DiLiGenT10²: A Photometric Stereo Benchmark Dataset with Controlled Shape and Material Variation Supplementary Material

Jieji Ren[1†], Feishi Wang[2], Jiahao Zhang[2], Qian Zheng[3], Mingjun Ren[1*], Boxin Shi[2,4,5*]

[1]School of Mechanical Engineering, Shanghai Jiao Tong University,

[2]National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

[3]College of Computer Science and Technology, Zhejiang University

[4]Institute for Artificial Intelligence, Peking University, [5]Beijing Academy of Artificial Intelligence

{jiejiren, renmj}@sjtu.edu.cn, {wangfeishi, 1900013264, shiboxin}@pku.edu.cn, qianzheng@zju.edu.cn

This supplementary material provides additional technical details to re-product the results and set up their own research, including "on the fly" calibration process of lighting direction, the details of uniformly re-sampling of lighting directions, "ground truth" normal calibration, and estimated surface normal results from all methods on all shapes with diverse materials. Besides, more discussions and result visualizations are presented on different perspectives.

A. Lighting Direction Calibration

To obtain diverse lighting directions, we adopt "on the fly" calibration strategy, which is different from existing calibration processes that are with fixed LEDs. As shown in Fig. 1, six specular balls are uniformly placed around the object to record the shiny reflect spots to calibrate lighting directions. We use specular balls with dark color to reduce inter-reflection.

We manually label the masks of calibration balls. The edge of each ball is calculated by fitting a circle in the image coordinate. Then, the vector from ball center to specular point can be obtained to calculate the lighting direction. Six calibration balls are used to jointly optimize the lighting direction. Because the center stage is fixed during the capture process, the manual fitting is processed once. Finally, based on the geometric relationship of the positions of specular spots, viewpoint and surface normal, the lighting direction vector is calculated.
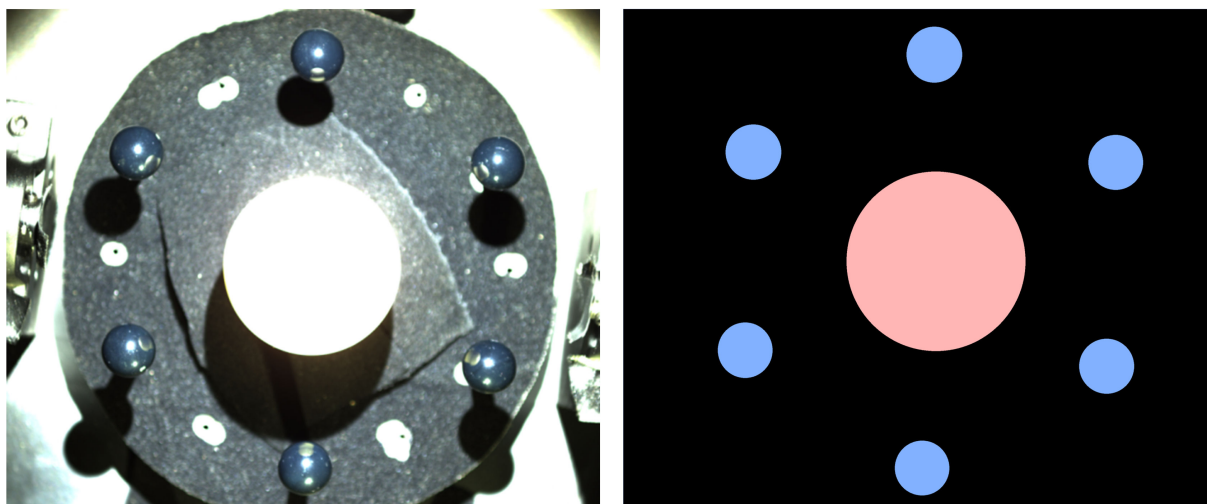

Figure 1. The left part shows the raw image to find the locations of calibration ball and object. Slightly overexposed image are captured to clearly show the dark calibration ball and help object mask detection. The masks of 6 calibration balls (blue) and object region (red) are manually labeled. Note that this overexposed image is only used to calculate the masks of balls.

---

*Corresponding authors.
†Part of this work was finished as a remote intern at Peking University.
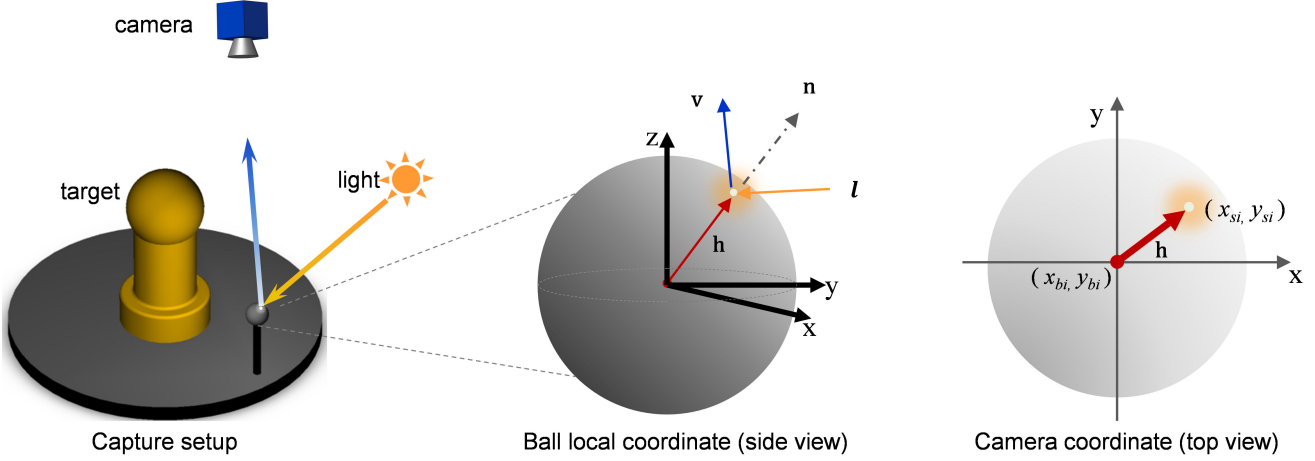
Figure 2. Geometrical relationship between the captured object and a calibration ball in our light calibration process. Left part shows our setup, middle part shows the 3D coordinate, and the right part shows the geometrical relationship in camera coordinate.

To be specific, the camera is placed on the top of the center stage, as shown in Fig. 2. For i-th ball ($i = 1, 2, 3, 4, 5, 6$), we first use its edge to fit a circle with radius r_i and obtain the coordinate of center (x_{bi}, y_{bi}). The location of specular point on the ball (x_{si}, y_{si}) is detected based on an adaptive threshold method. The half vector $\mathbf{h}$, which is between lighting direction and viewing direction, is determined by the coordinates of ball center and specular point. Finally, the lighting direction $\mathbf{l}$ is calculated according to the law of reflection:

$$\mathbf{l} = 2(\mathbf{h}^\top \cdot \mathbf{v})\mathbf{h} - \mathbf{v}, \tag{1}$$

where

$$\mathbf{h} = [\frac{x_{si} - x_{bi}}{r_i}, \frac{y_{si} - y_{bi}}{r_i}, \frac{z_{bi}}{r_i}]^\top,$$
$$\mathbf{v} = [0, 0, 1]^\top, \tag{2}$$
$$z_{bi} = \sqrt{r_i^2 - (x_{si} - x_{bi})^2 - (y_{si} - y_{bi})^2}.$$

As the distance between light source and calibration ball is far enough (far field illumination and orthographic camera), we assume $\mathbf{h}$ is the same for all calibration balls and $\mathbf{h} = [h_x, h_y, h_z]^\top, h_z = \sqrt{1 - h_x^2 - h_y^2}$ can be calculated by optimizing

$$\underset{h_y, h_x}{\mathrm{argmin}} \sum_{i=1}^{6} \left[(x_{si} - r_i h_x - x_{bi})^2 + (y_{si} - r_i h_y - y_{bi})^2 \right]. \tag{3}$$

The closed form solution to Eq. (3) is

$$h_x = \frac{\sum_{i=1}^{6}(x_{si} - x_{bi})r_i}{\sum_{i=1}^{6} r_i^2},$$
$$h_y = \frac{\sum_{i=1}^{6}(y_{si} - y_{bi})r_i}{\sum_{i=1}^{6} r_i^2}. \tag{4}$$

Finally, based on the relationship in Eq. (1), we can calculate the lighting directions for each frame: $\mathbf{l} = 2(\mathbf{h}^T \cdot \mathbf{v})\mathbf{h} - \mathbf{v}$.
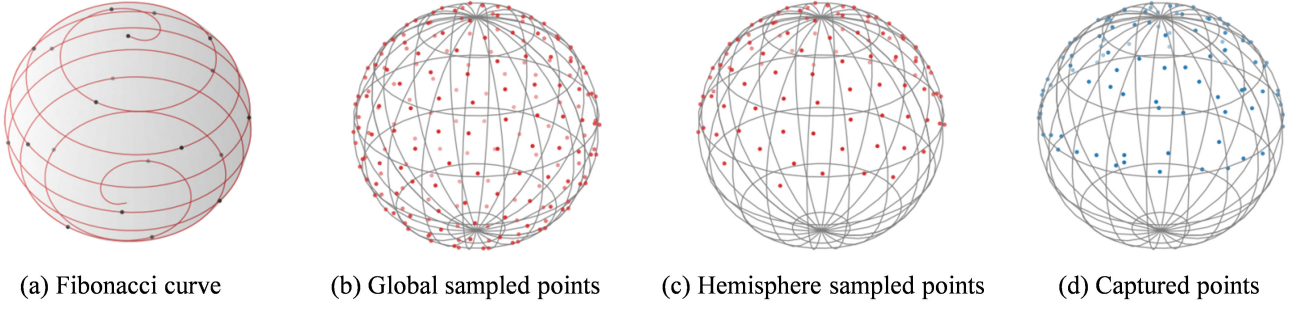
(a) Fibonacci curve   (b) Global sampled points   (c) Hemisphere sampled points   (d) Captured points

Figure 3. Uniform sampling process. (a) shows the Fibonacci curve in shpere space; (b) shows the uniformly sampled 201 points on the surface when $m = 100$, and we only use the 100 points on the upper hemisphere in (c); the finally re-sampled lighting directions are shown in (d).

B. Lighting Direction Re-Sampling

We adopt Fibonacci sphere sampling method [1] to obtain uniformly distributed lighting directions, for every object. As shown in Fig. 3 (a), the Fibonacci lattices can evenly choose positions on sphere surface. The coordinate of each point is defined by latitude and longitude as following:

$$lat_i = \arccos(\frac{2i}{2m+1}),$$
$$lon_i = 2\pi i \Phi^{-1}, i \in [-m, m],$$

(5)

where m controls the sampled density, Φ is golden ratio (1.618). This process can generate $2m + 1$ evenly spaced points (Fig. 3 (b)). We only consider the lighting direction in upper hemisphere and choose these m positive values as our re-sampled positions (Fig. 3 (c)).

Then we can find the nearest points of each re-sampled position from densely captured results. Suppose we have recorded c lighting directions as vector l_c, we select m positions as fixed re-sampled directions l_m. Based on matrix calculation, we can get the distance matrix:

$$\mathbf{D} = l_m^\top \times l_c.$$

(6)

We find the index of maximum value in each row to uniformly sample the lighting directions, which are closest to the generated uniform positions (Fig. 3 (d)). In our setup, we set $m = 100$, $c \in [300, 400]$, and we only select the first 100 positions on the upper hemisphere. After re-sampling, each object in the proposed dataset has 100 observed images with accurately calibrated lighting direction which approximates uniform distribution.

C. "Ground Truth" Normal Map Calibration

In this section, we present the details of obtaining the "ground truth" surface normal map of each object. Previous dataset [3] renders the surface normal map as "ground truth" by utilizing the scanner to obtain the shapes of objects and align them to images. The accuracy of shapes not only depends on the scanner's performance, but is also affected by the material of the surface being scanned. *i.e.*, some specular materials , such as metal and acrylic, are rather difficult to scan directly, as shown in Fig. 4.
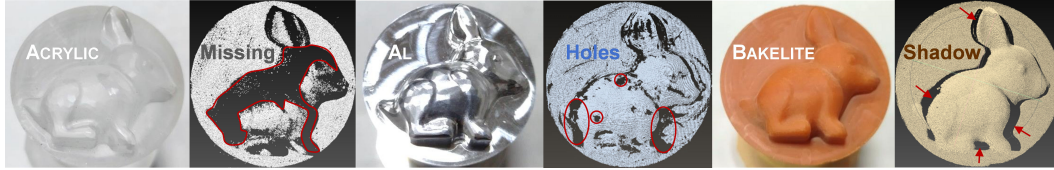
Figure 4. Scanned results of BUNNY on different challenge materials.

