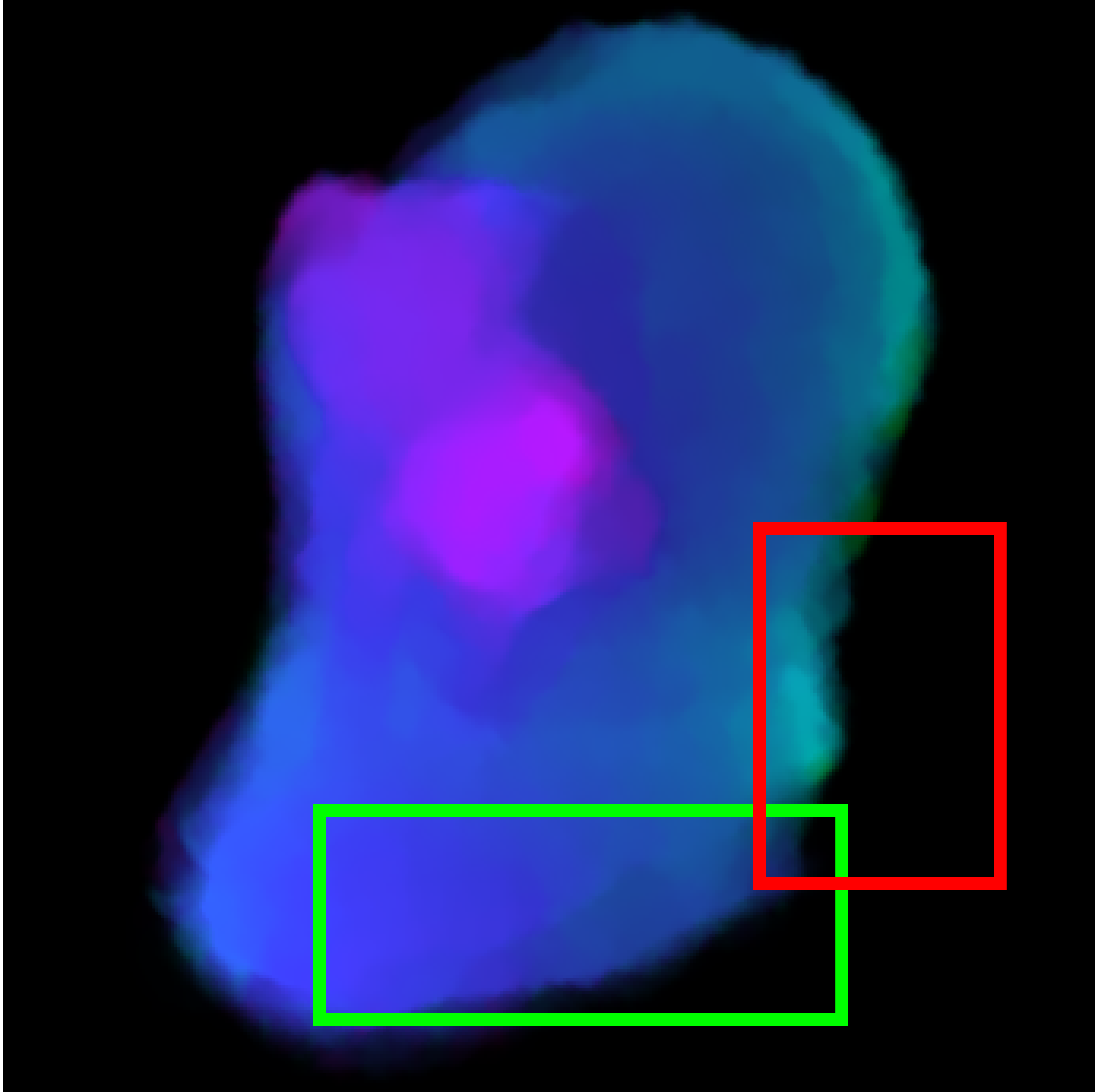


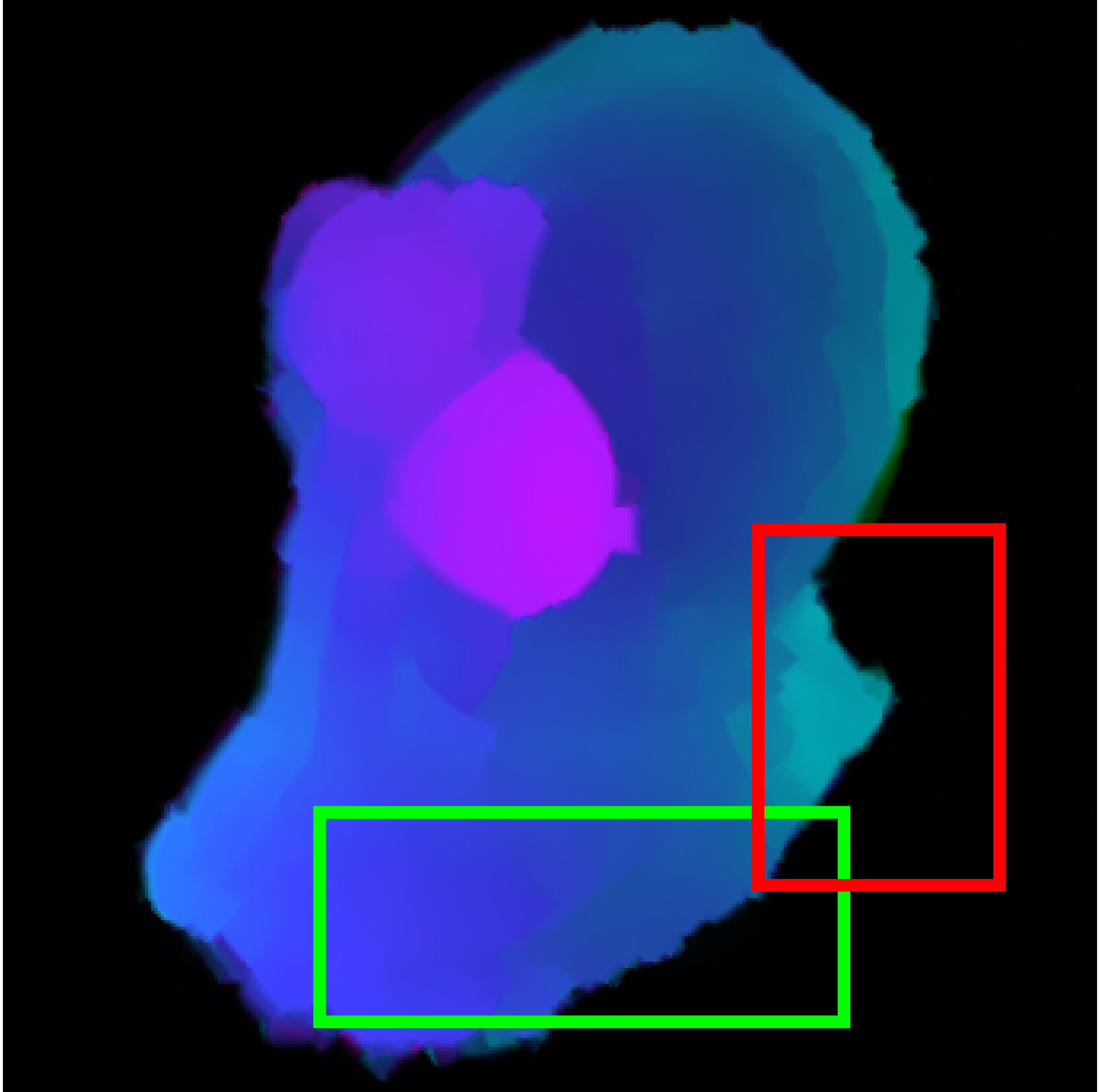


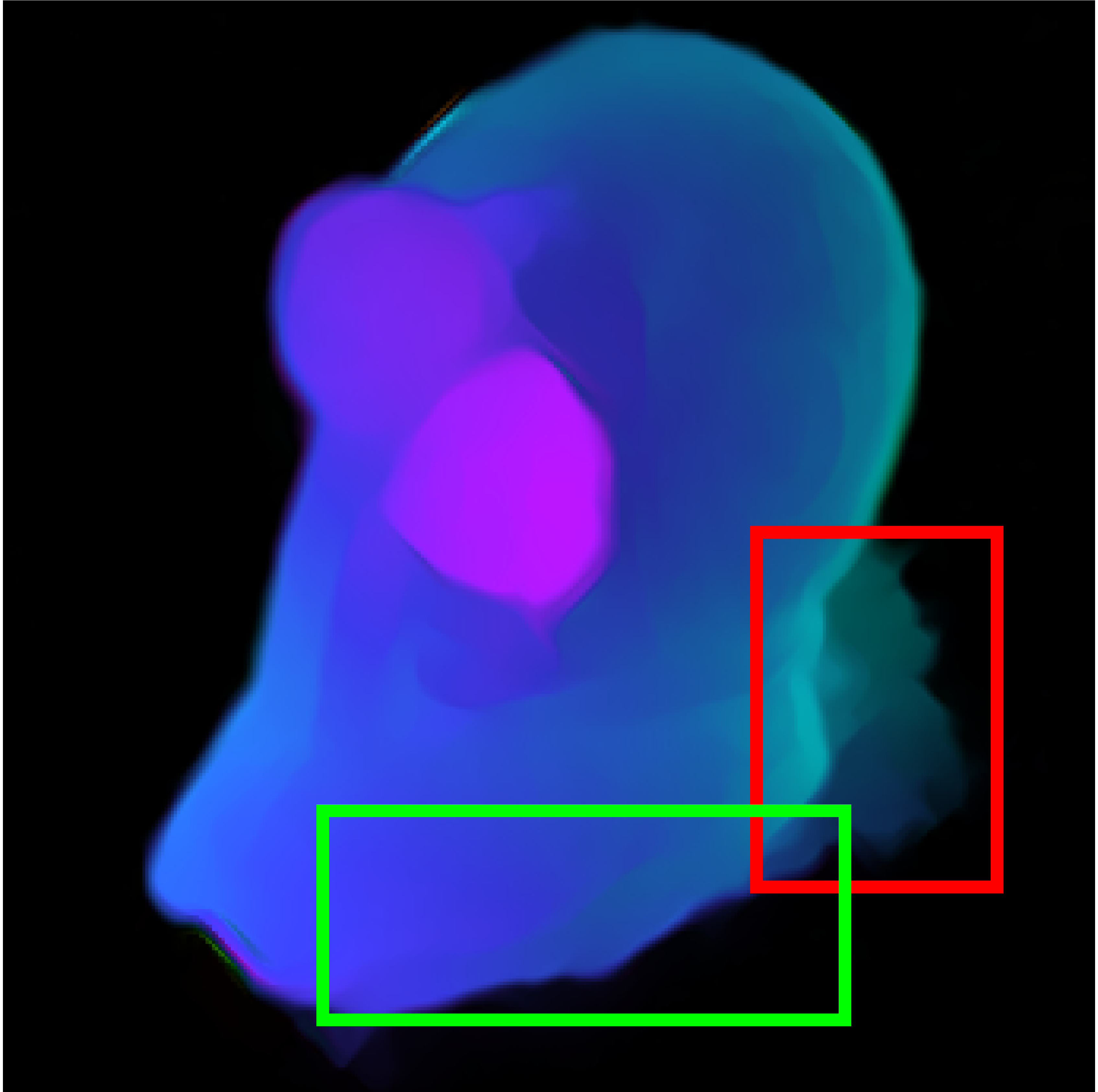


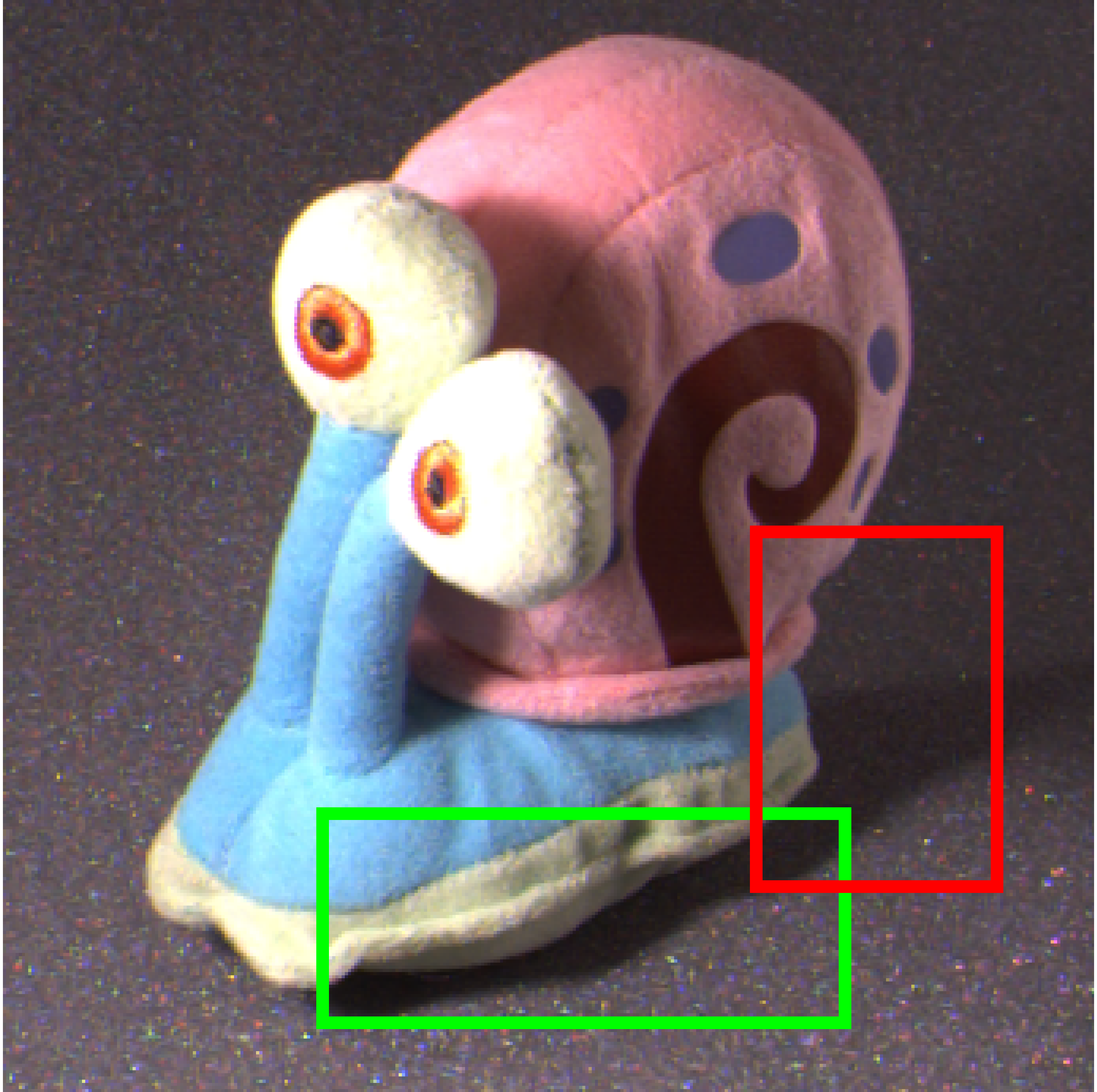


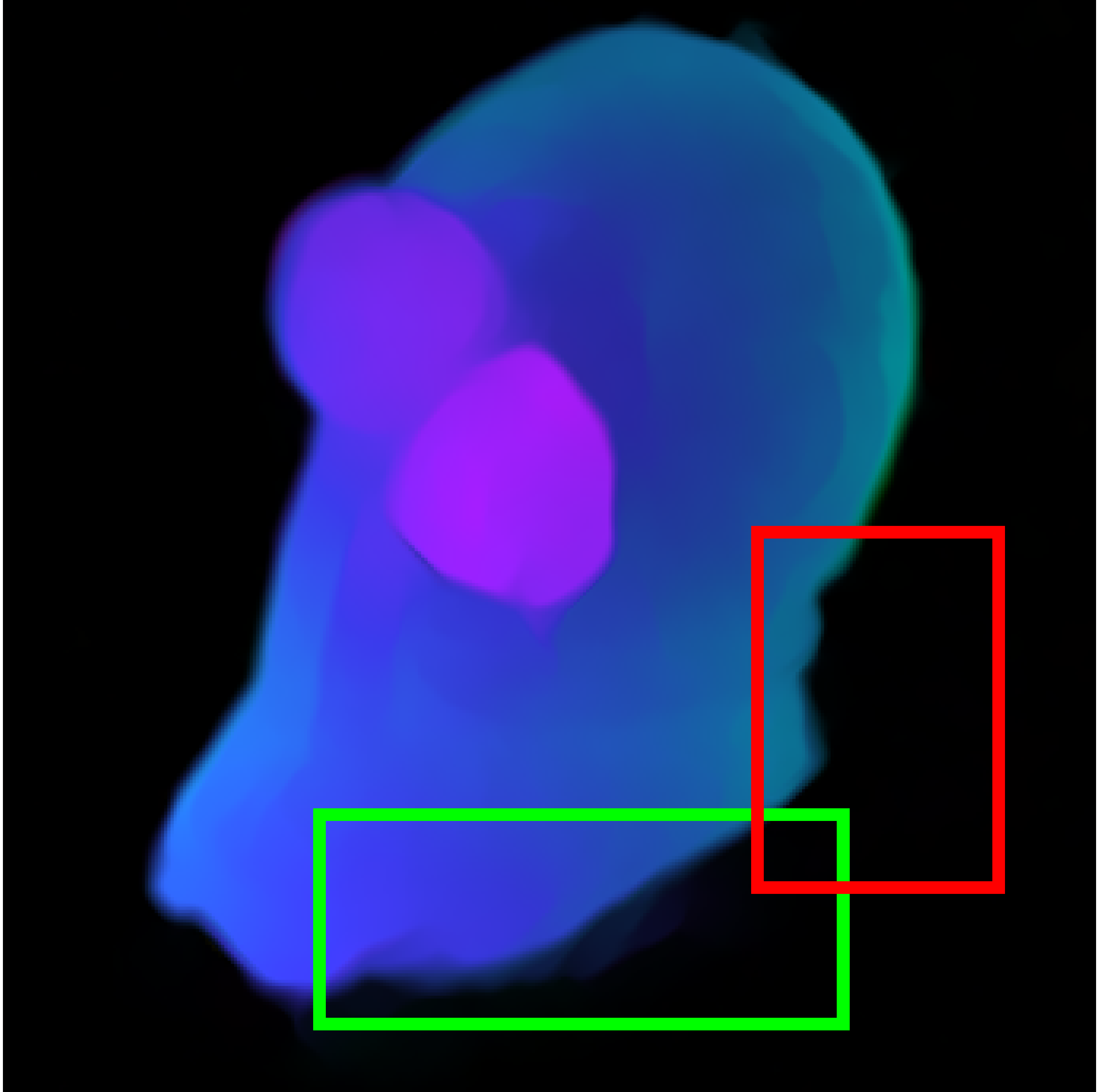


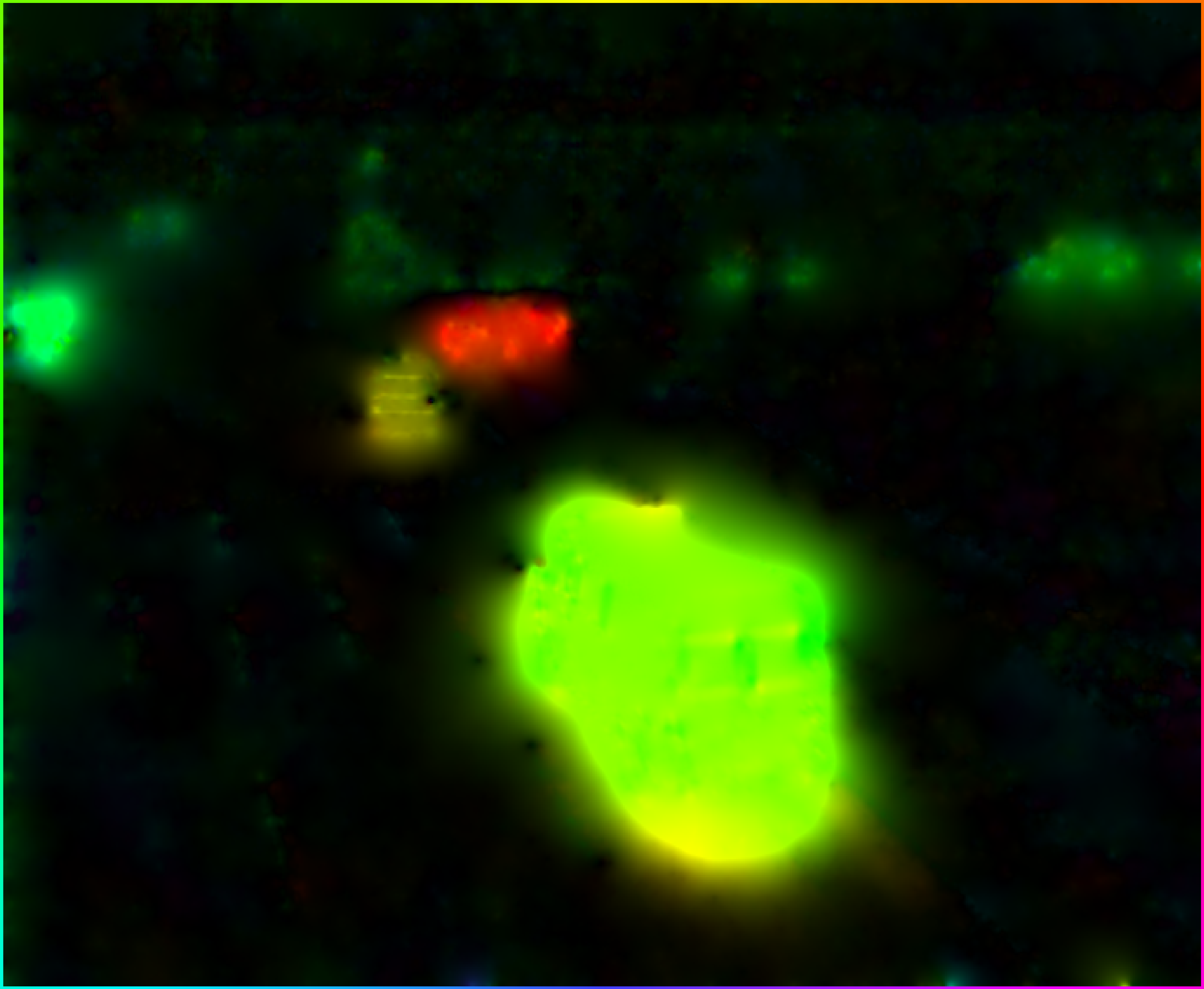


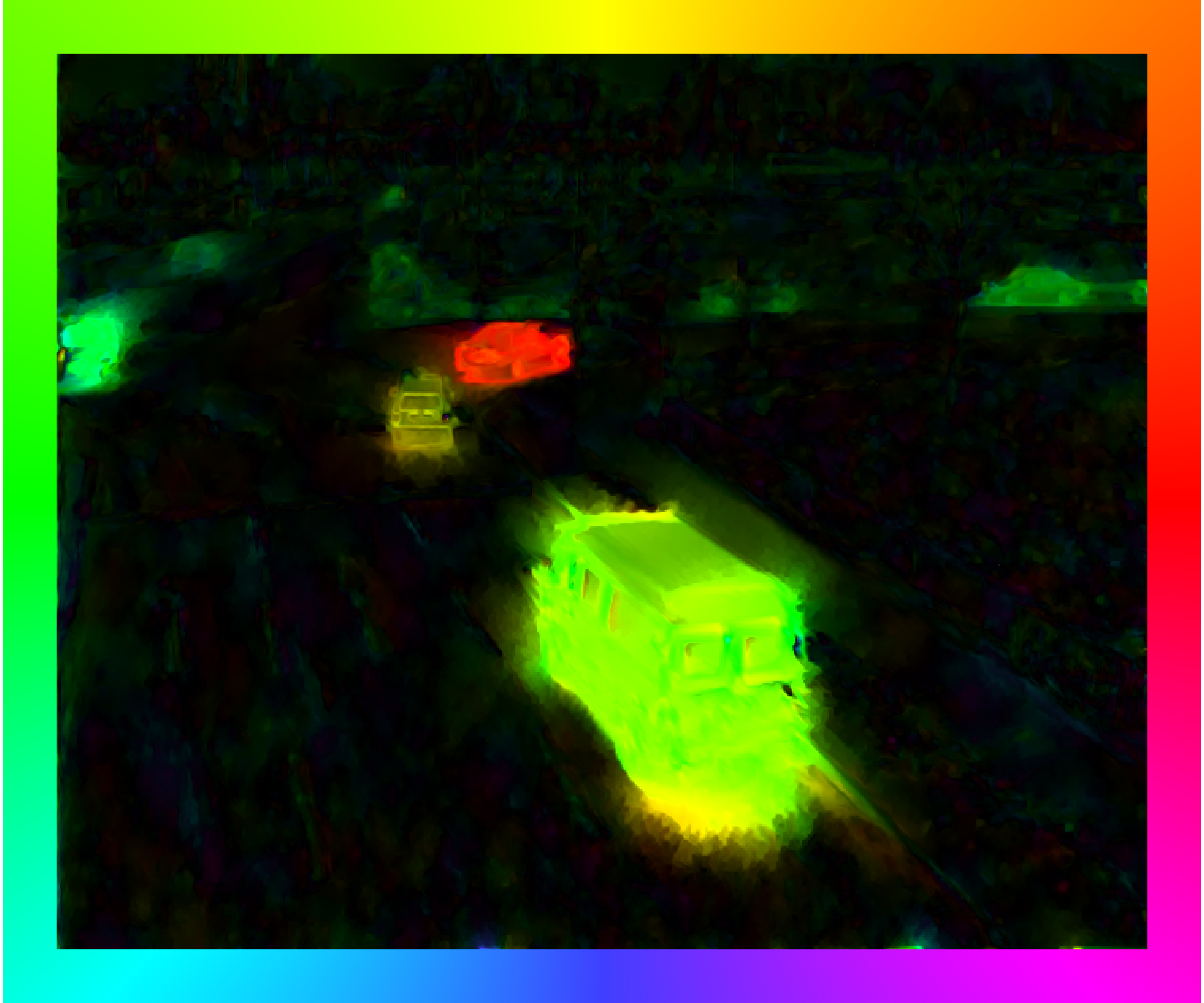


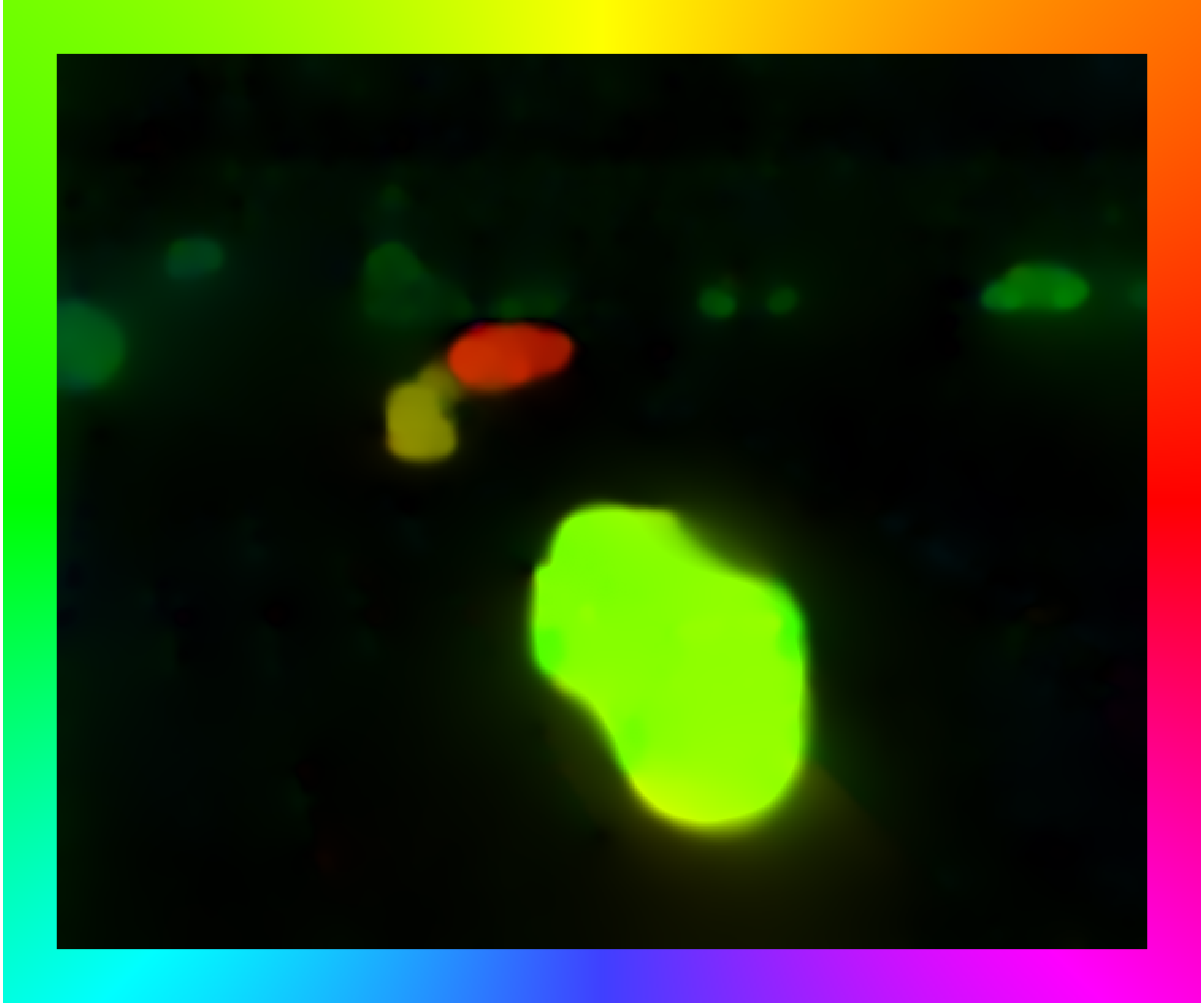


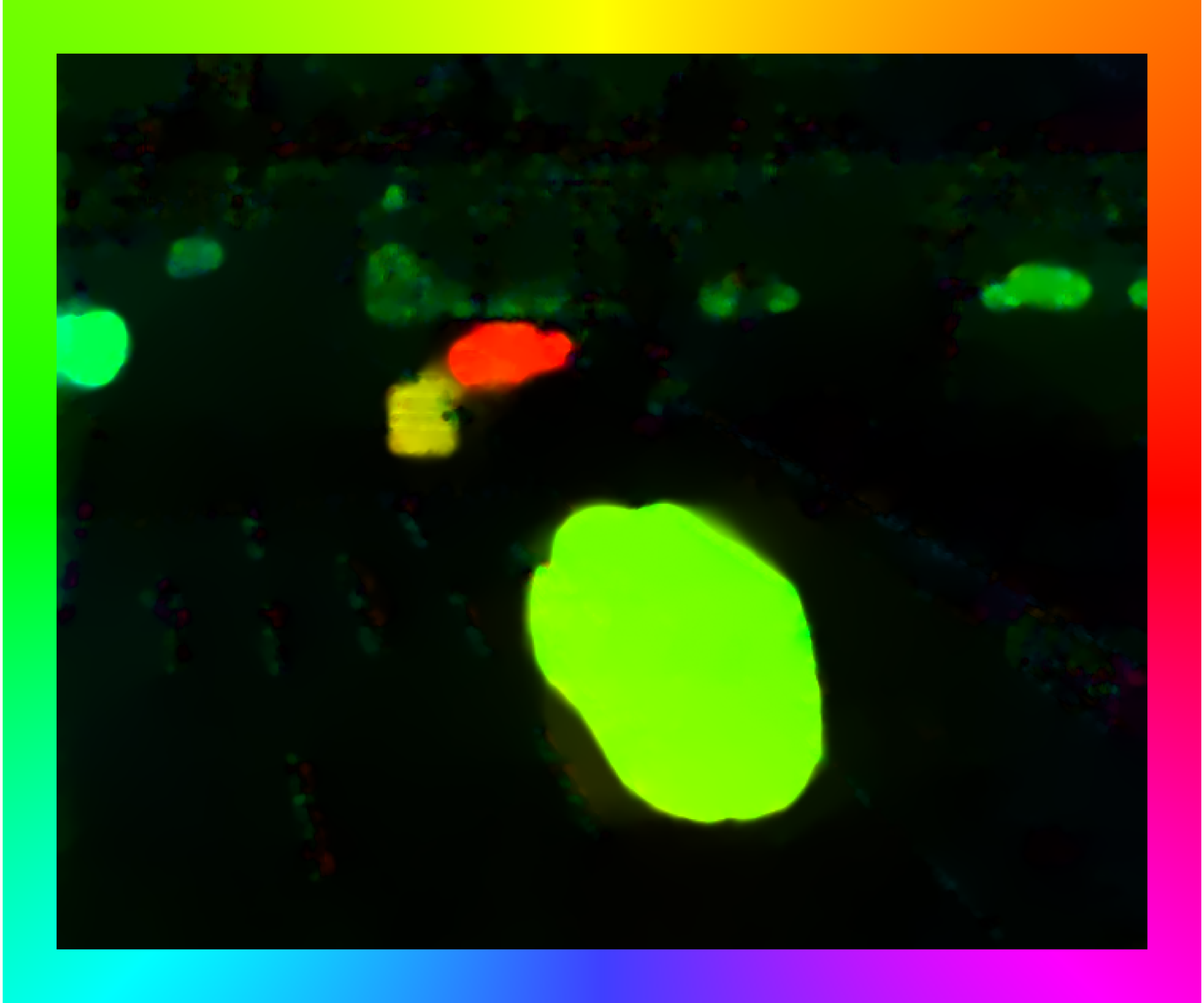




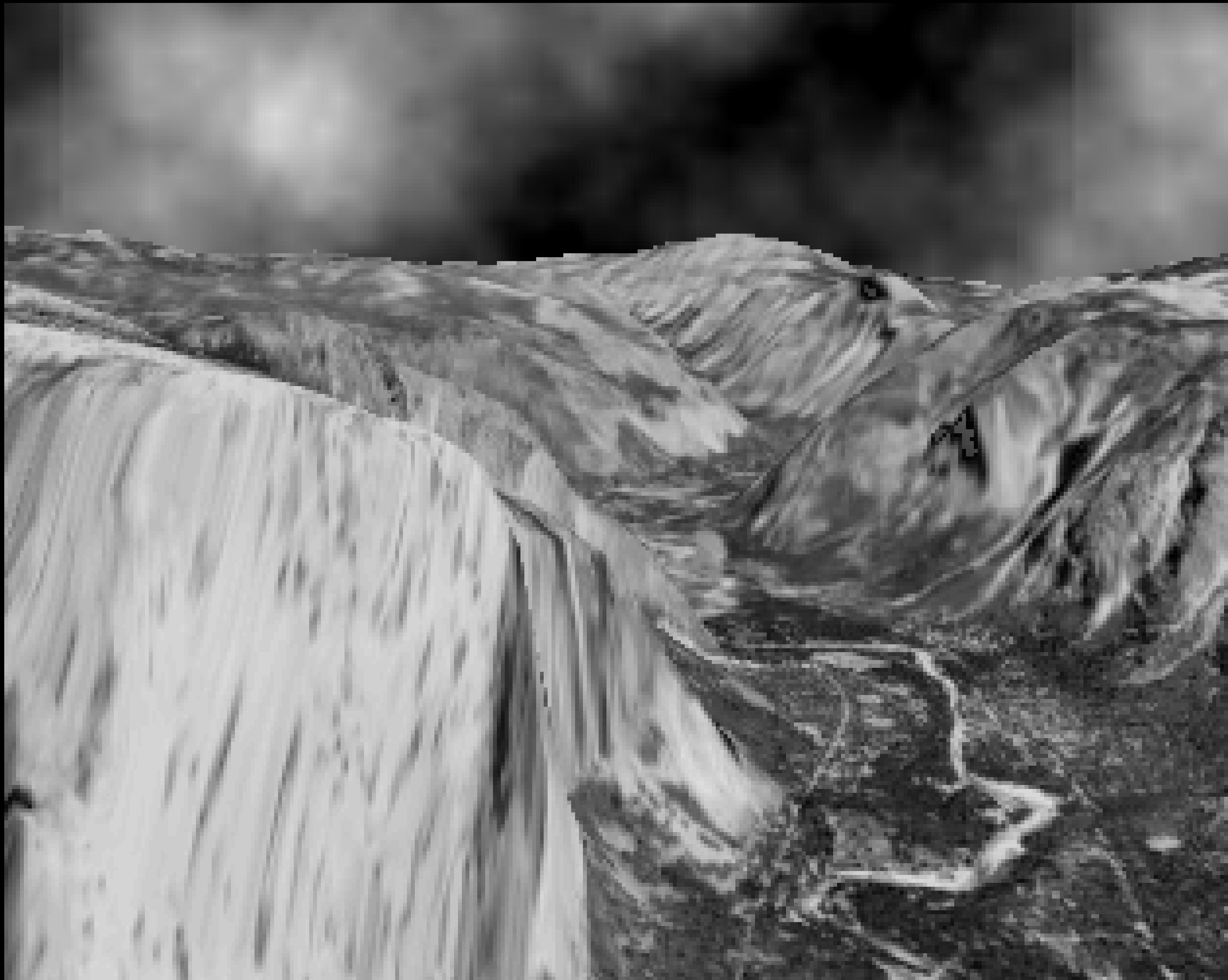


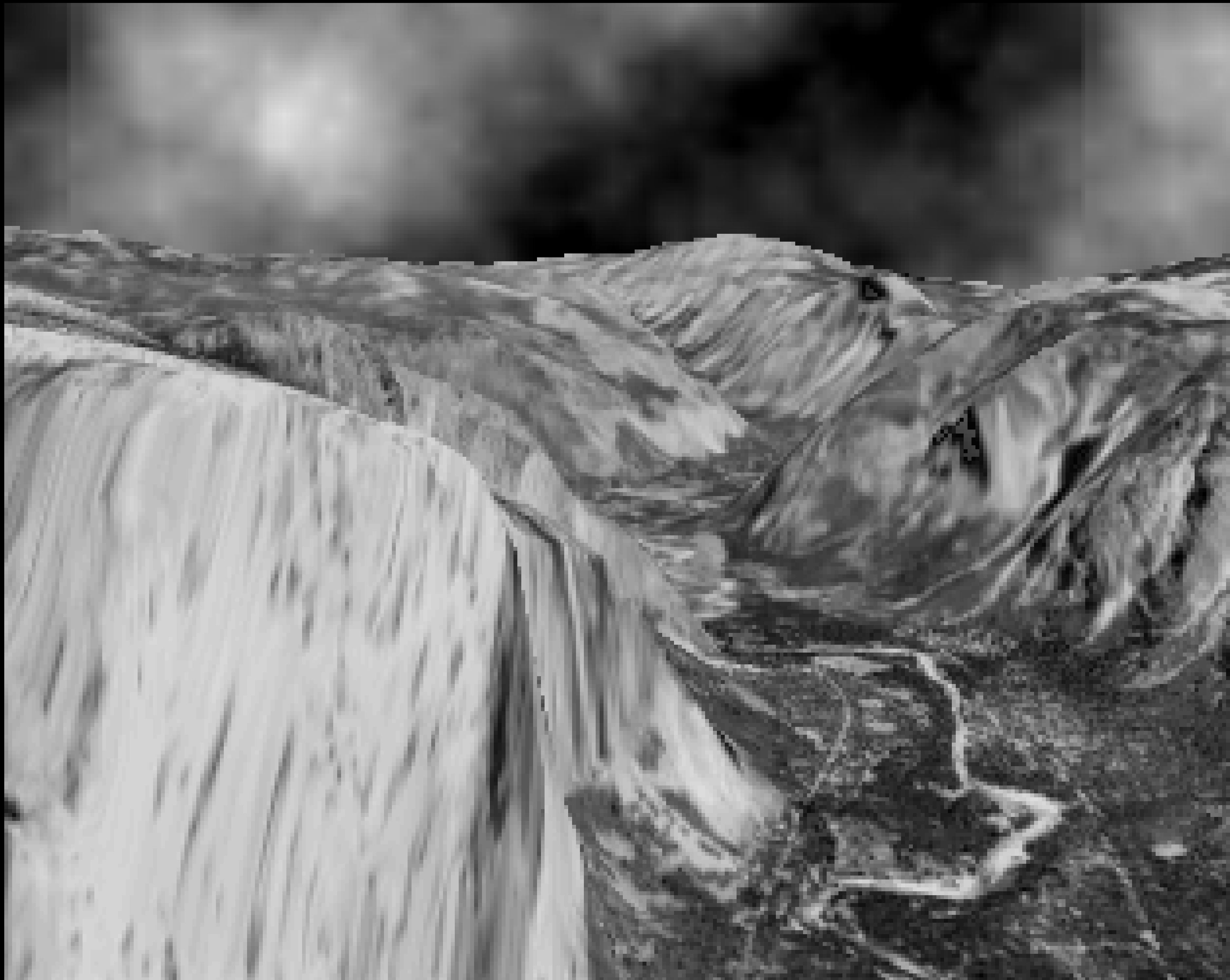


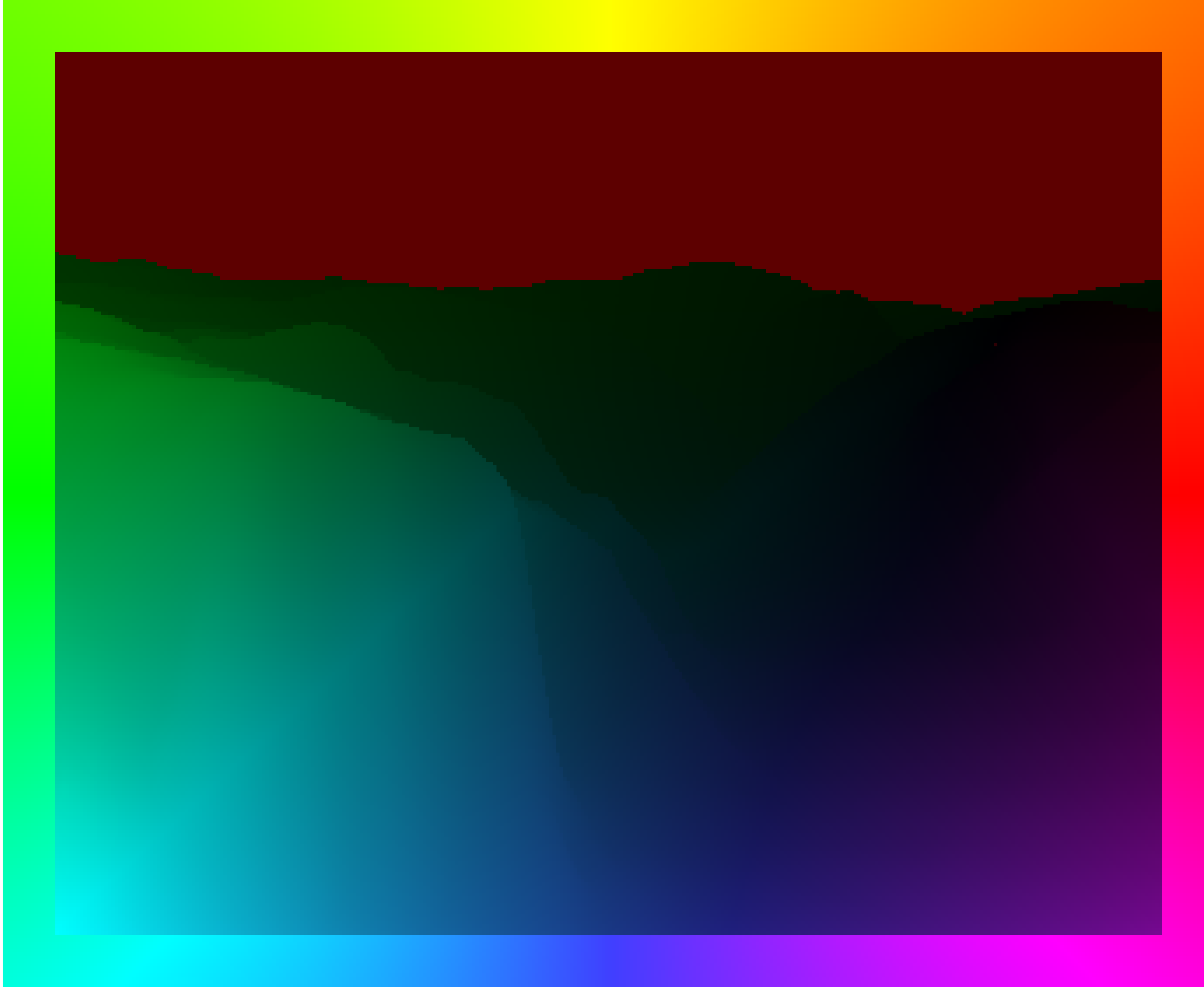


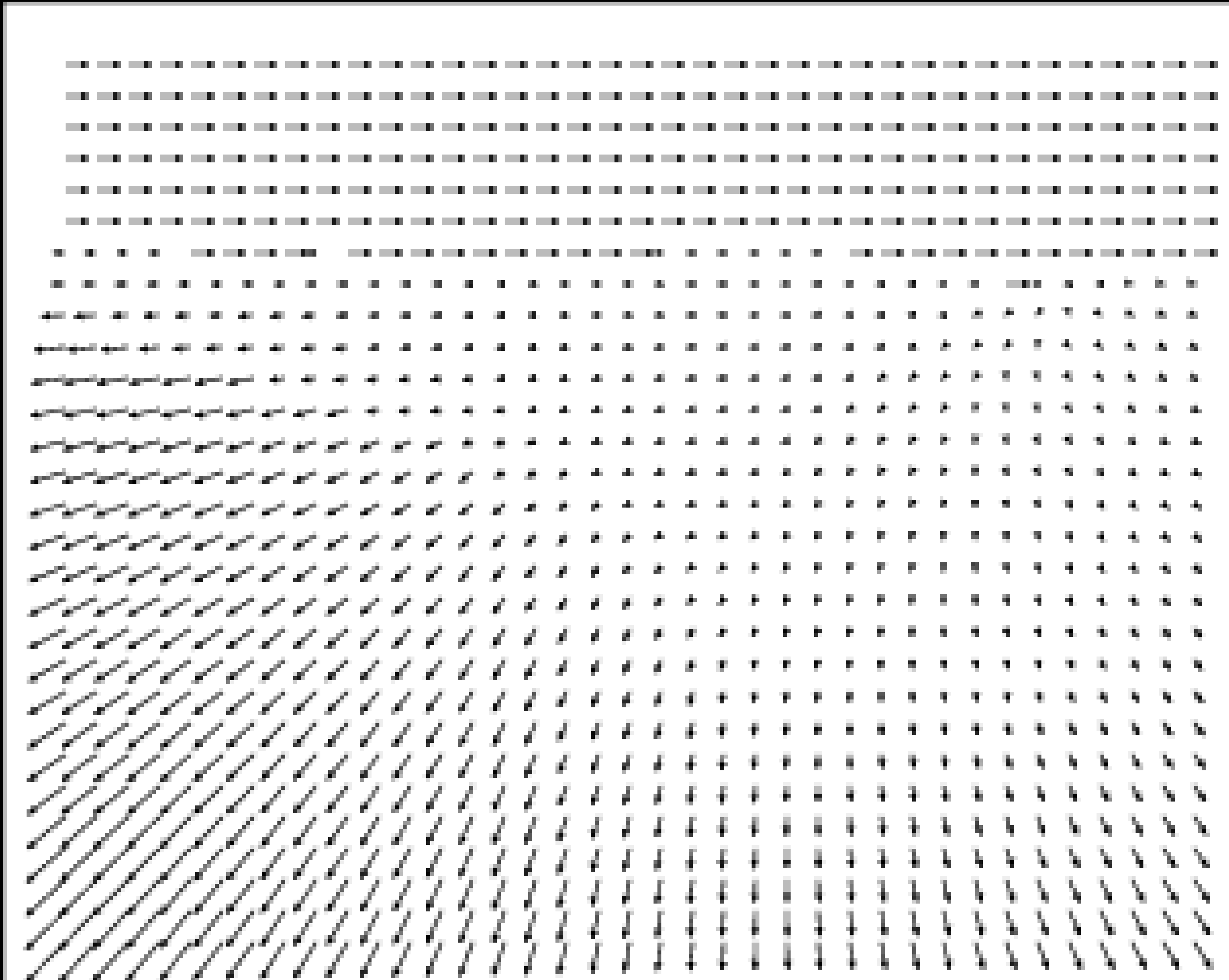


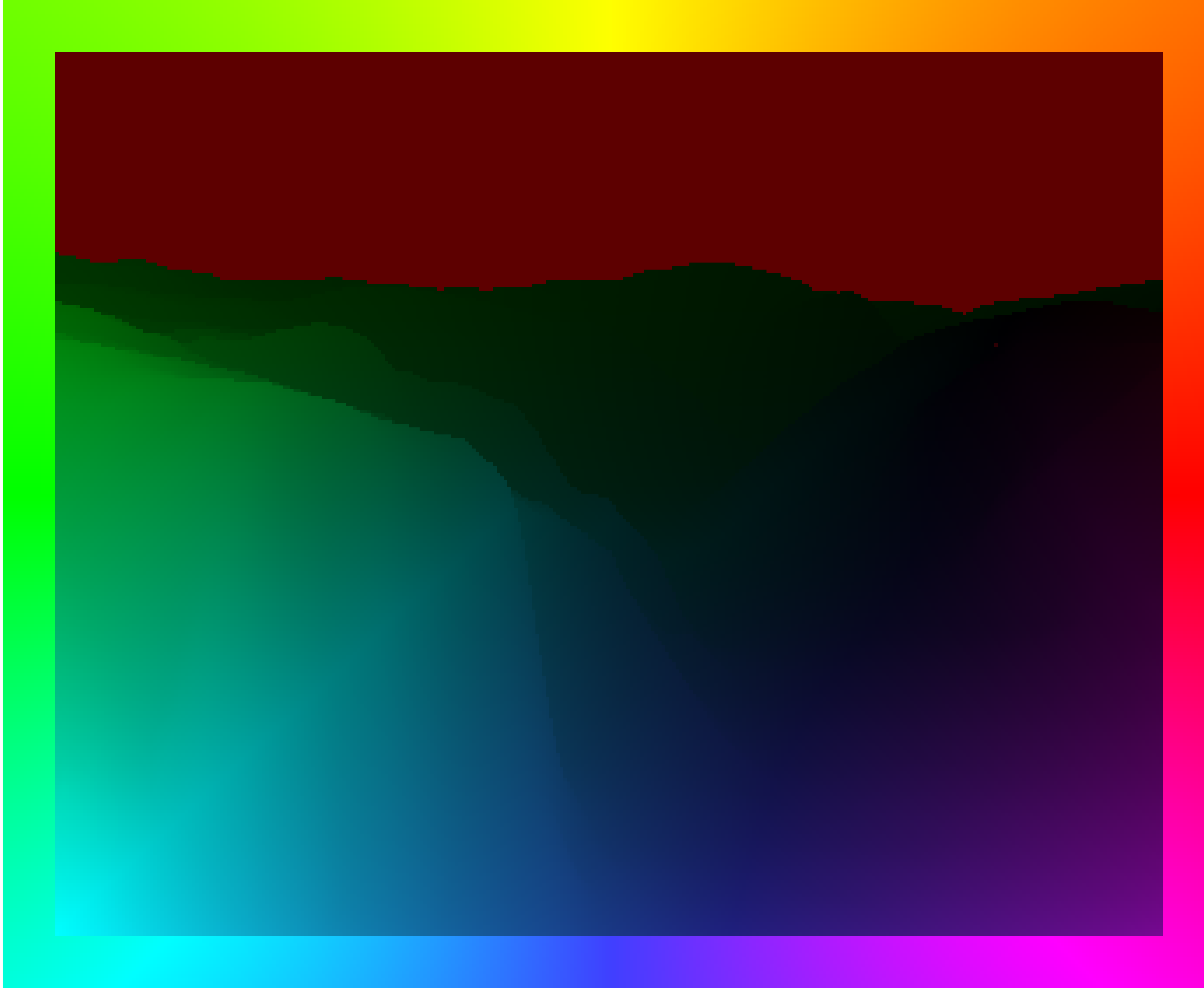

C05:03:05:02

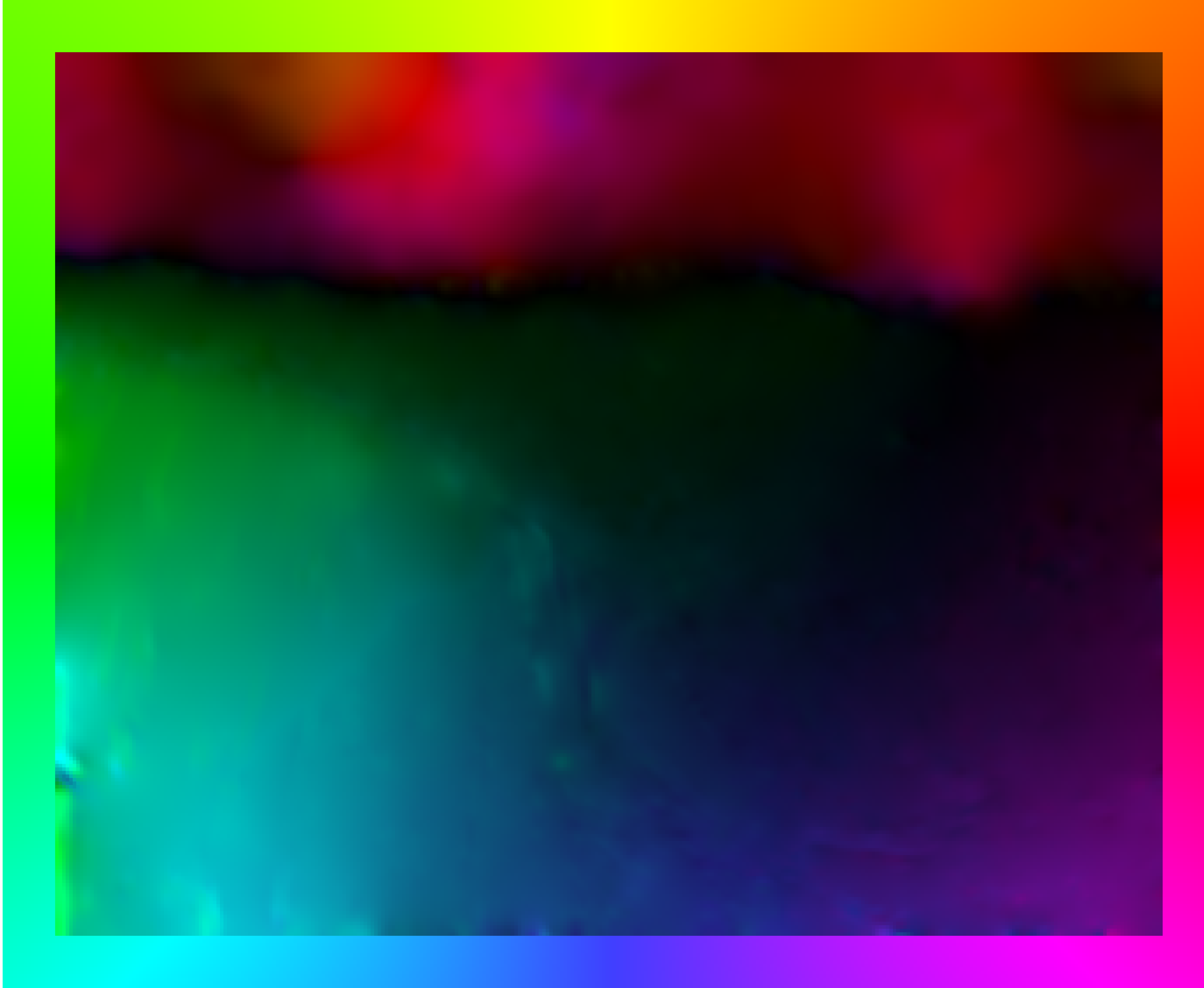


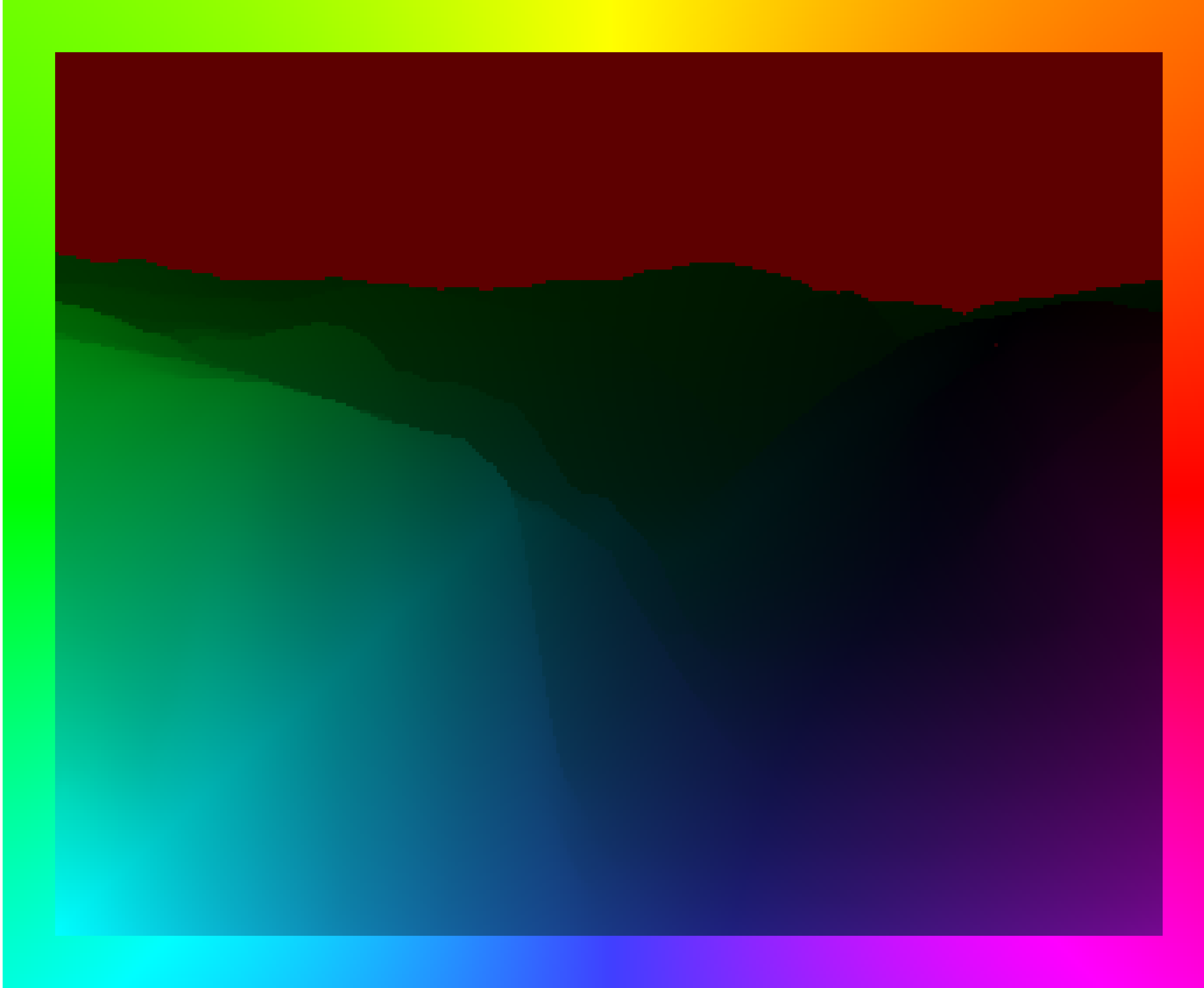


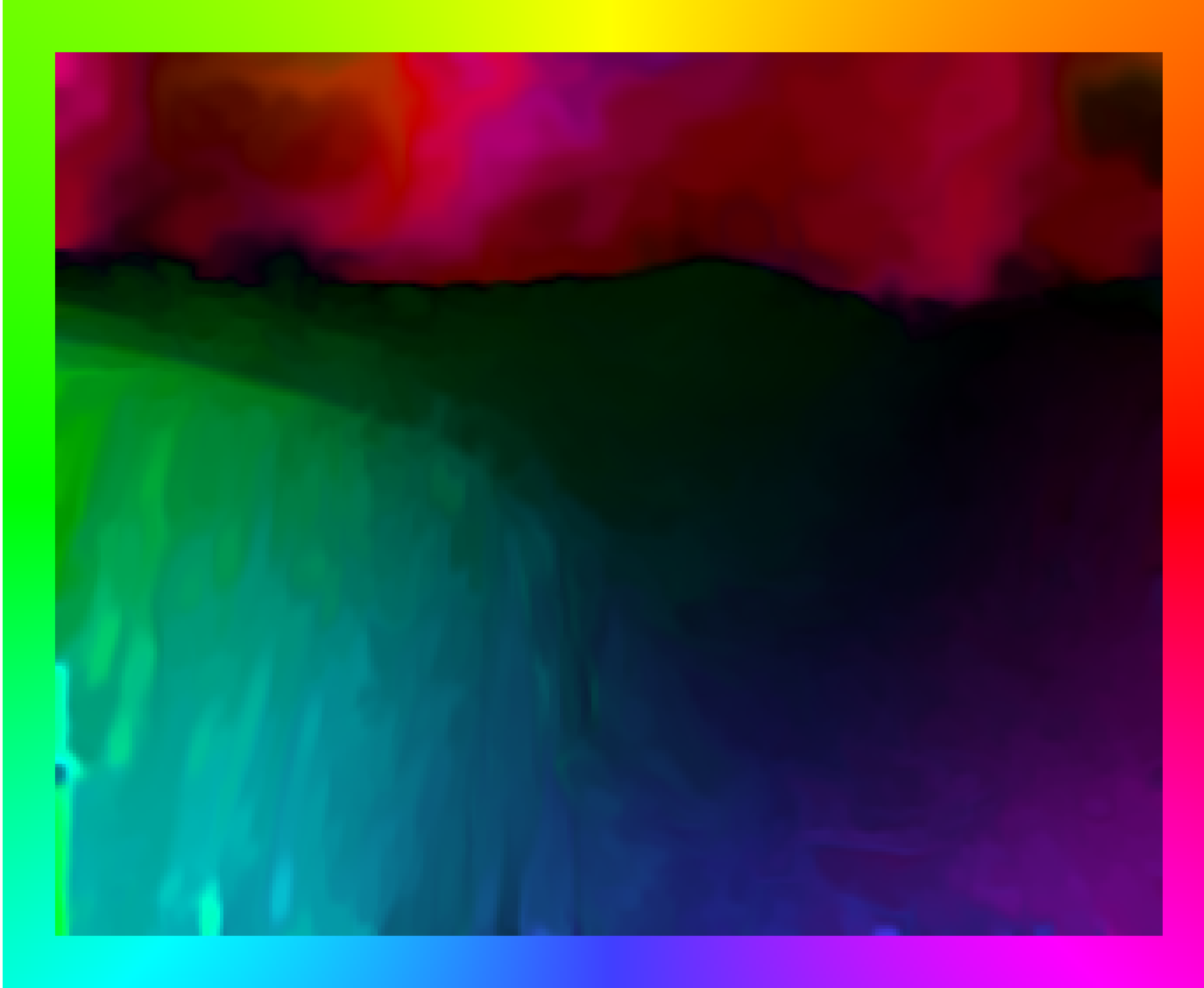


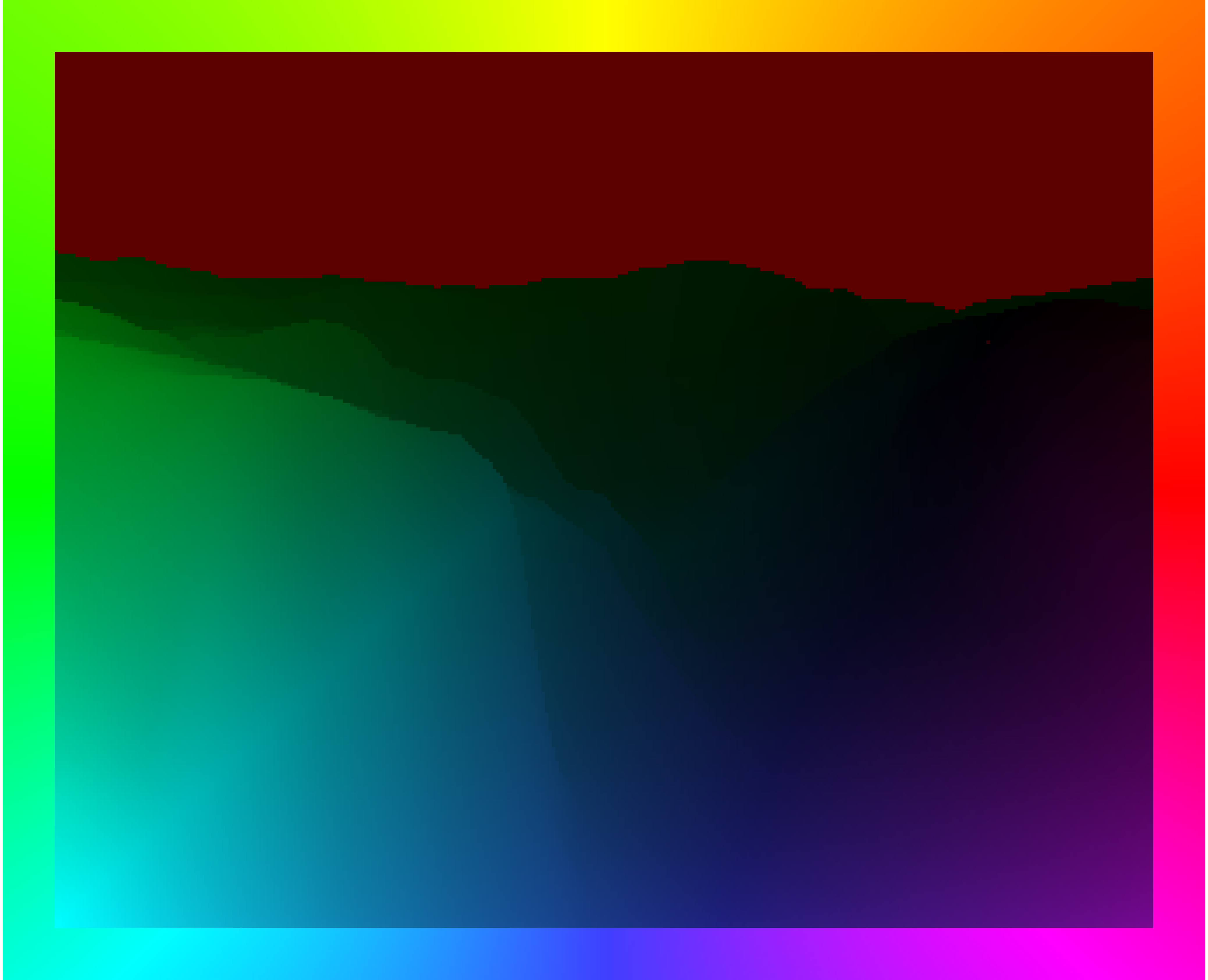


