Unsupervised Deep Unrolling Networks for Phase Unwrapping (Supplemental Material)

1. Algorithm Flow of Our DUN

For better understanding of our method, the algorithm flow of U3Net is provided in Algorithm 1.

Algorithm 1 Algorithm flow of U3Net.

Input: $\boldsymbol{Y}$: wrapped phase; σ: noise strength
Output: $\boldsymbol{X}^{(T)}$: unwrapped phase

$\boldsymbol{X}^{(0)} \leftarrow \mathbf{1}$, $\boldsymbol{E}^{(0)} \leftarrow \mathbf{0}$
$\{\lambda^{(t)}, w^{(t)}, d^{(t)}\}_{t=1}^{T} = \mathrm{CAM}(\sigma)$
for $t = 1$ to T **do**
  $\boldsymbol{X}^{(t)}_{(0)}, \boldsymbol{V}^{(t)}_{(0)} \leftarrow \boldsymbol{X}^{(t-1)}, \alpha_{(0)} \leftarrow 0$
  for $j = 1$ to J **do**
    $\boldsymbol{X}^{(t)}_{(j)} = \boldsymbol{V}^{(t)}_{(j-1)} + \lambda^{(t)} \mathrm{div}\big(\boldsymbol{\nabla} \boldsymbol{V}^{(t)}_{(j-1)} - (\mathcal{W}(\boldsymbol{\nabla} \boldsymbol{Y}) - \boldsymbol{E}^{(t-1)})\big)$
    $\alpha_{(j)} = \frac{1}{2}(1 + \sqrt{1 + 4\alpha^2_{(j-1)}})$
    $\boldsymbol{V}^{(t)}_{(j)} = \boldsymbol{X}^{(t)}_{(j)} + \frac{\alpha_{(j-1)} - 1}{\alpha_{(j)}}(\boldsymbol{X}^{(t)}_{(j)} - \boldsymbol{X}^{(t)}_{(j-1)})$
  end for
  $\boldsymbol{X}^{(t)} = \mathrm{NN}_{\phi}(\boldsymbol{X}^{(t)}_{(J)}, \mathcal{G}_t(\boldsymbol{X}^{(t)}_{(J)}), w^{(t)})$
  $\boldsymbol{E}^{(t)} = \mathrm{NN}_{\psi}(\mathcal{W}(\boldsymbol{\nabla} \boldsymbol{Y}) - \boldsymbol{\nabla} \boldsymbol{X}^{(t)}, d^{(t)})$
end for

2. Proof of Proposition 1

Consider $\overline{\boldsymbol{X}} = \mathcal{W}(\boldsymbol{X})$. We have

$$\boldsymbol{\nabla}_x \overline{\boldsymbol{X}}[m,n] = \boldsymbol{\nabla}_x \boldsymbol{X}[m,n] - 2\pi \boldsymbol{\nabla} \boldsymbol{K}[m,n],$$

where $\boldsymbol{K}[m,n] \in \mathbb{Z}$ for any point $[m,n]$. Thus, $\boldsymbol{\nabla}_x \boldsymbol{K}[m,n] = \boldsymbol{K}[m_1,n] - \boldsymbol{K}[m,n]$ remains an integer. Recall that $\boldsymbol{\nabla}_x \boldsymbol{X}[m,n] \in (-\pi,\pi)$ by 2D Itoh's continuity condition. We have then

$$\mathcal{W}(\boldsymbol{\nabla}_x \overline{\boldsymbol{X}})[m,n] = \boldsymbol{\nabla}_x \boldsymbol{X}[m,n].$$

The same derivation is applicable for $\boldsymbol{\nabla}_y(\overline{\boldsymbol{X}})$. Then, we have that

$$\mathcal{W}(\boldsymbol{\nabla} \mathcal{W}(\boldsymbol{X}))[m,n] = \boldsymbol{\nabla} \boldsymbol{X}[m,n], \text{ if } \|\boldsymbol{\nabla} \boldsymbol{X}[m,n]\|_\infty = \max\{|\nabla_x \boldsymbol{X}[m,n]|, |\nabla_y \boldsymbol{X}[m,n]|\} < \pi. \quad (1)$$

The proof is done.

