Supplementary Material: Mix-and-Match Holography

YIFAN PENG, The University of British Columbia, King Abdullah University Of Science And Technology
XIONG DUN, King Abdullah University Of Science And Technology
QILIN SUN, King Abdullah University Of Science And Technology
WOLFGANG HEIDRICH, King Abdullah University Of Science And Technology, The University of British Columbia

ACM Reference format:
Yifan Peng, Xiong Dun, Qilin Sun, and Wolfgang Heidrich. 2017. Supplementary Material: Mix-and-Match Holography. *ACM Trans. Graph.* 36, 6, Article 191 (November 2017), 13 pages.
DOI: 10.1145/3130800.3130839

1 MATRIX FACTORIZATION DETAILS

In the following we describe in detail the individual steps of the multiplexing method. We focus on two DOEs, with a rank 1, but extensions to a larger number of DOEs are possible with tensor facotrization, and dynamic DOEs with spatial light modulators can be modeled as higher rank matrices and tensors [Heide et al. 2014].

Newton's update for solving sub-problems. As has been briefly discussed in the main text, for rank $r = 1$, our complex matrix factorization problem simplifies to vector form, illustrated as:

$$\mathbf{b}_{\text{opt}}, \mathbf{a}_{\text{opt}} = \underset{\mathbf{a}\in\mathbb{C}^{m\times1}_{|\cdot|=1}, \mathbf{b}\in\mathbb{C}^{n\times1}_{|\cdot|=1}}{\operatorname{argmin}} \ \frac{1}{2}\left\|\mathbf{T} - \mathbf{W}\circ\mathbf{a}\mathbf{b}^{\dagger}\right\|^2_{\mathbf{W}}. \tag{1}$$

Here we derive the b-step in Alg.1 in main text, and the derivations for a-step follow from symmetry. Specifically, we first reformulate the optimization problem to derive f as:

$$\begin{aligned}\mathbf{b}_{\text{opt}} &= \underset{\mathbf{b}}{\operatorname{argmin}} \ \frac{1}{2}\left\|\mathbf{T} - \mathbf{a}\mathbf{b}^{\dagger}\right\|^2_{\mathbf{W}} \\ &= \underset{\mathbf{b}}{\operatorname{argmin}} \ \frac{1}{2}\left\|\operatorname{diag}(\mathbf{W})\,\mathbf{t} - \operatorname{diag}(\mathbf{W})\,\mathbf{O}_{\mathbf{a}}\mathbf{b}\right\|^2_2 \\ &= \underset{\mathbf{b}}{\operatorname{argmin}} \ \frac{1}{2} f(\mathbf{b}),\end{aligned} \tag{2}$$

and

$$\begin{aligned} f(\mathbf{b}) =& \mathbf{t}^{\dagger}\operatorname{diag}(\mathbf{W})^2\,\mathbf{t} - 2\mathbf{t}^{\dagger}\operatorname{diag}(\mathbf{W})^2\,\mathbf{O}_{\mathbf{a}}\mathbf{b} \\ &+ \mathbf{b}^{\dagger}\mathbf{O}_{\mathbf{a}}^{\dagger}\operatorname{diag}(\mathbf{W})\,\mathbf{O}_{\mathbf{a}}\mathbf{b},\end{aligned} \tag{3}$$

where $\operatorname{diag}(\cdot)$ puts the matrix from the subscript on the diagonal and $\mathbf{O}_{(\cdot)}$ corresponds to the outer vector product operation with the vector in the subscript and the right hand side, followed by vectorization. Then we can easily derive the gradient as:

$$\nabla_{\mathbf{b}} f = \mathbf{O}_{\mathbf{a}}^{\dagger}\operatorname{diag}(\mathbf{W})\,\mathbf{O}_{\mathbf{a}}\mathbf{b} - \mathbf{O}_{\mathbf{a}}^{\dagger}\operatorname{diag}(\mathbf{W})^2\,\mathbf{t}. \tag{4}$$

The operator $\mathbf{O}_{\mathbf{a}}^{\dagger}$ is the same as the outer vector product operation plus a subsequent summation over the rows of the resulting matrix. We finally get a diagonal matrix Hessian as:

$$\frac{\partial^2 f}{\partial b^2} = \mathbf{O}_{\mathbf{a}}^{\dagger}\operatorname{diag}(\mathbf{W})\,\mathbf{O}_{\mathbf{a}}. \tag{5}$$

Notice that since Hessian is diagonal, the matrix inversion in Newton's method becomes a point-wise division:

Algorithm 1 Newton Update to solve the **b-step**

1: $k = 0, \mathbf{b}^0_{opt} = \mathbf{a}_{init}, \mathbf{b}^0_{opt} = \mathbf{b}_{init}$
2: **repeat**
3: $\quad \mathbf{b}^{k+1}_{\text{opt}} := \mathbf{b}^{k}_{\text{opt}} - \frac{\nabla_{\mathbf{b}} f}{\frac{\partial^2 f}{\partial b^2}}$ ▷ Pointwise division
4: $\quad \mathbf{b}^{k+1}_{\text{opt}} := proj_{(|\mathbf{b}^{k+1}_{\text{opt}}|>1)}(\mathbf{b}^{k+1}_{\text{opt}})$
5: ▷ Conditional projection
6: $\quad k := k + 1$
7: **until** Optimality achieved

Conditional projection. In order to extract phase-only designs, a naive projection update process is usually implemented as:

$$\mathbf{b}_i^{k+1} := \frac{\mathbf{b}_i^{k+1}}{|\mathbf{b}_i^{k+1}|}. \tag{6}$$

In theory, it is reasonable to force the extracted phase component of every pixel to be unified for the amplitude all the time. While in practice, it is not necessary to normalize it at every step since we only need the final optimum to fit the physical model. Thus, during the optimization, our conditional projection (step 4 in Alg. 1) maintains as much of the data as possible while eliminating ill-conditioned cases, leading to not just faster but also better convergence. Once we approach convergence, the complex phase expression can then be forced to have unit amplitude.

2 FABRICATION DETAILS

As has been mentioned in the main text, the DOEs are fabricated by repeatedly applying photolithography and reactive ion etch (RIE) technology. Each round of fabrication results in a binary profile on the fused silica substrate. By repeating the process for n times, a 2^n-level phase surface can be fabricated finally.

