Decoupling Fine Detail and Global Geometry for Compressed Depth Map Super-Resolution

Anonymous CVPR submission

Paper ID 9659

1. Proof of Employing the Neumann Series for Approximating the Matrix Inverse

Theorem 1 (Neumann Series Approximation for $(A^TA)^{-1}$). *Let $A \in \mathbb{R}^{n\times m}$ be a full-rank matrix (with $m \geq n$), and let $\lambda_{\max}$ denote the largest eigenvalue of A^TA. Then, the inverse of A^TA can be approximated using the Neumann series as:*

$$(A^TA)^{-1} = \frac{1}{\lambda_{\max}} \sum_{k=0}^{\infty} \left(I - \frac{A^TA}{\lambda_{\max}}\right)^k, \tag{1}$$

where the series converges when $\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}} < 1$, and $\lambda_{\min}$ is the smallest eigenvalue of A^TA.

Proof. Let A^TA be the Gram matrix of A, which is symmetric and positive definite. Denote the eigenvalues of A^TA as $\lambda_1, \lambda_2, \ldots, \lambda_n$ with $\lambda_{\max} = \max_i \lambda_i$ and $\lambda_{\min} = \min_i \lambda_i$. We prove the theorem in the following steps.

Step 1: Normalizing A^TA. Write A^TA as:

$$A^TA = \lambda_{\max} I - B, \tag{2}$$

where $B = \lambda_{\max} I - A^TA$. The spectral radius of B is $\rho(B) = \lambda_{\max} - \lambda_{\min}$, and the normalized matrix $\frac{B}{\lambda_{\max}}$ satisfies:

$$\left\| \frac{B}{\lambda_{\max}} \right\| = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max}} < 1. \tag{3}$$

This ensures that the Neumann series expansion is valid for $\frac{B}{\lambda_{\max}}$.

Step 2: Neumann Series Expansion. The inverse of A^TA is:

$$(A^TA)^{-1} = \frac{1}{\lambda_{\max}} \left(I - \frac{B}{\lambda_{\max}}\right)^{-1}. \tag{4}$$

Using the Neumann series expansion $(I - C)^{-1} = \sum_{k=0}^{\infty} C^k$ for $\|C\| < 1$, we expand:

$$\left(I - \frac{B}{\lambda_{\max}}\right)^{-1} = \sum_{k=0}^{\infty} \left(\frac{B}{\lambda_{\max}}\right)^k. \tag{5}$$

Substituting $B = \lambda_{\max} I - A^TA$, we have:

$$(A^TA)^{-1} = \frac{1}{\lambda_{\max}} \sum_{k=0}^{\infty} \left(\frac{\lambda_{\max} I - A^TA}{\lambda_{\max}}\right)^k. \tag{6}$$

Step 3: Final Expression. Simplifying further, the inverse of A^TA is expressed as:

$$(A^TA)^{-1} = \frac{1}{\lambda_{\max}} \sum_{k=0}^{\infty} \left(I - \frac{A^TA}{\lambda_{\max}}\right)^k. \tag{7}$$

Step 4: Convergence Condition. The convergence of the Neumann series requires $\|I - \frac{A^TA}{\lambda_{\max}}\| < 1$. This is satisfied since:

$$\|I - \frac{A^TA}{\lambda_{\max}}\| = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max}} < 1. \tag{8}$$

Conclusion. The Neumann series approximation is valid under the given conditions, and the expression for $(A^TA)^{-1}$ is derived as stated.

2. The Process of Synthesizing a Compressed Depth Map

We adopt the methodology proposed in [1] to synthesize a novel dataset, referred to as Compressed-NYU, derived from the NYU Depth V2 dataset [2]. For a given depth map, we perform bit-depth compression by reducing four bits, eight-fold downsampling and noise injection, which is mathematically represented as follows:

$$\begin{cases} \bar{D} = H_{down}(D_{\downarrow}), \\ \hat{D} = \bar{D} + \bar{D} \odot N_1 + N_2, \end{cases} \tag{9}$$

where D, $\bar{D}$ and $\hat{D}$ denote the original depth map, the intermediate depth map and the degraded depth map, respectively. $\downarrow$, $H_{down}(\cdot)$ and $\odot$ represent the transformation of bit-depth compression and eight-fold downsampling,