In contrast, precise machining can avoid such difficulties during scanning and maintain consistency of the shape with varying materials in manufacturing. Therefore, we choose the computer numerical control (CNC) approach to machine the shape with varying materials, and select the CAD model to render surface normal map as "ground truth". The CAD models of ten shapes are shown in Fig. 5.
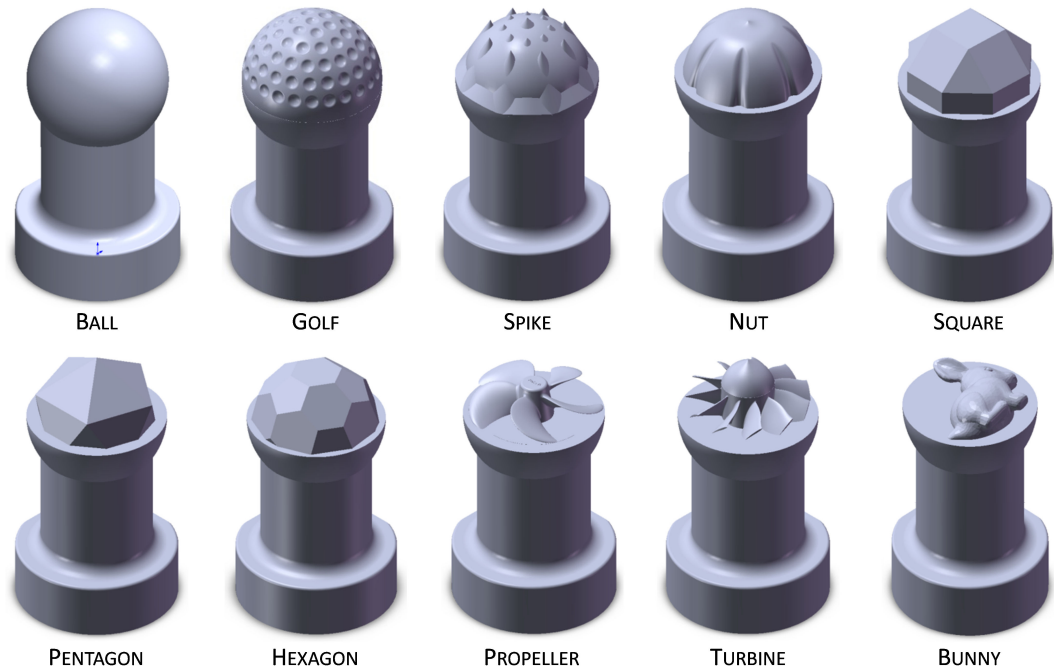

Figure 5. CAD models of ten shapes. The observed objects are mounted on the top of the based.

Replacing scanning model with CAD model can provide more accurate surface normal, because CNC can fabricate high-accuracy and consistent shapes. We first measure the SQUARE by precise CMM (coordinate measuring machine) to verify the accuracy on different materials. It has three orthogonal groups of plane to determine their machining errors on x, y, and z directions. Results show majority machining errors of our materials are smaller than 0.05mm in Fig. 6.

To verify the accuracy of "ground truth" surface normal, we numerically analyse the accuracy of SQUARE on different materials as we can sparsely measure their real surface normal with high-precision equipment (right figure). For each of SQUARE's face, we sample 9 points with CMM to calculate the precise normal measurement. We then calculate the accuracy between precisely measured normal and proposed "ground truth" normal (based on CAD). Mean angular errors are as follows:
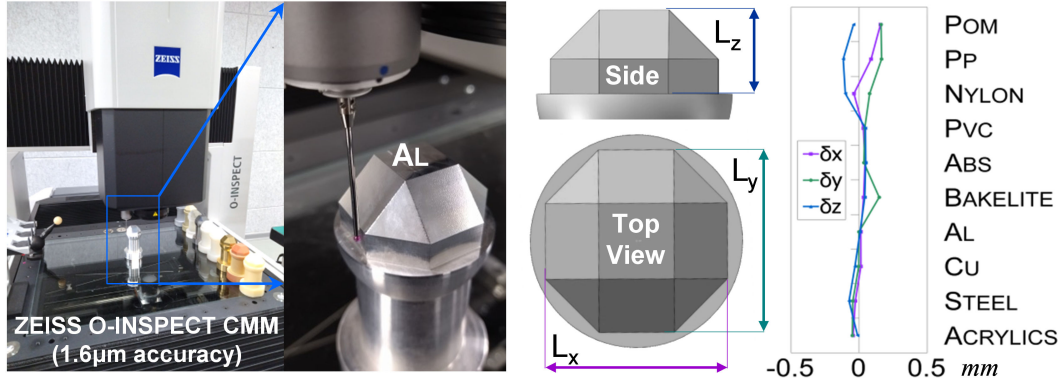
Figure 6. Precise CMM measured the fabricating accuracy of SQUARE shape on different materials.

POM: 0.011°, PP: 0.015°, NYLON: 0.002°, PVC: 0.007°, ABS: 0.027°, BAKELITE: 0.006° , AL: 0.012°, CU: 0.007°, STEEL: 0.005°, ACRYLIC: 0.007°.

By replacing the scanned meshes with CAD models, other procedure for calibrating the "ground truth" normal is the same as [3]. Please refer to their released toolbox and corresponding instructions for details[1].

D. Complete Benchmark Results

We provide complete benchmark results in the form of estimated normal maps and their corresponding error maps[2] in Fig. 7 – Fig. 26. In each page, we can observe the performance difference of 13[3] methods (vertically) and their variation with materials (horizontally) for each shape. Their mean and median of angular errors are labeled on the upper left.

---

[1]https://sites.google.com/site/photometricstereodata/single

[2]The errors are truncated at 45 degrees for better visualization.

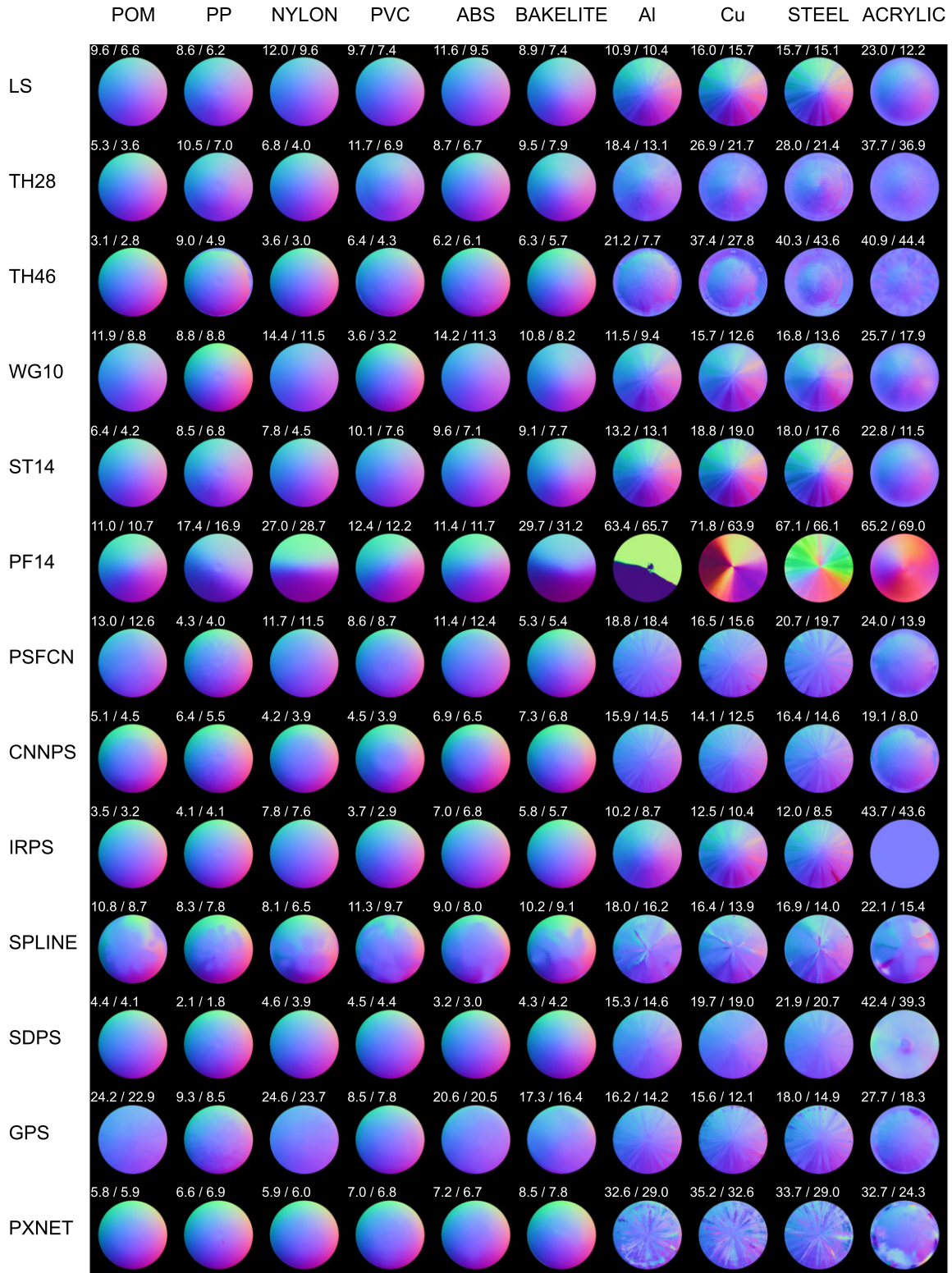[3]A more recent method PX-Net [2] is added in this document.

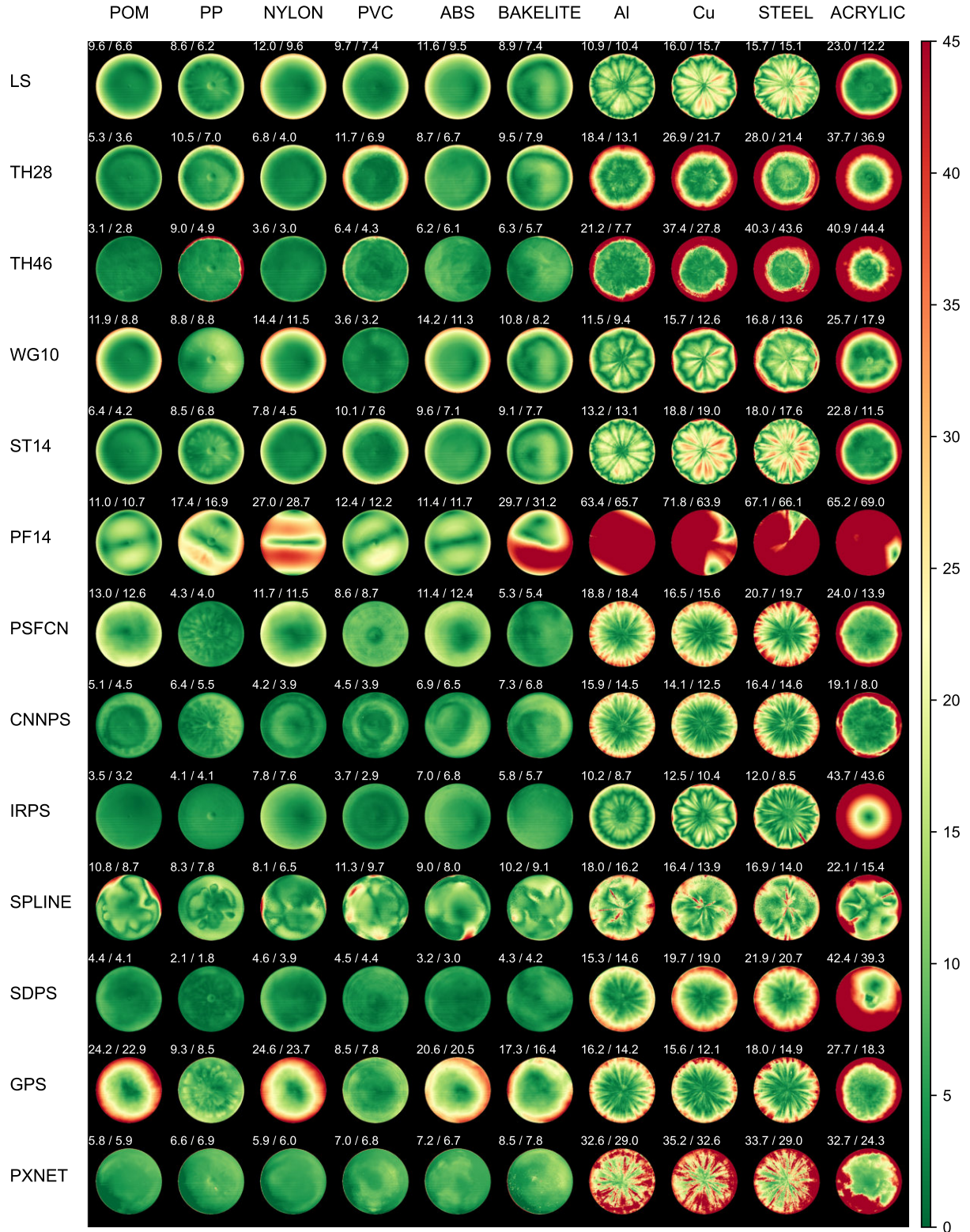Figure 7. BALL. Estimated normal map from 13 representative methods on various materials.