DOI: 10.1145/3130800.3130839

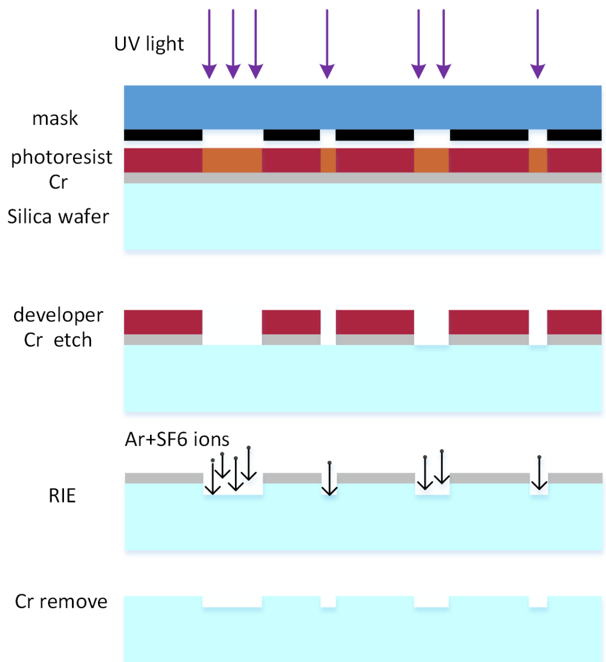

Fig. 1. Basic steps for fabricating DOEs. First, patterns are transferred from mask to photoresist on the fused silica wafer through exposure under UV illumination and following developer process. Then, the transferred patterns are converted to a binary profile on the wafer by Cr etching and reactive ion beam (Ar + SF6) bombardment. The final binary profile is obtained by removing the Cr layer.

Fig. 2. Comparison of synthetic results with different quantizations. Note that all images are resampled with 8μm on the DOE plane with the final intensity distributions integrated over 200μm on the receiver screen.

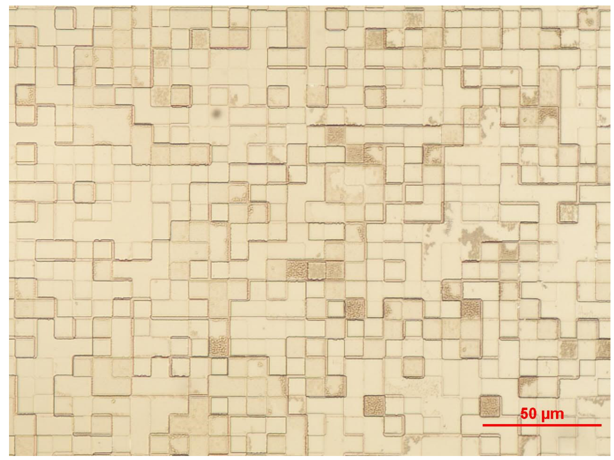

Fig. 3. Illustration of 20X microscope image of a fabricated DOE. Notice that the bright and dark changes indicate the physical height distribution.

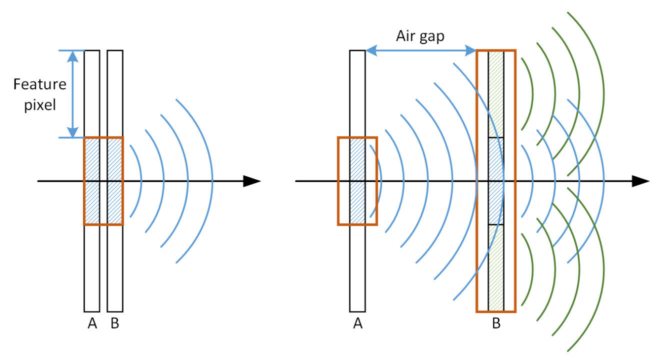

Fig. 4. Illustration of possible air gap when multiplexing DOEs.

In our work, the basic fabrication procedure is shown in Fig. 1. Four iterations are choose to produce 16-level phase surface, which is capable of providing sufficiently high diffraction efficiency, as shown in Fig. 2. The principle wavelength is λ = 550nm, and the corresponding total etching depth is 1195nm. A final depth error of ±10nm has been achieved. The microscope images of one of the fabricated DOE plates are shown in Fig. 3.

3 ADDITIONAL RESULTS

Analysis on air gap issue. To derive the encoding model, we make an assumption that the two DOE layers are in direct physical contact to enforce the factorization towards a well-conditioned optimization problem. In the following we analyze the air gap issue that is somehow inevitable in practice, especially where relative motion multiplexing is enclosed.

As shown in Fig. 4, the left one indicates a multiplexing configuration that the size of air gap is much smaller than a feature pixel so as to be reasonably neglected (e.g. static pairing scenarios). Further, the right one has an air gap whose size could be several times larger than that of a feature pixel. Based on Huygens–Fresnel diffraction principle, the wave emitted from a pixel on layer **A** is spread out a bit when arriving layer **B**. Consequently, the synthesized phase modulation for this pixel is no longer one-to-one mapping, but to involve additional phase delays for neighbouring pixels (illustrated by different color wavefronts in Fig. 4). The two layer patterns encoded under the closely stacked assumption may fail. Fortunately, we can still derive a constrained range where the principle diffractive “rays” still travel through the target pixel. Then the encoded image can still be decoded while suffering from diffraction efficiency loss of some sense.

We reasonably consider a square aperture’s diffraction (i.e. each cell in the DOEs is a square). The intensity of the diffraction of a square of layer **A** on layer **B** is mainly distributed in two orthogonal directions [Goodman 2005], and can be expressed as:

$$I(x,y) = I_0(x,y) \cdot sinc^2(\frac{k\omega x}{2z}) \cdot sinc^2(\frac{k\omega y}{2z}), \tag{7}$$

where (x, y) is the coordinate of diffraction pattern, $k = \frac{2\pi}{\lambda}$ is the wave number, λ is the principal wavelength of incident illumination, ω represents the feature size (diameter) of this pixel, and z is the size of air gap.

Now we want to limit the wavefront spread of layer **A** within one pixel's range on layer **B** by applying the condition that the half width of central order diffraction is with one pixel's size, such that we derive the constraint as:

$$z \leq \frac{\omega^2}{0.886\lambda}. \tag{8}$$

Accordingly, we note that it won't result in noticeable image degradation when the air gap size falls below $131\mu m$ in our implementations (where the feature size $8\mu m$, wavelength $550nm$). This value is feasible in engineering even for relative motion setups.

