IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 25, 2018, accepted November 18, 2018, date of publication November 27, 2018,
date of current version December 27, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2883489

Removing Atmospheric Turbulence Effects in Unified Complex Steerable Pyramid Framework

CHAO ZHANG[1], BINDANG XUE[1], FUGEN ZHOU[1], AND WEI XIONG[2]
[1]School of Astronautics, Beihang University, Beijing 100191, China
[2]Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632

Corresponding author: Chao Zhang (zhangchao.101@hotmail.com)

This work was supported by the Academic Excellence Foundation of BUAA for the Ph.D. students.

ABSTRACT Simultaneously removing atmospheric turbulence-induced geometric distortion and blurry degradation is a challenging task. In this paper, we propose an effective method to remove or at least reduce turbulence effects in unified complex steerable pyramid (CSP) framework. The proposed method first decomposes the degraded image sequence by CSP. Then, the local motion and the energy information of the image sequence can be represented by multiscale and multidirectional phases and amplitudes. To mitigate turbulence-induced random oscillation, we use temporal average phase as the initial reference phase. Then, the reference phase is iteratively corrected, using the proposed phase correction method which is capable of correcting the large displacement. To reduce blurry degradation, optimal amplitude selection and fusion methods are proposed to reduce blur variation and CSP reconstruction errors. Finally, the corrected phase and fused amplitude can be synthesized to generate a reconstructed image. To further enhance the image quality, a blind deconvolution approach is adopted to deblur the reconstructed image. Through a variety of experiments on the simulated and real data, experimental results show that the proposed method can effectively alleviate the turbulence effects, recover image details, and significantenhance visual quality.

INDEX TERMS Atmospheric turbulence, image restoration, complex steerable pyramid.

I. INTRODUCTION

The image captured by the long-distance imaging systems, such as those for visual surveillance or astronomical observations, can be severely degraded by atmospheric turbulence. The spatiotemporal distortion and varied blur are the main atmospheric turbulence effects on imaging.

In general, there are two categories of methods to restore the degraded image, one is hardware-based adaptive optics (AO) techniques [1], [2], and the other is software-based image processing methods. The AO system is expensive and difficult to implement without special devices. Therefore, several image processing methods have been proposed [3]–[21]. Supposing the camera and the scene are static, the main purpose of these methods is to reconstruct a single high-quality image from the degraded image sequence.

These methods [3]–[21] consist of three major steps. Step 1, stabilizing the degraded video through non-rigid image registration. Step 2, fusing the registered image sequence into one image. Step 3, using deconvolution to the fused image and generate a sharp image.

To stabilize the degraded video through non-rigid image registration, a reference image should be prepared in advance. The easiest method is to select the first frame of the video as the reference image [3]. However, the first frame may be of low signal-to-noise ratio (SNR) resulting in inaccurate registration. The temporal mean [4], [5] and the temporal median [6] of the image sequence are also widely used as the reference image. However, the temporal mean and the temporal median may also be blurry. This will yield poor registration results. To avoid using a potential blurry reference image, work [7] uses a blind image quality metric to select the sharpest frame from the degraded image sequence as the reference image. A similar work was done in [8] by using ''lucky frame selection''. The suitable reference image should be sharp and less distorted simultaneously [9]. Image sharpness can be measured by the first or second derivative of the image [11]. However, turbulence-degraded images inevitably contain high-frequency noise. Derivative-based measurements are sensitive to noise. The direct method for measuring the geometric distortion between two images is to measure the deformation field. However, the intensity

similarity between a frame and the average frame is often used to measure the geometric distortion. In the turbulence-degraded video, the average frame is blurred. The selected reference frame is likely to be similar to the average frame. In work [12], the low-rank image, obtained by low-rank matrix decomposition method, is considered as the initial reference image, which is then iteratively refined by using a variational optimization method. However, both low-rank matrix decomposition and variational optimization are time-consuming methods for high-resolution images.

Image registration and fusion are critical in high-quality image recovery from degraded videos. In these regards, the B-spline based non-rigid image registration method [4], [5], [13] has been widely used to suppress geometric distortion. In [4] and [13], it is used to align each frame to the chosen reference image, followed by a patch-wise kernel regression based near-diffraction-limited (NDL) image fusion. However, such a registration method may not accurately achieve pixel-level correspondence between two images due to the missed detail features in the degraded image sequence. Inevitably, this will lead to registration errors. To improve the quality of the registration, the deformation-guided fusion method [12] is proposed to further reduce the registration errors. In this work, image feature orientations are computed in advance; they are then used to compute a pixel-wise deformation-guided kernel to remove the residual geometric distortion. In [14], Sobolev sharpening flow based method was proposed to reduce the inter-frame geometric distortion. However, the image intensity is degraded due to the turbulence effects, the distribution of the image intensity is not preserved under Sobolev sharpening, and the results will generate cross-color artifacts. Work [15] proposes to estimate the geometric distortion and use the inverse operation to recover the original image. Optical flow based registration techniques [9], [16]–[19] are also widely used to suppress the geometric distortion. However, the optical flow techniques used in these methods are under the assumption of brightness consistency. This assumption does not hold here because of fluctuations in the wavefront angles of arrivals (AOA), when the light wave propagates through atmospheric turbulence. Therefore, it cannot generate accurate registration results. In [10], the dual-tree complex wavelet-based fusion method has been proposed. This method requires manual interaction to select the informative region of interest (ROI) and use a segmentation technique to implement region-level fusion. To improve the quality of image fusion, method [20] extends Fourier burst accumulation (FBA) and uses wavelet burst accumulation (WBA) to accumulate high-frequency information to restore the turbulent blur. Recently, a phase-based method has been proposed to stabilize the turbulence-distorted video [21]. It uses the temporal lowpass filtering to alleviate the temporal variation of local phases and amplitudes. However, the temporal lowpass filtering technology does not have the ability to identify the optimal phase and the optimal amplitude and cannot reconstruct satisfactory results.

The fused image is a blurry image. Thus, deblurring is required. Because the blur varies spatially and temporally, restoring high-quality image is a nontrivial problem. An efficient filter flow method [22] was proposed for spatially varied point spread functions (PSF). This method divides each frame into small overlapped patches which are considered containing space-invariant blurs. According to this method, near-diffraction-limited (NDL) image reconstruction [4] and near-stationary patch based fusion [12] transform the space-variant blur to space-invariant blur. Thus, the sharp image can be generated by global deconvolution.

All the above-mentioned methods consider the turbulence-induced geometric distortion and blurry degradation as two unrelated problems and solve them separately by using two independent algorithms. However, from the point of view of atmospheric optics, the turbulence effects caused by wavefront AOA fluctuations, which results in the optical wave propagating oscillation and energy degradation. They correspond to the turbulence-induced geometric distortion and blurry degradation respectively in the captured long-distance video.

Complex steerable pyramid (CSP) [23], [24] can decompose an image into multiscale and multidirectional phases and amplitudes information and has the characteristics of compact structure and perfect reconstruction. Just as stated in [21] and [25], the local phase variation corresponds to the local motion information. Local amplitude can be used to represent local energy information, which corresponds to the degree of blurry degradation. Inspired by the AO system (adjusting the deformable mirror to compensate for the wavefront fluctuations to suppress optical wave oscillation and reduce energy degradation) [1], [2], we can adjust the local phase and the amplitude to suppress geometric distortion and reduce blurry degradation. Therefore, we propose an effective method to remove or at least reduce turbulence effects in unified complex steerable pyramid (CSP) framework.

