Supporting Information

Deep-learning Enabled Multicolor Meta-Holography

Dina Ma, Zhancheng Li, Wenwei Liu, Guangzhou Geng, Hua Cheng, Junjie Li, Jianguo Tian, and Shuqi Chen**

Dr. D. Ma, Dr. Z. Li, Dr. W. Liu, Dr. G. Geng, Prof. H. Cheng, Prof. J. Tian, Prof. S. Chen
The Key Laboratory of Weak Light Nonlinear Photonics, Ministry of Education, School of Physics and TEDA Institute of Applied Physics, Nankai University, Tianjin 300071, China
Email: hcheng@nankai.edu.cn; schen@nankai.edu.cn;
Prof. S. Chen
The Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China
Collaborative Innovation Center of Light Manipulations and Applications, Shandong Normal University, Jinan 250358, China
Prof. J. Li
Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

This file includes:

S1. Learning curves of deep neural network (DNN)

Learning curves of each DNN are shown in **Figure** S1. As can be seen from Figure S1(a) and S1(b), the networks of nanorod datasets (real part and imaginary part sets) have MAE errors less than 0.04. And the MAE errors of dimer datasets are less than 0.45 (Figure S1(c) and S1(d)). It can be seen from the results in the Figure that after 380 epochs, the four networks can be effectively trained without obvious over fitting, which also shows the rationality of the hyperparameters setting. The hyperparameters used in the training of all the DNNs are shown in Table S1.

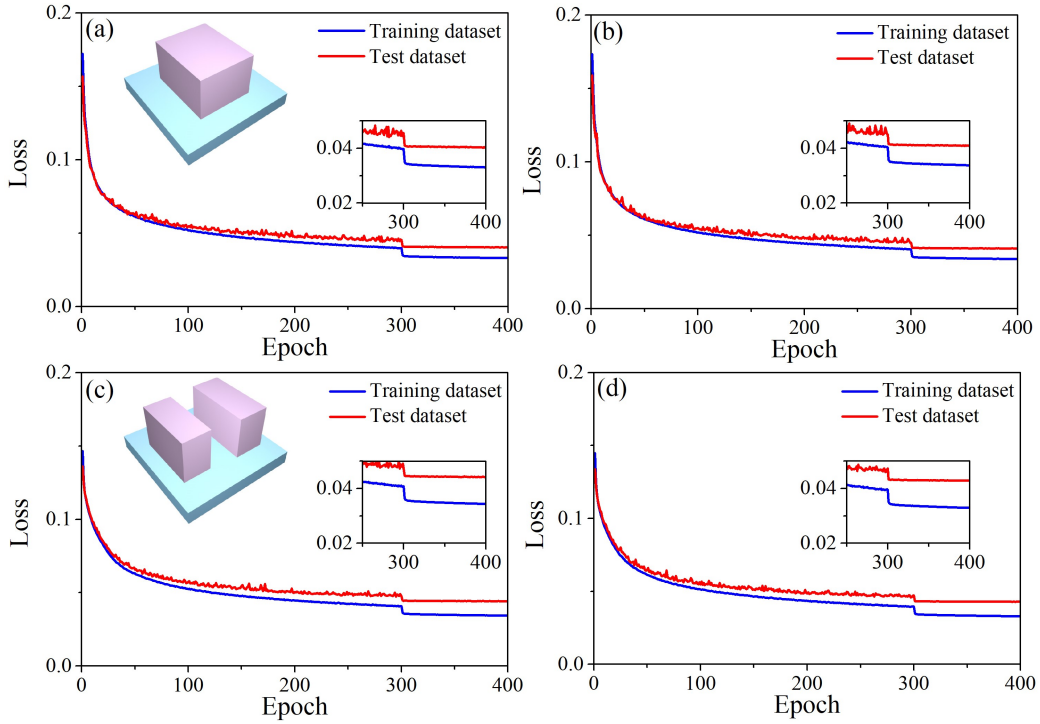

Figure S1. Learning curves of four networks. (a) The learning curves of nanorod real part feature DNN. (b) The learning curves of nanorod imaginary part feature DNN. (c) The learning curves of dimer real part feature DNN. (d) The learning curves of dimer imaginary part feature DNN.