Figure 8. Error map of BALL of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
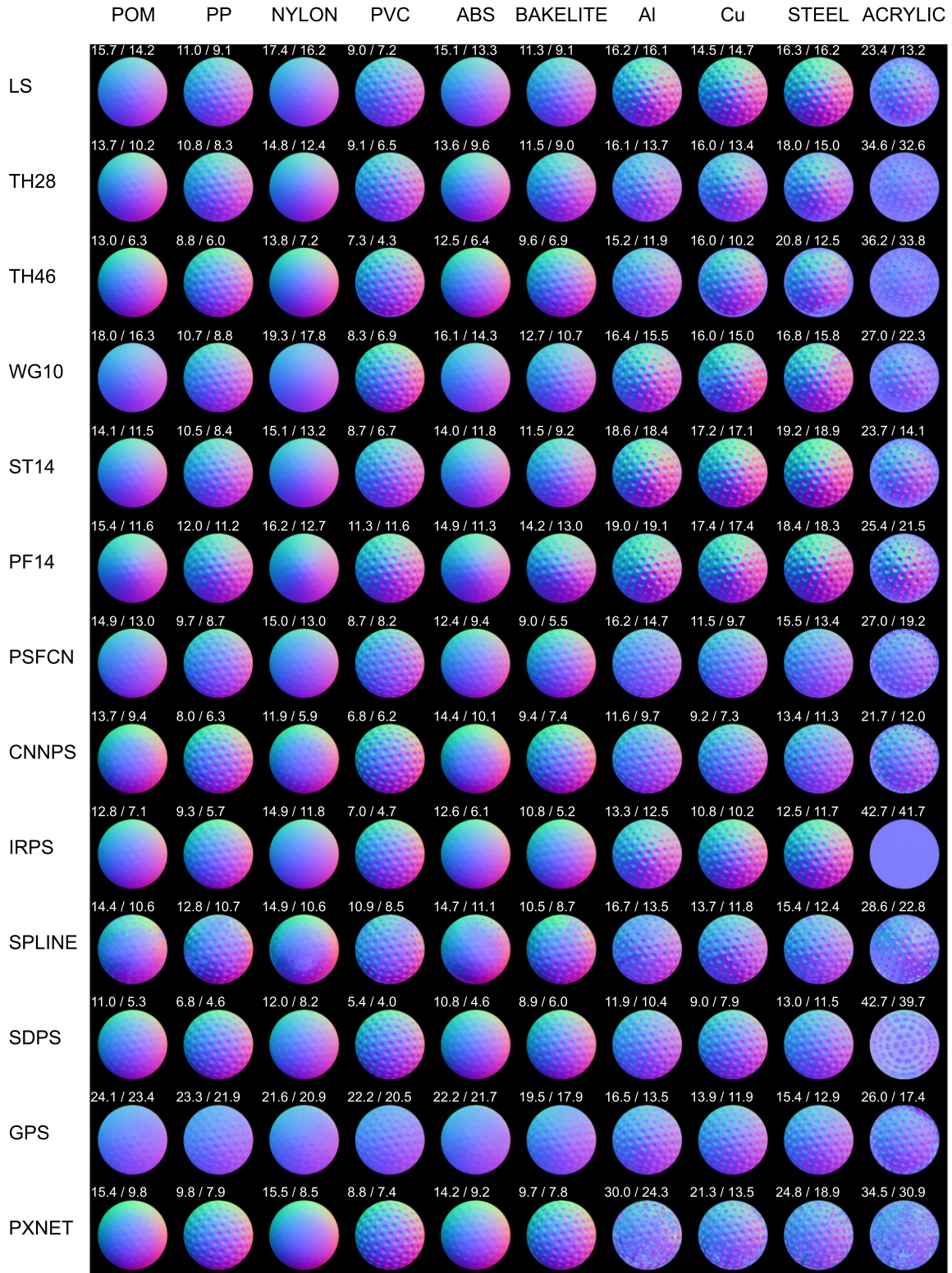
Figure 9. GOLF. Estimated normal map from 13 representative methods on various materials.

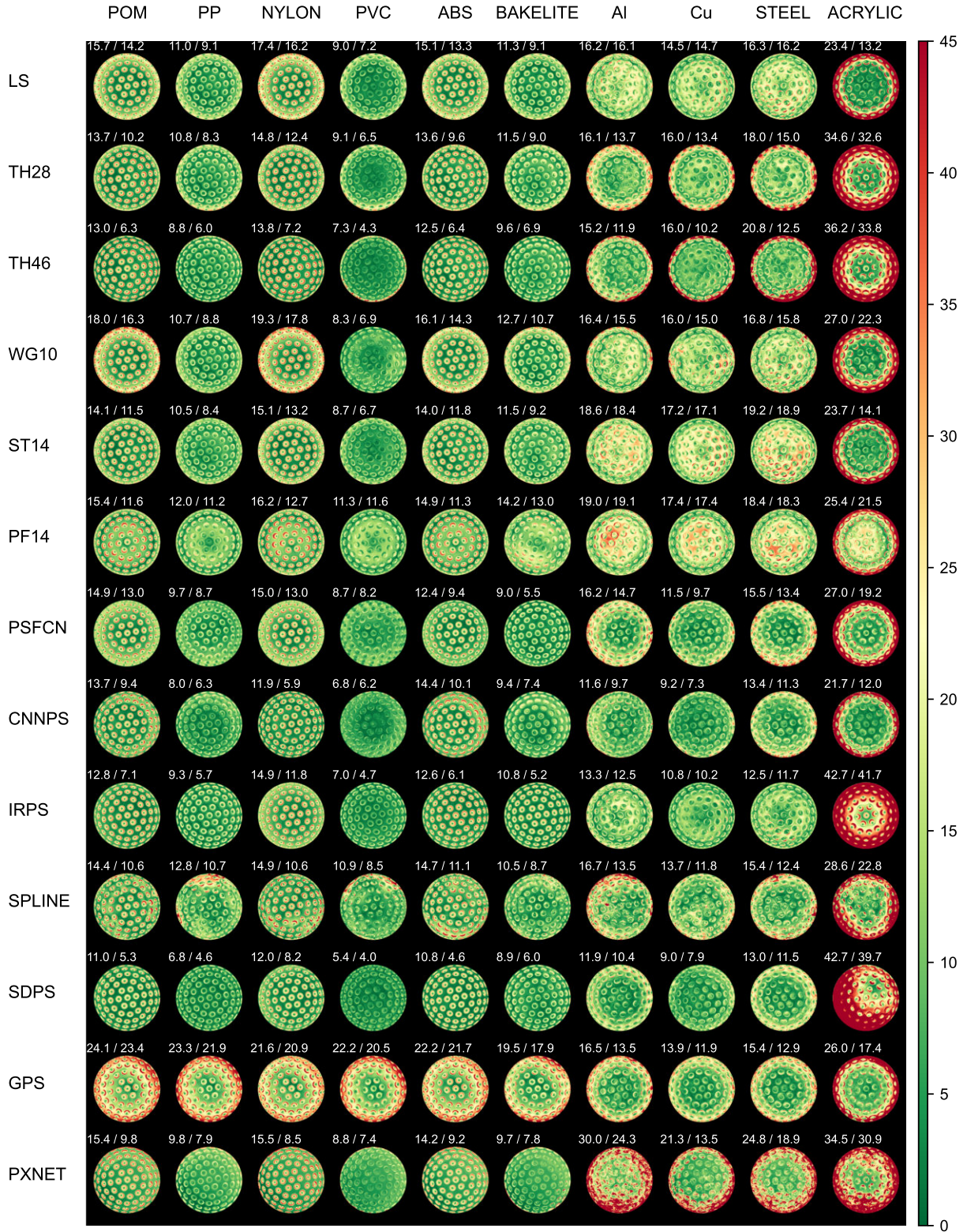Figure 10. Error map of GOLF of 13 representative methods on various materials. Errors increase from green (via yellow) to red.

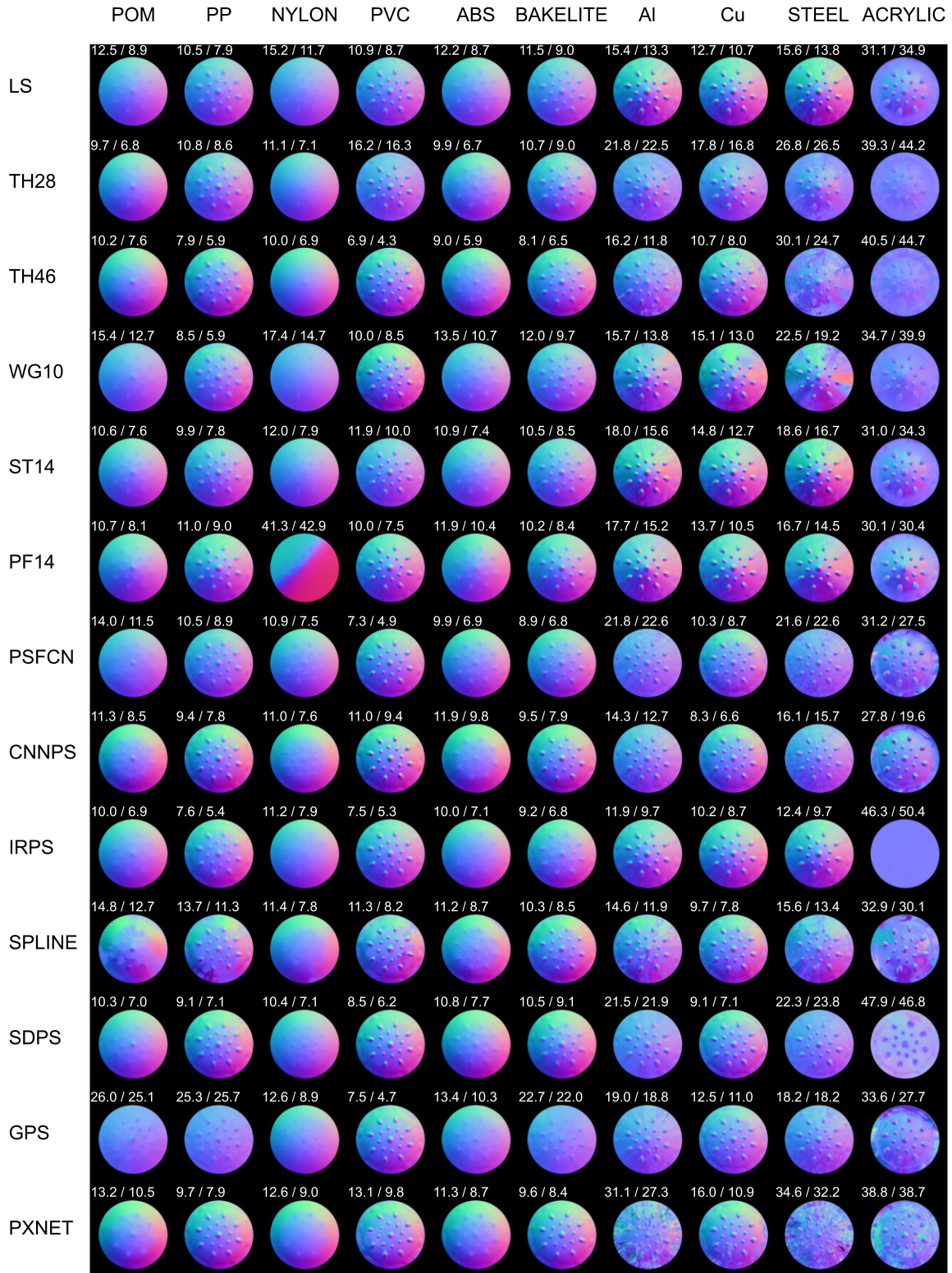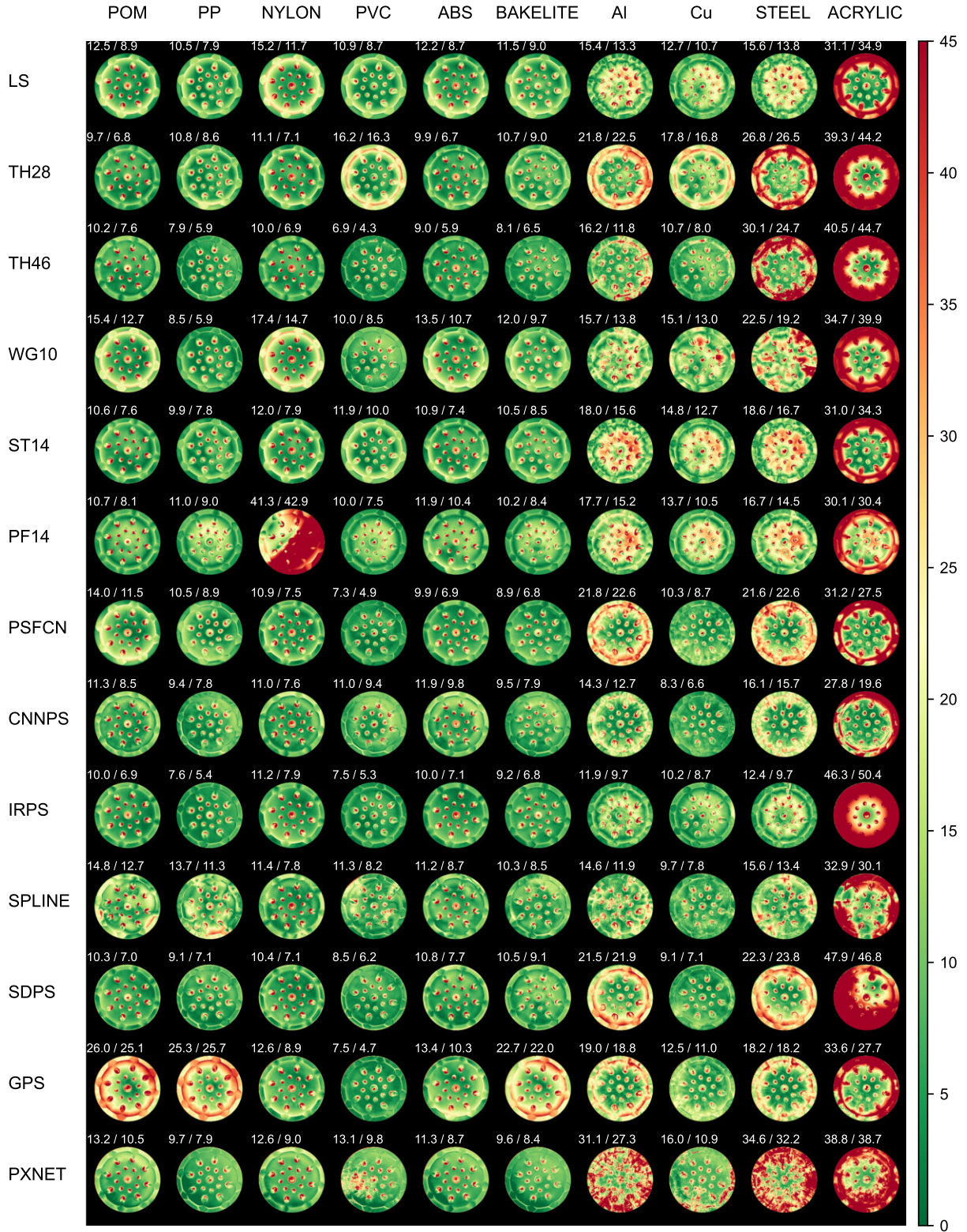Figure 11. Sᴘɪᴋᴇ. Estimated normal map from 13 representative methods on various materials.