The proposed method is implemented in the CSP domain. Therefore, we first use CSP to decompose each frame into multiscale and multidirectional phases and amplitudes. To remove the random oscillation induced by turbulence, we use the temporal average phase as the initial reference phase. Then, the local phase difference between the phase of each frame and the reference phase can be computed. The local phase difference over time corresponds to the turbulence induced local motion displacement. Because the mean of the geometry deformation field corresponds to the ground truth [18], we can derive the mean of the phase difference of all the frames closes to the phase of the ground truth. In other words, if the corrected phase is close to the phase of the ground truth, the mean of the phase difference of all the frames is a small value closing to zero. Under this constraint, the reference phase is iteratively corrected to make the corrected phase close to the phase of ground truth. To reduce blurry degradation, optimal amplitude selection and fusion methods are proposed by considering local movement, local energy, and CSP reconstruction error. Finally, the corrected

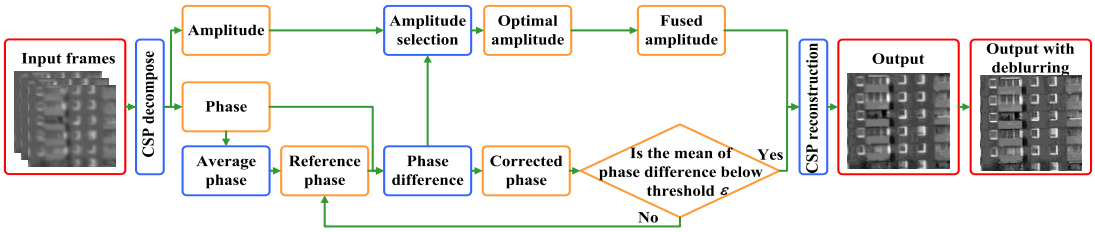

FIGURE 1. Block diagram for the proposed method.

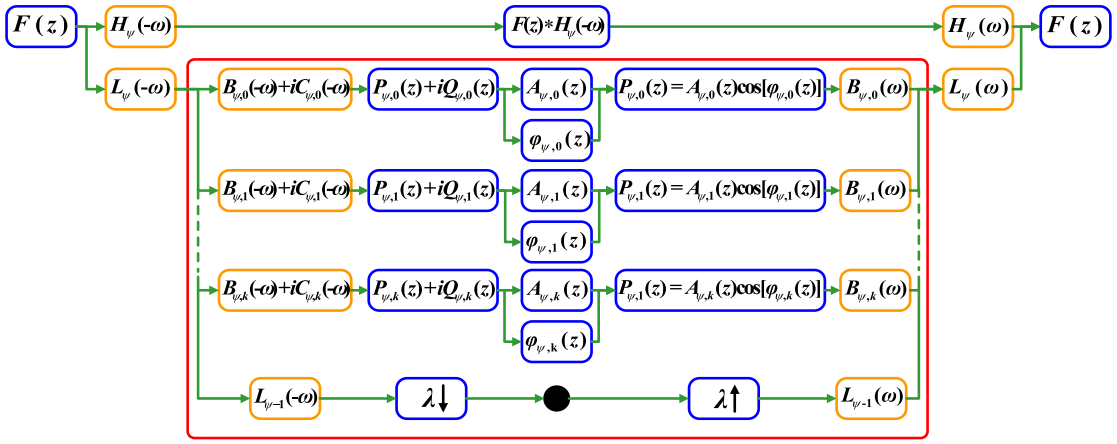

FIGURE 2. Decomposition and reconstruction block diagram for CSP.

phase and the fused amplitude can be used to generate a reconstructed image with reduced turbulence effects significantly. To further enhance the image quality, a blind deconvolution algorithm is applied on the reconstructed image to generate a high-quality result.

The main contributions of this work are as follows:

(1) An effective method is proposed to remove the turbulence effects, which combines removing turbulence-induced random oscillation and reducing blurry degradation in a unified CSP framework.

(2) The proposed phase correction method has the ability to rectify the phases even with large displacement induced by the strong turbulence.

(3) The importance of amplitude has been analyzed in details. The proposed optimal amplitude selection and fusion method can reduce the blur variation and CSP reconstruction errors.

The rest of the paper is organized as follows: Section II describes the proposed restoration framework. Experimental results are given in Section III to show the performance comparisons with other methods. Finally, we conclude the proposed method in Section IV.

II. PROPOSED METHOD

The proposed method includes three stages (see the diagram in Fig. 1): (1) removing turbulence-induced random oscillation by phase correction; (2) reducing blurry degradation by optimal amplitude selection and fusion; (3) CSP reconstruction and deblurring to further enhance the quality of the reconstructed image. Details of each stage are described in the following sections.

A. CSP DECOMPOSITION AND PHASE AND AMPLITUDE REPRESENTATION

Steerable pyramid [23], [24] uses steerable even filters to decompose the frequency of an image into several oriented frequency bands. CSP is the extension of the steerable pyramid [24]. It adopts steerable even filters and steerable odd filters (the Hilbert transform of the steerable even filters) to form the complex steerable filter (steerable analytic filters) and uses it to decompose an image into a real part steerable pyramid and an imaginary part steerable pyramid. The combination of the real part steerable pyramid and the imaginary part steerable pyramid forms the CSP.

The CSP decomposes an image into multiscale and multidirectional analytic signals. The complex-valued coefficients in the CSP can be used to obtain multiscale and multidirectional phases and amplitudes information, which will be exploited in the following to remove the turbulence effects. The decomposition and reconstruction block diagram for the CSP is shown in Fig. 2. For an input image $F(z)$, it is first decomposed into highpass and lowpass bands. Then, the lowpass band is further decomposed into oriented subbands and a lower frequency band. Next, the lower frequency is downsampled by a scale factor of λ. Such a decomposition

(indicated by the red region) can be recursively continued at the location of the solid circle to form the CSP pyramid.

After the CSP decomposition, we can compute the multiscale and multidirectional phases and amplitudes. Suppose $V_{\psi,\theta}(z)$ is the complex steerable response of image $F(z)$, indexed by scale ψ and orientation θ. It can be computed by convolution with the complex steerable filter $B_{\psi,\theta} + iC_{\psi,\theta}$, i.e.,

$$V_{\psi,\theta}(z) = F(z) * (B_{\psi,\theta} + iC_{\psi,\theta}) = P_{\psi,\theta}(z) + iQ_{\psi,\theta}(z) \quad (1)$$

where $P_{\psi,\theta}$ is the even filter response, $Q_{\psi,\theta}$ is the odd filter response. Based on Eq. (1), we can compute multiscale and multidirectional amplitudes and phases.

$$A_{\psi,\theta}(z) = \sqrt{P_{\psi,\theta}(z)^2 + Q_{\psi,\theta}(z)^2} \quad (2)$$

$$\varphi_{\psi,\theta}(z) = \text{atan2}[Q_{\psi,\theta}(z), P_{\psi,\theta}(z)] \quad (3)$$

where $A_{\psi,\theta}(z)$ is the amplitude component describing the local energy information at z, $\varphi_{\psi,\theta}(z)$ is the phase component corresponding to the local motion information and atan2 returns the arc tangent of the two variables in $[-\pi, \pi]$. Based on Eq. (2) and (3), the even filter response $P_{\psi,\theta}$ can also be expressed with $A_{\psi,\theta}(z)$ and $\varphi_{\psi,\theta}(z)$ as

$$P_{\psi,\theta}(z) = A_{\psi,\theta}(z)\cos[\varphi_{\psi,\theta}(z)] \quad (4)$$

After the amplitude $A_{\psi,\theta}(z)$ and phase $\varphi_{\psi,\theta}(z)$ have been processed, based on Eq. (4), the restored amplitude and the restored phase can be synthesized to generate the restored real component, which will be convoluted with the conjugate steerable even filter to generate the reconstructed image. The reconstruction process is opposite to the decomposition process. It is shown in Fig. 2.