At current stage of proof-of-concept, we reasonably rely on the closely stacked assumption. Experimentally, we find that although involving a small air gap around the thickness of two regular A4 papers (i.e. $200\mu m$), the target image can still be decoded with a visually pleasing quality except for a slight loss of diffraction efficiency. Fig. 5 shows the synthetic images reconstructed from the multiple discrete pairings (second application) with different thickness settings. We found that when the air gap is set $131\mu m$ as calculated above, the image exhibits almost the same quality as the one without any air gap. When increasing the air gap up to 262 μm, even to $500\mu m$, the images still exhibit visually pleasing quality, except for the occurrence of speckle noise and the diffraction efficiency loss. Consequently, we can release the constraint on Eq. 8 up to 2 physical pixels, then the tolerance of air gap size for a $8\mu m$ feature size's design is $262\mu m$. Nevertheless, this is an engineering issue that can be fixed with a modest amount of work.

However, to enforce a more robust image reconstruction, the extra phase modulation derived from noticeable air gap shall be considered. Consequently, the model $\mathbf{T} = \mathbf{AB}^\dagger$ presented in Section 4 of the main text now becomes $\mathbf{T} = (\mathbf{A} * \mathbf{F}_{(x,y)})\mathbf{B}^\dagger$. Notice that the left term involves a spatial variant convolution corresponding to delays caused by the occurrence of air gap. Thus, the proposed factorization updates shall involve additional convolution and deconvolution steps, which may yield an ill-conditioned optimization. Notice that in phase modulation domain, the inverse formation model becomes more complex. This extension for arbitrary geometric alignment setting remains an open question and deserves further study.

Simulated results approximating pixelated misalignment. We quantitatively evaluate the perceptual quality of a reconstructed image using the offset pairing design as detailed in Section 5 of the main text. In the following we present the corresponding full resolution images (Fig. 7).

Full resolution experimental results of offset pairing scenarios. We have fabricated two pairs of offset layering DOEs that are encoded with two different gray-scale images, one "Einstein" and the other "Lena". Partial results have been presented in Fig.6 of main text. Here we present the corresponding full resolution synthetic and experimental results of both target images, shown in Fig. 8.

Artifacts analysis of multiview projection. As mentioned in Section 5 of the main text, readers notice the residual image of the mask which can be accounted from diffraction efficiency loss. As shown in Fig. 6, we see that the fabrication error (i.e. incorrect etching depth) results in a noticeable diffraction efficiency loss, which further leads to a fact that 0-order image of the mask can be observed. Despite this diffraction efficiency loss, the barrier-based pairing scheme also suffers from the crosstalk effect when changing from one view to another. This is very similar to the crosstalk in conventional multiview stereoscopic displays. However, the holographic image formation in our case generates a relatively blurry but smooth transition of perspective images. Refer to the accompanying video to have an intuitive evaluation.

Evaluation on individual layer decoding. As mentioned in Section 5 of the main text, the multiple discrete pairing has promising applications in privacy or security. Here we present the synthetic reconstruction results when illuminating each individual layer of the fabricated masks (Fig.8 of the main text) with the desirable setting. In Fig. 9, we see that the factorized holograms exhibit random height field distribution. As expected, illuminating each individual layer creates only random noise energy distribution at the target distance. That said, when used as the individual keys for security purpose, our design exhibits good randomness. We have in the accompany video shown the experimental result of illuminating a single layer fabricated hologram.

Evaluation on two factorization methods. In addition to the quantitative evaluation presented in Section 4.3 of the main text, the corresponding simulation results of 6 testing images subject to two different factorization update rules are presented in Fig. 10. For each pair, the top one is obtained via well-known alternating multiplicative update [Ho 2008], while the bottom one is obtained via fast Newton's update [Heide et al. 2016]. Both of these two kinds of update are followed by a conditional projection as discussed in Section 4.3 of main text and in Alg. 1. The value provided in the bottom-left of each sub-figure indicates the normalized energy preserved within the visualized patch. We observe that both methods lead to visually pleasing visualization results. Particularity, Newton's update generally outperforms the other. We recommend the readers to zoom-in for a clear view on background speckle variance of each pair.

We also test our offset pairing scheme on factorizing a continuous target phase map (i.e. a Fresnel lens phase profile) rather than random pattern holograms (Fig. 11). Empirically, the relatively continuous distribution benefits the factorization. Our method outperforms the alternatives also in this case, indicating that our framework can be generalized to a variety of phase modulation designs.

Full frame results of translational pairing application. are shown in Fig. 12. As well, we qualitatively compare the results of multiplexing different target data amount onto the same DOEs in Fig. 13. We observe that adding more frames results in a fact that some of the frames may suffer from noised background, especially for those frames that are involved with more physical overlap with others. This is reasonable since the available bandwidth that can be used for