Figure 12. Error map of SPIKE of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
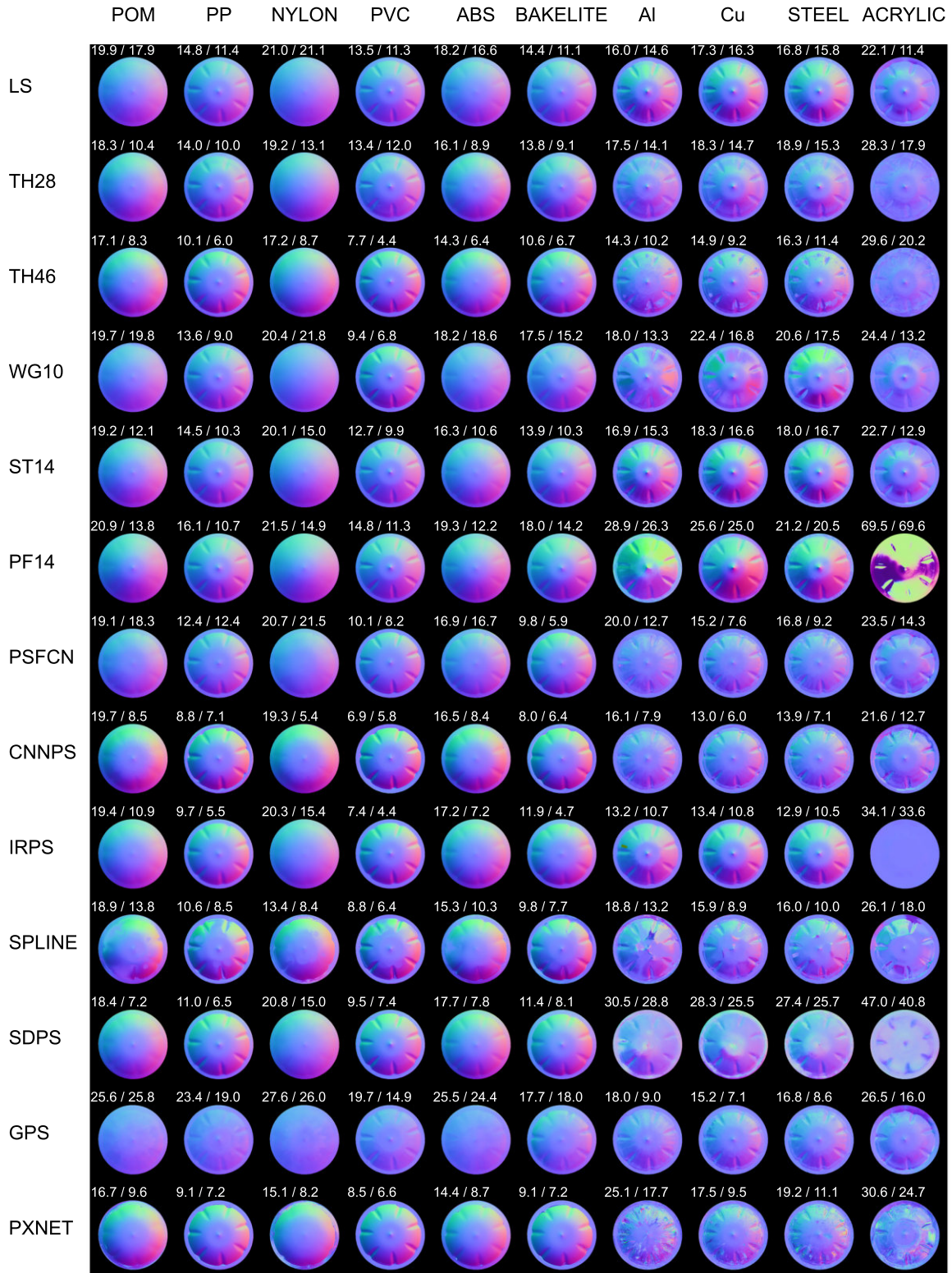
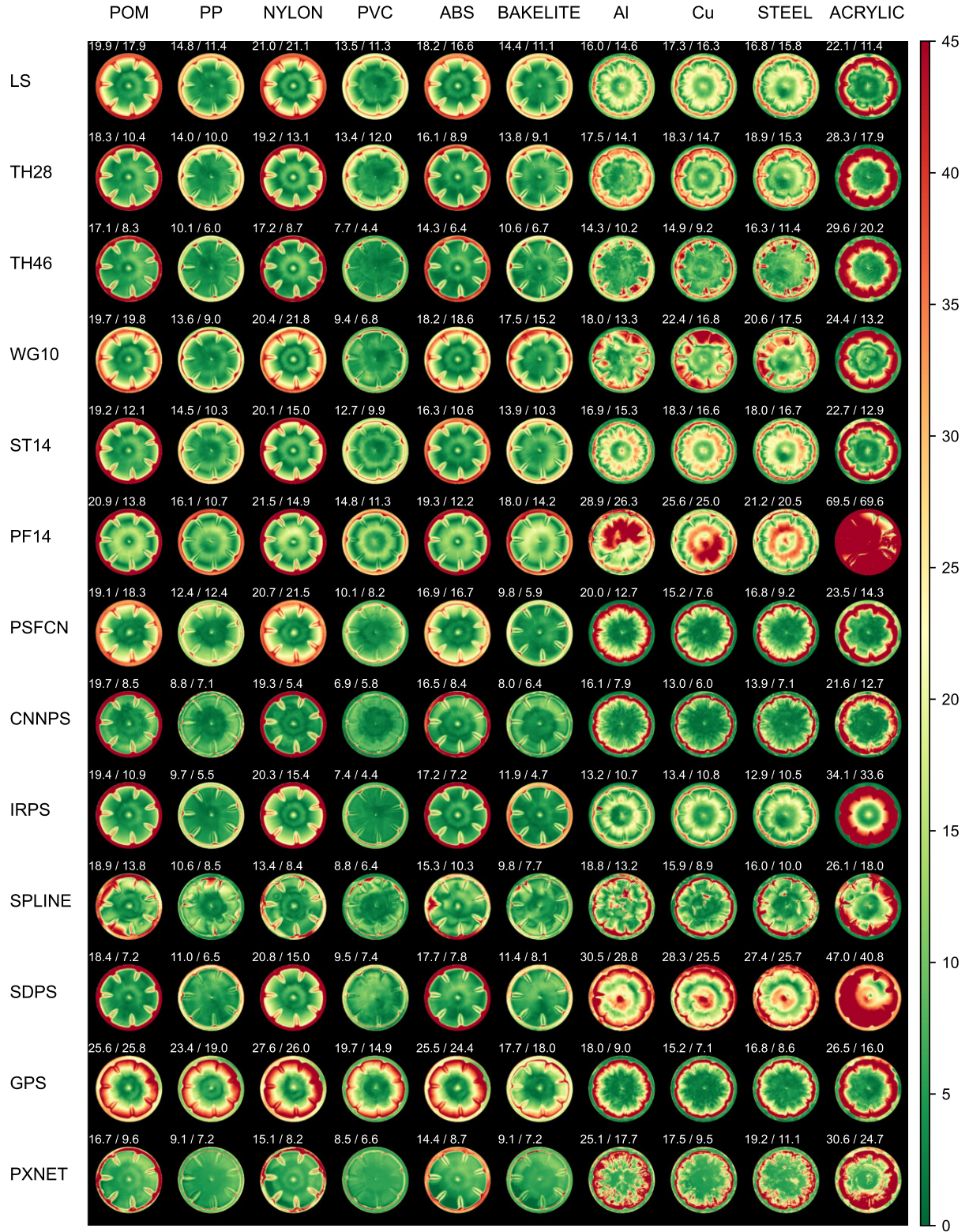Figure 13. NUT. Estimated normal map from 13 representative methods on various materials.

Figure 14. Error map of NUT of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
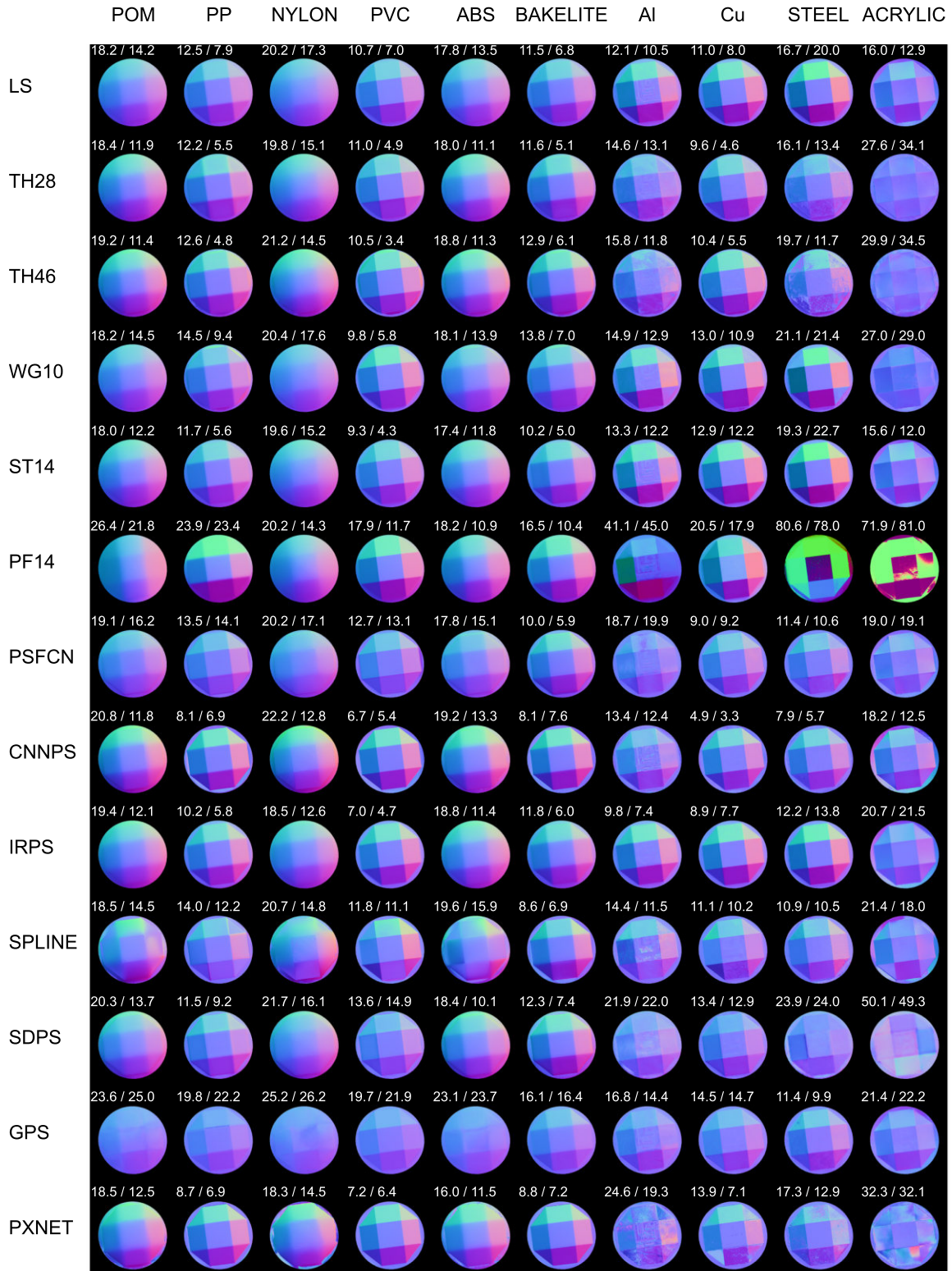
SQUARE

Figure 15. SQUARE. Estimated normal map from 13 representative methods on various materials.