B. PHASE CORRECTION

1) PHASE DIFFERENCE COMPUTATION

According to the phase-based signal processing techniques [25], the local phase difference corresponds to the local motion. Therefore, the turbulence induced random oscillation can be transformed to analyze the local phase difference. After the degraded image sequence has been decomposed with CSP, the phase difference between the phase of each frame and the reference phase can be computed as

$$\Delta\varphi_{\psi,\theta,t}(z) = \text{atan2}[\sin(\varphi_{\psi,\theta}(z) - \varphi_{\psi,\theta,t}(z)), \cos(\varphi_{\psi,\theta}(z) - \varphi_{\psi,\theta,t}(z))] \quad (5)$$

where $\varphi_{\psi,\theta}(z)$ and $\varphi_{\psi,\theta,t}(z)$ are respectively the reference phase and the phase of t-th frame on the scale ψ and orientation θ. $\Delta\varphi_{\psi,\theta,t}(z)$ is their phase difference. A large phase difference corresponds to a large displacement. But due to the periodicity of the phase difference, $\Delta\varphi_{\psi,\theta,t}(z)$ is limited in the range of $[-\pi, \pi]$. The motion displacement $d_{\psi,\theta,t}$ and the motion range can be computed as

$$d_{\psi,\theta,t} = \frac{\left|\Delta\varphi_{\psi,\theta,t}(z)\right|}{\omega} \leq \frac{\pi}{\omega} \quad (6)$$

where ω is the spatial frequency and each scale in the CSP represents a band of spatial frequencies. From Eq. (6), we can find both the phase difference and spatial frequency determine the motion range. The lower-frequency layers (coarse layers) can support the larger displacement, while the high-frequency layers limit the large motion. In addition, the motion of two adjacent layers should be close, which means the phase difference between two adjacent layers should be similar. Therefore, we intend to correct the phase difference on the high frequency layer when the phase difference between two adjacent layers is larger than a threshold. The difference between two adjacent layers can be expressed as

$$\upsilon_{\psi,\theta,t} = \text{atan2}[\sin(\Delta\varphi_{\psi,\theta,t}(z) - \lambda\Delta\varphi_{\psi-1,\theta,t}(z)), \cos(\Delta\varphi_{\psi,\theta,t}(z) - \lambda\Delta\varphi_{\psi-1,\theta,t}(z))] \quad (7)$$

where λ is the scale factor of CSP. The scales of two adjacent layers are respectively ψ and $\psi - 1$. If $\left|\upsilon_{\psi,\theta,t}\right|$ is larger than a threshold, the phase difference on the high frequency layer is unreliable, we should correct $\Delta\varphi_{\psi,\theta,t}(z)$ and set $\Delta\varphi_{\psi,\theta,t}(z) = \lambda\Delta\varphi_{\psi-1,\theta,t}(z)$. In this paper, we set the threshold to be 0.5π. If the threshold is too low, the high frequency layer will ignore their own motion and lose the detail of the motion at the high frequency. On the contrary, the high value does not benefit to correct the large displacement.

2) REMOVING TURBULENCE INDUCED RANDOM OSCILLATION BY PHASE CORRECTION

We first use the temporal mean phase as the initial reference phase $\varphi_{\psi,\theta}(z)$. Then, we use the above-proposed method to compute the phase difference $\Delta\varphi_{\psi,\theta,t}(z)$ between the phase of each frame and the reference phase. However, in some cases, a few of phase difference $\Delta\varphi_{\psi,\theta,t}(z)$ may differ significantly from the majority. These outliers may be originated from abrupt image motion due to large tilts in the wave-front angle-of-arrival. These outliers may result in the inaccurate geometric correction. To solve this problem, the outlier identification method [26] is applied. Because turbulence induced temporal random oscillation follows Gaussian distribution approximately, the outlier can be identified and eliminated if

$$\left|\Delta\varphi_{\psi,\theta,t}(z) - \mu\right| > \rho_{1-\alpha/2}\sigma \quad (8)$$

where ρ_c is the c quantile of the standard normal distribution, α is the confidence level, σ and μ are respectively the standard deviation and the mean of $\Delta\varphi_{\psi,\theta,t}(z)$. Here, we set $\rho_{1-\alpha/2} = 5$ to suppress the outlier. Then, the mean of the remaining phase differences can be computed as

$$\overline{\Delta\varphi_{\psi,\theta}(z)} = \frac{\sum_{t_r} \Delta\varphi_{\psi,\theta,t_r}(z)}{N_r} \quad (9)$$

where $\Delta\varphi_{\psi,\theta,t_r}(z)$ is the remained phase difference between the t-th frame and the reference phase, N_r is the number of the remaining frames, $\overline{\Delta\varphi_{\psi,\theta}(z)}$ is the mean of the phase difference. Finally, reference phase $\varphi_{\psi,\theta}(z)$ can be corrected

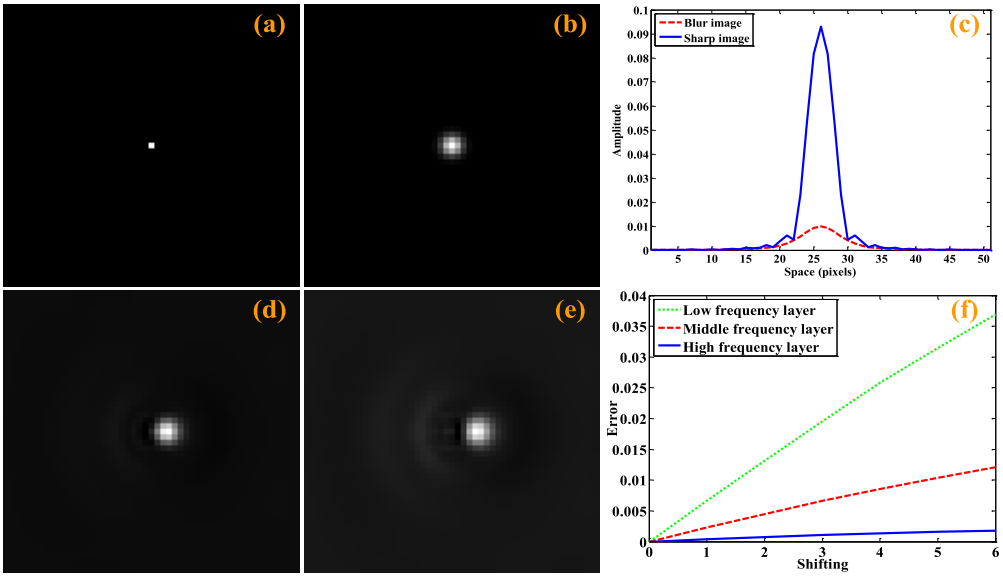

FIGURE 3. Amplitude degradation and reconstruction error. (a) a sharp image, (b) a Gaussian blur image of (a), (c) the amplitude intensity of the centerline profile in the high frequency layer, (d) reconstructed image by shifting three pixels to the right, (e) reconstructed image by shifting five pixels to the right (f) reconstruction error of different movement and different frequency layer.

with $\overline{\Delta\varphi_{\psi,\theta}(z)}$ to remove the geometric distortion. This can be expressed as

$$\varphi'_{\psi,\theta}(z) = \varphi_{\psi,\theta}(z) + \overline{\Delta\varphi_{\psi,\theta}(z)} \tag{10}$$

where $\varphi'_{\psi,\theta}(z)$ denotes the corrected reference phase. If the corrected phase $\varphi'_{\psi,\theta}(z)$ is close to the phase of ground truth of undistorted image, the mean of the phase difference $\overline{\Delta\varphi_{\psi,\theta}(z)}$ should be a small value close to zero. In order to generate the accurate results, we set $\varphi'_{\psi,\theta}(z)$ as the new reference phase $\varphi_{\psi,\theta}(z)$ and repeat the above method to correct iteratively the reference phase until $\left|\overline{\Delta\varphi_{\psi,\theta}(z)}\right| < \varepsilon$, where ε is a threshold for controlling the accuracy.

C. OPTIMAL AMPLITUDE SELECTION AND FUSION

After we get the corrected phase, the local amplitude is needed to reconstruct an image. The average amplitude of the image sequence and temporal lowpass filtering of amplitude [21] are two common amplitude processing methods. However, these methods can not generate satisfactory results because the reconstructed amplitude is not the optimal amplitude, which ignores the amplitude degradation induced by turbulence blur and movement. Fig. 3 reveals this problem. To solve this problem, we propose a novel strategy to select the optimal amplitude.

Fig. 3(a) is a sharp image with an impulse signal at the image center. In the local patch, the turbulence blur approximate to the Gaussian blur [27]. Therefore, we use Gaussian blur to degrade the local impulse signal. Fig. 3(b) is a Gaussian blurred image of the Fig. 3(a). We then use CSP to decompose these two images into three frequency layers (high frequency layer, middle frequency layer and low frequency layer). Fig. 3(c) shows the amplitude intensity of the centerline profile in the high frequency layer. Also, we can find the amplitude intensity of the sharp image is larger than that of the blurred image in all the frequency layers. The reason of this phenomenon is because the blurring degrades the local energy. Therefore, we should select the large amplitude.