CVPR
#9659

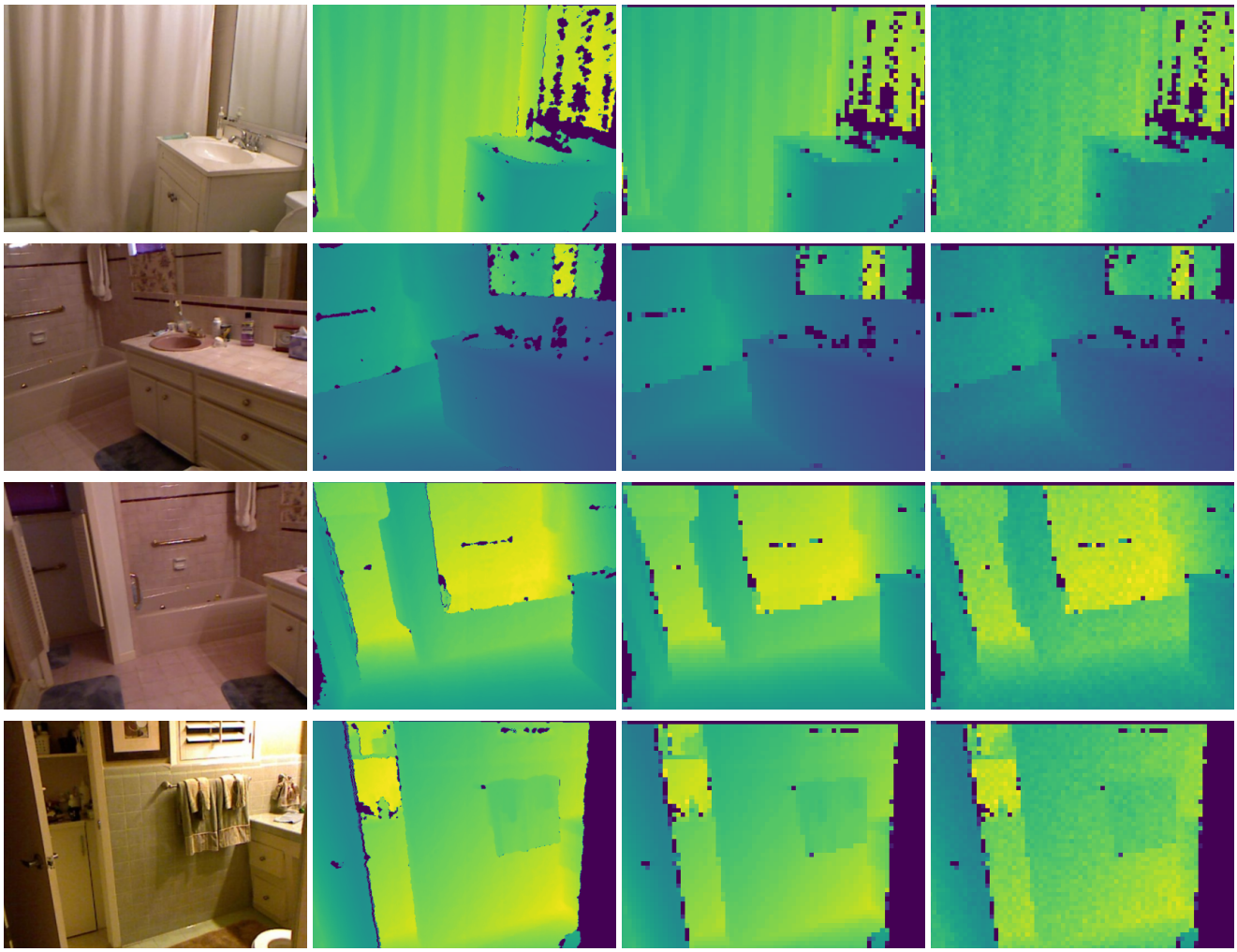

Figure 1. **Visual Results of Several Samples in the Process of Synthesizing a Compressed Map.**

element-wise multiplication, respectively. $N_1 \sim \mathcal{N}(0, \sigma_1^2)$ and $N_2 \sim \mathcal{N}(0, \sigma_2^2)$ denotes two distinct types of noise, where σ_1 and σ_2 are set to 0.02 and 0.05, respectively. We present the visual results of several samples in the synthesis process, as shown in Figure 1.