Hyperparameters	DNN (nanorod real part features)	DNN (nanorod imaginary part features)	DNN (dimer real part features)	DNN (dimer imaginary part features)
Learning rate	From 10^{-4} to 10^{-5}	From 10^{-4} to 10^{-5}	From 10^{-4} to 10^{-5}	From 10^{-4} to 10^{-5}
Optimizer	Adam	Adam	Adam	Adam
Batch size	16	16	16	16
Nonlinear activations	ReLU	ReLU	ReLU	ReLU
Epochs	400	400	400	400
Error (train)	0.033	0.034	0.034	0.033
Error (test)	0.040	0.041	0.044	0.043
Time taken	30 min	30 min	30 min	30 min

Table S1. Hyperparameters used in the training of DNNs

S2. Principal component analysis (PCA) method

We use PCA method to reduce the dimension of the array we need to fit. Take the dimension reduction process of $\boldsymbol{S}_{real}$ as an example, the specific steps can be described as follows:

(1) Calculate the covariance matrix Σ_{real}:

$$\mu_{real} = \frac{1}{N}\sum_{n=1}^{N} S_{real}^{(n)}, \tag{S1}$$

$$\boldsymbol{\Sigma}_{real} = \frac{1}{N}\sum_{n=1}^{N}(S_{real}^{(n)} - \mu_{real})(S_{real}^{(n)} - \mu_{real})^{T}. \tag{S2}$$

(2) Calculate the eigenvalues and eigenvectors of the covariance matrix Σ_{real}.

(3) The eigenvectors are arranged into a matrix according to the corresponding eigenvalues, and the first m rows are taken to form a matrix $\boldsymbol{W}$.

(4) Calculate the PCA feature array $\boldsymbol{f}_{\text{real}}$.

$$\boldsymbol{f}_{real} = \boldsymbol{W}\boldsymbol{S}_{real}. \tag{S3}$$

We performed the above dimensionality reduction method on all the data collected by simulation. In order to show more real part and imaginary part

information with fewer PCA feature numbers, we evaluated the reconstruction results of different number of features. As shown in **Figure** S2(a), when the number of features is greater than 50, the mean absolute error (MAE) between the reconstruction results (real and imaginary part values of two structures) and the real results is less than 0.012. The comparison results from Figure S2(c) to S2(f) also show that the error between the reconstructed results and the real results mainly occurs in the peak with very narrow bandwidth, and the lack of these peak information does not affect our overall training, so our PCA feature number is set to 50.

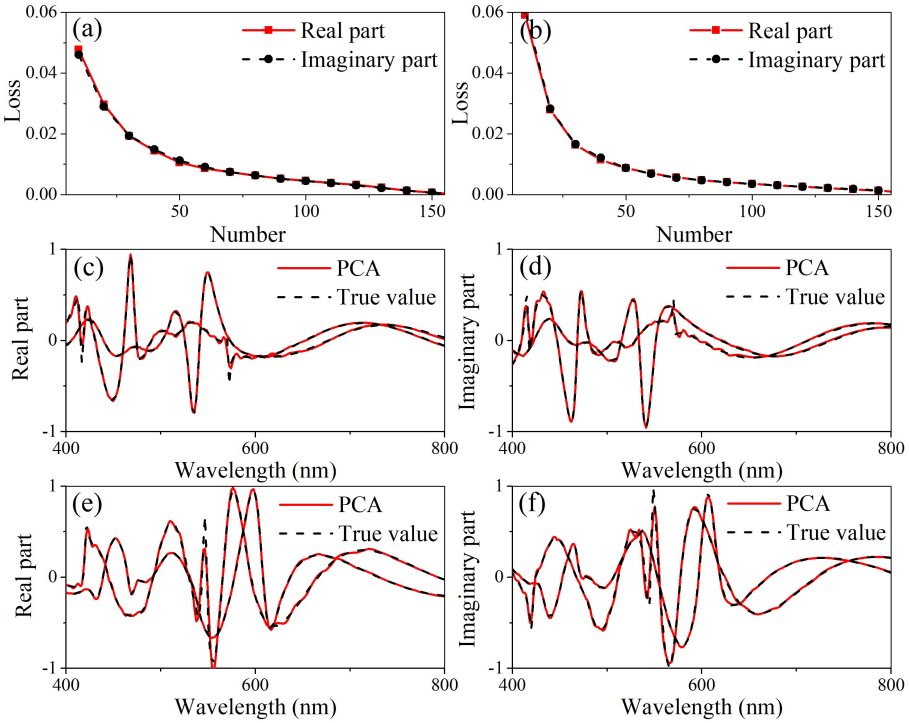

Figure S2. (a) and (b) show the MAE error curves of reconstructed values and actual simulation results with PCA feature number. The red and black lines in (a) represent the loss curves corresponding to the real part data and imaginary part data of all nanorod training sets. The red and black lines in (b) represent the loss curves corresponding to the real part data and imaginary part data of all dimer training sets.