Based on this, we further analyze the effect of turbulence motion on the result of CSP reconstruction. For the better explanation, we use the local impulse signal to simulate turbulence motion. We use phase correction method to shift the degraded impulse signal in Fig. 3(b) to the right. Then, the shifted impulse signal can be reconstructed with the corrected phase (phase of the ground truth, there is no error in the corrected phase) and the amplitude of the original Fig. 3(b). Fig. 3(d) and (e) are respectively the reconstructed images by shifting three pixels and five pixels to the right. From Fig. 3(e), we can visually find the reconstructed image is degraded. To reveal reconstruction error more intuitively, the reconstruction error has been analyzed. With the increase of the movement, we compute the mean of the reconstruction error on each frequency layer and show the result in Fig. 3(f). From Fig. 3(f), we can see that, for all the frequency layers, the reconstruction error increases as the movement increases. Therefore, we should select the amplitude which has less movement.

Based on the above analysis, we should select the local amplitude which is large with less movement. To reduce blur variation, we select the optimal amplitude in the local patch. The proposed method is illustrated in Fig. 4. This strategy corresponds to detect near-stationary patches. Near-stationary patch based fusion can be deblurred and enhanced using a space-invariant blind deconvolution method [12].

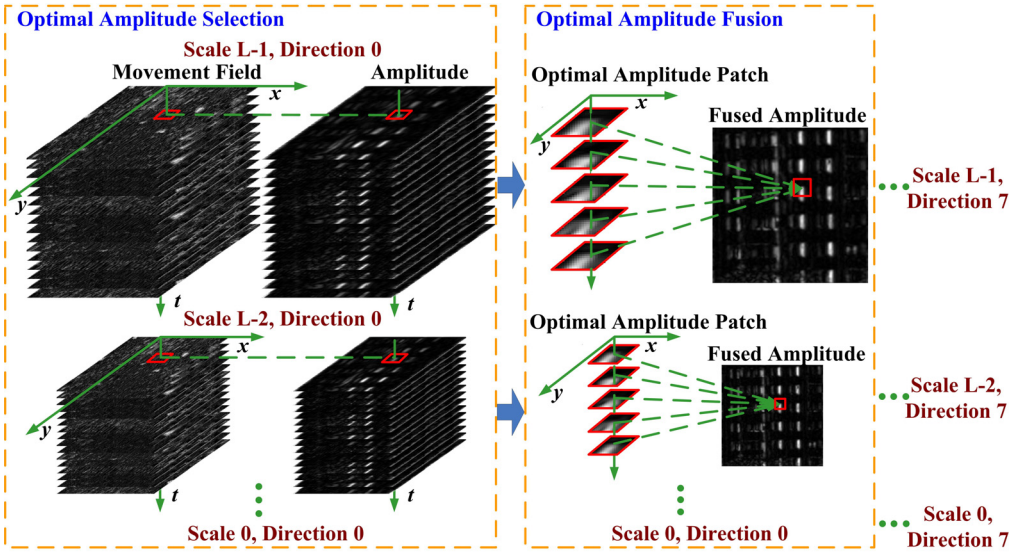

FIGURE 4. Illustrate of optimal amplitude selection and fusion.

In a multiscale and multidirectional amplitude sequence $\{A_{\psi,\theta,t}\}$, each amplitude image is divided into $K \times K$ overlapped patches centered at each pixel and the mean of the amplitudes is considered as the amplitude of local patch Ω. It can be defined as

$$a_{\psi,\theta,t}(z) = \frac{1}{K \times K} \sum_{z \in \Omega} A_{\psi,\theta,t}(z) \tag{11}$$

where $a_{\psi,\theta,t}(z)$ is the local amplitude of patch Ω centered at location z, indexed by scale ψ, orientation θ and frame number t.

In the same way, the local movement of patch Ω centered at location z can be derived from Eq. (5) and Eq. (9), which is

$$m_{\psi,\theta,t}(z) = \frac{1}{K \times K} \sum_{z \in \Omega} \left| \overline{\Delta\varphi_{\psi,\theta}(z)} - \Delta\varphi_{\psi,\theta,t}(z) \right| \tag{12}$$

To select the optimal amplitude, we normalize $a_{\psi,\theta,t}(z)$ and $m_{\psi,\theta,t}(z)$ to the range of [0, 1]. Then, the energy functional $E_{\psi,\theta,t}(z)$ for finding the optimal amplitude can be defined as

$$E_{\psi,\theta,t}(z) = \eta_\psi a_{\psi,\theta,t}(z) + (1 - \eta_\psi)(1 - m_{\psi,\theta,t}(z)) \quad \eta_\psi \in [0, 1] \quad t \in [1, N] \tag{13}$$

where t is the frame number, η_ψ is a weight coefficient for controlling the importance of the local amplitude and the local movement. We propose to maximize the energy functional $E_{\psi,\theta,t}(z)$ over t. Assuming $A_{\psi,\theta,t^\wedge}(z)$ is the largest value of $E_{\psi,\theta,t}(z)$ at the frame $t^\wedge$, the amplitude $A_{\psi,\theta,t^\wedge}(z)$ is chosen as the optimal amplitude at z for the fusion. From Fig. 3(f), we can see that when the movement is constant, the reconstruction error of the low frequency layer is larger than that of the high frequency layer. To reduce reconstruction error, we set a low weight value on the low frequency layer and a large weight value on the high frequency layer. According to the total number of the CSP layers, the weight coefficient of each scale is proportional to the scale number. It can be defined as

$$\eta_\psi = \frac{\psi}{L - 1} \quad \psi = 0, 1, 2, ...L - 1 \tag{14}$$

where L is the total number of layers, ψ is the scale, the larger ψ, the higher the frequency.

In Fig. 3, we have analyzed the effect of the image motion on the reconstruction error. The reconstruction needs the corrected phase and the amplitude. In Fig. 3(d) and (e), the corrected phase is replaced by the phase of the ground truth. Therefore, the reconstruction error comes from the amplitude. In this experiment, the original amplitude (the amplitude of the original image Fig. 3(b)) is used to reconstruct the shifted image. Because there is no movement in the original amplitude, the reconstruction error will appear.

The suitable amplitude should be close to the amplitude of an image which is without geometric distortion. In the turbulence-degraded videos, the temporal average is an easy method to remove the geometric distortion. Therefore, we use Eq. (13) to select T most optimal amplitudes and fuse them by setting their average amplitude as the reconstructed amplitude. This operation also can avoid local outlier induced by interframe noise and brightness changing. T is selected according to the turbulence strength. Compared with the turbulence effects, if T is too small, the reconstructed image may contain noise. On the contrary, a high value may contain the degraded amplitude and generate blurry results.

D. RECONSTRUCTION THE STABLE IMAGE AND DEBLURING

1) RECONSTRUCTING THE STABLE IMAGE

Based on Eq. (4), the corrected phase and the fused amplitude can be synthesized to generate the restored real component.

Then, the restored real component and temporal average highpass and lowpass residues can be used to generate the reconstructed image with reduced turbulence effects significantly. The reconstruction process is opposite to the decomposition process. It is shown in Fig. 2.

2) DEBLURRING

Although the optimal amplitude can be used to reconstruct a relatively clear result, the turbulence induced blur still exists in the reconstructed image. To further enhance the image quality, a blind deconvolution algorithm is adopted.

The degradation is generally modeled as

$$Y = F * h + n \tag{15}$$

where Y denotes the reconstructed stable image, F and h are respectively the sharp image and the blur kernel, n is the noise. The deconvolution model is described as follows

$$(\hat{F}, \hat{h}) = \arg\min_{F,h} \|Y - F \otimes h + n\|^2 + \beta_1 U_F(F) + \beta_2 U_h(h) \tag{16}$$

Based on the prior knowledge of F and h, where $U_F(F)$ and $U_h(h)$ are respectively their regularization terms used to restrain them. As suggested in [28], the sparse regularization term for F is defined as

$$U_F(F) = \left\| \varsigma\left(\frac{\partial F}{\partial x}\right) + \varsigma\left(\frac{\partial F}{\partial y}\right) \right\|_1 \tag{17}$$

where $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ are respectively denote the derivatives of F in horizontal and vertical directions, and $\varsigma(\cdot)$ is defined as

$$\varsigma(s) = \begin{cases} -\gamma_1 |s| & \text{s} \le l_t \\ -(\gamma_2 s^2 + \gamma_3) & \text{s} > l_t \end{cases} \tag{18}$$

where l_t, γ_1, γ_2 and γ_3 are fixed parameters. We use the suggested parameter settings described in [28]. The sparse regularization term for h is defined as

$$U_h(h) = \|h\|_1 \tag{19}$$

The details to optimize Eq. (16) can be found in [28].