3. More Visual Results

We provide additional visual results to further demonstrate the effectiveness of the proposed method, as illustrated in 2, 3, 4, 5, 6, 7. In comparison to other methods, our approach yields reconstructed depth maps of superior quality, with enhanced recovery of fine geometric details.

References

[1] Marcos V Conde, Florin-Alexandru Vasluianu, Jinhui Xiong, Wei Ye, Rakesh Ranjan, and Radu Timofte. Compressed depth map super-resolution and restoration: Aim 2024 challenge results. *arXiv preprint arXiv:2409.16277*, 2024. 1

[2] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *Proceedings of European Conference on Computer Vision*, 2012. 1

[3] Zhengxue Wang, Zhiqiang Yan, and Jian Yang. Sgnet: Structure guided network via gradient-frequency awareness for depth map super-resolution. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024. 3, 4, 5, 6, 7, 8

[4] Lihe Yang, Bingyi Kang, Zilong Huang, Xiaogang Xu, Jiashi Feng, and Hengshuang Zhao. Depth anything: Unleashing the power of large-scale unlabeled data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2024. 3, 4, 5, 6, 7, 8

CVPR
#9659

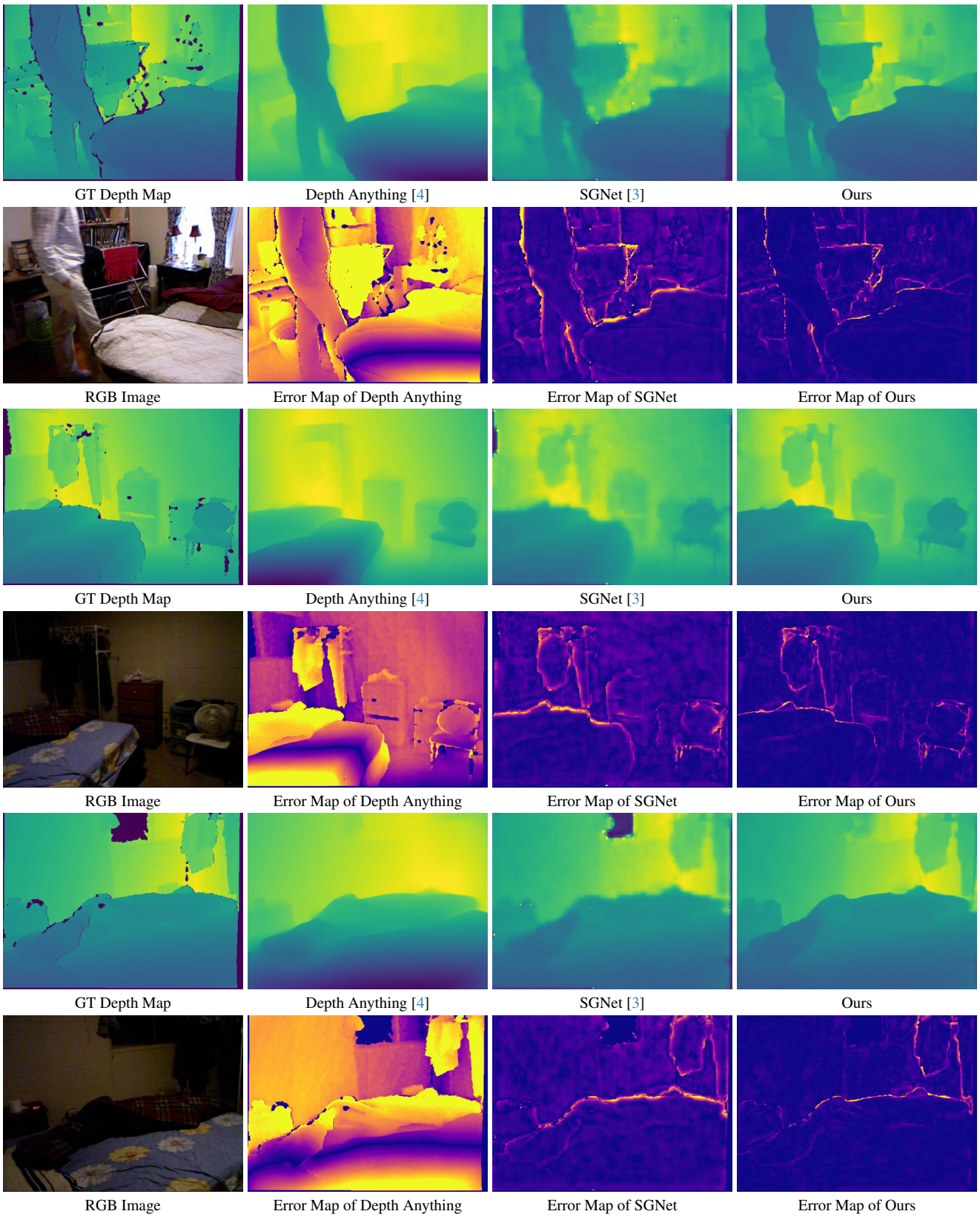

Figure 2. **Additional Visual Comparisons on Compressed-NYU Dataset between Our GDNet and Other Well-known Methods.** The first row displays the ground truth depth alongside the predicted results of various approaches, while the second row shows the corresponding RGB image and error maps. For error maps, darker areas indicate smaller errors.