(c) and (d) show the comparison between the PCA reconstruction results (red solid lines) and true simulation (black dotted line) results of two groups of nanorod training set. (e) and (f) show the comparison between the PCA reconstruction results (red solid lines) and true simulation (black dotted line) results of two groups of dimer training set.

S3. Home-built experimental setup

The experimental results of the reflection spectra and captured color images are measured by a home-built setup. As shown in **Figure** S3, a polarizer (P1) was used to control the polarization state of incident light. The lens set (L1 and L2) allowed the light to normally incident on the sample. The 5× (OBJ1) and 10× objectives (OBJ2) were employed to collect the incident light onto the sample and magnify reflected light, respectively. The lens L3 can make the enlarged images clearly captured by the CCD camera. Two unpolarized beam splitters (BS1 and BS2) can effectively change the direction of the light path and realize the two functions of spectra measurement and image capture simultaneously.

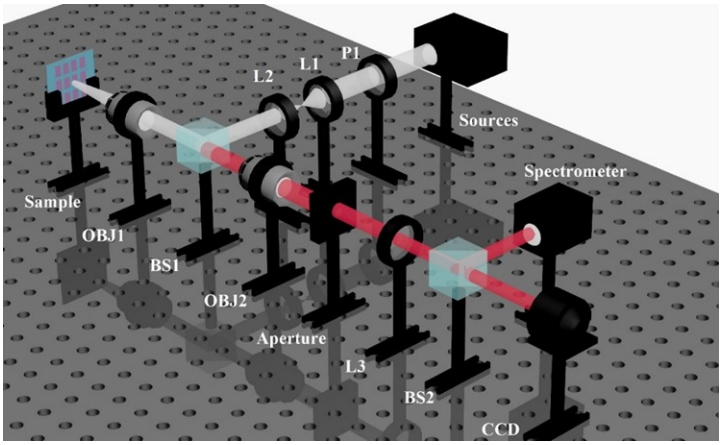

Figure S3. Schematic illustrating the home-built setup to measure spectra and capture color images.

S4. The fitness function of evolution strategy (ES) algorithm

For different inverse design problems, we need to choose different fitness functions. In the work, we choose two fitness functions (Equation 1 and Equation 2) to solve the design of reflection spectra and the design of amplitude and phase at specific wavelength. In order to explain the influence of fitness function in detail, we test the inverse design results obtained by different fitness functions (**Figure** S4). It can be seen from the figure that our method can obtain reflection spectra with different bandwidths by fitting Gaussian curves with different bandwidths. When an efficiency term is added to the loss function, the fitness function can not only meet the spectral bandwidth, but also ensure the reflectivity intensity at a specific wavelength.