III. EXPERIMENTAL RESULTS AND ANALYSIS

We elaborate on the processes of our experiments in this section. We first describe the parameter settings for the proposed method. Then, we illustrate the importance of amplitude and compare the results generated using the proposed amplitude processing method with those using the traditional methods. Next, we illustrate the detail process and demonstrate the validity of the proposed method in the case of strong turbulence. Finally, we compare and evaluate the proposed method respectively on real and simulated turbulence-degraded videos.

To quantitatively evaluate the proposed method, we input a sharp image (as shown in Fig. 5) and strong turbulence parameters into the turbulence simulation method [27] to generate 100 frames degraded image sequences.

FIGURE 5. Original sharp image.

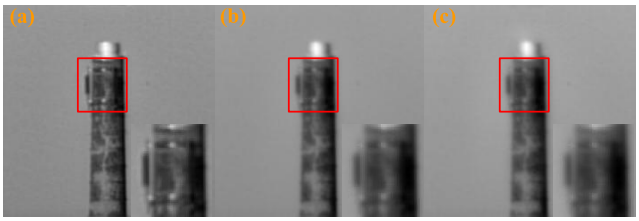

FIGURE 6. Reconstruction results with different amplitude processing methods. (a) proposed method, (b) average amplitude of the image sequence, (c) amplitude lowpass filtering [21].

The proposed method is compared with seven different methods: the two-stage method for image restoration [29] (Twostage), the centroid method [18] (Centroid), Sobolev gradient-Laplacian method [14] (SGL), non-local total variance for turbulence stabilization [15] (NLTV), Laplacian-Riesz pyramid for turbulence stabilization [21] (LRP), near-diffraction-limited based image reconstruction [4] (NDL) and complex wavelet-based fusion method [10] (CW). The outputs of all compared methods are generated using the authors' codes, with the related parameters unchanged.

A. PARAMETER SETTINGS

For CSP decomposition, the original CSP uses octave bandwidth filters and decomposes the frequency band into four orientations, but the impulse response of this filter is narrow in space, which cannot support large motions [25]. In addition, the more orientations are, the more accurate motion estimation can be computed [25]. Therefore, we should choose wide filters and decompose the frequency band into more (than four) orientations. But it will increase data computational cost. To balance between the performance and the computational cost, we decompose the frequency band into eight orientations with octave bandwidth filters (scale factor $\lambda = 2$) and half-octave bandwidth filters (scale factor $\lambda = 1.5$) respectively for weak turbulence and strong turbulence. For more about CSP, we refer interested readers to [23] and [24] for details. We set a small threshold $\varepsilon = 0.1$ as the phase correction iteration stop threshold. The size of the patch Ω is set as $K = 5$. We empirically set the blur kernel range and noise level range as [5] and [11] and [0.01, 0.06] to match with the actual turbulence condition.

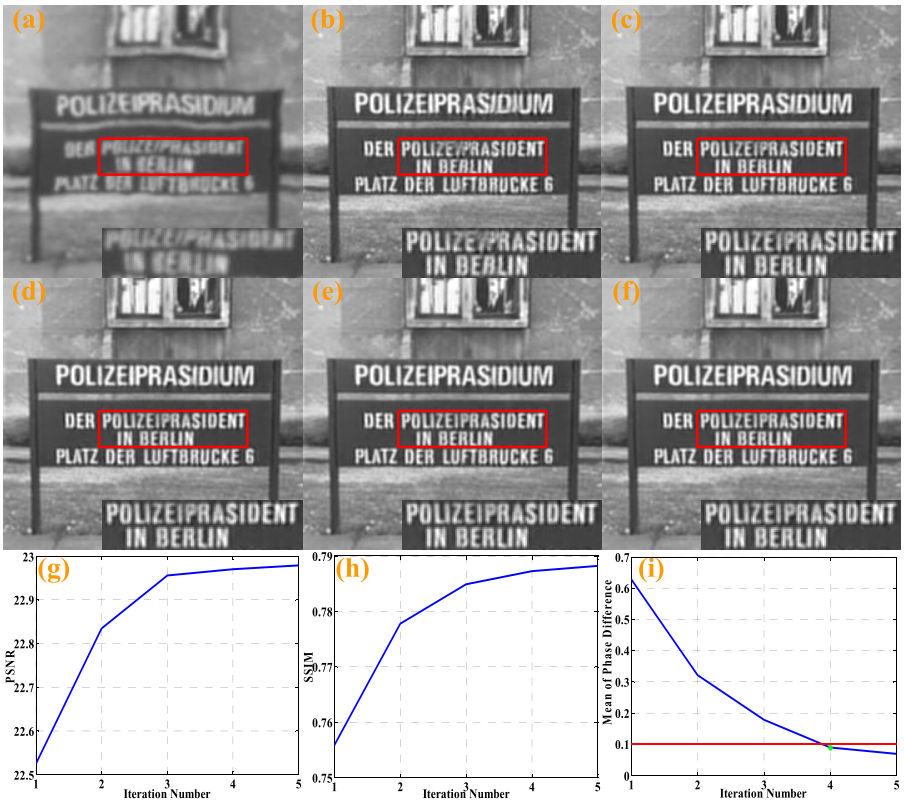

FIGURE 7. Image restoration results and quantitatively evaluate after each iteration. (a) a distorted frame, (b)-(f) are respectively the restored results after one to five iterations of phase correction, (g)-(i) are respectively the PSNR, SSIM and the mean of phase difference for each iteration.

The deblurring weight is chosen within the interval [0.2, 0.5]. Other parameters use the default settings described in [28].

B. IMPORTANCE OF OPTIMAL AMPLITUDE SELECTION

This section illustrates the importance of amplitude selection in reconstruction. We first use the proposed method to correct phase. Then, we use the proposed amplitude processing method, the average amplitude of the image sequence and amplitude lowpass filtering method [21] to produce the reconstruction results, respectively. As can be seen by comparing the reconstructed results with different amplitude processing methods in Fig. 6, the reconstructed result obtained from the proposed amplitude processing method can generate sharper result than the other two methods. There are two factors can explain this phenomenon. (1) Because the amplitudes of most frames in the image sequence are degraded by the turbulence blur, the average amplitude of the image sequence will also generate degraded amplitude. (2) The purpose of amplitude lowpass filtering is to alleviate the temporal variation of amplitudes caused by turbulent oscillation and to generate the stable amplitude. However, due to lowpass filtering technology only considers the temporal variation of amplitudes and cannot identify the degree of amplitude degradation, the generated amplitude cannot be guaranteed to be the optimal amplitude. The proposed method solves the above two problems. Our method not only selects and fuses optimal amplitudes but also considers the effects of the turbulent oscillation on the reconstruction error. In summary, the amplitude plays an important role in the complex analysis based turbulence removal method. Compared to the phase-based method [21], the proposed method improves noticeably the quality of the reconstruction result.

C. EFFECTIVENESS OF THE PROPOSED METHOD

In this section, we illustrate the detail process of the proposed method and demonstrate the validity of the proposed method. We first use the average phase as the initial phase. Then the restored images can be generated after each iteration of phase correction. We employ Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [30] to quantify the quality of the restored images. Fig. 7 shows the experimental results. As shown, Fig. 7 (b), (c), (d), (e) and (f) are respectively the restored results after one to five iterations of phase correction, in which the red regions are enlarged and shown in the bottom right corner. Fig. 7(g), (h) and (i) are respectively the PSNR, SSIM and the mean phase difference for each iteration.