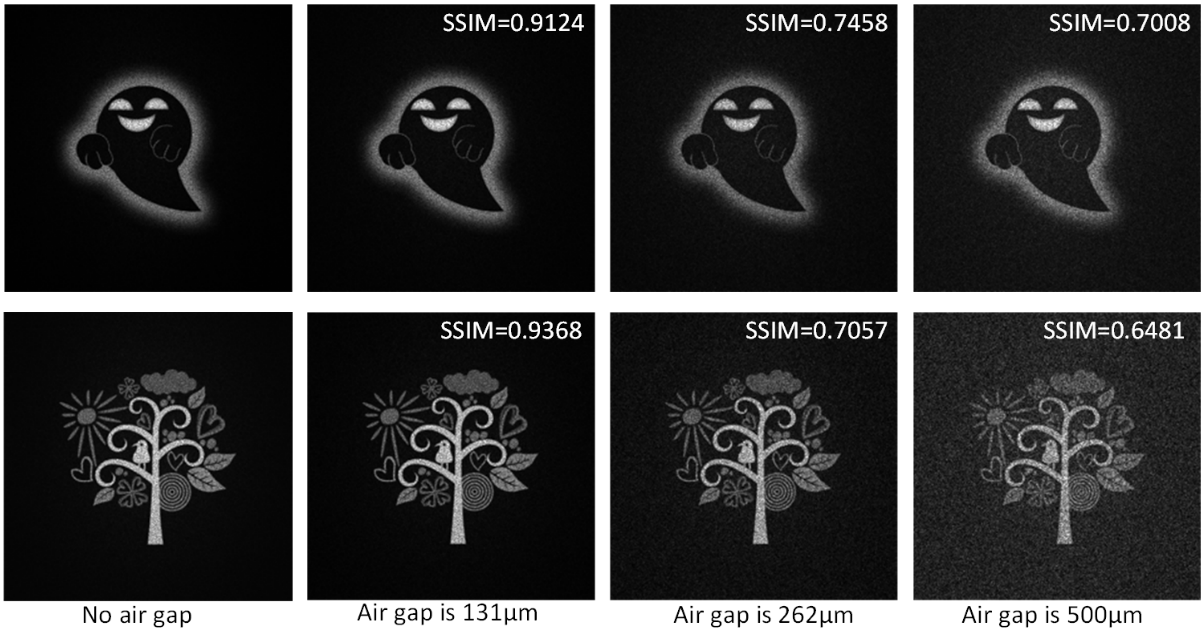

Fig. 5. Comparisons of synthetic images reconstructed from the 3 by 3 pairing examples subject to different air gap settings: 0 (left column), 131μm (center left column), 262μm (center right column), and 500μm (right column), respectively. The SSIMs are evaluated with respect to the image reconstructed without air gap.

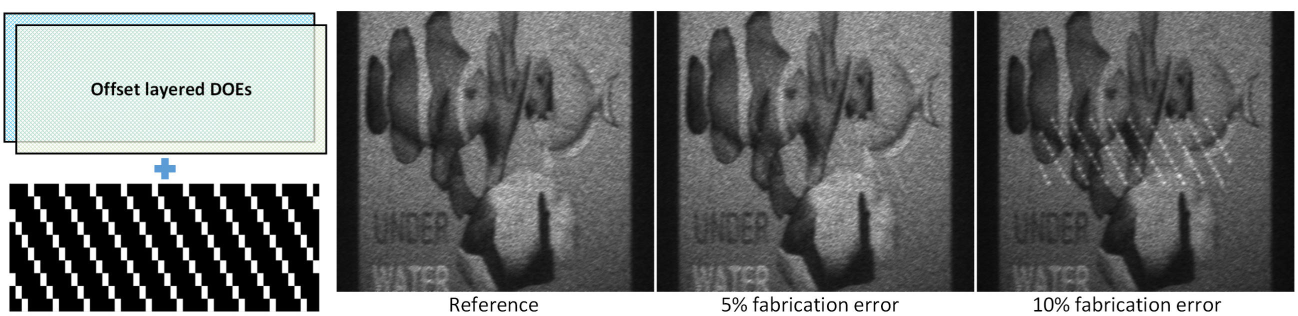

Fig. 6. The synthetic comparison results of reference (ideal fabrication), 5% fabrication error, and 10% fabrication error, respectively.

multiplexing is constant once the physical specification of DOEs are determined. Notice that the labels on bottom-left of each sub-figure indicates the frame number. At current stage, how to evaluate the maximum data amount that can be multiplexed subject to given hardware data capacity is still unclear. We yield this to future work.

One last thing, the frames are photographed after manually translating the top DOE layer to each target position. We further apply a small translation stage to record a video stream indicating well-defined animation effect. Compared to the discrete frame photographed above, the reconstructed running character in the video suffers from two constant "flare" due to the direct light leakage of the two alignment marks on the top and bottom sides of the patterns (refer to the photographs of the fabricated DOEs). Limited by the motion mechanics, we are unable to add additional cover to block these two marks when using a control motion stage. Nevertheless, this is not a fundamental issue of our method.

Possible color imaging solution. Conventional holographic projection suffers from the wavelength-dependency problem such that reconstructing a color image with a single pattern design is very tough. An alternative solution to obtain color projection is illustrated in Fig. 14. Again, following the holographic design scheme provides the design flexibility to merge or fusion multiple target holograms into one pattern design. Benefiting from the multiplexing scheme, the data bandwidth that can be encoded is drastically increased.

Application briefing for QR code encryption. We have mentioned that our multiplexing design can further boost new applications in security. Benefiting from the property that holograms intuitively exhibit like random patterns, under some sense of customized multiplexing, the two phase DOEs can be thought of as an encrypted message and a key: both are required to decode target information. In this market, current shortcomings of wavelength-dependency and diffraction efficiency loss are not that critical. Here we briefly demonstrated an instance to reconstruct a QR code for message encryption (shown in Fig. 15). Notice that it can be multiplexed into the 4 application scenarios, especially the offset one (1-to-1) and the multiple pairing one (many-to-many), that have been detailed in the main tex. Please refer to the accompanying video.

Application briefing for emissive transparent display. One unique advantage that diffractive optics has over refractive optics is its higher light efficiency when used in short wavelength's spectra (e.g. ultraviolet (UV)). Most glass medium have inevitable high absorption on UV spectrum (i.e. $\lambda < 400$nm). To guarantee relative high transmittance, expensive material and special coating need to be involved when designing UV-based refractive optics.

We have tested one of our DOE pair on an emissive fluorescent film (provided by [Sun et al. 2013]). This specially designed fluorescent film absorbs UV light and emits visible light, making it particularly promising in a wide range of VR/AR applications. As shown in Fig. 16, directly illuminating a pair of our DOEs with 405nm UV light source, and accordingly adjusting the projection distance, we successfully observe a red color reconstructed "cat" on an optical-clear film. Notice that the relative low image contrast and the noisy background are due to the large wavelength deviation (i.e. our DOEs are optimized at 550nm). Here we only show an example application. Combining the investigated multiplexing schemes with this emissive transparent display technique, a wide range of applications can be enabled.

Accompanying video. We recommend the readers to refer to the accompanying video to comprehensively catch the spirit of our mix-and-match holography solution.

REFERENCES

Joseph W Goodman. 2005. *Introduction to Fourier optics.* Roberts and Company Publishers.

Felix Heide, Qiang Fu, Yifan Peng, and Wolfgang Heidrich. 2016. Encoded diffractive optics for full-spectrum computational imaging. *Scientific Reports* 6 (2016).