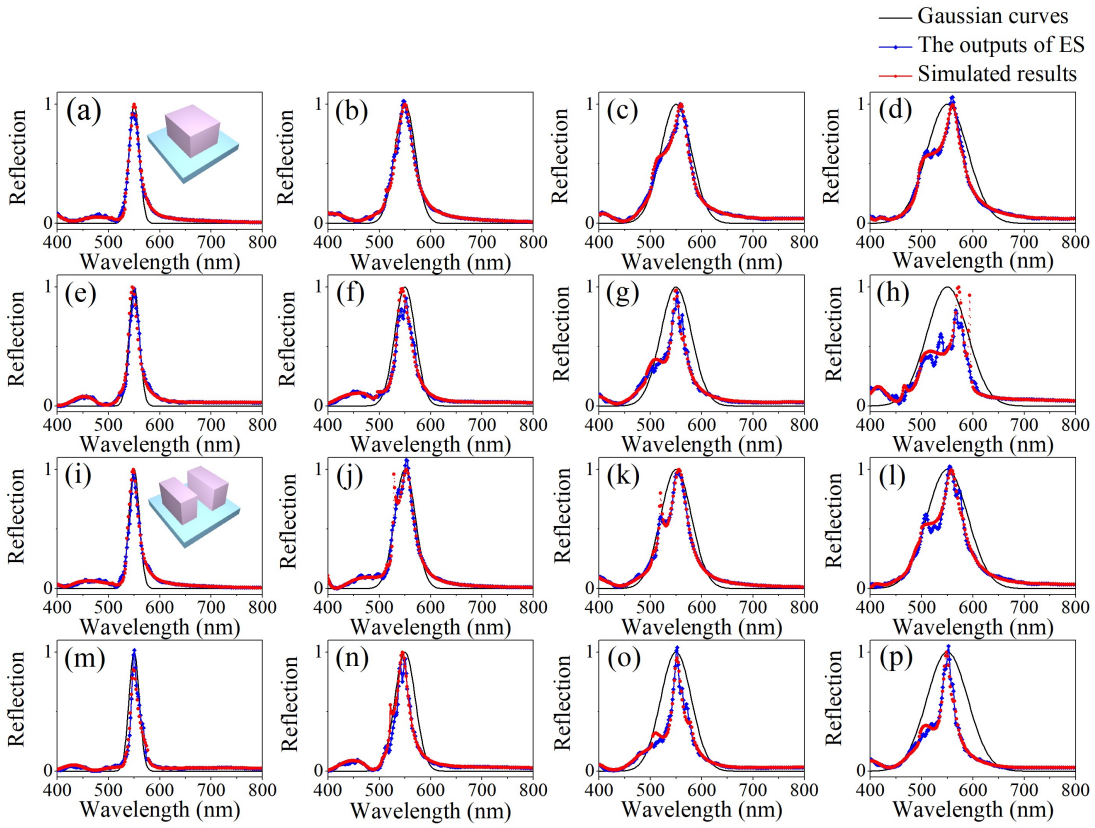

Figure S4. The inverse design results obtained by different fitness functions. (a) to (d) are the inverse design results of the nanorods obtained by fitness function

$F(\boldsymbol{S},\boldsymbol{g})=-\frac{1}{201}\sum_i(\boldsymbol{S}_i-g_i)^2$. The black lines are the Gaussian curves $\boldsymbol{g}$, the blue lines correspond to the predicted spectra $\boldsymbol{S}$ of the structure obtained by ES algorithm, and the red dash lines correspond to the simulation results of the structures obtained by inverse design. (e) to (h) are the inverse design results of the nanorods obtained by fitness function $F(\boldsymbol{S},\boldsymbol{g})=-\frac{1}{201}\sum_i|\boldsymbol{S}_i-g_i|-\frac{1}{3}\left|reflection_{pre}-reflection_{target}\right|$. (i) to (l) are the inverse design results of the dimers obtained by fitness function $F(\boldsymbol{S},\boldsymbol{g})=-\frac{1}{201}\sum_i(\boldsymbol{S}_i-g_i)^2$. (m) to (p) are the inverse design results by using the fitness function $F(\boldsymbol{S},\boldsymbol{g})=-\frac{1}{201}\sum_i|\boldsymbol{S}_i-g_i|-\frac{1}{3}\left|reflection_{pre}-reflection_{target}\right|$. In this figure, the μ values of the Gaussian curves of each column are 550 nm, and the σ values of each column are 10 nm, 20 nm, 30 nm, 40 nm, respectively.

S5. Structural parameters of the structures in Figure 4 and 5

The structural parameters of the structures in Figure 4 and 5 are shown in Table S2 and S3, respectively.

Unit: nm

Case	Parameters
b	[278, 280, 300, 189, 141]
c	[282, 271, 300, 212, 190]
d	[331, 362, 300, 215, 167]
e	[397, 359, 300, 48, 46, 35, 50, 73, 285]
f	[398, 397, 300, 64, 61, 64, 61, 79, 267]
g	[382, 400, 300, 54, 50, 51, 61, 329, 263]

Table S2. Design Parameters of Structures in Figure 4

Unit: nm.

	B (470)	G (550)	R (610)
p_1	[375, 300, 300, 43, 36, 57, 37, 190, 163]	[375, 375, 300, 50, 59, 59, 42, 70, 249]	[400, 400, 300, 37, 65, 56, 57, 329, 176]
p_2	[400, 250, 300, 230, 175]	[375, 375, 300, 59, 58, 56, 59, 80, 216]	[400, 400, 300, 59, 36, 56, 36, 239, 291]
p_3	[250, 300, 300, 170, 146]	[375, 375, 300, 56, 36, 40, 37, 232, 109]	[300, 375, 300, 115, 305]
p_4	[250, 300, 300, 73, 206]	[375, 375, 300, 35, 60, 60, 35, 215, 76]	[300, 375, 300, 214, 219]