CVPR
#9659

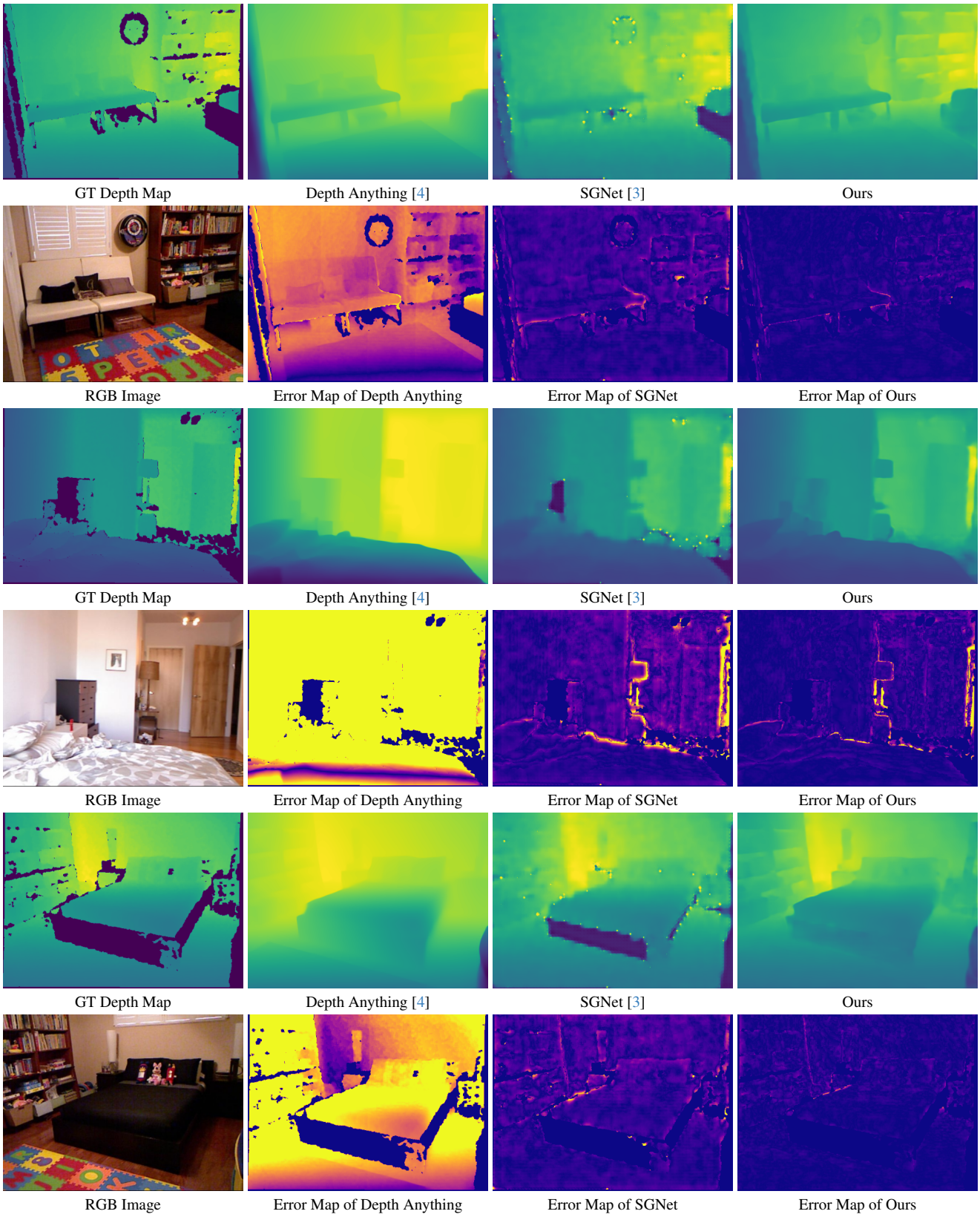

Figure 3. **Additional Visual Comparisons on Compressed-NYU Dataset between Our GDNet and Other Well-known Methods.** The first row displays the ground truth depth alongside the predicted results of various approaches, while the second row shows the corresponding RGB image and error maps. For error maps, darker areas indicate smaller errors.