Felix Heide, Douglas Lanman, Dikpal Reddy, Jan Kautz, Kari Pulli, and David Luebke. 2014. Cascaded displays: spatiotemporal superresolution using offset pixel layers. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 60.

Ngoc-Diep Ho. 2008. *Nonnegative matrix factorization algorithms and applications.* Ph.D. Dissertation. ÉCOLE POLYTECHNIQUE.

Ted Sun, Stephen Wu, and Botao Cheng. 2013. 54.4: Novel Transparent Emissive Display on Optic-Clear Phosphor Screen. In *SID Symposium Digest of Technical Papers*, Vol. 44. Wiley Online Library, 755–758.

Fig. 7. Comparisons of synthetic images reconstructed from the offset pairing subject to different implementation conditions— (a) perfect condition, (b) weak noise, (c) strong noise, (d) small rotation, (e) large rotation, (f) strong noise + small rotation one (g) directional shift, (h) dual directional shift, (i) strong noise + large rotation, respectively. Notice that all images have been resampled and integrated over 200μm per pixel subject to human eye's angular resolution at a viewing distance of 50cm. (Lena image source from Wikipedia)

Fig. 8. Comparisons of synthetic (left) images and experimental images (right two) reconstructed from the offset pairing. Notice that all the images have been resampled and integrated over 200μm per pixel subject to human eye's angular resolution at a viewing distance of 50cm. We also note that the intensity distribution of a target image has an impact on the encoding performance, further reflected as the speckle noise level in the decoded image. Specifically, the hair of Einstein exhibits a nosier distribution than that of Lena's hat. Overall, the two decoded gray-scale images are visually recognizable with details preserved. (Albert Einstein Photograph by Orren Jack Turner, Lena image source from Wikipedia)

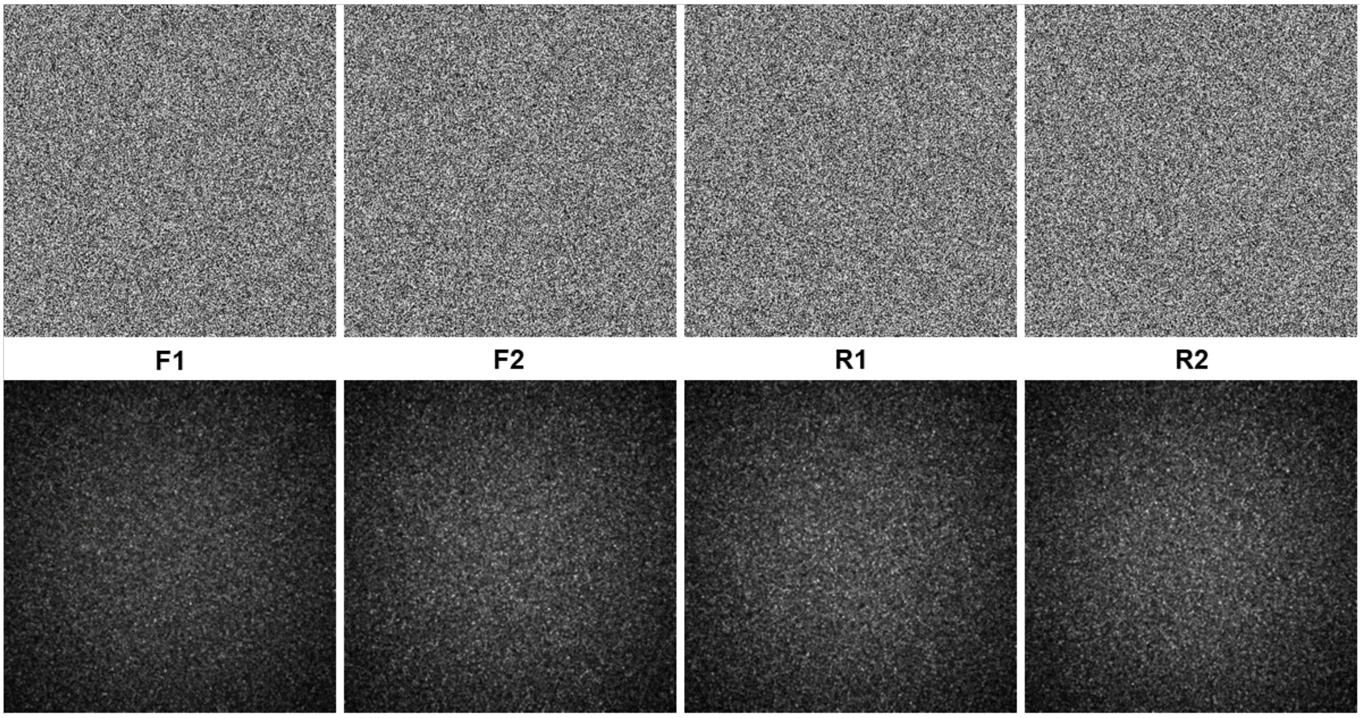

Fig. 9. Hologram visualization results (top row) and their corresponding synthetic reconstruction results (bottom row) of the banking account encoding application scenario. The parameters are the same as in Fig.8 of the main text.