Table S3. Design Parameters of Structures in Figure 5

S6. Target, predicted and simulated reflection phases for structures in Figure 5

The target, predicted and simulated reflection phases for structures in Figure 5 are shown in Table S4.

	target phase			predicted phase			simulated phase		
	B	G	R	B	G	R	B	G	R
p_1	0°	-155°	-20°	-1.1°	-166.0°	-21.1°	5.2°	-150.1°	-25.2°
p_2	-100°	-115°	90°	-108.6°	-115.6°	95.9°	-104.9°	-104.3°	90.0°
p_3	-40°	-50°	140°	-40.8°	-52.7°	129.1°	-41.5°	-37.2°	143.8°
p_4	60°	15°	50°	56.7°	24.6°	56.2°	65.3°	25.8°	58.4°

Table S4. Target, Predicted and Simulated Reflection Phases for Structures in Figure 5

S7. A more detailed phase mask for multicolor hologram

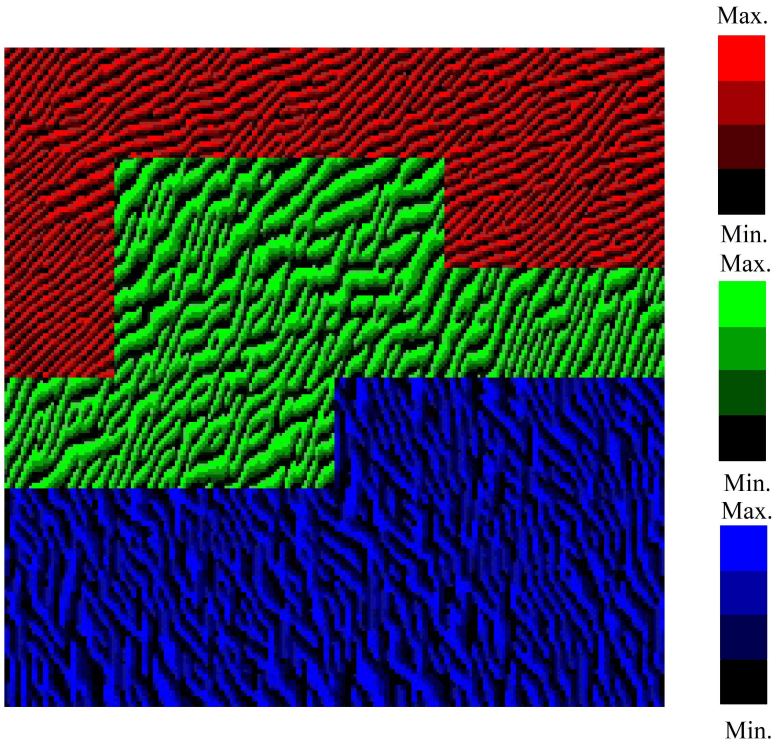

Figure S5. The phase distribution of the designed metasurface for multicolor hologram. The maximum and minimum values of the color bars represent the maximum and minimum phase values shown in Figure 4 at wavelength channels of 470 nm (blue color), 550 nm (green color) and 610 nm (red color).

S8. A detail discussion on the maximum phase and amplitude coverage of the nanostructures designed by the proposed inverse design method

To a certain extent, since the data in the data set are all randomly generated within a given value range, the phase and amplitude coverage of the data set can reflect the maximum phase and amplitude coverage of the inversely designed nanostructures. Therefore, in order to show the maximum phase and amplitude coverage of the nanostructures designed by the proposed inverse design method, we analyzed all data in the data set. The reflection intensities at different wavelength channels (470, 550 and 610 nm) are firstly considered. Here, the reflection intensity at