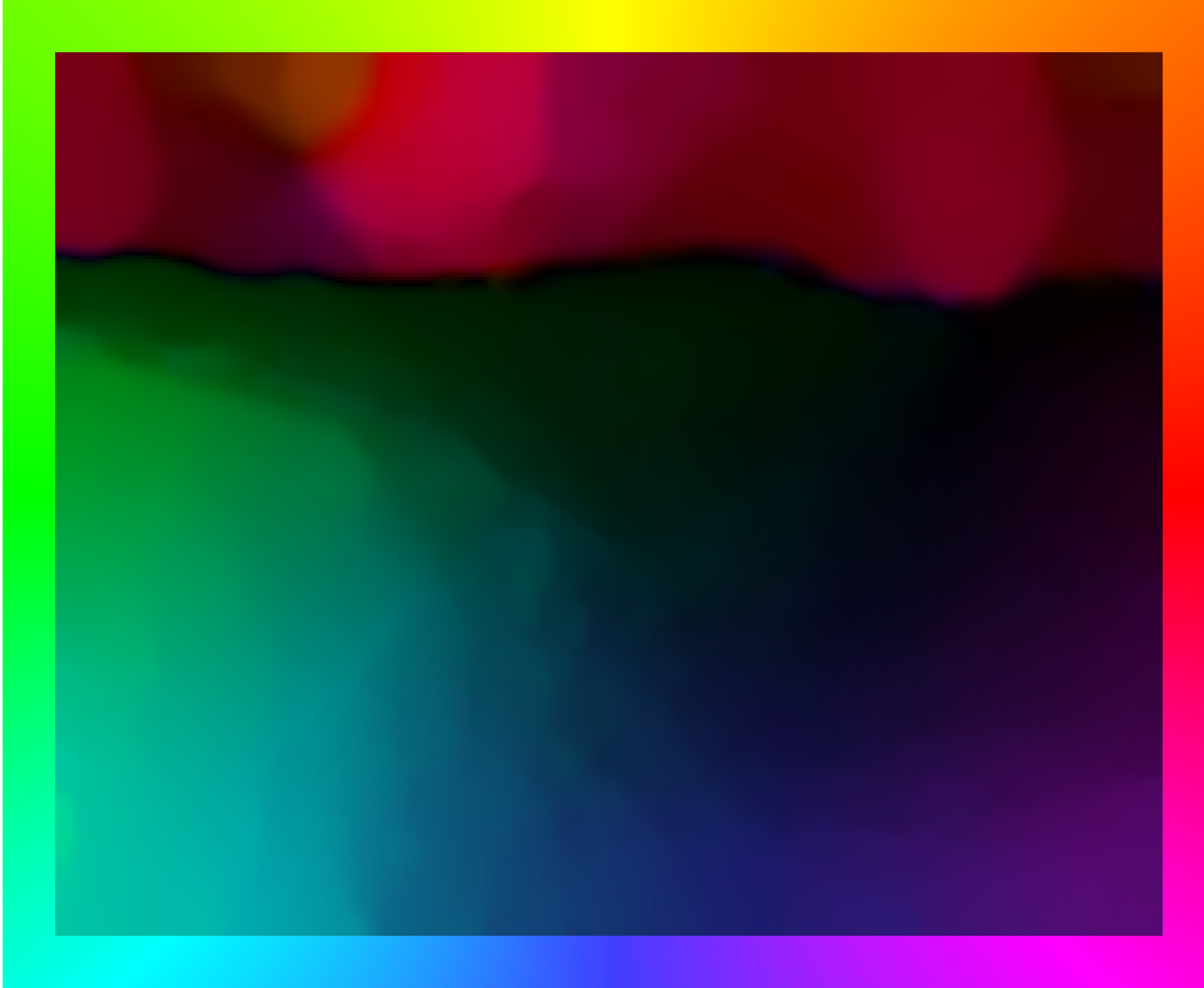


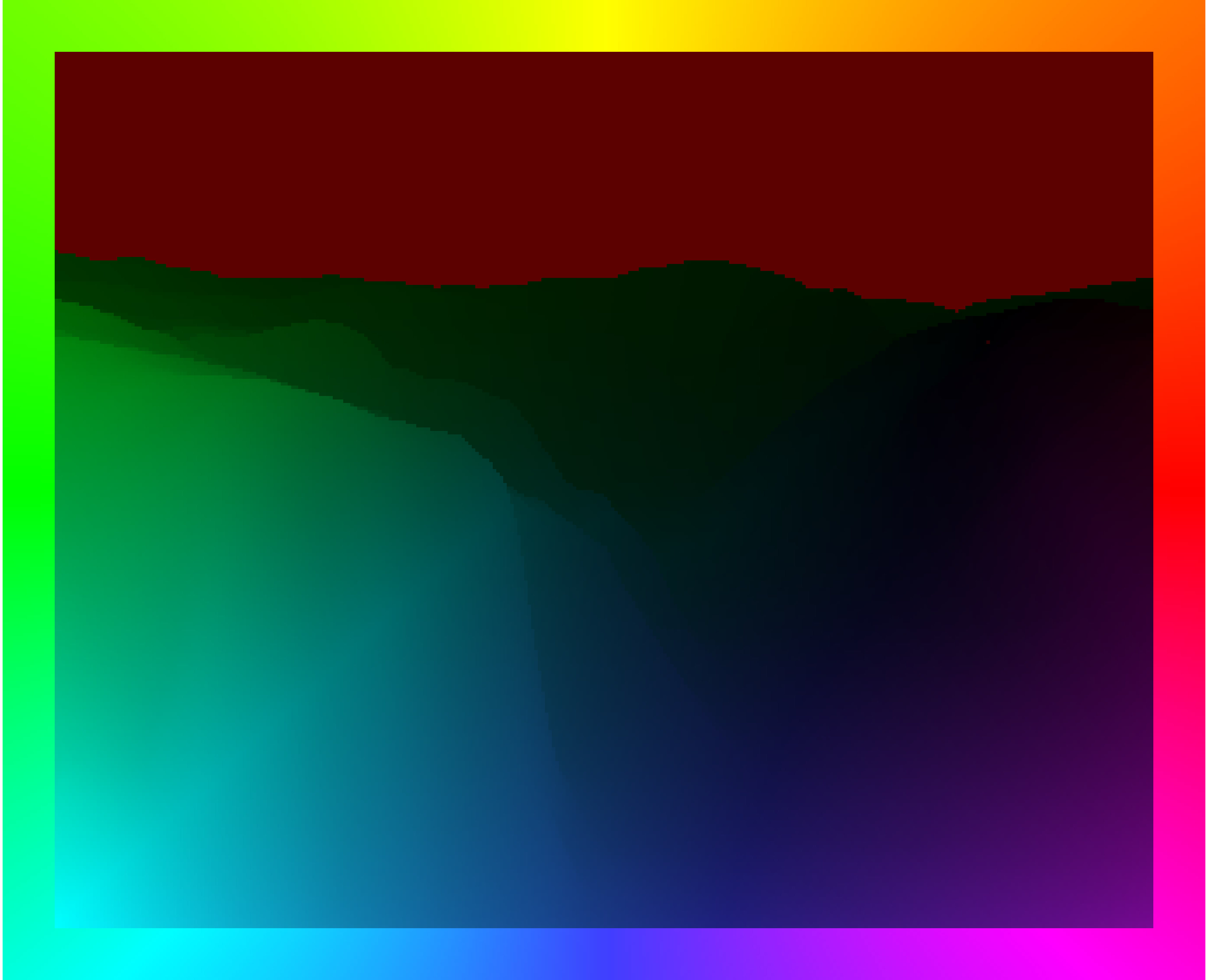


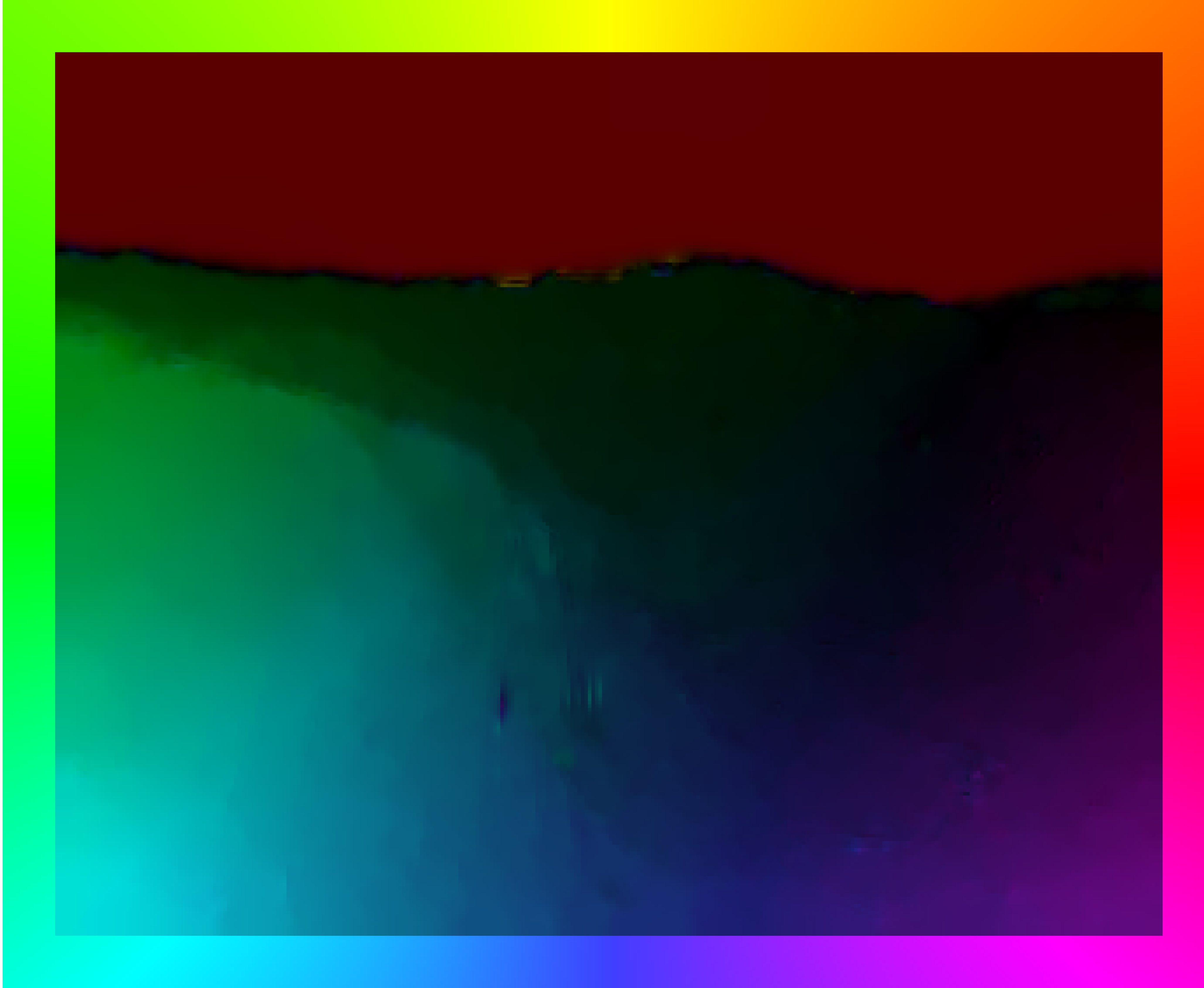




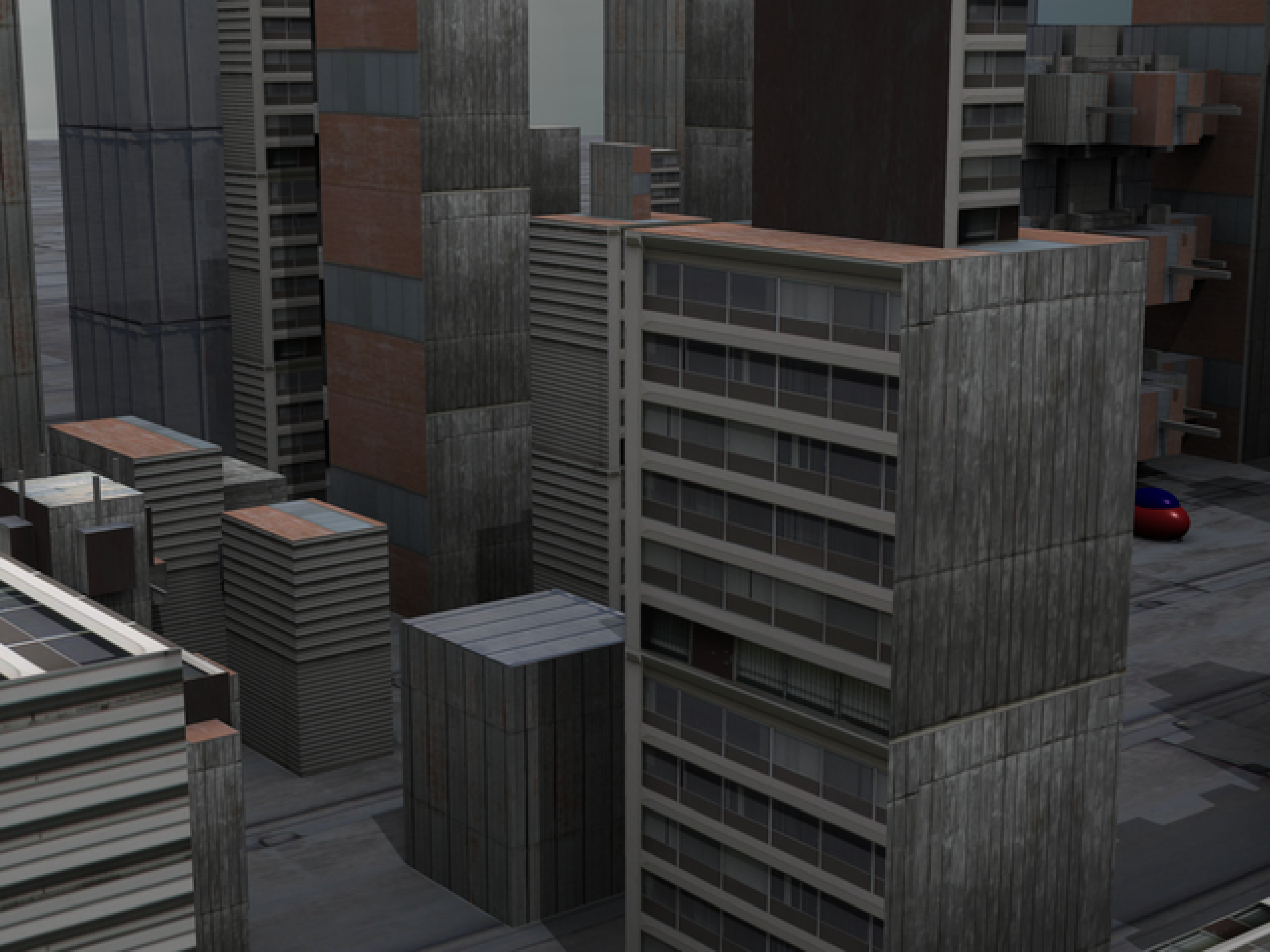


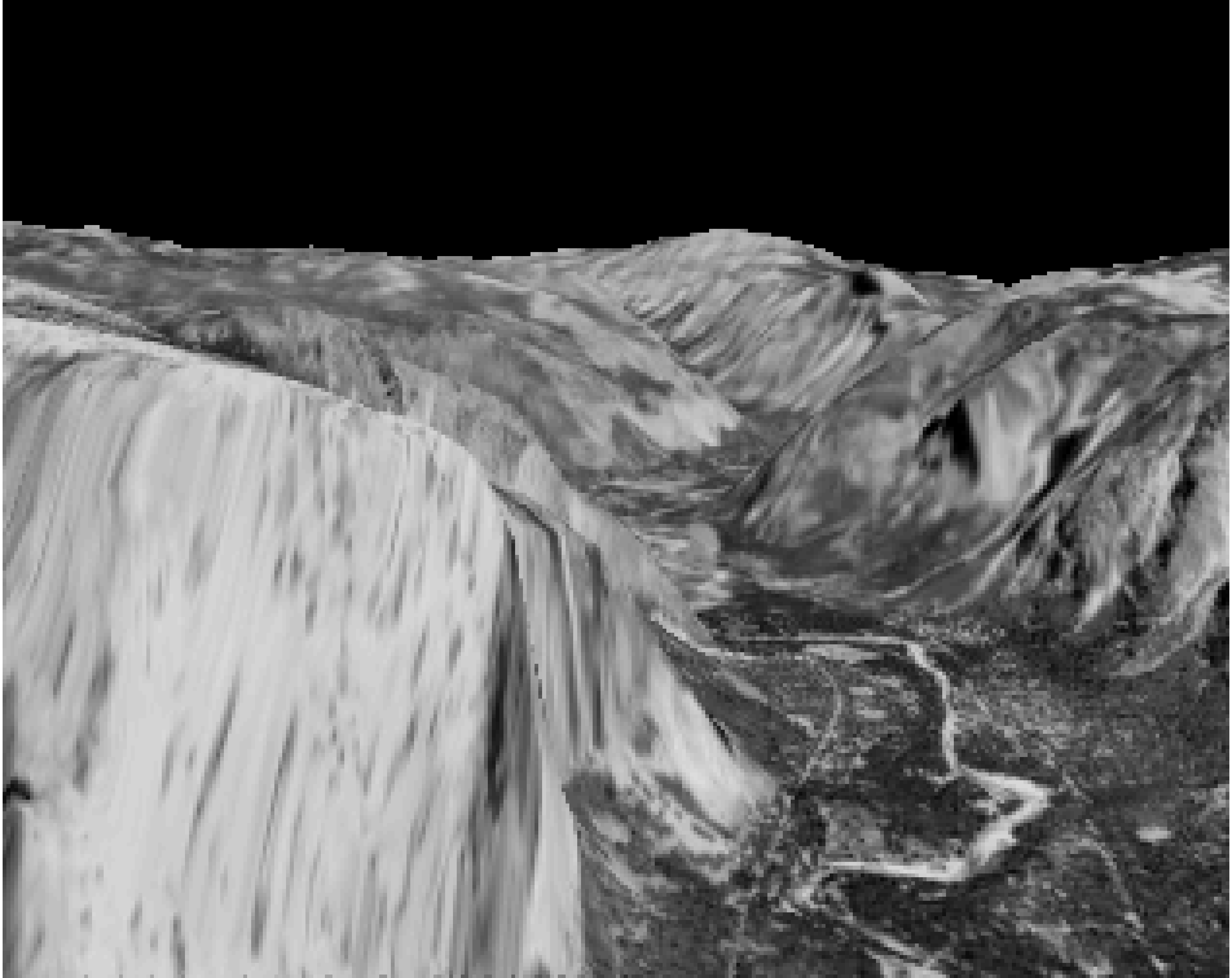




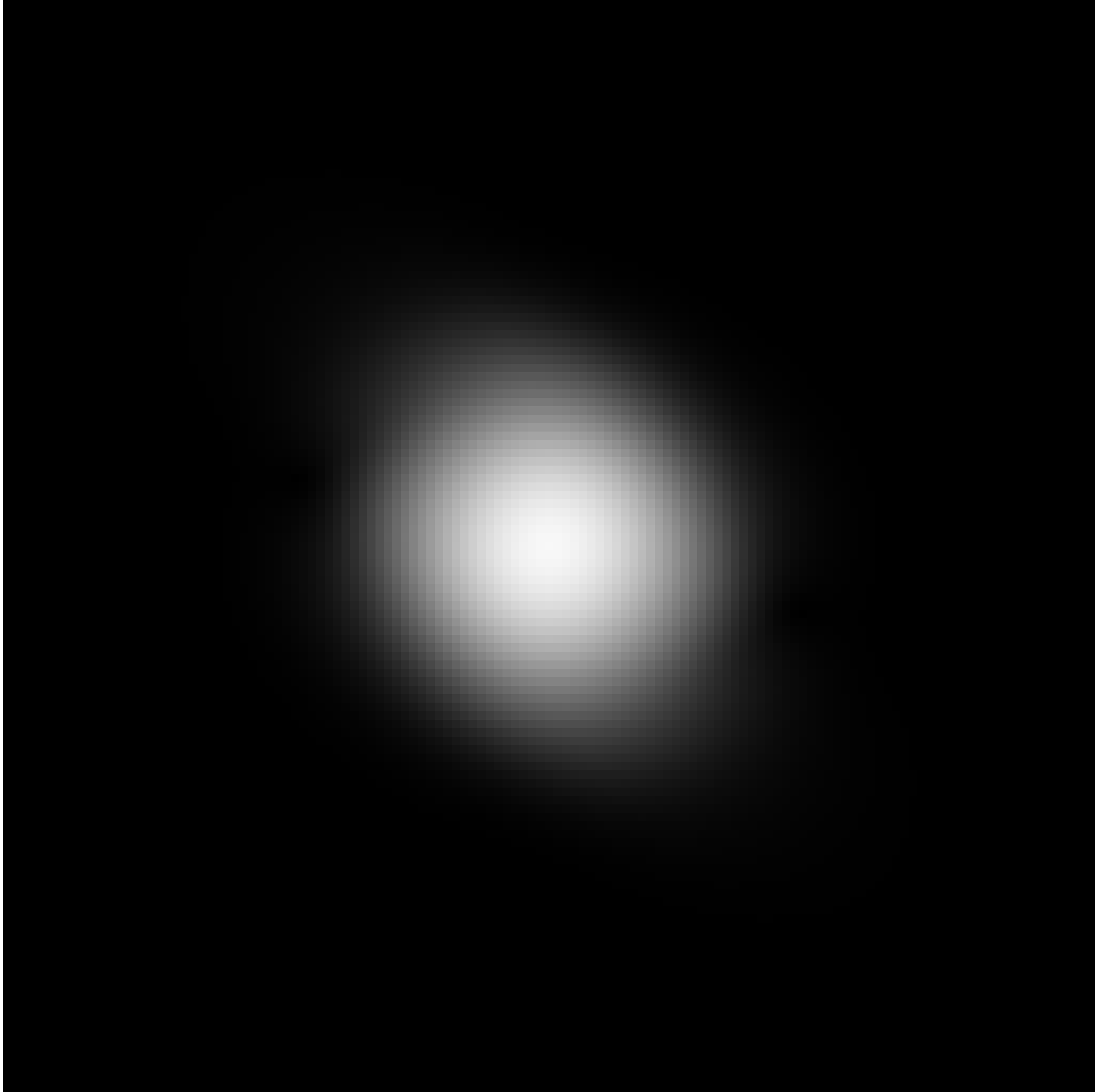


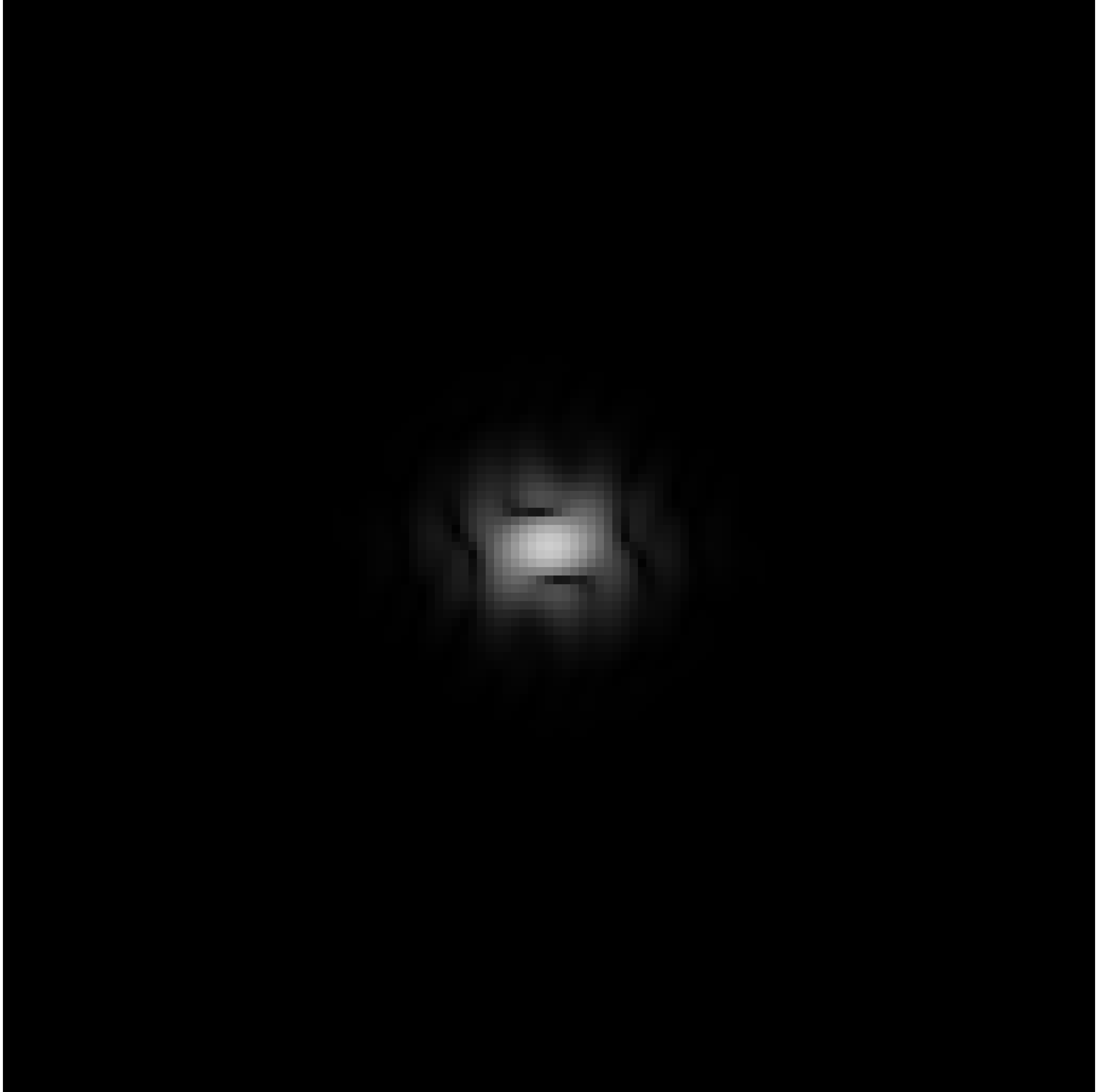

UNITED

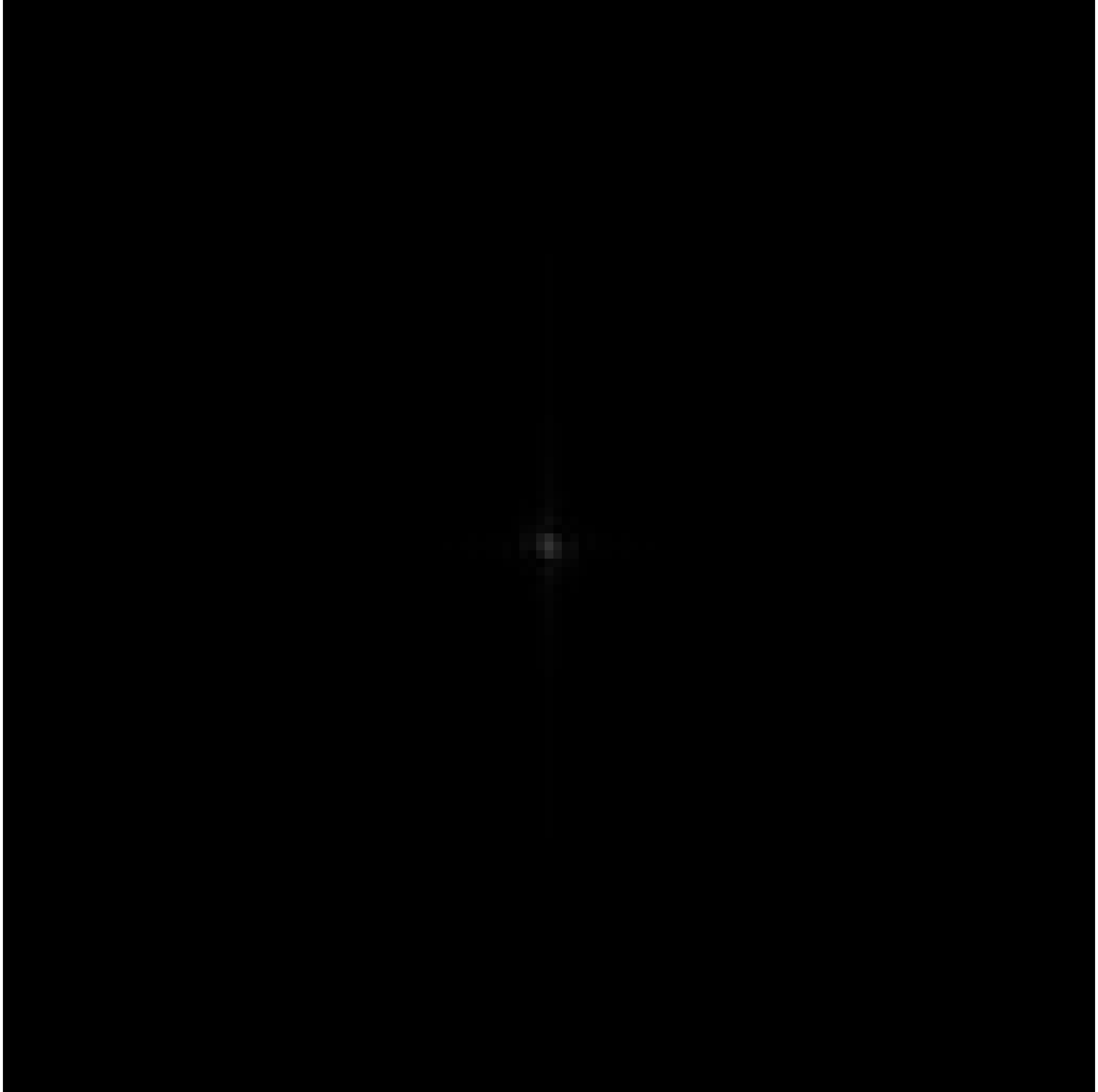




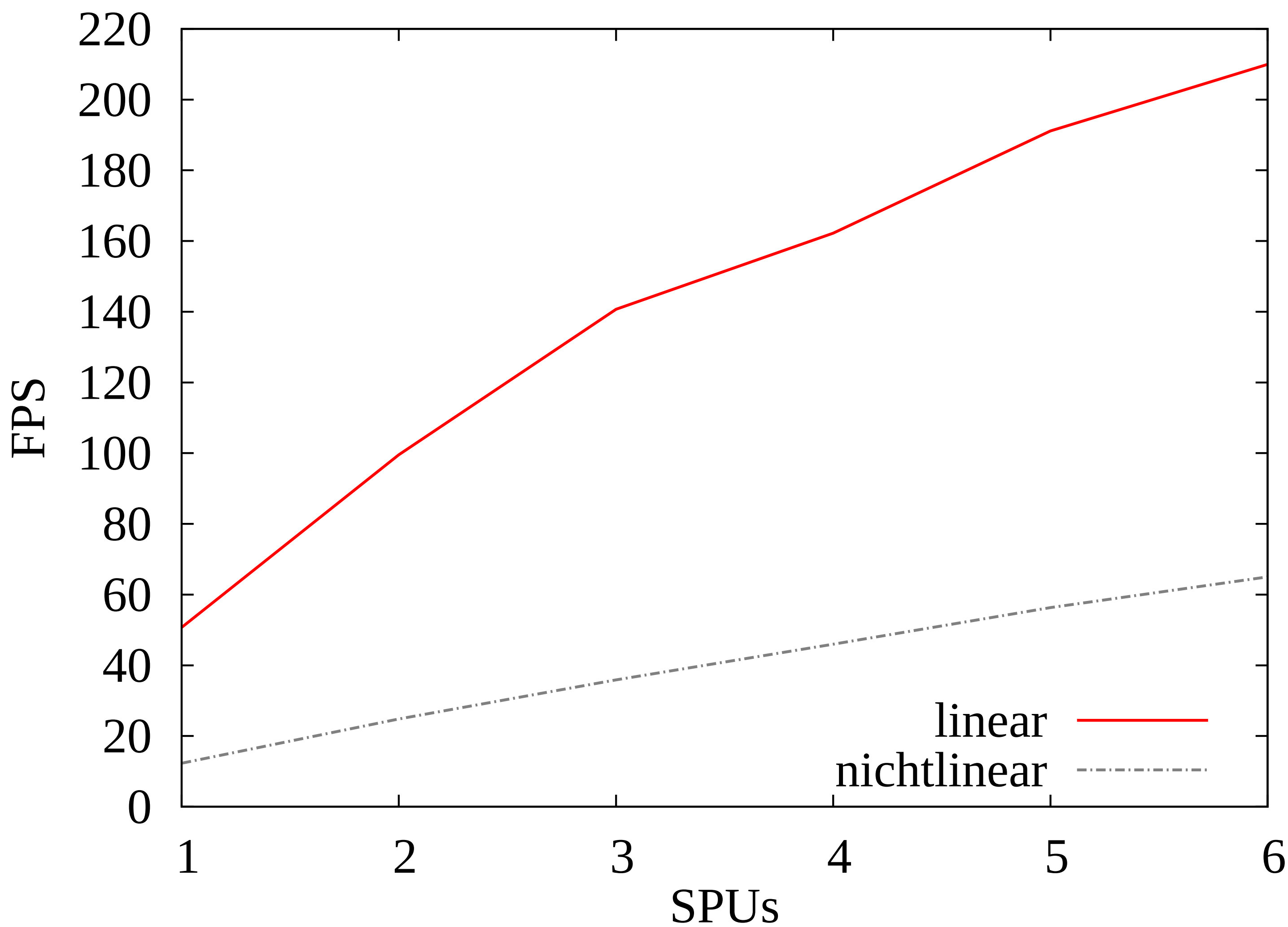
Average angle error	avg. rank	Army (Hidden texture) GT im0 im1			Mequon (Hidden texture) GT im0 im1			Schefflera (Hidden texture) GT im0 im1			Wooden (Hidden texture) GT im0 im1			Grove (Synthetic) GT im0 im1			Urban (Synthetic) GT im0 im1			Yosemite (Synthetic) GT im0 im1			Teddy (Stereo) GT im0 im1		
		all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext