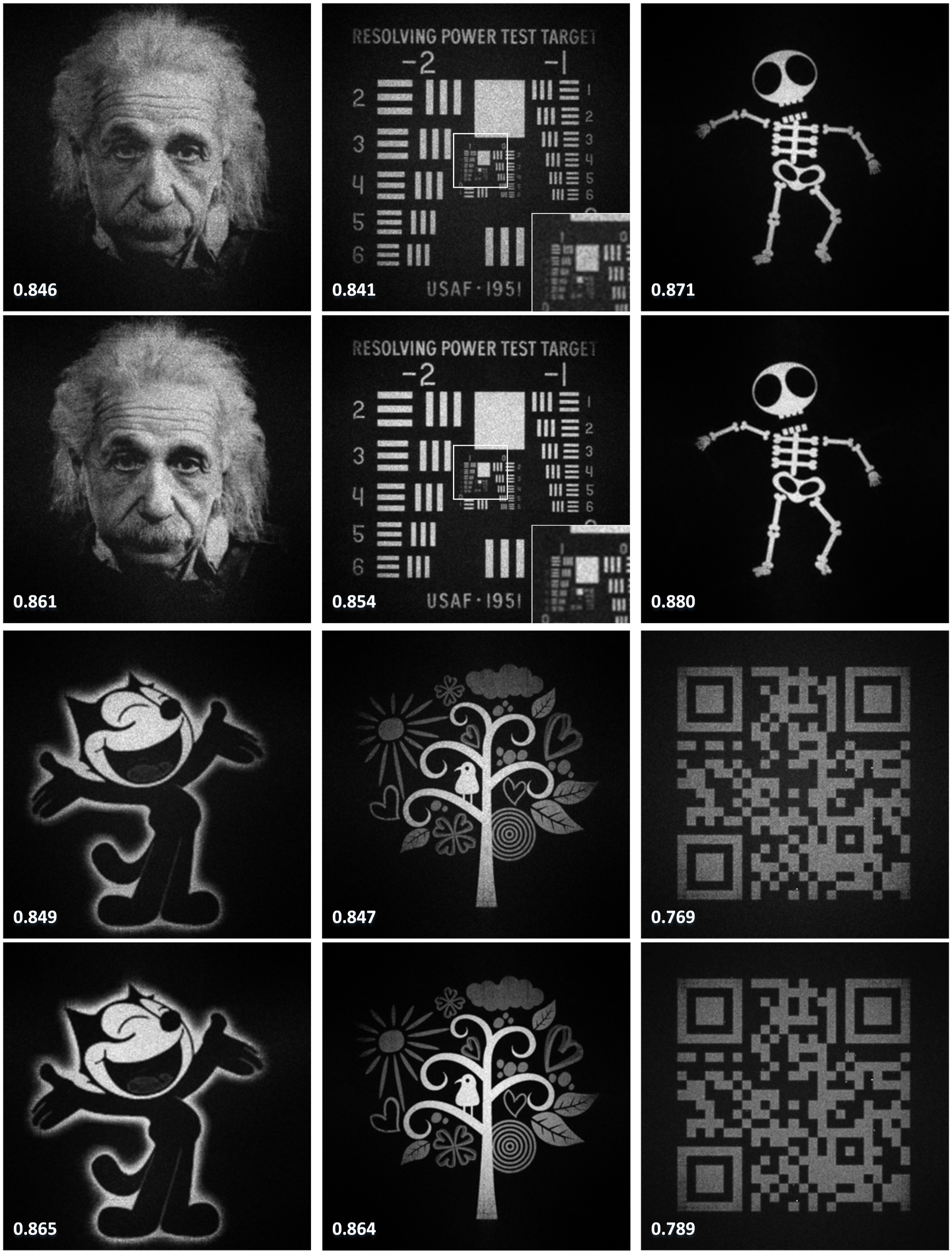

Fig. 10. Comparisons of synthetic images reconstructed from the offset pairing subject to two factorization update rules. Specially, we provide zoom-in close-ups of the second pair, as an instance. Notice that all images have been resampled and integrated as in Fig. 7.

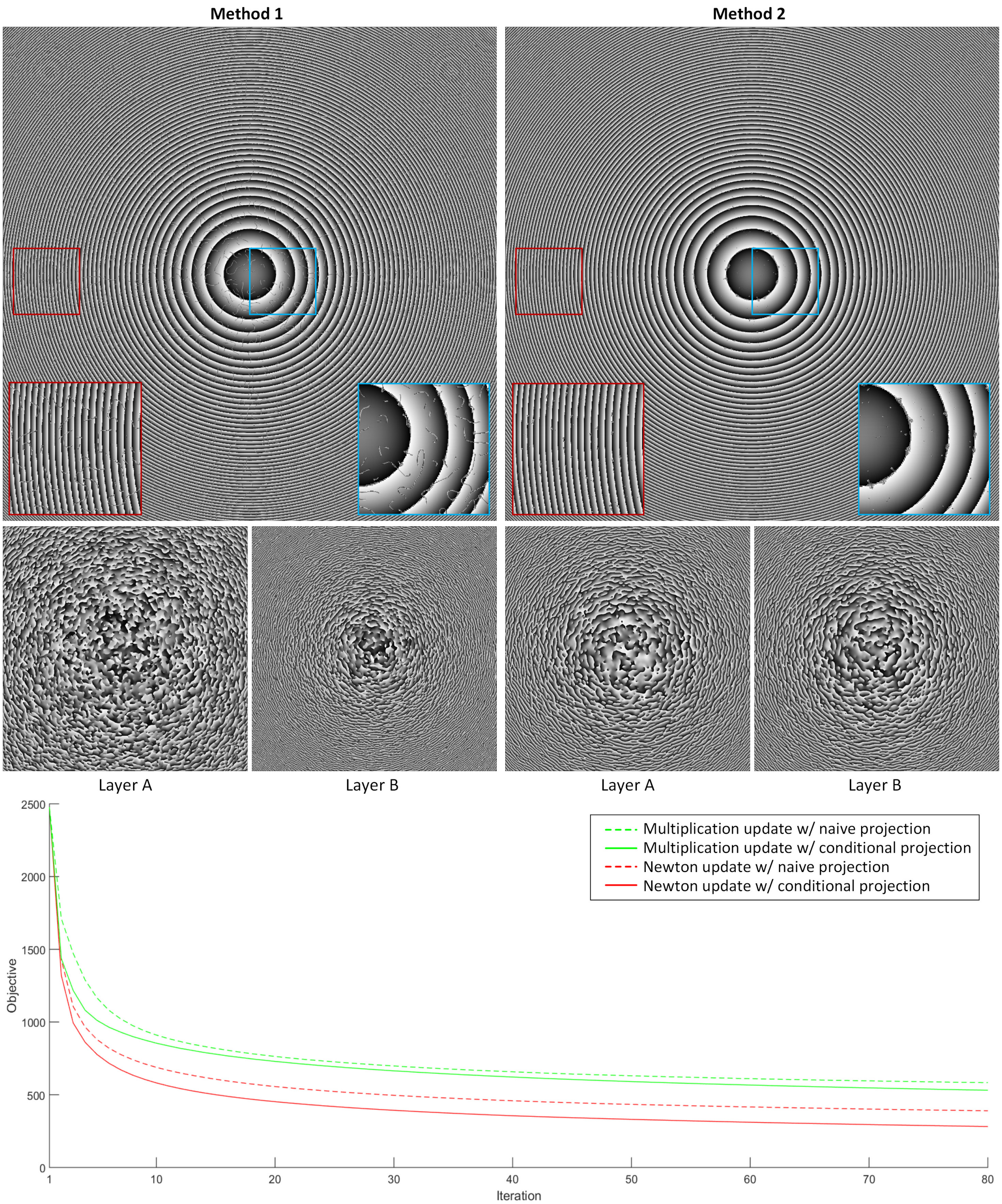

Fig. 11. Synthetic results (top) reconstructed from the offset pairing of factorizing the phase profile of a Fresnel lens, and its factorization convergence evaluation (bottom).