CVPR
#9659

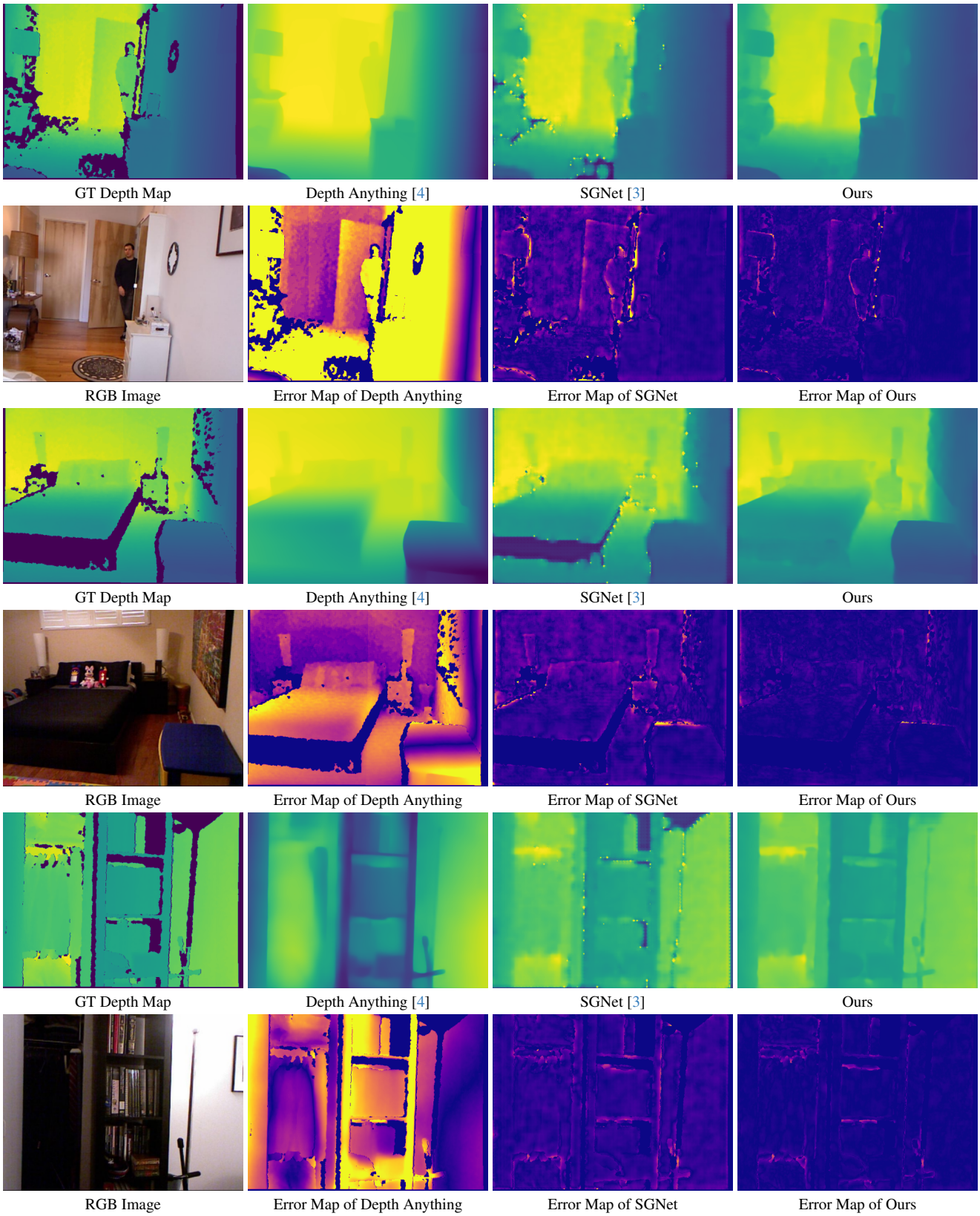

Figure 4. **Additional Visual Comparisons on Compressed-NYU Dataset between Our GDNet and Other Well-known Methods.** The first row displays the ground truth depth alongside the predicted results of various approaches, while the second row shows the corresponding RGB image and error maps. For error maps, darker areas indicate smaller errors.