ComplOF [27]	4.2	4.44 8	11.2 7	4.04 8	2.51 2	9.77 3	**1.74** 1	3.93 4	10.6 4	**2.04** 1	3.87 7	18.8 4	2.19 6	**3.17** 1	**4.00** 1	2.92 2	4.64 5	13.8 2	3.64 4	2.17 6	3.36 3	2.51 12	3.08 2	7.04 2	3.65 5
Adaptive [26]	4.3	**3.29** 1	**9.43** 1	**2.28** 1	3.10 6	11.4 7	2.46 7	6.58 9	15.7 8	2.52 5	**3.14** 1	**15.6** 1	**1.56** 1	3.67 6	4.46 5	3.48 6	**3.32** 1	**13.0** 1	**2.38** 1	2.76 11	4.39 10	1.93 7	3.58 3	8.18 3	2.88 2
Spatially variant [22]	5.9	3.73 3	10.2 5	3.33 4	3.02 5	11.0 6	2.67 8	5.36 5	13.8 6	2.35 2	3.67 2	19.3 6	1.84 3	3.81 8	4.81 13	3.69 9	4.48 4	16.0 7	3.90 5	2.11 4	3.26 2	2.12 9	4.66 8	9.41 8	4.35 10
TV-L1-improved [20]	6.8	3.36 2	9.63 2	2.62 2	2.82 4	10.7 5	2.23 3	6.50 8	15.8 9	2.73 6	3.80 5	21.3 11	1.76 2	3.34 2	4.38 4	**2.39** 1	5.97 7	18.1 13	5.67 11	3.57 15	4.92 16	3.43 17	4.01 7	9.84 9	3.44 3