3. Proof of Proposition 2

Once $\boldsymbol{\nabla} \boldsymbol{Y} = \boldsymbol{\nabla} \boldsymbol{X} + \boldsymbol{\nabla} \boldsymbol{N}$ is satisfied, rewrite $\mathbb{E}_{\boldsymbol{Y}} \mathcal{L}$ by

$$\begin{aligned} \mathbb{E}_{\boldsymbol{Y}} \mathcal{L} &= \mathbb{E}_{\boldsymbol{X},\boldsymbol{N},\boldsymbol{U}} \|\boldsymbol{\nabla} \mathcal{F}\big(\mathcal{W}(\boldsymbol{\nabla} \boldsymbol{Y} + \boldsymbol{\nabla} \boldsymbol{U})\big) - (\boldsymbol{\nabla} \boldsymbol{X} + \boldsymbol{\nabla} \boldsymbol{N} - \boldsymbol{\nabla} \boldsymbol{U})\|_{\mathrm{F}}^2 \\ &= \mathbb{E}_{\boldsymbol{X},\boldsymbol{N},\boldsymbol{U}} \|\boldsymbol{\nabla} \mathcal{F}\big(\mathcal{W}(\boldsymbol{\nabla} \boldsymbol{Y} + \boldsymbol{\nabla} \boldsymbol{U})\big) - \boldsymbol{\nabla} \boldsymbol{X}\|_{\mathrm{F}}^2 \\ &\quad + \langle \boldsymbol{\nabla} \mathcal{F}\big(\mathcal{W}(\boldsymbol{\nabla} \boldsymbol{Y} + \boldsymbol{\nabla} \boldsymbol{U})\big) - \boldsymbol{\nabla} \boldsymbol{X}, \boldsymbol{\nabla} \boldsymbol{U} - \boldsymbol{\nabla} \boldsymbol{N} \rangle + \|\boldsymbol{\nabla} \boldsymbol{U} - \boldsymbol{\nabla} \boldsymbol{N}\|_{\mathrm{F}}^2. \end{aligned} \quad (2)$$

Since the $\boldsymbol{U}|\boldsymbol{X}$ and $\boldsymbol{N}|\boldsymbol{X}$ are independent and follow the same distribution $\mathcal{P}$, we have that

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{X},\boldsymbol{N},\boldsymbol{U}}(\nabla\boldsymbol{N})^\top\big(\nabla\mathcal{F}\big(\mathcal{W}(\nabla\boldsymbol{Y}+\nabla\boldsymbol{U})\big)-\nabla\boldsymbol{X}\big)\\
=&\,\mathbb{E}_{\boldsymbol{X}}\mathbb{E}_{\boldsymbol{N}|\boldsymbol{X}}\mathbb{E}_{\boldsymbol{U}|\boldsymbol{X}}(\nabla\boldsymbol{N})^\top\big(\nabla\mathcal{F}\big(\mathcal{W}(\nabla\boldsymbol{X}+\nabla\boldsymbol{N}+\nabla\boldsymbol{U})\big)-\nabla\boldsymbol{X}\big)\\
=&\int_{\boldsymbol{X}}\int_{\boldsymbol{N}|\boldsymbol{X}}\int_{\boldsymbol{U}|\boldsymbol{X}}\begin{array}{l}P_{\boldsymbol{X}}(\boldsymbol{X})P_{\boldsymbol{N}}(\boldsymbol{N}|\boldsymbol{X})P_{\boldsymbol{U}}(\boldsymbol{U}|\boldsymbol{X})(\nabla\boldsymbol{N})^\top\cdot\\ \big(\nabla\mathcal{F}\big(\mathcal{W}(\nabla\boldsymbol{X}+\nabla\boldsymbol{N}+\nabla\boldsymbol{U})\big)-\nabla\boldsymbol{X}\big)\end{array}\\
=&\int_{\boldsymbol{X}}\int_{\boldsymbol{U}|\boldsymbol{X}}\int_{\boldsymbol{N}|\boldsymbol{X}}\begin{array}{l}P_{\boldsymbol{X}}(\boldsymbol{X})P_{\boldsymbol{U}}(\boldsymbol{U}|\boldsymbol{X})P_{\boldsymbol{N}}(\boldsymbol{N}|\boldsymbol{X})(\nabla\boldsymbol{U})^\top\cdot\\ \big(\nabla\mathcal{F}\big(\mathcal{W}(\nabla\boldsymbol{X}+\nabla\boldsymbol{U}+\nabla\boldsymbol{N})\big)-\nabla\boldsymbol{X}\big)\end{array}\\
=&\,\mathbb{E}_{\boldsymbol{X},\boldsymbol{N},\boldsymbol{U}}(\nabla\boldsymbol{U})^\top\big(\nabla\mathcal{F}\big(\mathcal{W}(\nabla\boldsymbol{Y}+\nabla\boldsymbol{U})\big)-\nabla\boldsymbol{X}\big).
\end{aligned}$$