Fig. 7(i) shows the mean of phase difference decreases gradually as the iteration number increases. It means the phase of the result is gradually close to the phase of ground truth. Therefore, the quality of restoration (PSNR and SSIM)

FIGURE 8. Comparison of results on simulated sequence. (a) ground truth, (b)-(c) two distorted frames, (d) Centroid, (e) NLTV, (f) Twostage, (g) LRP, (h) SGL, (i) NDL, (j) CW, (k) proposed method without deblurring, (l) proposed method with deblurring.

increases with the increase of the iteration numbers. Although the more iteration numbers, the better the results, after four iterations, we can find the PSNR and SSIM increase slowly. Therefore, we set a small stop threshold $\varepsilon = 0.1$ in this paper, namely, the iteration is stopped when the mean of the phase difference is below 0.1. The experimental results demonstrate the validity of the proposed method.

D. SIMULATED EXPERIMENTS

In this section, simulated experiments are implemented on the simulated strong turbulence sequence to test the proposed method. The restoration results are shown in Fig. 8. The red regions are enlarged and shown in the bottom right. Because blind deconvolution in the proposed method and method [4] is post-processing step, other compared methods [10], [18], [14], [15], [21], [29] do not include the deblurring step. To be objective and fair, we only compare the CSP reconstruction image. The deblurring result is only used to show the final result visually.

The SGL method (Fig. 8(h)) produces cross-color and shadowy artifacts during restoration. The LRP method (Fig. 8(g)) produces more blurry results than other methods since the low pass filtering in this method cannot generate optimal amplitude and phase. In the Centroid method, the deformation field computed by intensity-based optical flow is not accurate for the degraded turbulence sequence. All the frames are warped with the mean deformation field to generate the inaccurate centroid images. The temporal mean of the centroid images is taken as the final result, which operation will further blur the result. Therefore, the Centroid method (Fig. 8(d)) also produces the blurred result. NLTV, Twostage, and NDL produce similar results but artifacts and blur still exist. The result generated by CW contains cross-color and noise. In comparison, the proposed method outperforms other methods. The reason includes three factors: (1) Multiscale and multidirectional phase information is insensitive to the image contrast and can be effectively leveraged to represent motion information of the degraded image. (2) The phase is iteratively corrected to suppress the geometric distortion and to make the phase of the restored results gradually close to the optimal phase. (3) Optimal amplitude selection not only selects the optimal amplitude but also considers the CSP reconstruction error, which provides the guarantee for generating a clear result.

To quantitatively evaluate and compare the performance, Table 1 shows the Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Multiscale Similarity Index (MS-SSIM) [31], Spectral Residual-based Similarity Index (SR-SIM) [32] and Feature Similarity Index (FSIM) [33] values for the outputs of the eight different methods.

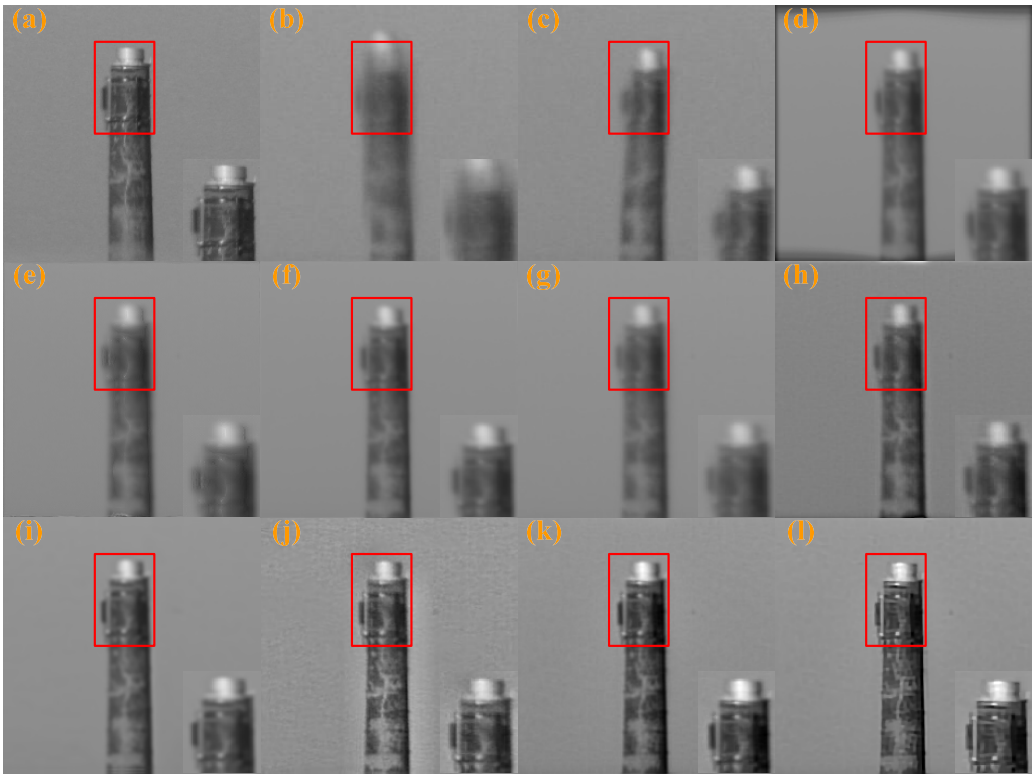

FIGURE 9. Comparison of results on Chimney sequence. (a) ground truth, (b)-(c) two distorted frames, (d) Centroid, (e) NLTV, (f) Twostage, (g) LRP, (h) SGL, (i) NDL, (j) CW, (k) proposed method without deblurring, (l) proposed method with deblurring.

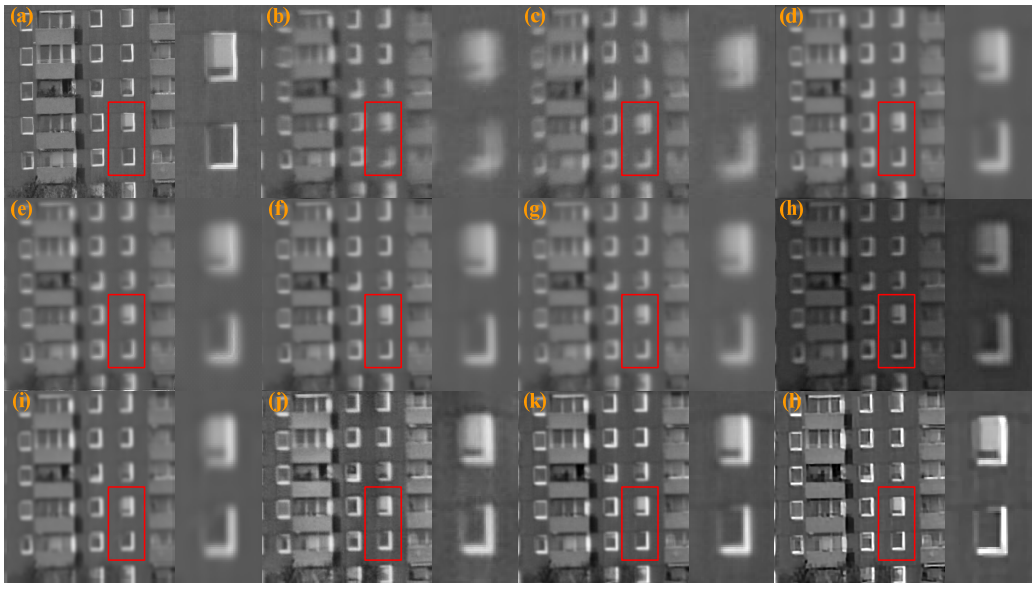

FIGURE 10. Comparison of results on Building sequence. (a) ground truth, (b)-(c) two distorted frames, (d) Centroid, (e) NLTV, (f) Twostage, (g) LRP, (h) SGL, (i) NDL, (j) CW, (k) proposed method without deblurring, (l) proposed method with deblurring.

To objectively and fairly compare the proposed method against other methods, here blind deconvolution is not applied to all the outputs. As shown, our method outperforms the other methods.