CVPR
#9659

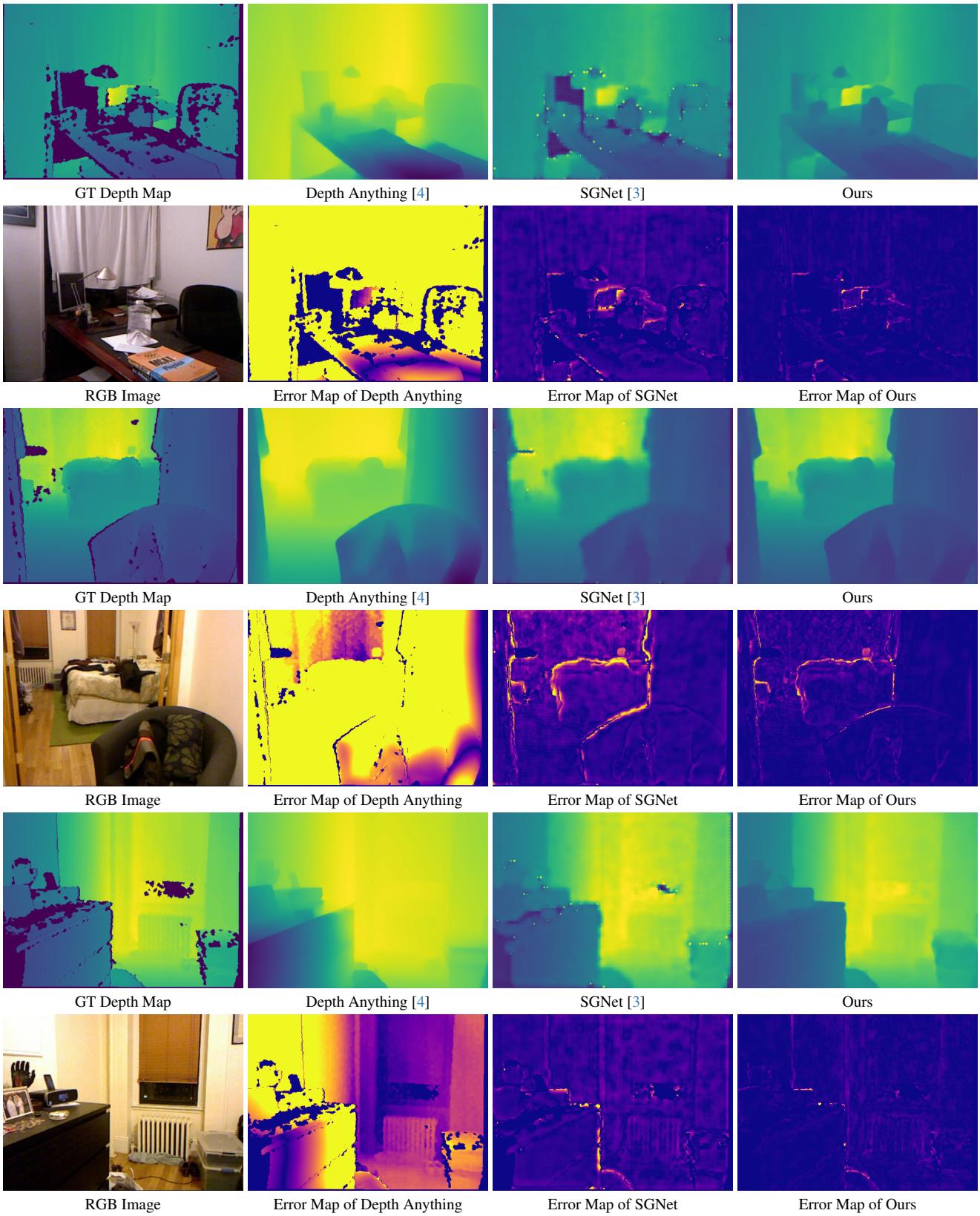

Figure 5. **Additional Visual Comparisons on Compressed-NYU Dataset between Our GDNet and Other Well-known Methods.** The first row displays the ground truth depth alongside the predicted results of various approaches, while the second row shows the corresponding RGB image and error maps. For error maps, darker areas indicate smaller errors.