Multicue MRF [24]	7.9	4.50 9	10.1 3	4.18 12	2.52 3	**7.07** 1	2.36 6	**3.09** 1	**7.41** 1	2.36 3	4.46 10	20.8 10	2.73 9	3.51 4	4.11 2	4.06 14	6.08 9	15.6 6	5.40 9	5.25 22	5.36 18	9.02 22	3.63 4	8.39 4	4.15 8
F-TV-L1 [18]	8.5	5.44 12	12.5 11	5.69 16	5.46 13	15.0 13	4.03 13	7.48 13	16.3 11	3.42 11	5.08 12	23.3 15	2.81 10	3.42 3	4.34 3	3.03 3	4.05 3	15.1 4	3.18 2	2.43 8	3.92 8	1.87 6	3.90 6	9.35 7	**2.61** 1
DPOF [21]	10.2	5.63 13	10.9 6	4.16 11	4.05 10	12.1 8	3.31 9	3.87 3	8.82 2	3.17 9	4.34 8	16.2 2	3.13 12	3.95 12	4.78 12	4.17 16	6.69 15	15.2 5	6.27 14	5.62 23	6.89 24	6.60 21	**2.44** 1	**4.83** 1	3.74 7
Brox et al. [8]	10.2	4.80 11	14.4 15	4.29 13	4.05 10	13.5 10	3.71 11	6.63 10	16.0 10	7.26 13	5.22 13	22.7 14	3.22 13	4.56 17	6.09 23	3.40 4	3.97 2	17.9 11	3.41 3	2.07 3	3.76 6	1.18 2	5.14 10	11.9 12	4.28 9
Fusion [9]	10.7	4.43 7	13.7 13	4.08 9	**2.47** 1	8.91 2	2.24 4	3.70 2	9.68 3	3.12 8	3.68 3	19.8 7	2.54 8	4.26 15	5.16 14	4.31 18	6.32 11	16.8 9	6.15 13	4.55 19	5.78 20	3.10 16	7.12 18	13.6 18	7.86 19
SegOF [13]	11.2	5.85 14	13.5 12	3.98 7	7.40 15	14.9 12	8.13 19	8.55 15	17.3 15	9.01 14	6.50 16	18.1 3	5.14 16	3.90 11	4.53 6	4.81 21	6.57 14	21.7 19	6.81 17	**1.65** 1	3.49 5	**1.08** 1	3.71 5	9.23 6	3.63 4
Dynamic MRF [10]	11.2	4.58 10	12.4 10	4.14 10	3.25 8	13.9 11	2.27 5	6.02 7	16.8 12	2.36 3	4.39 9	22.6 13	2.51 7	3.61 5	4.55 7	3.46 5	6.81 16	22.2 21	6.78 16	2.41 7	3.48 4	3.69 18	9.26 22	17.8 22	10.2 22