E. REAL EXPERIMENTS

We have tested the proposed method on three real turbulence-distorted videos, namely the Chimney, Building [22] and Door sequences [34]. The size of Chimney and Building

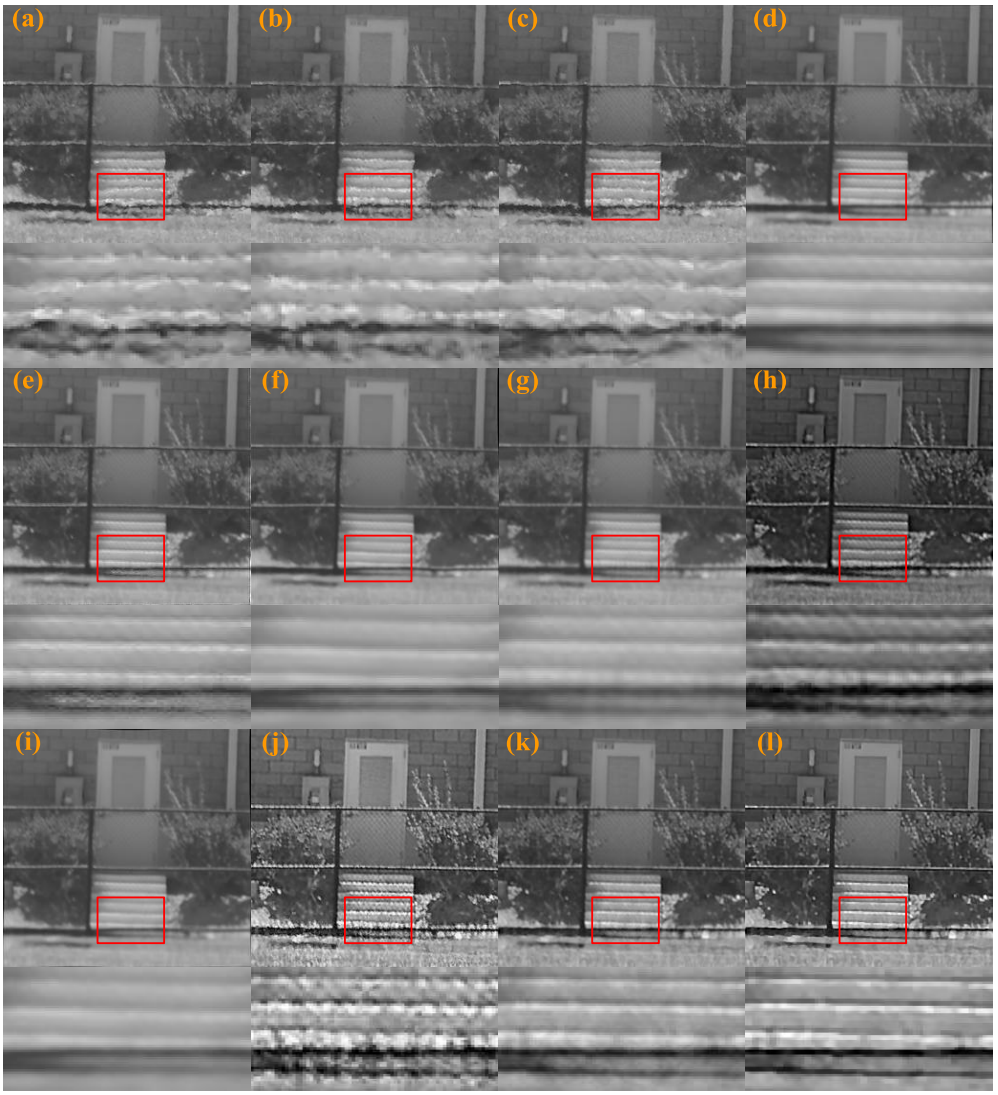

FIGURE 11. Comparison of results on Door sequence. (a)-(c) three distorted frames, (d) Centroid, (e) NLTV, (f) Twostage, (g) LRP, (h) SGL, (i) NDL, (j) CW, (k) proposed method without deblurring, (l) proposed method with deblurring.

sequences is 237 × 237 × 100. They were captured through the hot air exhausted by a building's vent, while the approximate ground truths were captured without through the hot air. The size of Door sequence is 520 × 520 × 300 frames, here we resize it to 260 × 260, and use the first 100 frames. Real Door sequence does not have the corresponding ground truth; therefore, we only compare its visual quality.

Fig. 9, Fig. 10 and Fig. 11 respectively show the outputs of different methods. The marked regions are enlarged in each figure. As shown, the geometric distortion and blurring are more severe in real sequences, which increase the difficulty of preserving geometric structure and texture details. Because the temporal averaging of inaccurate centroid images in the Centroid method, the weighted averaging of non-local self-similar patterns in the NLTV method, the low-rank matrix decomposition in the Twostage method, and the lowpass filtering in the LRP method restrict the preservation of geometric structure and texture details. Therefore, Centroid, NLTV, Twostage, and LRP methods all generate blurry results. NDL method performs slightly better, but some artifacts and blur still exist due to the registration error generated from symmetry constraint B-spline based registration. CW method can preserve the structure and details, but the results contain a lot of noise. SGL method degrades the overall intensity distribution and produces cross-color results. Compared with other methods, it is obvious to observe from Fig. 9(k), 10(k), and 11(k) that the proposed method preserves the geometric structure and texture details and produces much clearer results, which indicates the superiority of the proposed method. The PSNR, SSIM, MS-SSIM, SR-SIM, and FSIM values in Table 1 also indicate the proposed method outperforms other methods for Chimney and Building sequences.

TABLE 1. Performance of Different Methods Evaluated by PSNR, SSIM, MS-SSIM, SR-SIM and FSIM.

Image sequence	Index	Centroid	NLTV	Twostage	LRP	SGL	NDL	CW	Proposed
Simulated	PSNR	21.5829	21.4186	21.8265	21.0393	12.0210	21.7964	22.4322	**22.9791**
	SSIM	0.7060	0.7026	0.7203	0.6387	0.5576	0.6992	0.7441	**0.7882**
	MS-SSIM	0.9250	0.9268	0.9348	0.8992	0.8467	0.9263	0.9381	**0.9471**
	SR-SIM	0.8785	0.8811	0.8844	0.8495	0.8719	0.8698	0.8987	**0.9153**
	FSIM	0.7898	0.7864	0.8010	0.7485	0.7678	0.7832	0.8216	**0.8518**
Chimney	PSNR	24.0133	30.6256	30.5382	30.5658	28.4552	30.7381	30.5270	**31.0484**
	SSIM	0.8736	0.9016	0.9073	0.9089	0.9065	0.9087	0.8887	**0.9155**
	MS-SSIM	0.9401	0.9618	0.9539	0.9587	0.9636	0.9661	0.9633	**0.9707**
	SR-SIM	0.9132	0.8434	0.8831	0.8707	0.8662	0.9227	0.9275	**0.9356**
	FSIM	0.7881	0.8780	0.7569	0.8816	0.8868	0.8984	0.9101	**0.9157**
Building	PSNR	24.0312	23.7054	24.8694	23.8499	15.8501	23.2042	24.1054	**25.3721**
	SSIM	0.7214	0.6989	0.7344	0.6850	0.6449	0.6900	0.7280	**0.7926**
	MS-SSIM	0.8838	0.8681	0.8944	0.8553	0.8608	0.8460	0.8910	**0.9267**
	SR-SIM	0.8709	0.7864	0.8801	0.8305	0.7803	0.8897	0.8951	**0.9055**
	FSIM	0.7552	0.7558	0.7688	0.7388	0.7546	0.7694	0.7926	**0.8262**

After the deblurring step, the texture details will be further enhanced with significant improvements in visual quality.

IV. CONCLUSION

In this work, we have proposed a novel framework to reconstruct a high-quality image from turbulence-distorted video. Performance increase comes from the proposed method is implemented in the CSP domain. CSP consists of multiscale and multidirectional phases and amplitudes, which can be effectively leveraged to represent degraded image's geometric distortion and blurry degradation in a unified framework. We first decompose the degraded image sequence by CSP and use the average phase as the initial reference phase. Then the reference phase is iteratively corrected to remove the turbulence induced geometric distortion. To reduce blurry degradation and preserve image detail, optimal amplitude selection and fusion methods are proposed by considering the local energy, local movement, and CSP reconstruction errors. Finally, the reconstruction result can be deblurred using the blind deconvolution method for further enhancing image quality. Compared with seven different methods, the proposed method obtains excellent visual quality.