CVPR
#9659

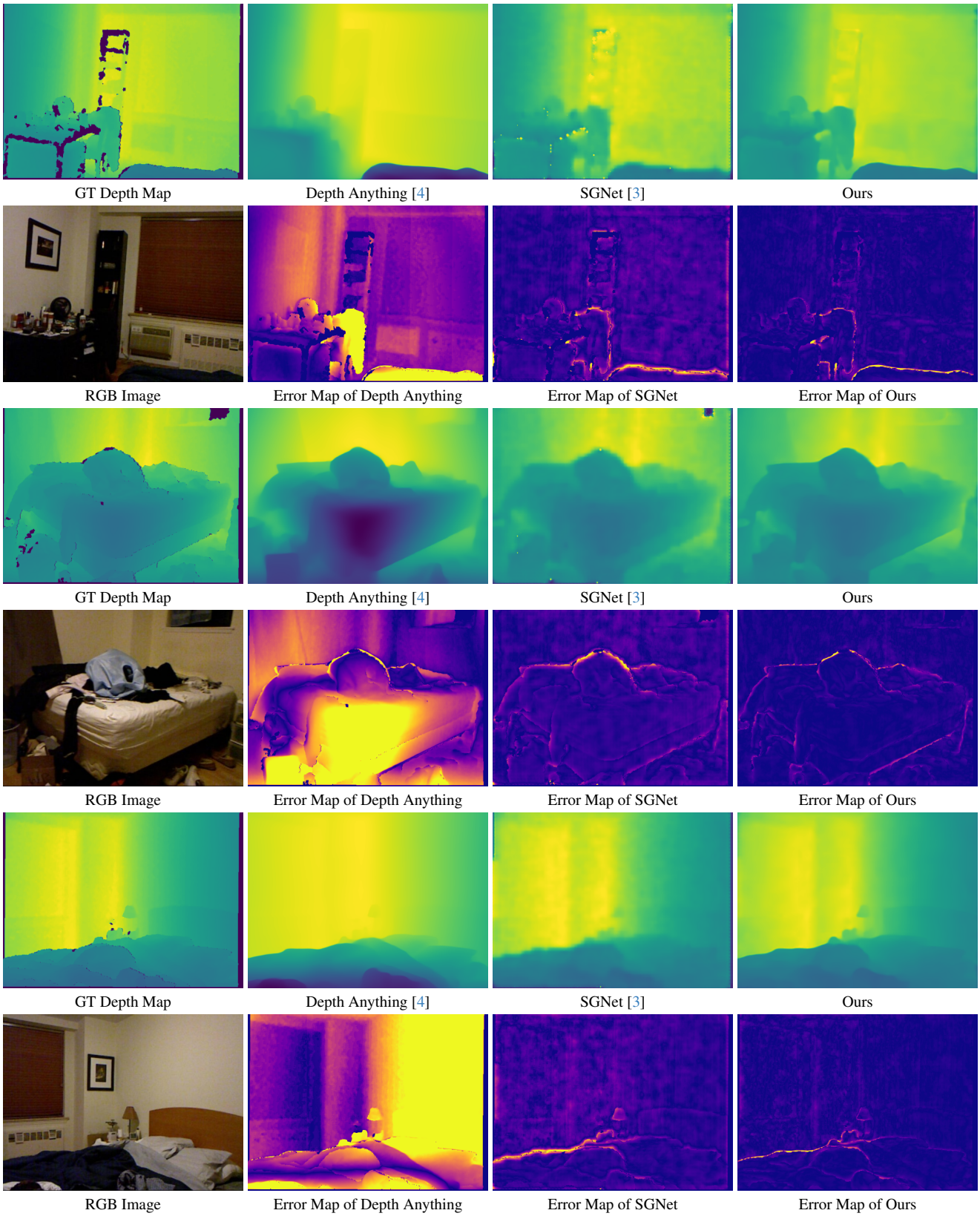

Figure 6. **Additional Visual Comparisons on Compressed-NYU Dataset between Our GDNet and Other Well-known Methods.** The first row displays the ground truth depth alongside the predicted results of various approaches, while the second row shows the corresponding RGB image and error maps. For error maps, darker areas indicate smaller errors.