CBF [15]	12.1	3.95 5	10.1 3	3.44 6	3.70 9	10.6 4	3.85 12	5.64 6	13.5 5	3.34 10	3.71 4	21.5 12	1.99 4	4.36 16	5.50 16	3.55 7	11.3 23	19.1 15	9.05 22	6.79 25	7.37 26	11.6 25	5.50 11	11.8 11	5.66 13
GraphCuts [17]	13.2	6.25 15	14.3 14	5.53 15	8.60 17	20.1 19	6.61 15	7.91 14	15.4 7	10.9 15	4.88 11	19.0 5	3.05 11	3.78 7	4.71 10	3.94 12	8.74 19	16.4 8	5.39 8	4.04 18	4.87 14	4.85 20	6.35 14	12.2 13	6.05 16
Learning Flow [14]	13.2	4.23 6	11.7 9	3.41 5	4.16 12	15.3 14	3.42 10	6.78 11	16.9 13	3.83 12	6.41 15	25.3 17	4.25 14	4.66 20	6.01 22	4.00 13	6.33 13	20.7 17	5.30 6	3.09 13	4.84 13	2.91 14	7.08 17	15.0 20	5.27 12
Second-order prior [11]	13.5	3.84 4	11.2 7	3.11 3	3.12 7	12.9 9	2.17 2	6.96 12	17.2 14	2.83 7	3.84 6	20.5 9	2.09 5	4.83 22	5.83 20	3.90 11	14.0 25	21.8 20	8.28 19	7.74 26	6.88 23	11.7 26	6.74 16	13.4 17	5.80 14