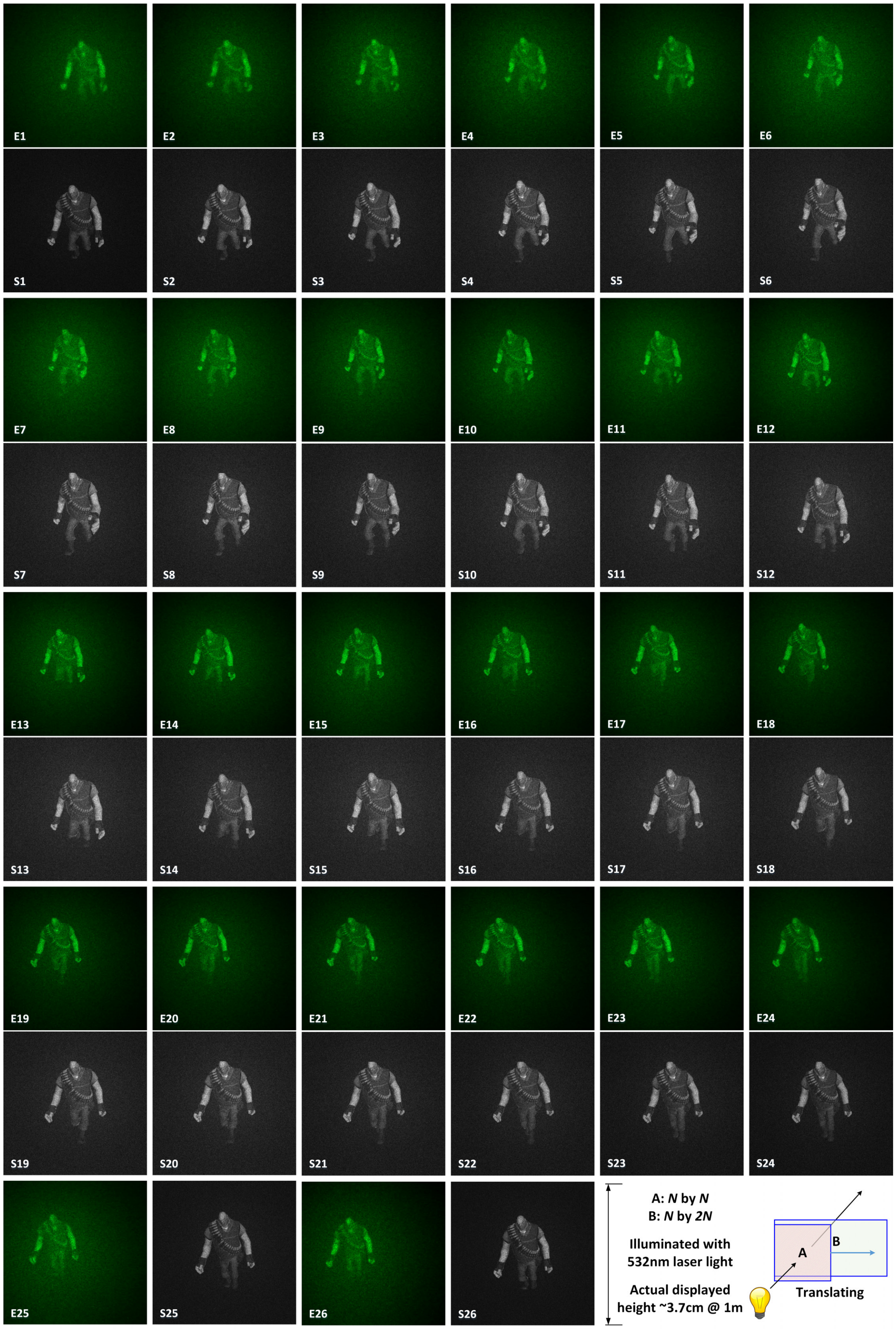

Fig. 12. The experimental (top) and corresponding simulated (bottom) results of projecting 26 animated frames through translational pairing of two DOEs.