the operation wavelength of 470 nm is constrained to be larger than 0.4 while the reflection intensity at the other two wavelength is constrained to be smaller than 0.3. In this case, most dimers in the data set can fulfill the requirement, as verified by Figure S5(a). We further limit the height of the nanostructures in a value range of 280 to 320 nm. Figure S5(b) shows the phase coverage of the dimers in the data set that fulfill the two above-mentioned restricted conditions. Results indicate that the phase range covered by the data gradually decreases, which covers about 56% of the phase space at this stage. Next, we limit the period in the *x* direction in a value range of 350 to 390 nm. The results in Figure S5(c) indicates that only 38% of the phase space can be covered by the data that fulfill all the three restricted conditions. We further analyze the phase coverage of the nanostructures that fulfill the first and the second restricted conditions at 550 nm and 610 nm, as shown in Figures S5(d) to S5(g). Results indicate that the phase delay of the nanostructures in the data set can cover about 50% phase space at these two wavelengths, which is in good consist with the results we proposed in the main text. We also analyze the amplitude coverage of the dataset. We select seven phase values between -3 and 0 at an interval of 0.5. Then, we choose the data in Figure S5(h) that have the same phase values as we selected, and present the corresponding amplitude values in Figure S5(i). Results indicate that the maximum reflection intensity is 1 and the minimum one is 0.7. It should be emphasized that the above analysis is based on very broad restrictions. For the inverse design of nanostructures for multicolor hologram, the restricted conditions are more stringent. Therefore, the amplitude and phase area that can be covered by the designed

nanostructures are further limited. The results in Figure S5 also validate that over 70% phase space can be covered when the nanostructures can be designed with different height.

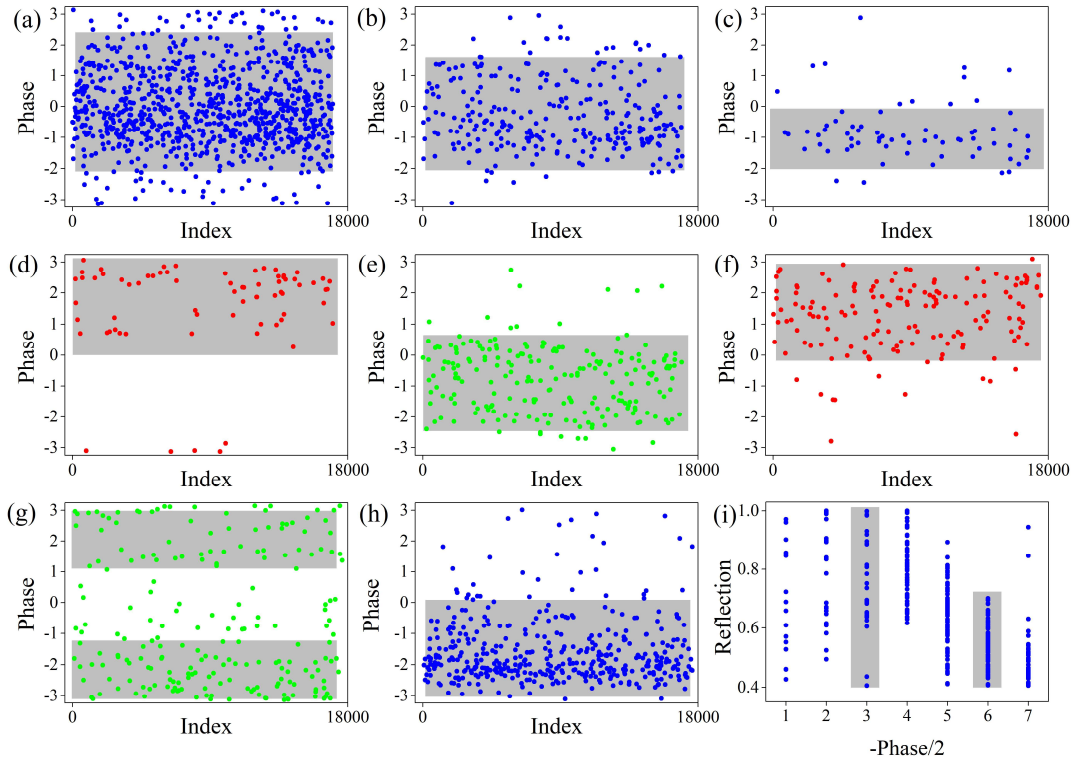

Figure S6. Phase delays and reflection intensities of the nanostructures satisfying different restricted conditions in the data set. (a) Phase delays of dimer structures satisfying the first restricted condition and the wavelength channel is 470 nm. (b) Phase delays of dimer structures satisfying the first and the second restricted condition, the wavelength channel is 470 nm. (c) Phase delays of dimer structures satisfying the first, the second and the third restricted condition, the wavelength channel is 470 nm. The phase delays of dimer structures satisfying the first and the second restricted condition, and the wavelength channels are (d) 550 nm and (e) 610 nm. The phase delays of nanorods satisfying the first and the second restricted condition, the wavelength channels are (f) 470 nm, (g) 550 nm and (h) 610 nm. (i) Reflection intensity of the nanostructures with different phase delays selected from (h).