Thus, the second term (inner product term) in the right-hand side of Eq. (2) is zero. The last term in Eq. (2) $\mathbb{E}_{\boldsymbol{X},\boldsymbol{N},\boldsymbol{U}}\|\nabla\boldsymbol{U}-\nabla\boldsymbol{N}\|_{\mathrm{F}}^2=(\nabla\boldsymbol{U}-\nabla\boldsymbol{N})^\top(\nabla\boldsymbol{U}-\nabla\boldsymbol{N})$ is a constant determined by the distribution of the noise and is independent of the parameter values of $\mathcal{F}$. Consequently, we have

$$\mathbb{E}_{\boldsymbol{Y}}\mathcal{L}=\mathbb{E}_{\boldsymbol{X},\boldsymbol{N},\boldsymbol{U}}\|\nabla\mathcal{F}\big(\mathcal{W}(\nabla\boldsymbol{Y}+\nabla\boldsymbol{U})\big)-\nabla\boldsymbol{X}\|_{\mathrm{F}}^2+C_0, \tag{3}$$

where C_0 is a constant. The proof is done.

4. More Implementation Details of Our DUN

We set the number of stages as $T=3$ and the number of steps for AGD as $J=10$. The first convolutional layer of the Sub-NN for estimating $\boldsymbol{X}$ outputs a feature map with 6 channels. As the spatial dimension of the features is reduced by half, the channel number is progressively increased to 12, 24, 48 and 96. The convolutional layers in the up-scaling decoder generate the symmetrical feature maps as the ones in the down-scaling encoder. The number of convolutional layers l in the Sub-NN for estimating $\boldsymbol{E}$ and its hidden channel number are set to 6 and 32, respectively. The hidden channel number of CAM is set to 128. The total loss $\mathcal{L}_{\text{total}}$ is applied at each stage, weighted by $\gamma_t=\frac{1}{T-t+1}$ for the t-th stage.

5. Analysis of Figure 5 of Main Paper

Due to space limitation in the main paper, we provide the analysis of Figure 5 of the main paper here. Recall that in the ablation study using $\mathcal{L}_{\text{sr}}\rightarrow\mathcal{L}$ ($\mathcal{L}_{\text{sr}}$ is replaced by $\mathcal{L}$), the outer wrapping operator is ablated from $\mathcal{L}_{\text{sr}}$. As a result, it fails to mitigate the negative effects caused by the outlier points that do not conform Eq. (6) in the main paper. As seen from Figure 5 of the main paper, the results of $\mathcal{L}_{\text{sr}}\rightarrow\mathcal{L}$ focus on addressing the 2π-jump outlier points, while neglecting the other areas with more subtle changes. The incorrect gradients of these outlier points result in further unsatisfying reconstruction results. As for the ablation study using w/o $\nabla\boldsymbol{U}$ (*i.e.*, the outer wrapping operation is included to preserve the correct phase structure), the measurement noise remains unaddressed due to the exclusion of the noise-resistant mechanism in training.

6. More Analysis on Limitations of Our Approach

Our experimental results in the main paper show that, when the noise is severe (*i.e*. SNR = 0), our U3Net performs worse than some supervised methods. This is probably due to that severe noise increases the number of outlier points, lowering the effectiveness of the training function. However, U3Net still outperforms other supervised methods. Indeed, severe noise presents a challenge to all compared GT-free methods, as seen in the experimental results. In comparison to those GT-free methods, U3Net demonstrates its better noise robustness as its performance advantage is more noticeable for heavier noise. Our research will focus on how to further improve the noise robustness.