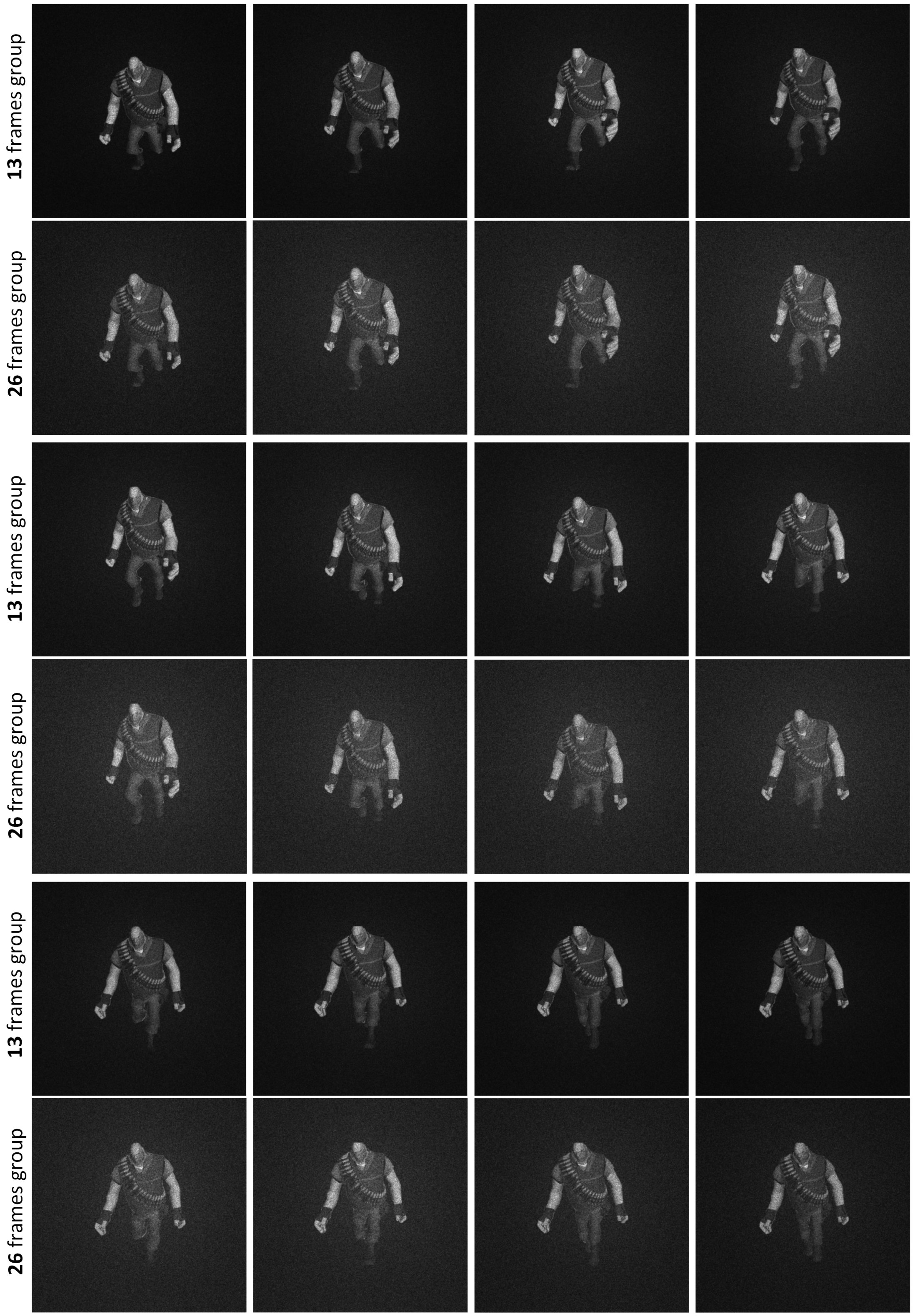

Fig. 13. The synthetic comparison results of multiplexing 13 frames versus 26 frames onto translational pairing of two DOEs with the same data bandwidth. Readers may notice the difference of image contrast and background noise level.

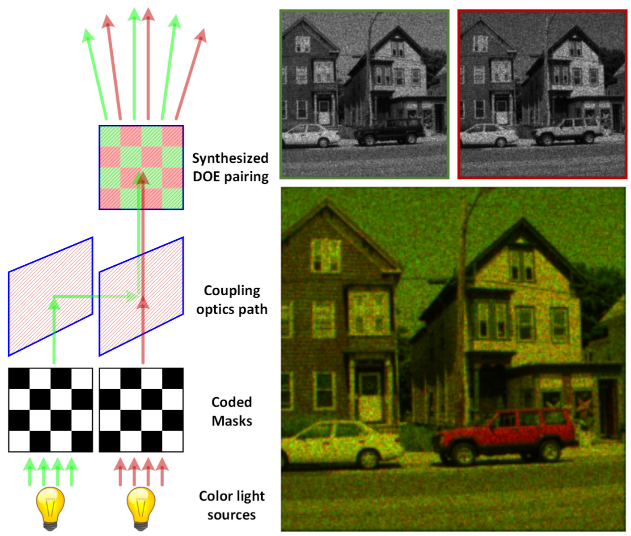

Fig. 14. An alternative solution of projecting a dual-color image through the offset pairing of DOEs with coded light sources. Left: diagram of set-up, where two binary masks are to be embedded in the light path to encode the incident illumination such that only those sub-regions designed for green light are illuminated by green light, meanwhile only those sub-regions designed for red light are illuminated by red light; Right: green channel, red channel, and dual-color reconstruction, respectively. Notice that this fusion can be multiplex into other application scenarios that have been detailed in the main text, by carefully constructing the target matrix incorporating the "filtering" process of coded light sources.

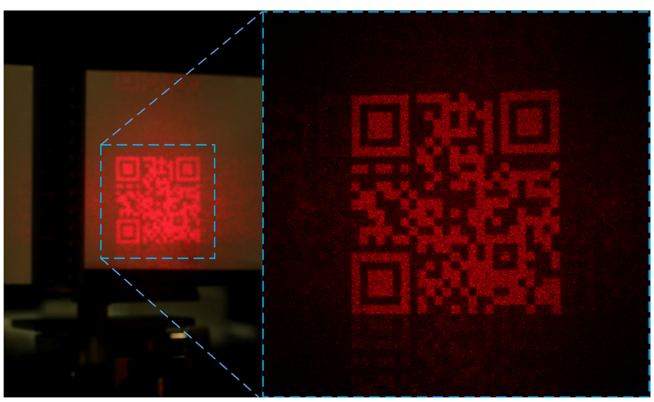

Fig. 15. An alternative application proof of incorporating mix-and-match multiplexing holograms with QR code encryption. Notice that the two DOEs are individually with the volume of 8mm by 8mm by 0.5mm only, so as to be easily assembled onto any kind of electronic device. With spot light or planar light illumination, a visualization with much larger size (5cm by 5cm at 1m distance) can be decoded.

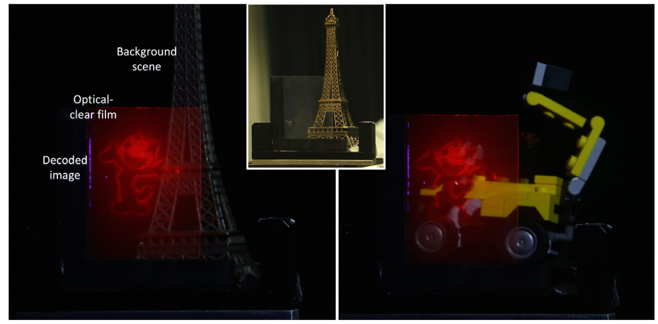

Fig. 16. A specific application proof of incorporating mix-and-match multiplexing holograms with UV emissive display. From the working status shown we observe a reconstructed scene as well as the background scene. The photographed scene is shown in the center close-up as a reference.