Figure 16. Error map of SQUARE of 13 representative methods on various materials. Errors increase from green (via yellow) to red.

Figure 17. PENTAGON. Estimated normal map from 13 representative methods on various materials.
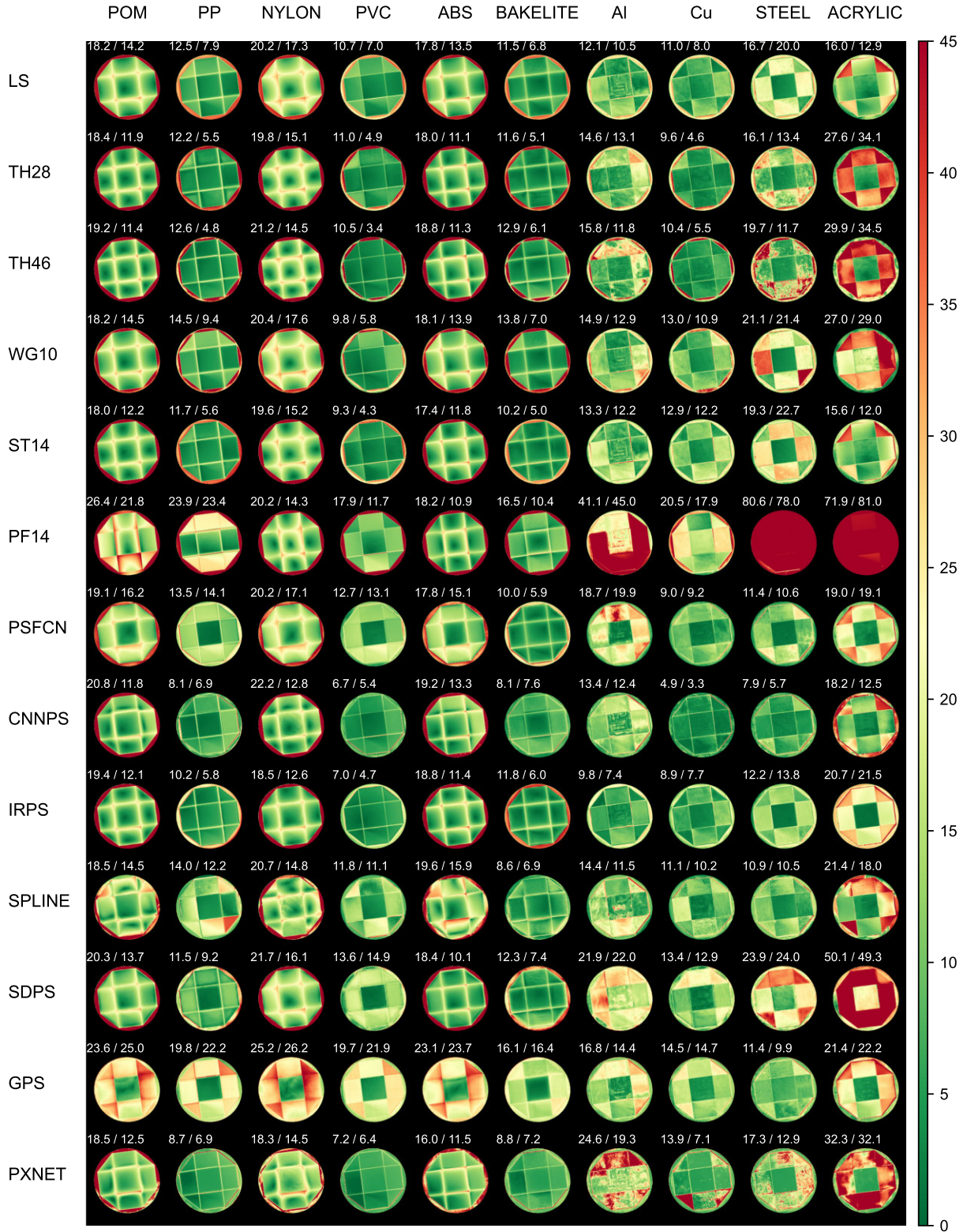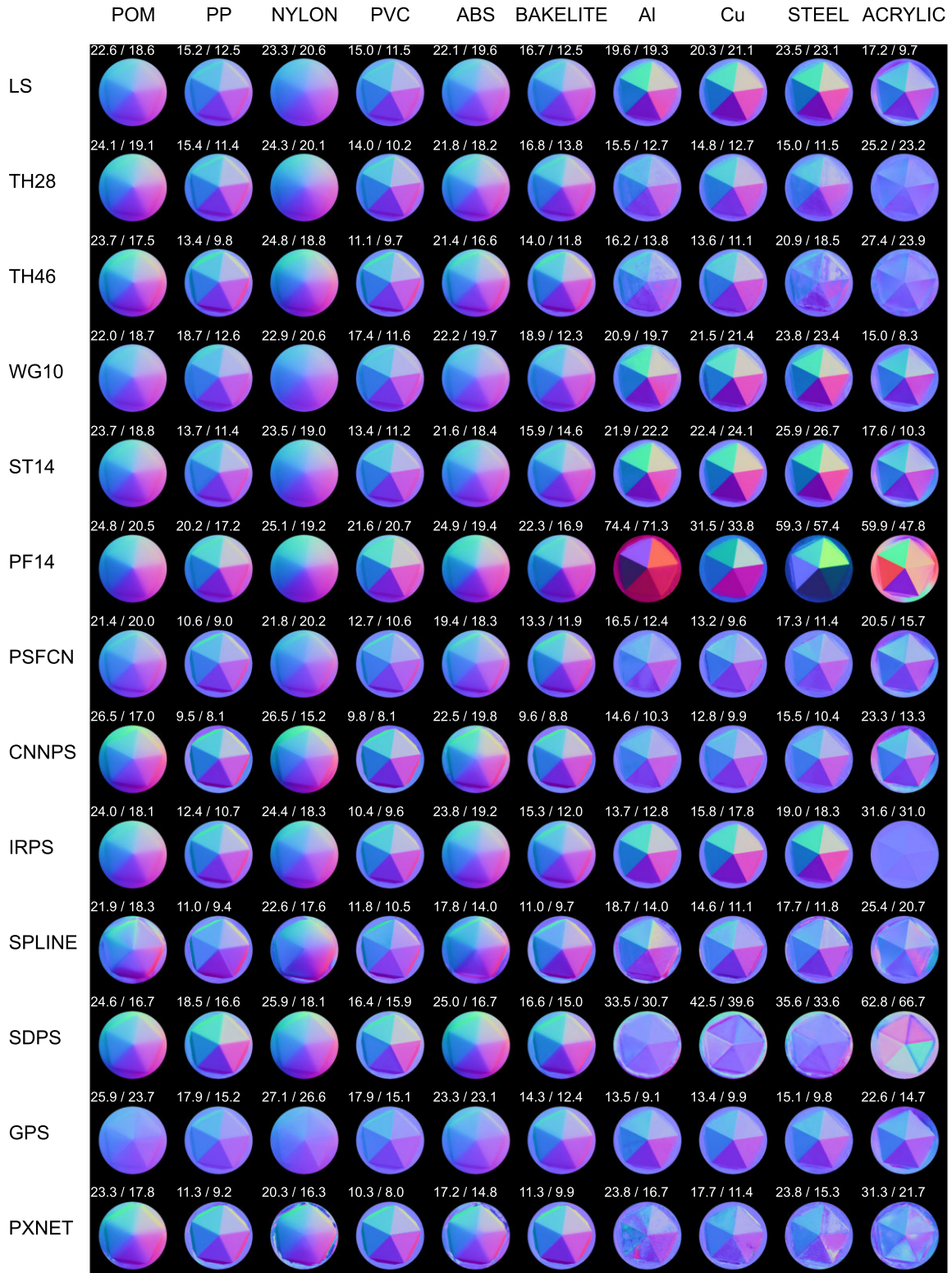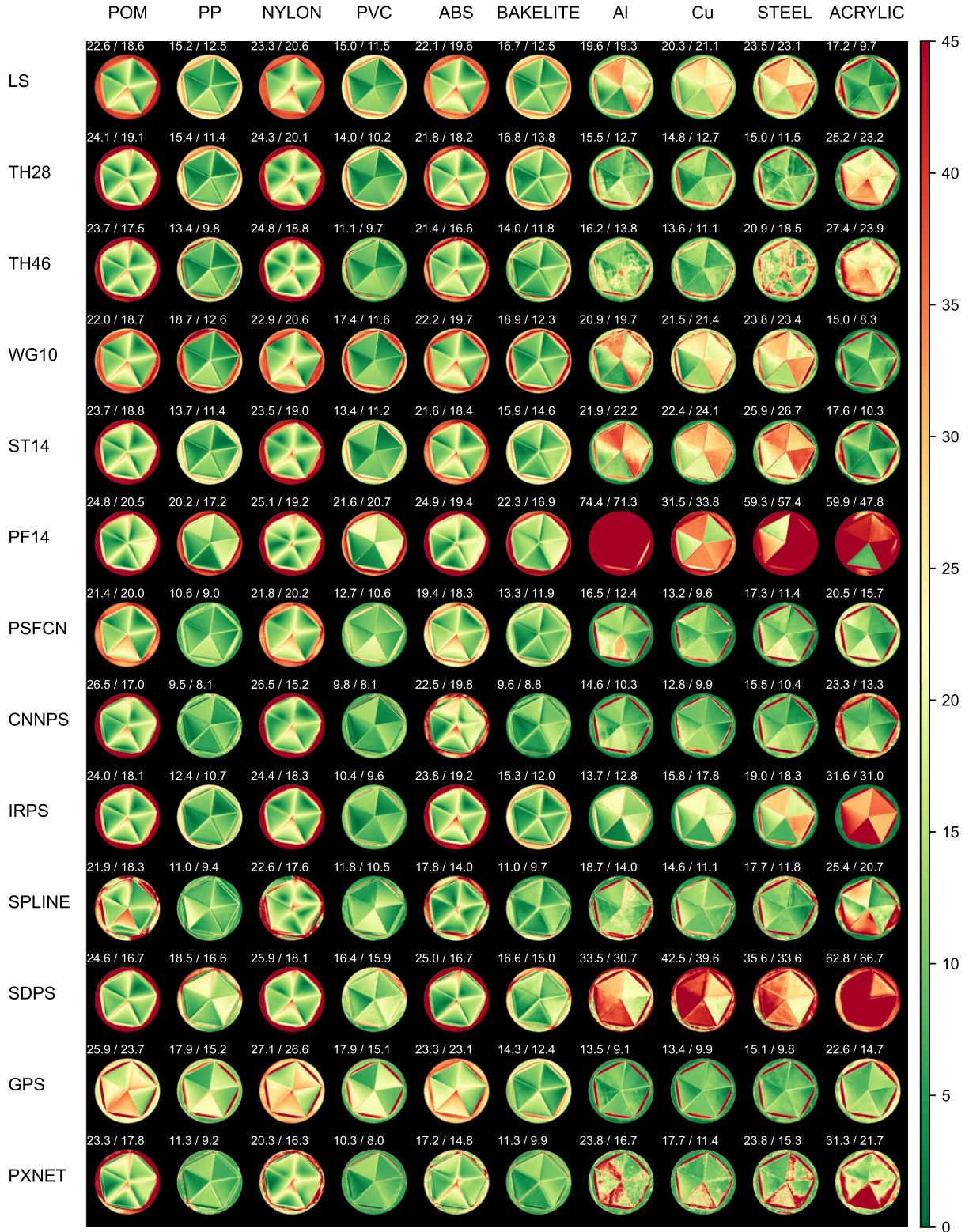
Figure 18. Error map of PENTAGON of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
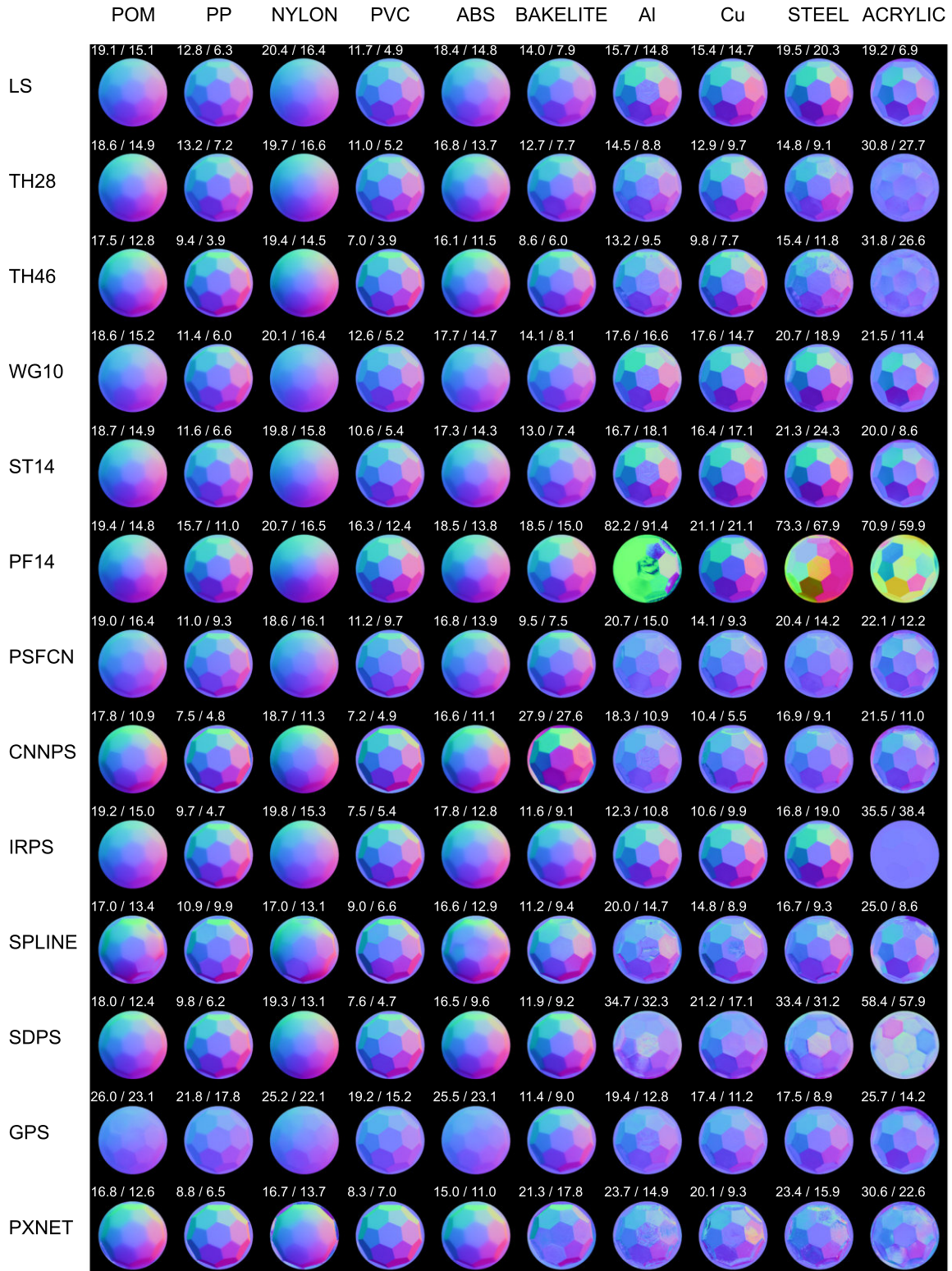
HEXAGON