CVPR
#9659

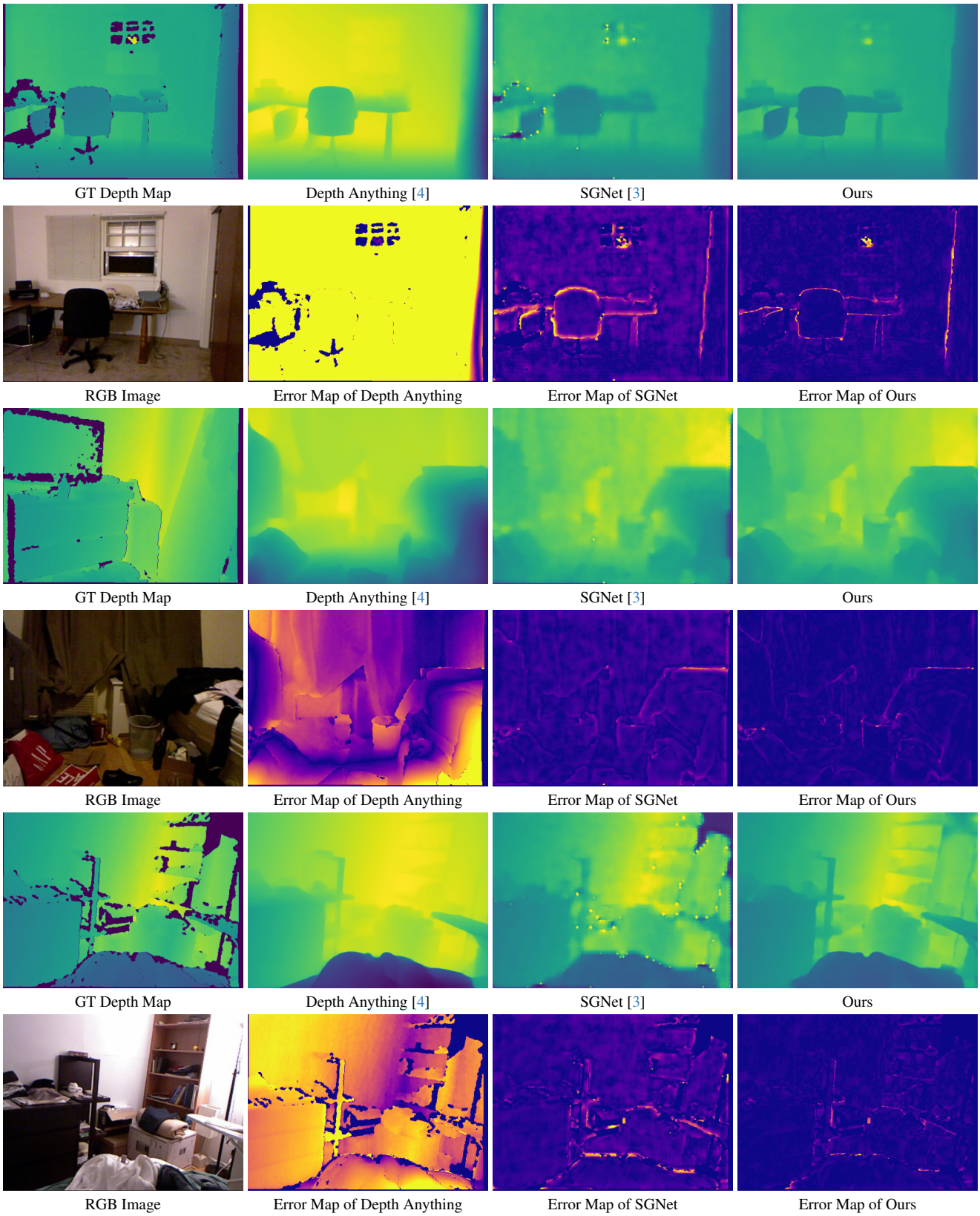

Figure 7. **Additional Visual Comparisons on Compressed-NYU Dataset between Our GDNet and Other Well-known Methods.** The first row displays the ground truth depth alongside the predicted results of various approaches, while the second row shows the corresponding RGB image and error maps. For error maps, darker areas indicate smaller errors.