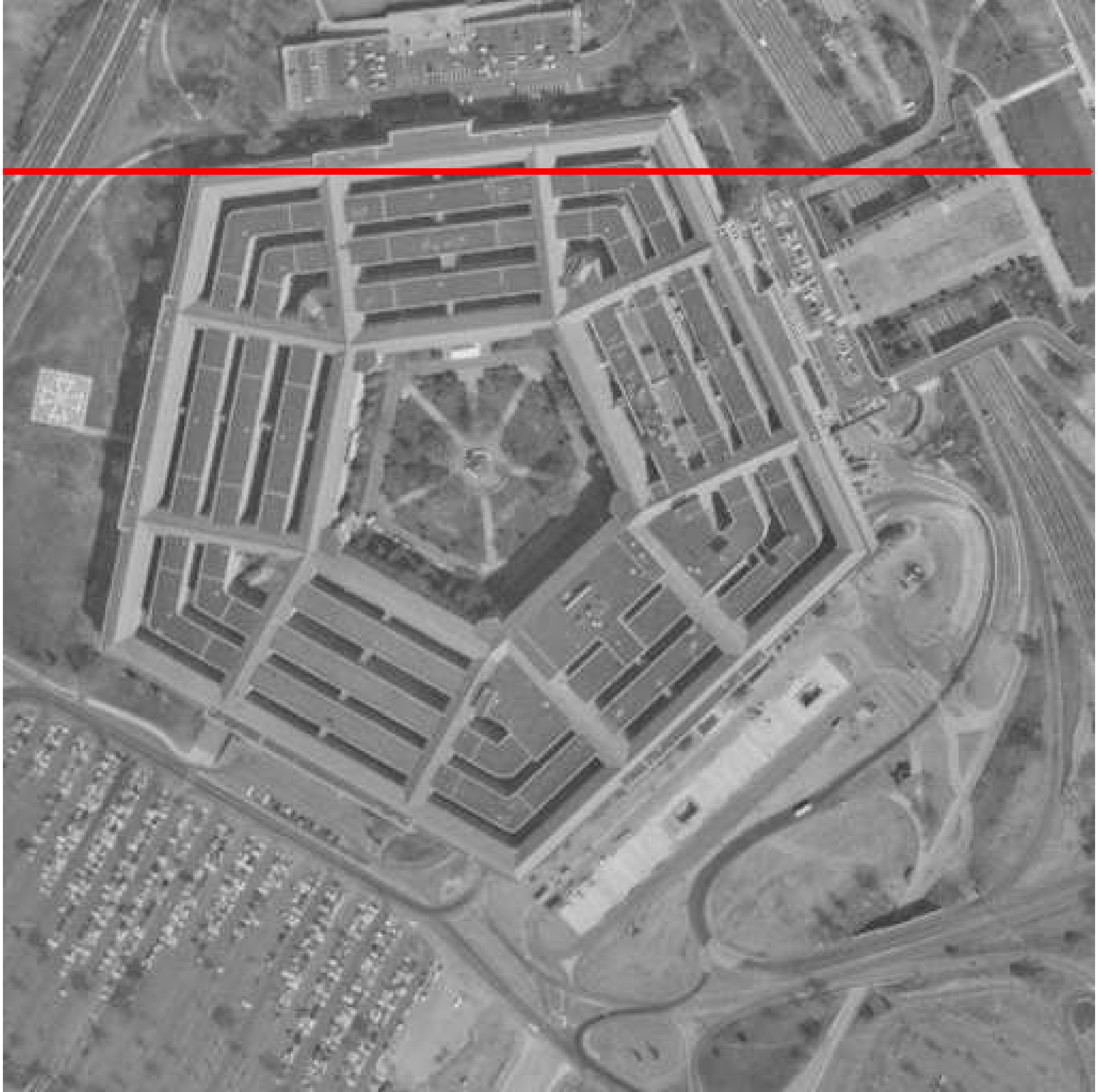


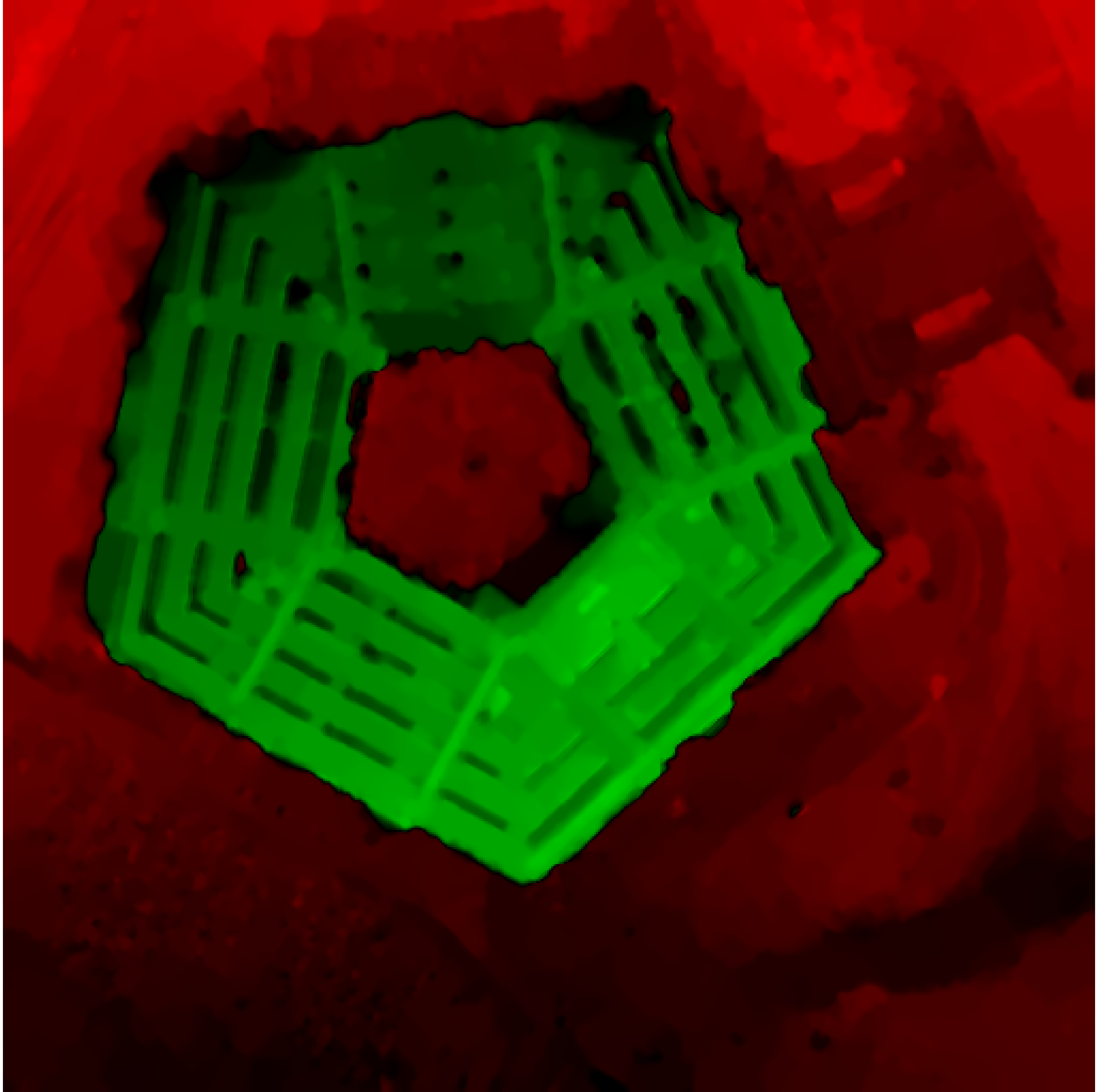


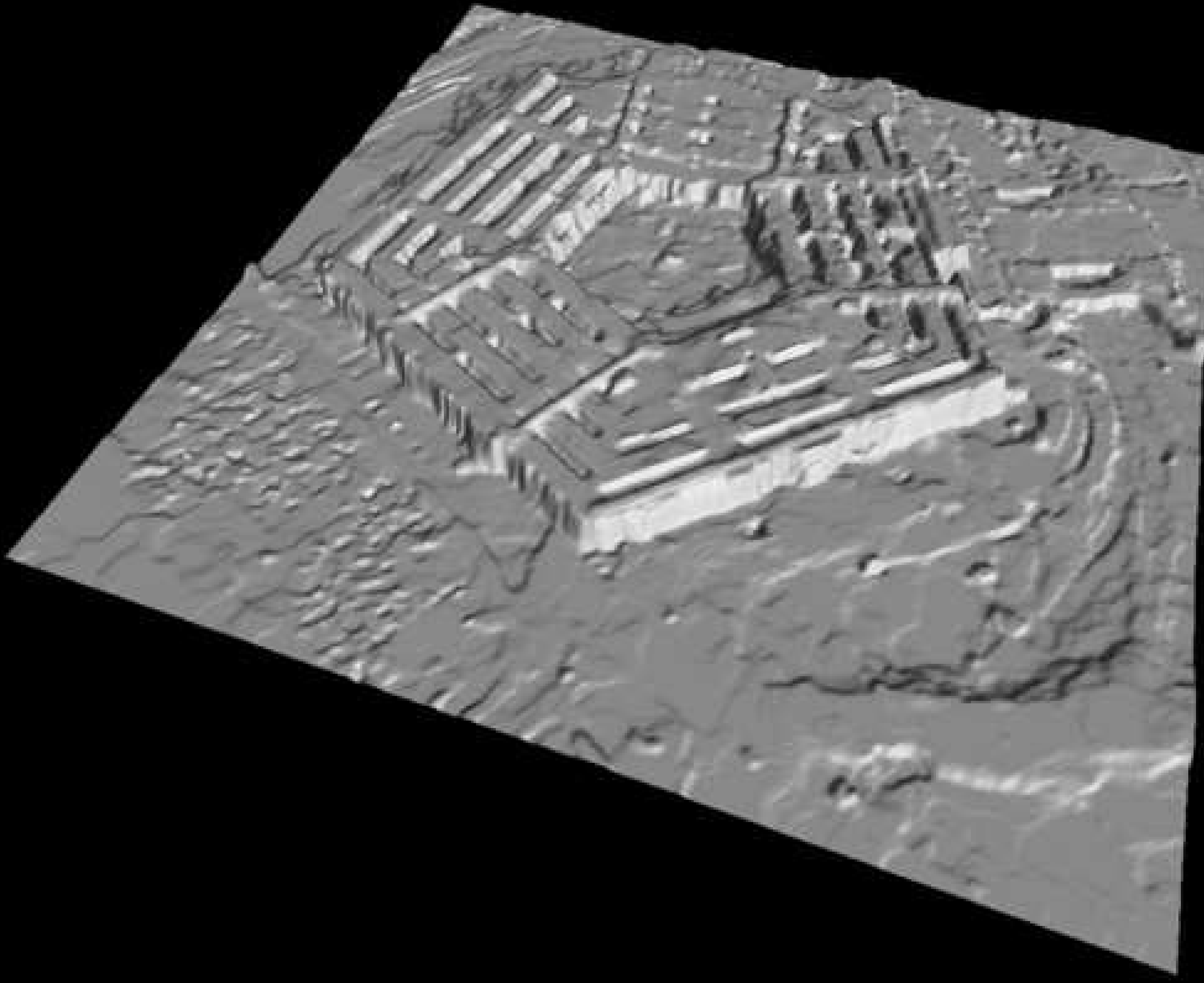

PLAYSTATION 3
SONY

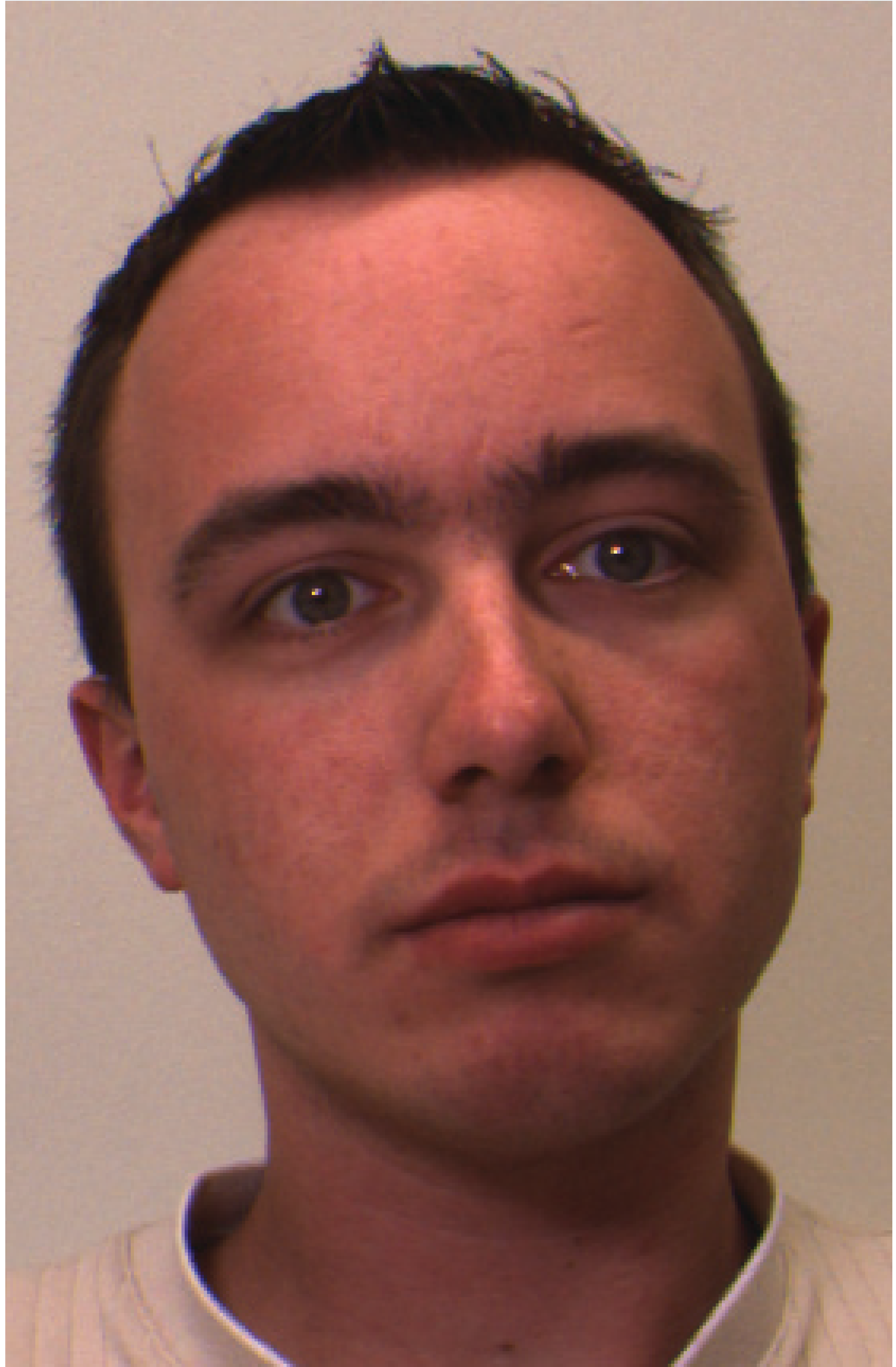


FPS
220
200
180
160
140
120
100
80
60
40
20
0
1
2
3
4
5
6
SPUs
linear
nichtlinear