	POM	PP	NYLON	PVC	ABS	BAKELITE	Al	Cu	STEEL	ACRYLIC
LS	19.1 / 15.1	12.8 / 6.3	20.4 / 16.4	11.7 / 4.9	18.4 / 14.8	14.0 / 7.9	15.7 / 14.8	15.4 / 14.7	19.5 / 20.3	19.2 / 6.9
TH28	18.6 / 14.9	13.2 / 7.2	19.7 / 16.6	11.0 / 5.2	16.8 / 13.7	12.7 / 7.7	14.5 / 8.8	12.9 / 9.7	14.8 / 9.1	30.8 / 27.7
TH46	17.5 / 12.8	9.4 / 3.9	19.4 / 14.5	7.0 / 3.9	16.1 / 11.5	8.6 / 6.0	13.2 / 9.5	9.8 / 7.7	15.4 / 11.8	31.8 / 26.6
WG10	18.6 / 15.2	11.4 / 6.0	20.1 / 16.4	12.6 / 5.2	17.7 / 14.7	14.1 / 8.1	17.6 / 16.6	17.6 / 14.7	20.7 / 18.9	21.5 / 11.4
ST14	18.7 / 14.9	11.6 / 6.6	19.8 / 15.8	10.6 / 5.4	17.3 / 14.3	13.0 / 7.4	16.7 / 18.1	16.4 / 17.1	21.3 / 24.3	20.0 / 8.6
PF14	19.4 / 14.8	15.7 / 11.0	20.7 / 16.5	16.3 / 12.4	18.5 / 13.8	18.5 / 15.0	82.2 / 91.4	21.1 / 21.1	73.3 / 67.9	70.9 / 59.9
PSFCN	19.0 / 16.4	11.0 / 9.3	18.6 / 16.1	11.2 / 9.7	16.8 / 13.9	9.5 / 7.5	20.7 / 15.0	14.1 / 9.3	20.4 / 14.2	22.1 / 12.2
CNNPS	17.8 / 10.9	7.5 / 4.8	18.7 / 11.3	7.2 / 4.9	16.6 / 11.1	27.9 / 27.6	18.3 / 10.9	10.4 / 5.5	16.9 / 9.1	21.5 / 11.0
IRPS	19.2 / 15.0	9.7 / 4.7	19.8 / 15.3	7.5 / 5.4	17.8 / 12.8	11.6 / 9.1	12.3 / 10.8	10.6 / 9.9	16.8 / 19.0	35.5 / 38.4
SPLINE	17.0 / 13.4	10.9 / 9.9	17.0 / 13.1	9.0 / 6.6	16.6 / 12.9	11.2 / 9.4	20.0 / 14.7	14.8 / 8.9	16.7 / 9.3	25.0 / 8.6
SDPS	18.0 / 12.4	9.8 / 6.2	19.3 / 13.1	7.6 / 4.7	16.5 / 9.6	11.9 / 9.2	34.7 / 32.3	21.2 / 17.1	33.4 / 31.2	58.4 / 57.9
GPS	26.0 / 23.1	21.8 / 17.8	25.2 / 22.1	19.2 / 15.2	25.5 / 23.1	11.4 / 9.0	19.4 / 12.8	17.4 / 11.2	17.5 / 8.9	25.7 / 14.2
PXNET	16.8 / 12.6	8.8 / 6.5	16.7 / 13.7	8.3 / 7.0	15.0 / 11.0	21.3 / 17.8	23.7 / 14.9	20.1 / 9.3	23.4 / 15.9	30.6 / 22.6

Figure 19. HEXAGON. Estimated normal map from 13 representative methods on various materials.
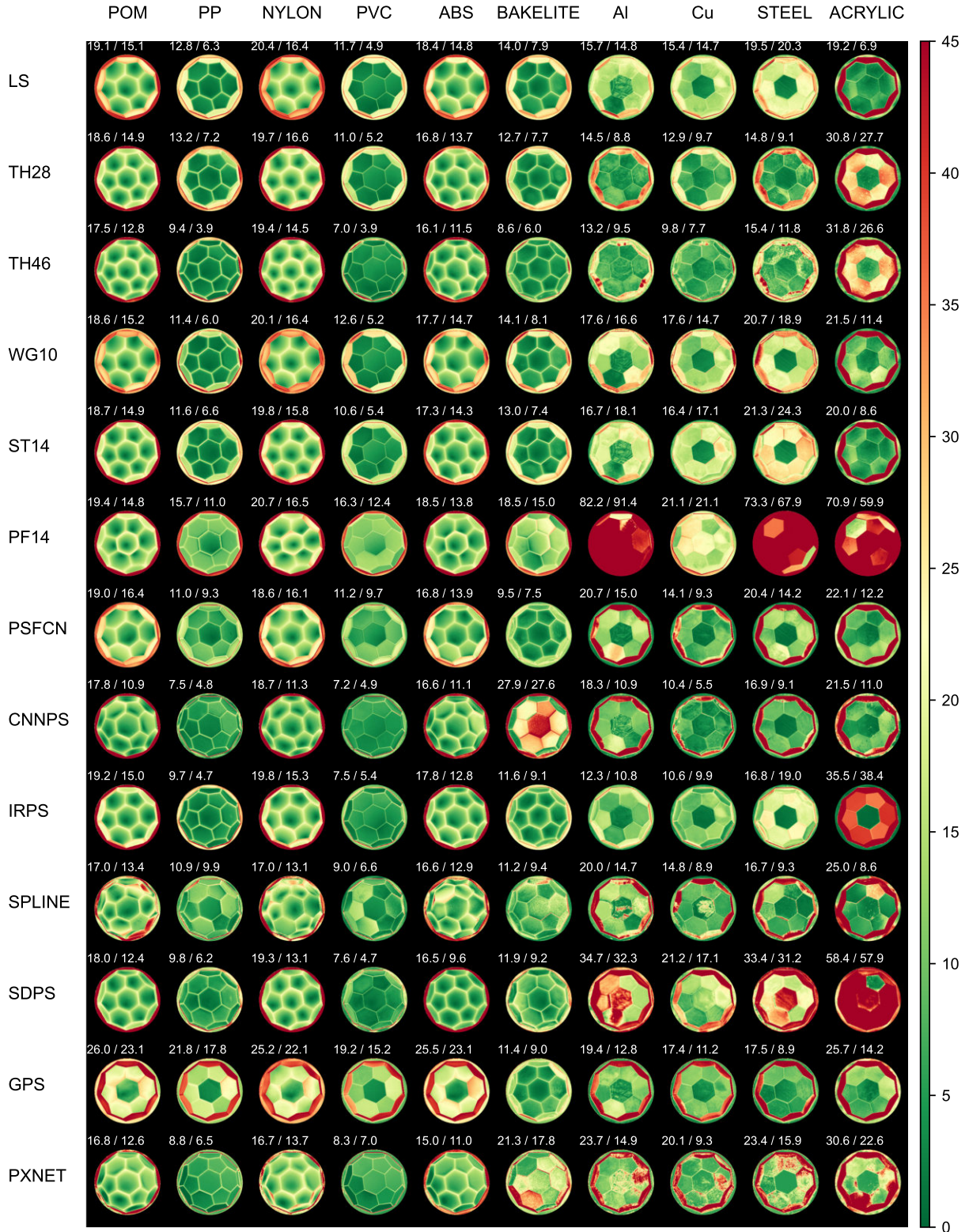
Figure 20. Error map of HEXAGON of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
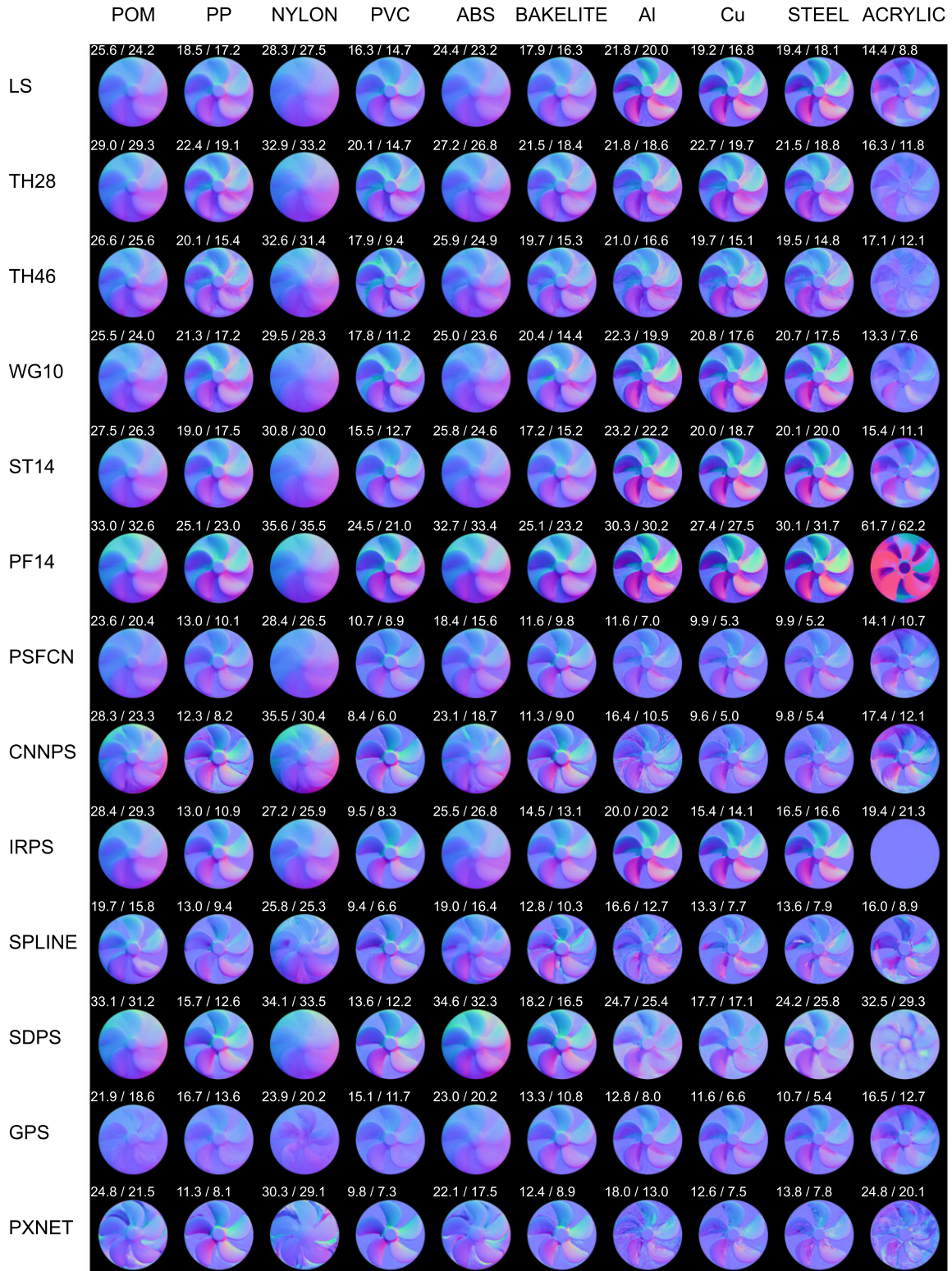
PROPELLER

Figure 21. PROPELLER. Estimated normal map from 13 representative methods on various materials.