ACKNOWLEDGMENT

The authors would like to thank Prof. M. Hirsch and Prof. S. Harmeling for providing Chimney and Building image sequences. They also would like to thank Prof. J. Gilles for sharing Door image sequence and the code of NLTV [15], Dr. N. Anantrasirichai for providing the code of CW [10], Prof. Y. Lou for providing the code of SGL [14], Prof. E. Meinhardt-Llopis and M. Micheli for sharing the code of Centroid [18], Dr. X. Zhu for sharing the code of NDL [4], and Dr. O. Oreifej for sharing the code of Twostage [29].

REFERENCES

[1] R. K. Tyson, "Adaptive optics compensation of atmospheric turbulence: The past, the present, and the promise," *Proc. SPIE*, vol. 2222, pp. 404–412, Jun. 1994.

[2] J. E. Pearson, "Atmospheric turbulence compensation using coherent optical adaptive techniques," *Appl. Opt.*, vol. 15, no. 3, pp. 622–631, Mar. 1976.

[3] K. K. Halder, M. Tahtali, and S. G. Anavatti, "Simple and efficient approach for restoration of non-uniformly warped images," *Appl. Opt.*, vol. 53, no. 25, pp. 5576–5584, Sep. 2014.

[4] X. Zhu and P. Milanfar, "Removing atmospheric turbulence via space-invariant deconvolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 157–170, Jan. 2013.

[5] J. Li, J. Zhang, and Z. Sui, "Image restoration from sequences under atmospheric turbulence effects," *Proc. SPIE*, vol. 10462, p. 104625G, Oct. 2017.

[6] M. Shimizu, S. Yoshimura, M. Tanaka, and M. Okutomi, "Super-resolution from image sequence under influence of hot-air optical turbulence," in *Proc. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.

[7] K. K. Halder, M. Paul, M. Tahtali, S. G. Anavatti, and M. Murshed, "Correction of geometrically distorted underwater images using shift map analysis," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 34, no. 4, pp. 666–673, Apr. 2017.

[8] M. A. Vorontsov and G. W. Carhart, "Anisoplanatic imaging through turbulent media: Image recovery by local information fusion from a set of short-exposure images," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 18, no. 6, pp. 1312–1324, Jun. 2001.

[9] C. P. Lau, Y. H. Lai, and L. M. Lui. (2017). "Restoration of atmospheric turbulence-distorted images via RPCA and quasiconformal maps." [Online]. Available: https://arxiv.org/pdf/1704.03140.pdf

[10] N. Anantrasirichai, A. Achim, N. G. Kingsbury, and D. R. Bull, "Atmospheric turbulence mitigation using complex wavelet-based fusion," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2398–2408, Jun. 2013.

[11] J. E. Ospina-Borras and H. D. B. Restrepo, "Non-reference assessment of sharpness in blur/noise degraded images," *J. Vis. Commun. Image Respresent.*, vol. 39, pp. 142–151, May 2016.

[12] Y. Xie, W. Zhang, D. Tao, W. Hu, Y. Qu, and H. Wang, "Removing turbulence effect via hybrid total variation and deformation-guided kernel regression," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4943–4958, Oct. 2016.

[13] M. H. Furhad, M. Tahtali, and A. Lambert, "Restoring atmospheric-turbulence-degraded images," *Appl. Opt.*, vol. 55, no. 19, pp. 5082–5090, Jun. 2016.

[14] Y. Lou, S. H. Kang, S. Soatto, and A. L. Bertozzi, "Video stabilization of atmospheric turbulence distortion," *Inverse Problems Imag.*, vol. 7, no. 3, pp. 839–861, 2013.
[15] Y. Mao and J. Gilles, "Non rigid geometric distortions correction application to atmospheric turbulence stabilization," *Inverse Probl. Imag.*, vol. 6, no. 3, pp. 531–546, 2012.
[16] K. K. Halder, M. Tahtali, and S. G. Anavatti, "Geometric correction of atmospheric turbulence-degraded video containing moving objects," *Opt. Express*, vol. 23, no. 4, pp. 5091–5101, Feb. 2015.
[17] K. K. Halder, M. Tahtali, and S. G. Anavatti, "Moving object detection and tracking in videos through turbulent medium," *J. Model Opt.*, vol. 63, no. 11, pp. 1015–1021, Nov. 2016.
[18] E. Meinhardt-Llopis and M. Micheli, "Implementation of the centroid method for the correction of turbulence," *Image Process. Line*, vol. 4, pp. 187–195, Jul. 2014.
[19] T. Caliskan and N. Arica, "Atmospheric turbulence mitigation using optical flow," in *Proc. IEEE Int. Conf. Pattern Recog.*, Aug. 2014, pp. 883–888.
[20] J. Gilles and S. Osher, "Wavelet burst accumulation for turbulence mitigation," *J. Electron. Imag.*, vol. 25, no. 3, p. 033003, May 2016.
[21] B. Xue, Y. Liu, L. Cui, X. Bai, X. Cao, and F. Zhou, "Video stabilization in atmosphere turbulent conditions based on the Laplacian-Riesz pyramid," *Opt. Express*, vol. 24, no. 24, pp. 28092–28103, Nov. 2016.
[22] M. Hirsch, S. Sra, B. Scholkopf, and S. Harmeling, "Efficient filter flow for space-variant multiframe blind deconvolution," in *Proc. Comput. Vis. Pattern Recognit.*, 2010, pp. 607–614.
[23] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *Proc. Int. Conf. Image Process.*, 1995, pp. 444–447.
[24] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 49–70, Oct. 2000.
[25] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, "Phase-based video motion processing," *ACM Trans. Graph.*, vol. 32, no. 4, pp. 1–10, Jul. 2013.
[26] L. Davies and U. Gather, "The identification of multiple outliers," *J. Amer. Stat. Asssoc.*, vol. 88, no. 423, pp. 782–792, Sep. 1993.
[27] E. Repasi and R. Weiss, "Analysis of image distortions by atmospheric turbulence and computer simulation of turbulence effects," *Proc. SPIE*, vol. 6941, p. 69410S, Apr. 2008.
[28] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–10, 2008.
[29] O. Oreifej, G. Shu, T. Pace, and M. Shah, "A two-stage reconstruction approach for seeing through water," in *Proc. Comput. Vis. Pattern Recognit.*, 2011, pp. 1153–1160.
[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[31] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, pp. 1398–1402.
[32] L. Zhang and H. Li, "SR-SIM: A fast and high performance IQA index based on spectral residual," in *Proc. Int. Conf. Image Process.*, Sep. 2012, pp. 1473–1476.
[33] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
[34] J. Gilles and N. B. Ferrante, "Open turbulent image set (OTIS)," *Pattern Recognit. Lett.*, vol. 86, pp. 38–41, Dec. 2017.

CHAO ZHANG received the M.S. degree in pattern recognition and intelligent system from Jilin University, Changchun, China, in 2013. He is currently pursuing the Ph.D. degree with the School of Astronautics, Beihang University, Beijing, China. He was a Visiting Student with the Department of Visual Computing, Agency for Science, Technology and Research, Singapore. He was supported by the Beihang Graduate Scholarship Council in 2015. His research interests include atmospheric optics, target detection and recognition, and computer vision.

BINDANG XUE received the Ph.D. degree from the Dalian University of Technology, Dalian, China, in 2003. He held a post-doctoral position at the Department of Control Science and Engineering, Beihang University, from 2003 to 2005. He is currently an Associate Professor with the School of Astronautics, Beihang University. He has authored or co-authored over 40 publications in international journals and conference proceedings. His research interests include atmospheric optics, target detection and recognition, and computer vision.

FUGEN ZHOU received the B.S. degree in electronic engineering and the M.S. and Ph.D. degrees in pattern recognition and intelligent system from Beihang University, Beijing, China, in 1986, 1989, and 2006, respectively. He is currently a Professor with the School of Astronautics, Beihang University. He has authored or co-authored over 70 publications in international journals and conference proceedings. He is also a Standing Member of the Chinese Society for Optical Engineering. His research interests include atmospheric optics and computer vision.

WEI XIONG received the B.S. degree from the Huazhong University of Science and Technology, Wuhan, China, and the M.S. and Ph.D. degrees from the National University of Singapore, Singapore. He is currently with the Institute for Infocomm Research, Agency for Science Technology and Research. He has authored or co-authored over 50 publications in international journals and conference proceedings. His research interests include computer vision, pattern recognition, and image processing.