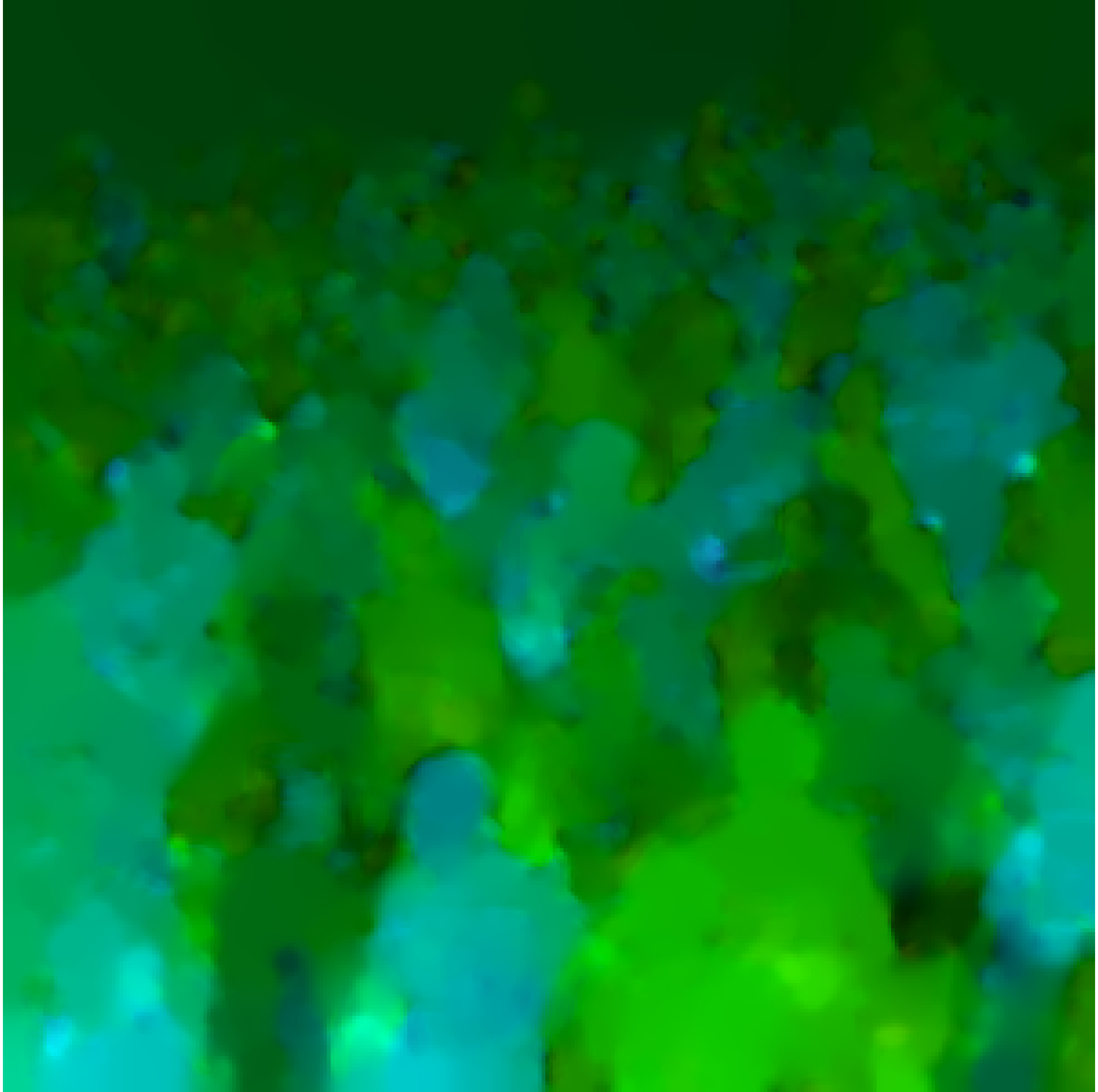


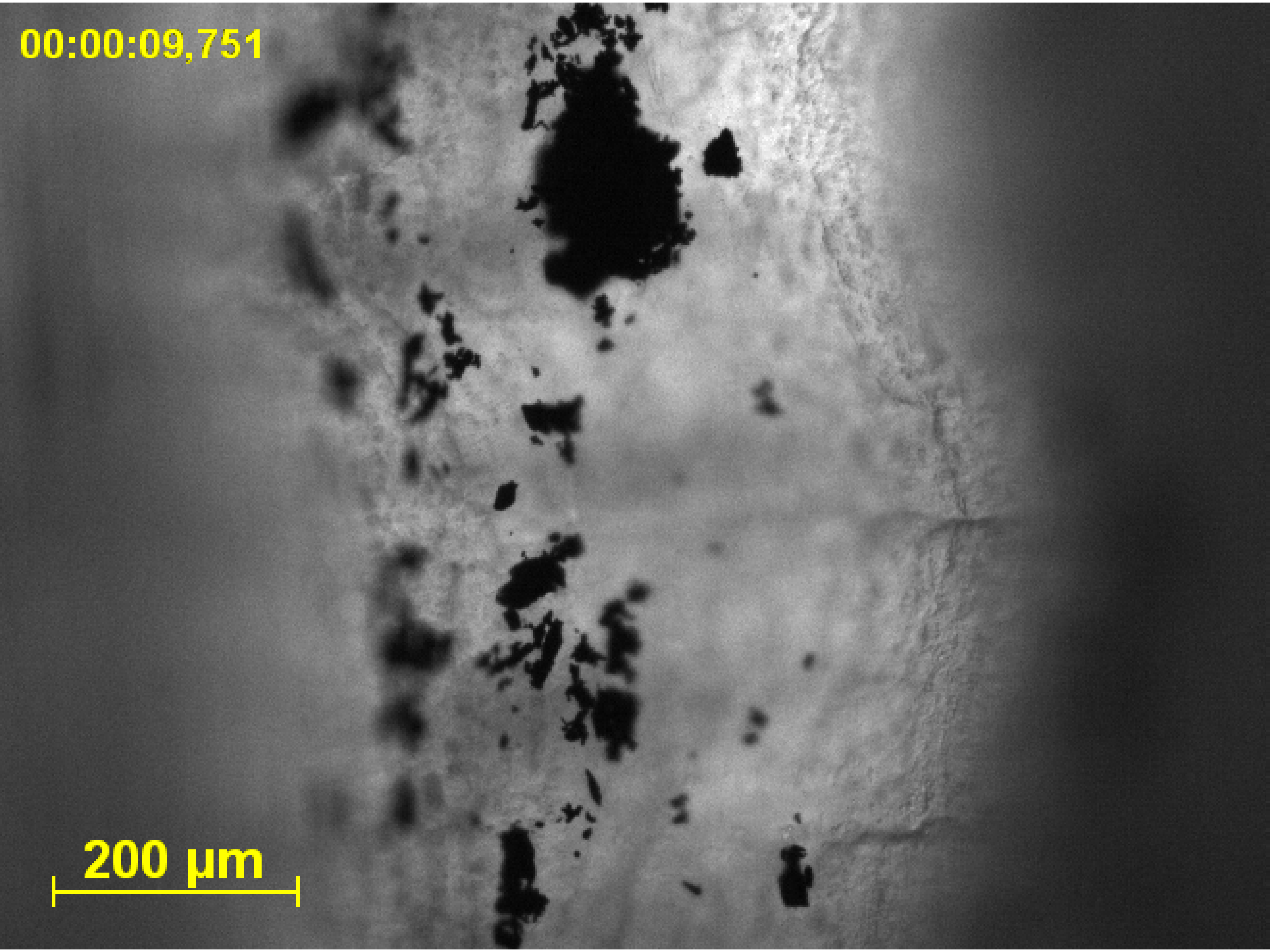


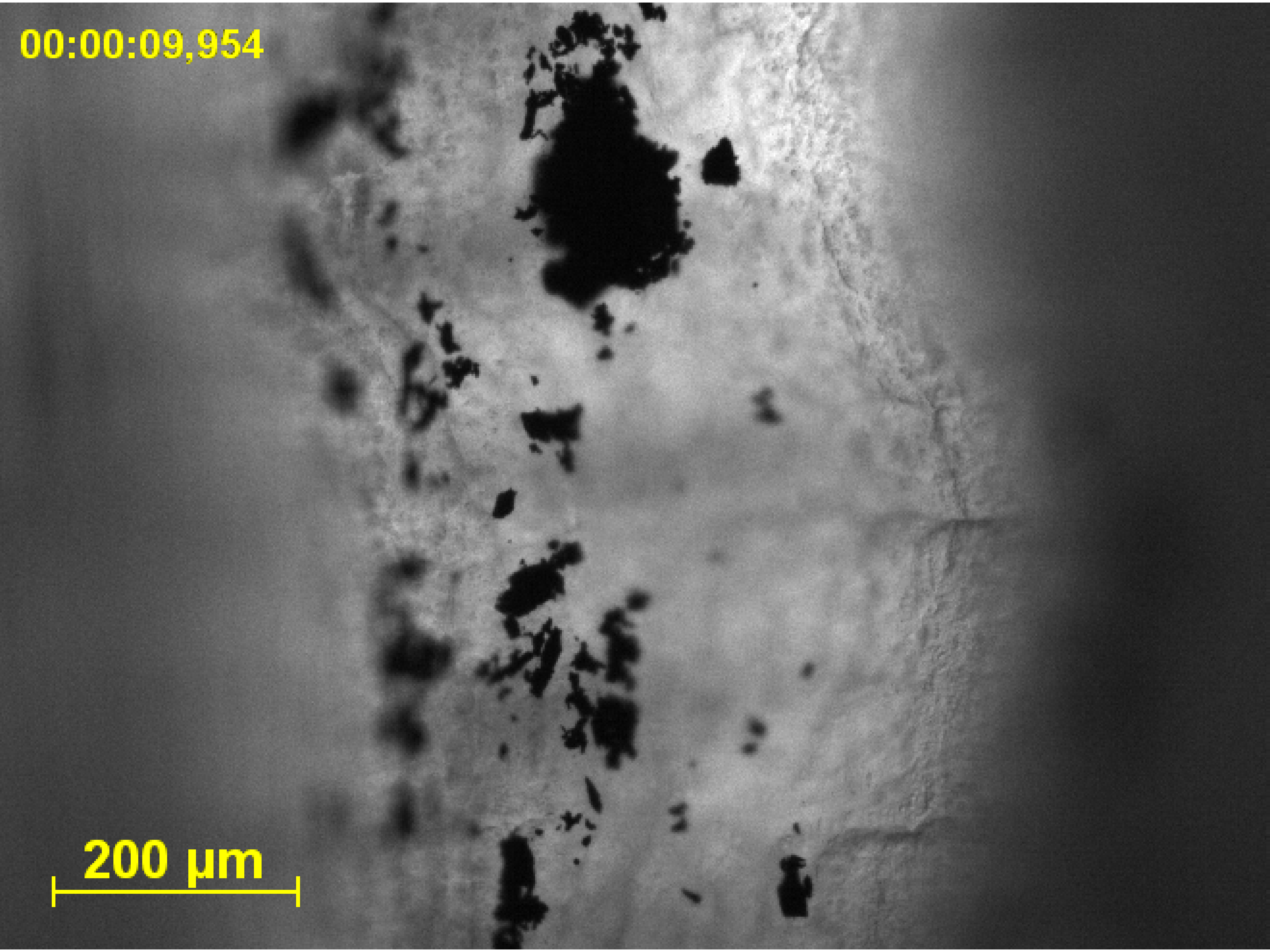


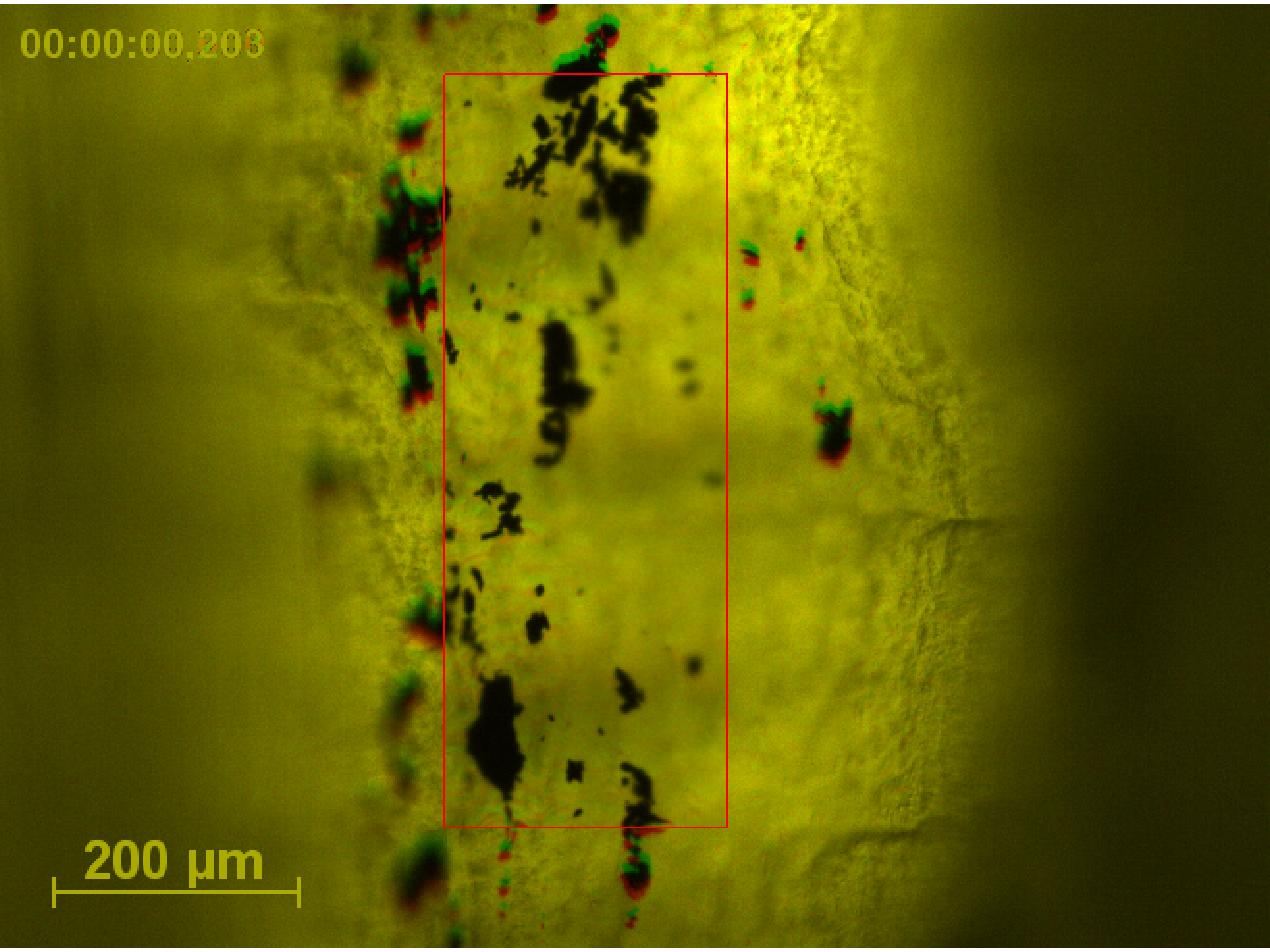


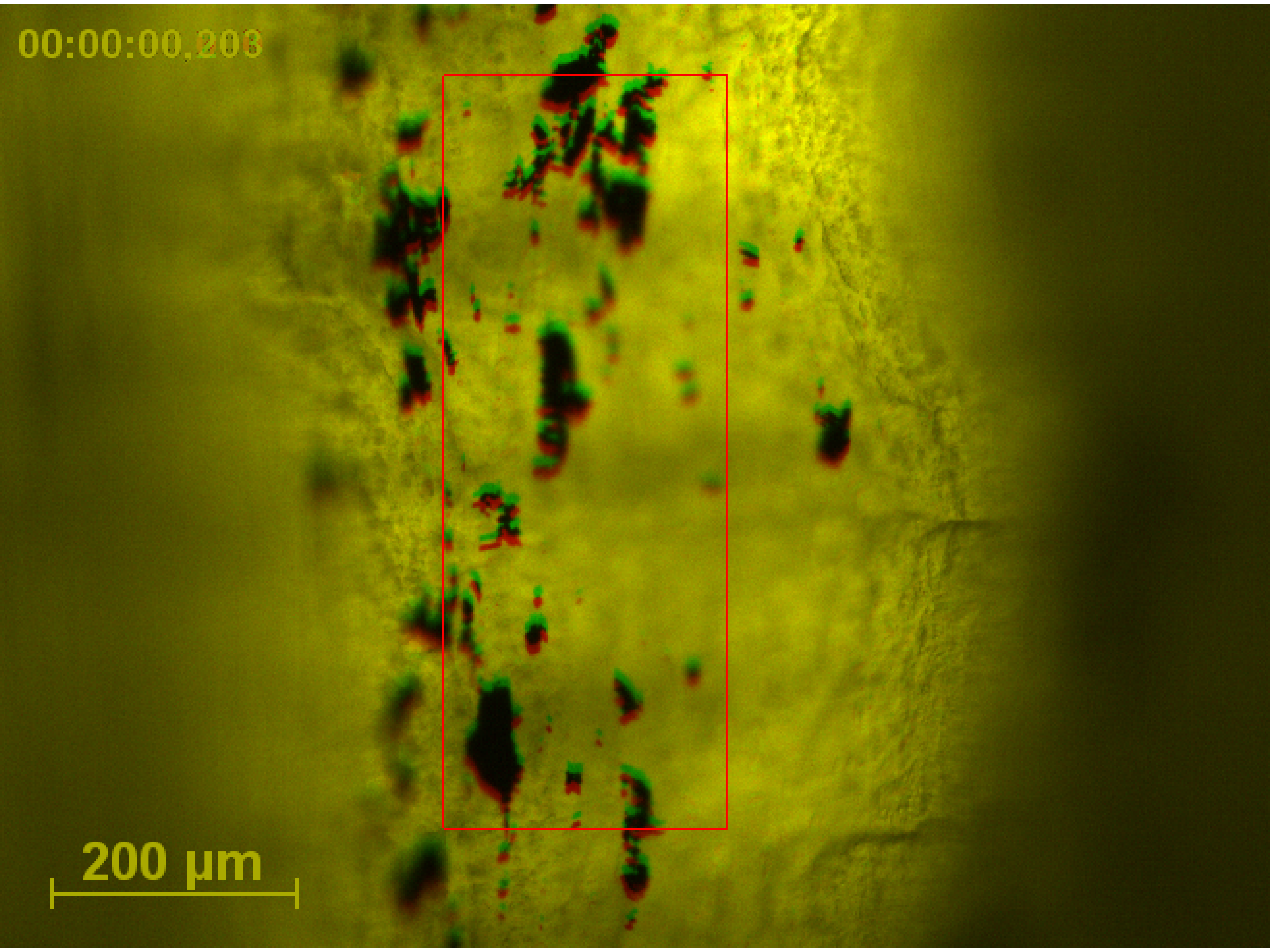

30953
31114
31269
30941




00:00:09,751
200 µm


00:00:09,954
200 µm


200 µm


200 µm