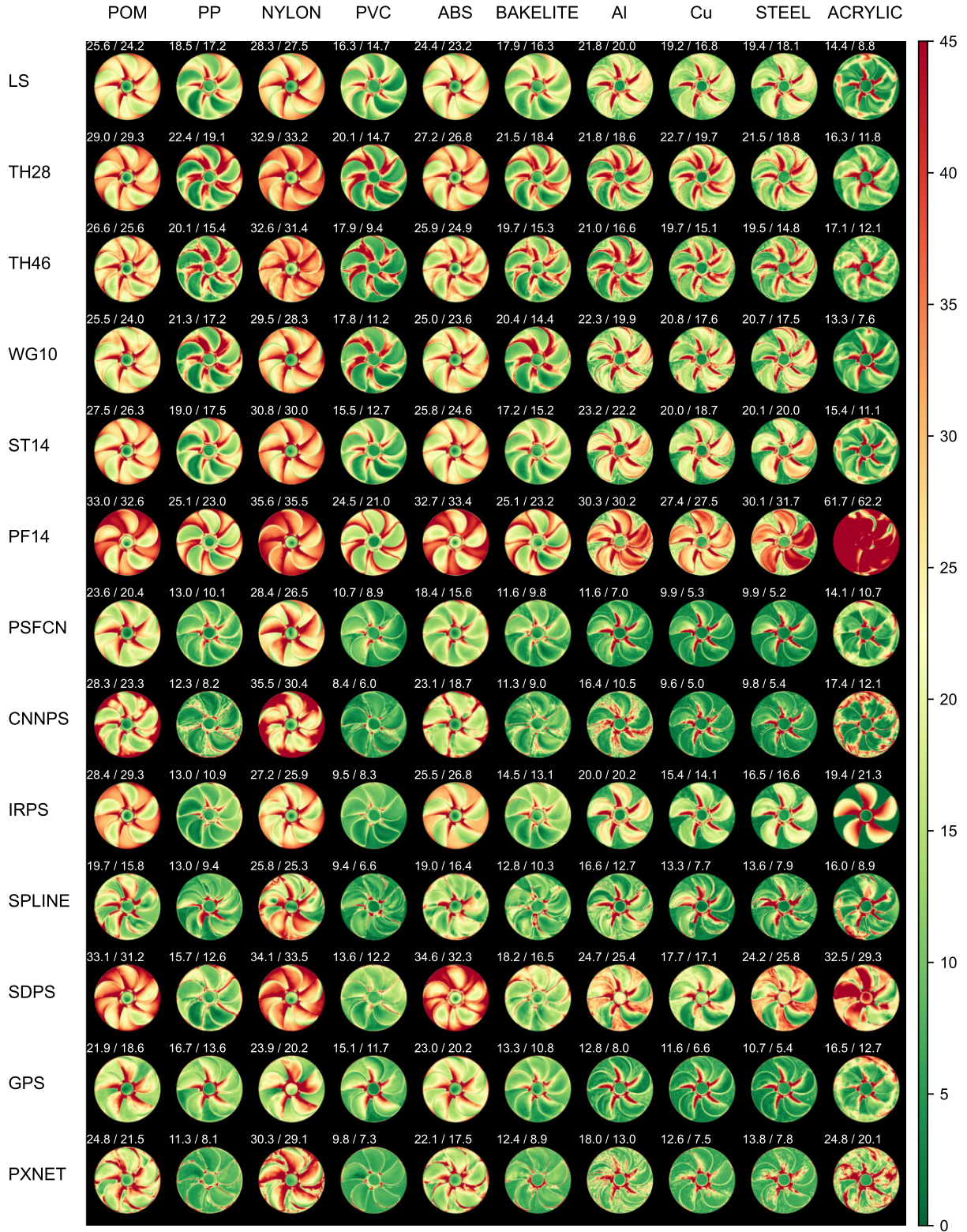Figure 22. Error map of PROPELLER of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
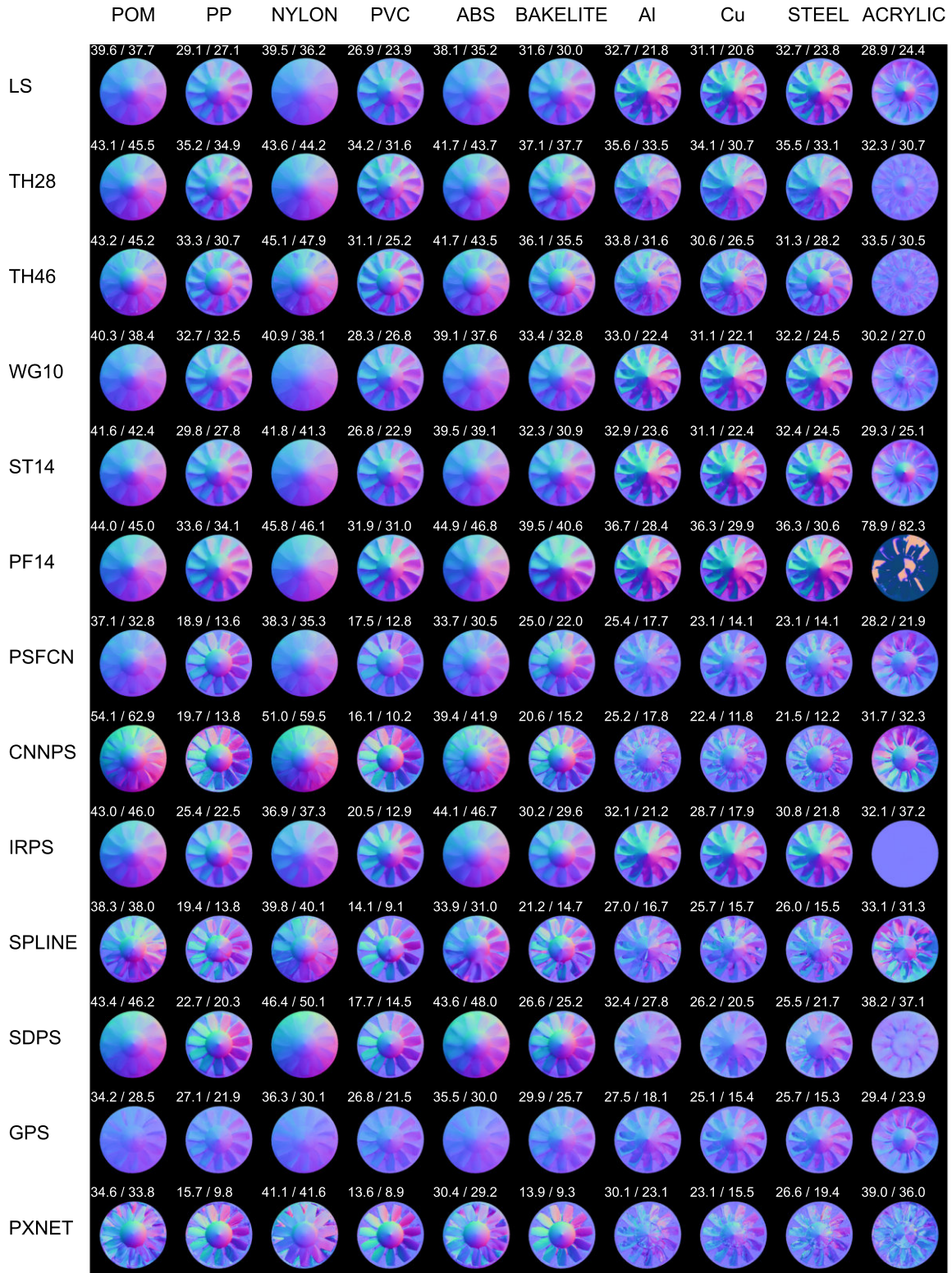
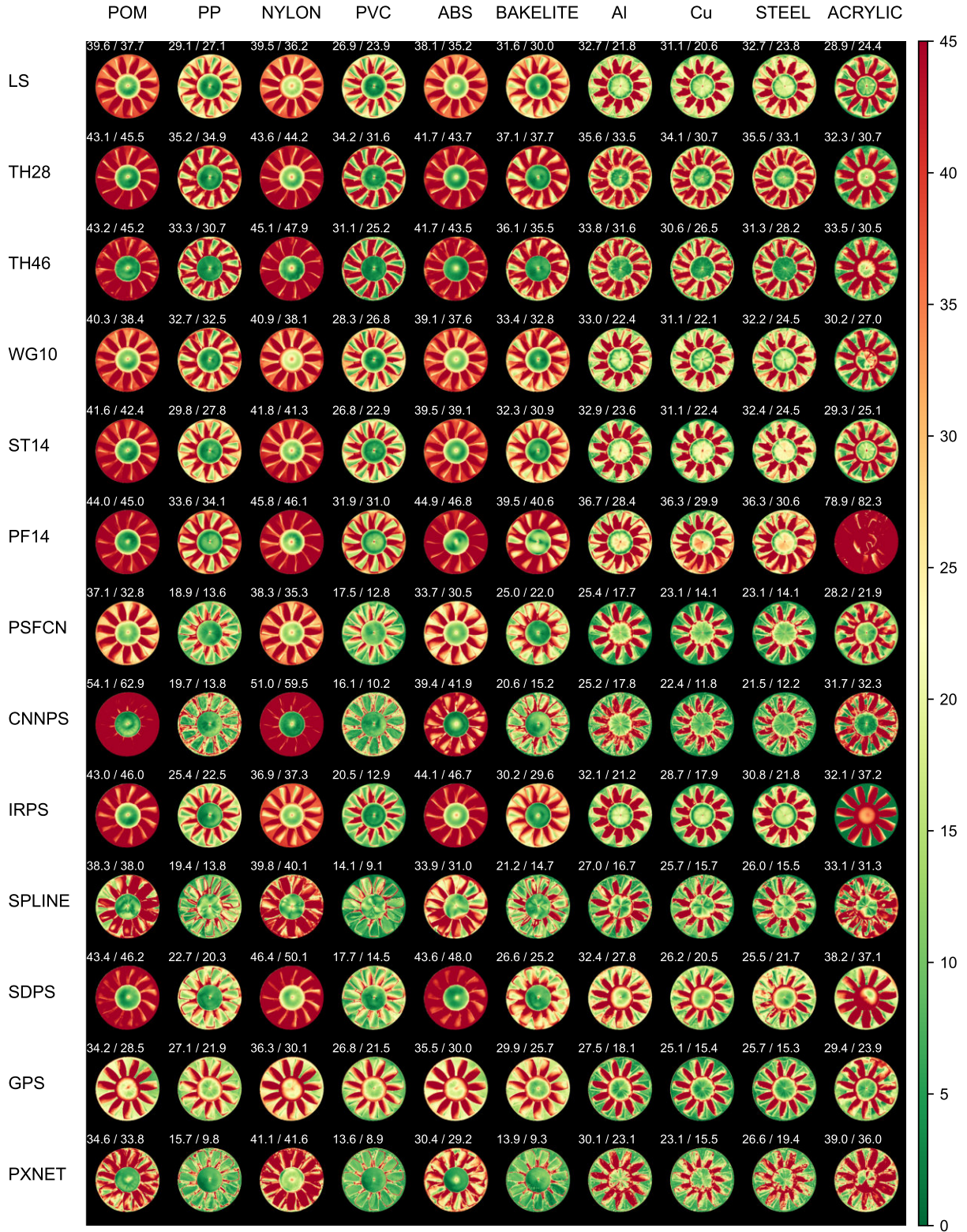Figure 23. TURBINE. Estimated normal map from 13 representative methods on various materials.

Figure 24. Error map of TURBINE of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
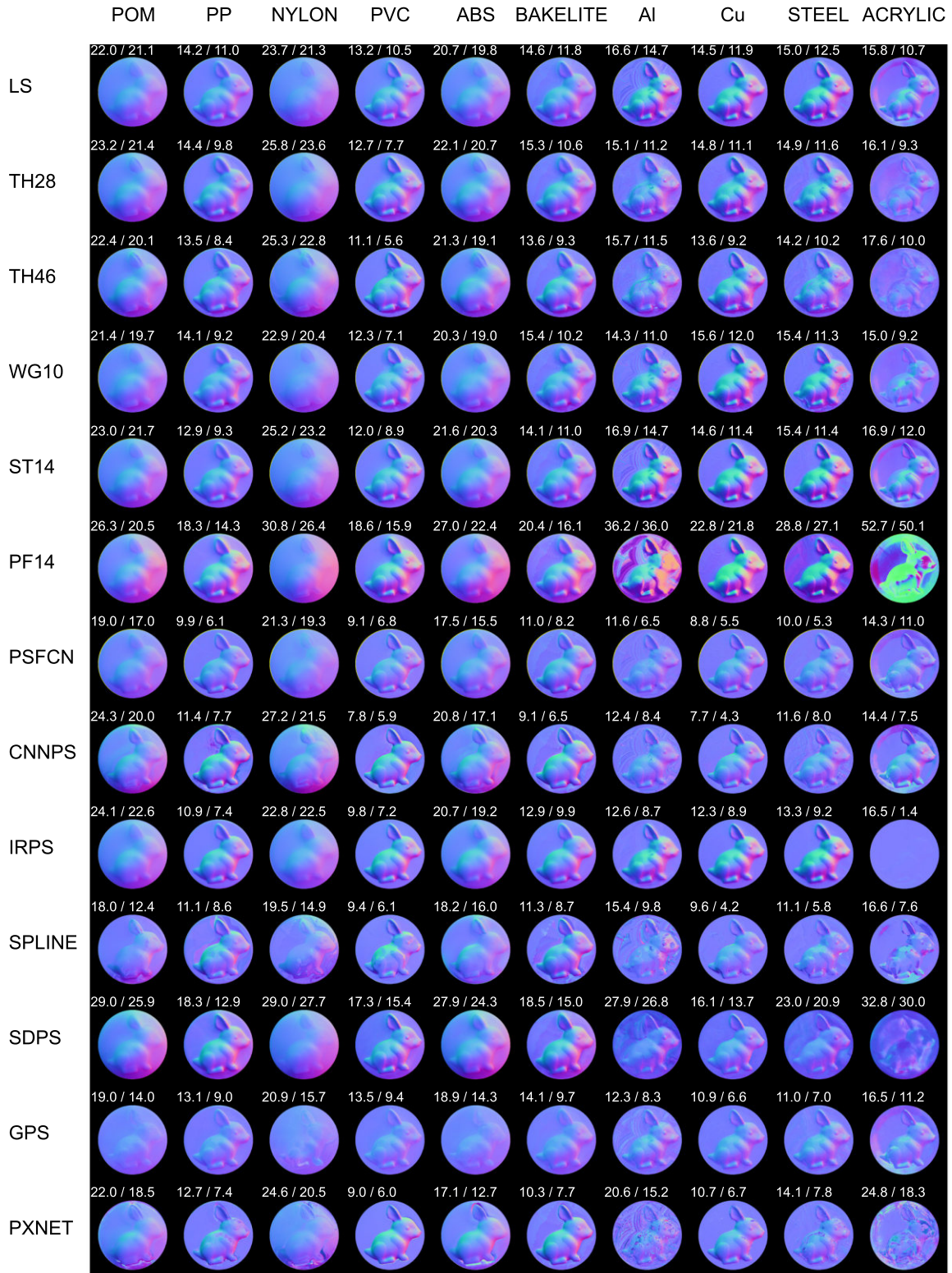
BUNNY

Figure 25. BUNNY. Estimated normal map from 13 representative methods on various materials.
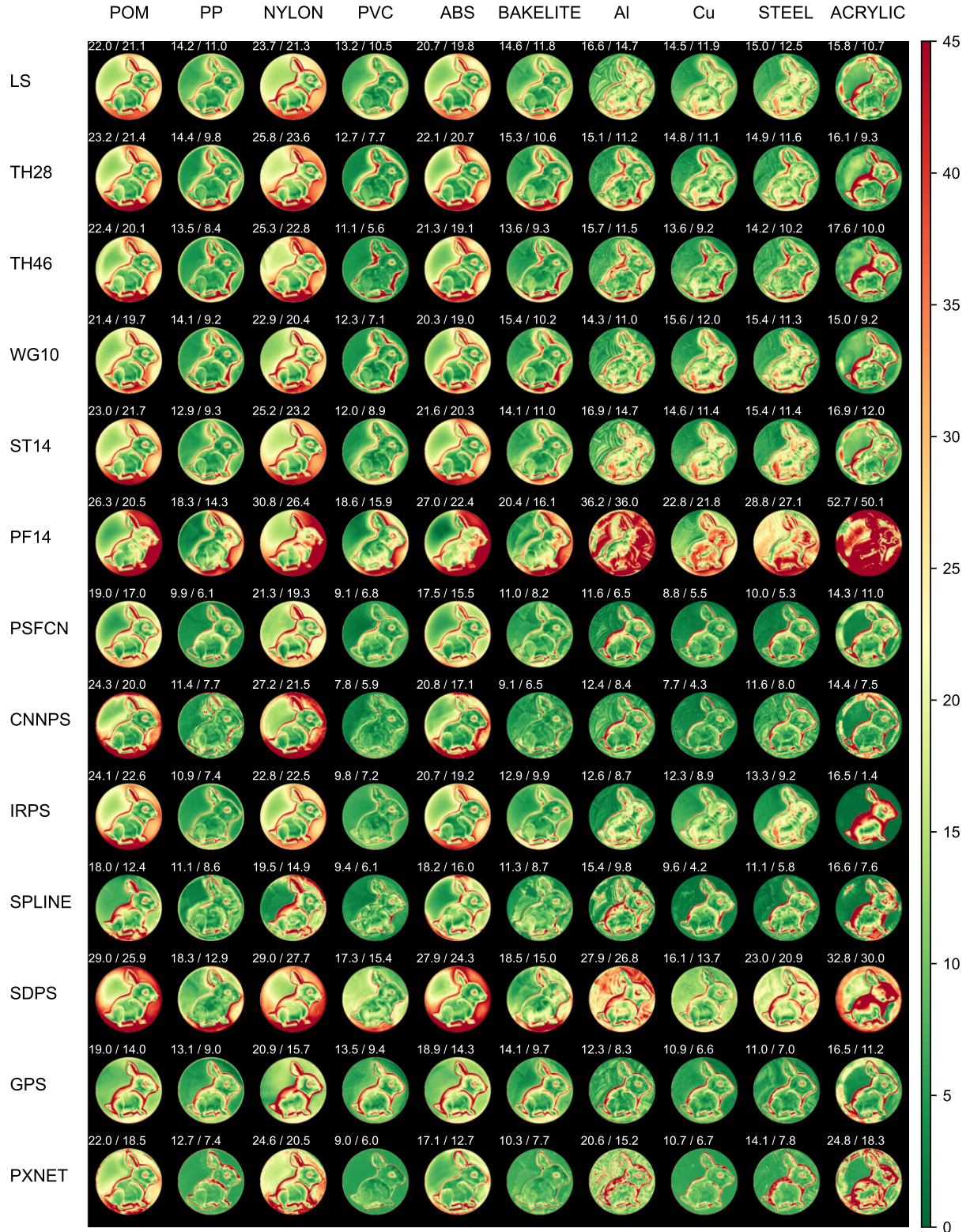
Figure 26. Error map of BUNNY of 13 representative methods on various materials. Errors increase from green (via yellow) to red.
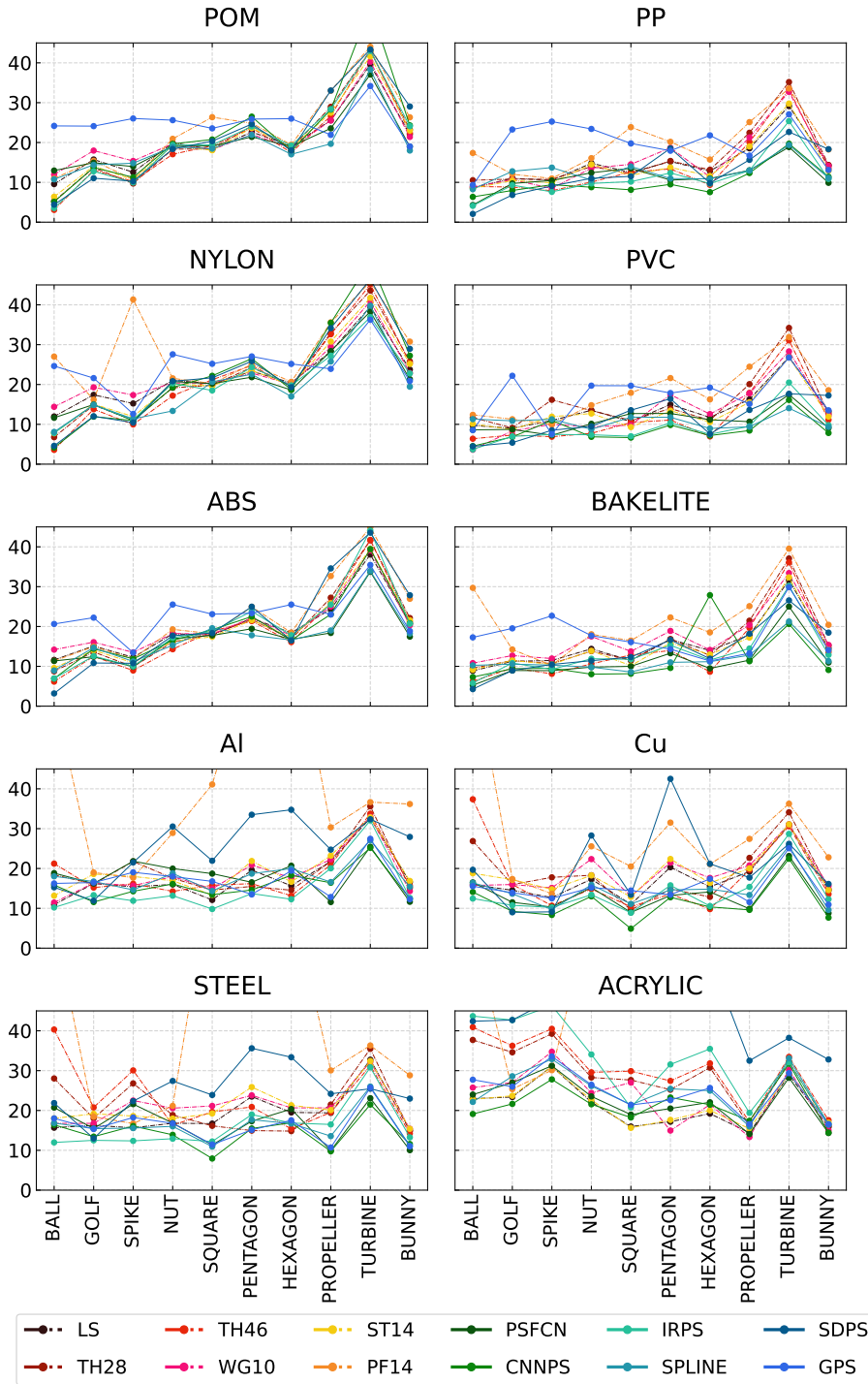
Figure 27. Material (per sub-figure)-shape (X-axis) error (Y-axis) plot for all methods. Best viewed in color and electronic version for magnified details.

We further provide "Material-Shape Error Plot for All Methods" in Fig. 27. This is the "dual" version of Fig. 5 analyzed in Sec. 4.1 of the main paper. The complete error plot for all evaluated methods on 100 sets of data in our dataset is provided in Fig. 28
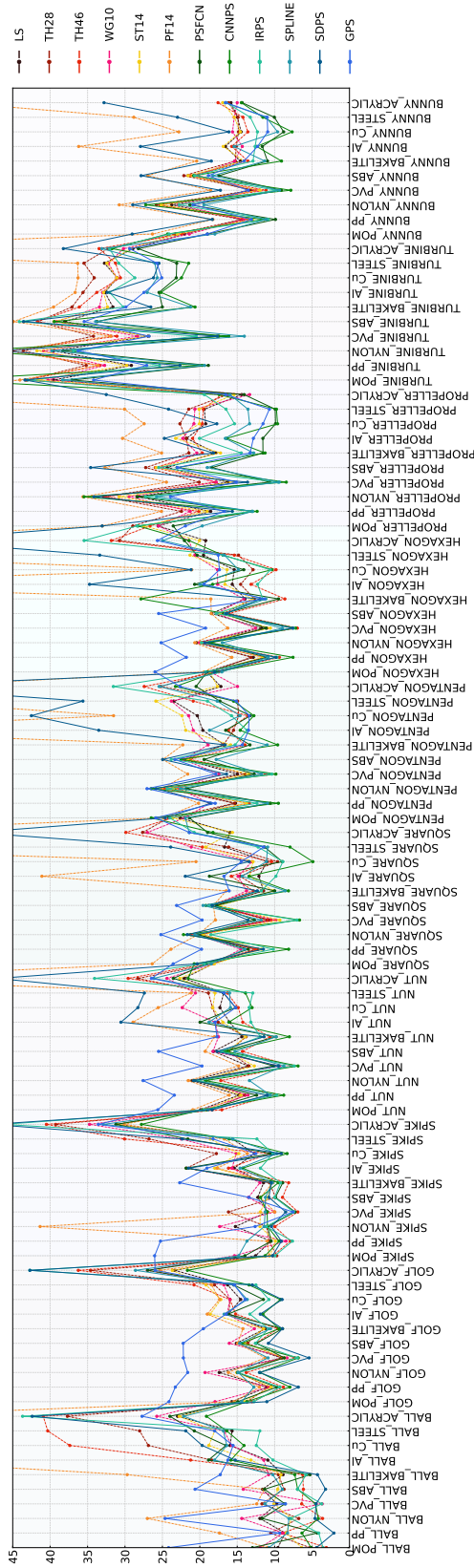
Figure 28. Complete error plot for all 100 sets of data for all evaluated methods. Best viewed in color and electronic version for magnified details.

References

[1] Álvaro González. Measurement of areas on a sphere using fibonacci and latitude–longitude lattices. *Mathematical Geosciences*, 42(1):49–64, 2010. 3

[2] Fotios Logothetis, Ignas Budvytis, Roberto Mecca, and Roberto Cipolla. PX-NET: Simple and efficient pixel-wise training of photometric stereo networks. In *Proc. ICCV*, 2021. 5

[3] Boxin Shi, Zhipeng Mo, Zhe Wu, Dinglong Duan, Sai-Kit Yeung, and Ping Tan. A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. *IEEE transactions on pattern analysis and machine intelligence*, 41(2):271–284, 2019. 4, 5