Multipath Interference Suppression in Time-of-Flight Sensors by Exploiting the Amplitude Envelope of the Transmission Signal

AYUSHYA PARE [ID]1, SHUFANG ZHANG [ID]2, (Member, IEEE), AND ZHICHUN LEI1
1 Institute of Measurement and Sensor Technology, University of Applied Sciences Ruhr West, 45479 Mülheim an der Ruhr, Germany
2 School of Electrical and Information Engineering, Tianjin University, Tianjin 300000, China

Corresponding author: Shufang Zhang (shufangzhang@tju.edu.cn)

ABSTRACT Time-of-flight (ToF) cameras provide depth information and they have been applied in many areas. However, multipath interference (MpI) affects the depth measurements. This article proposes an algorithm to suppress MpI during ToF camera depth measurements by exploiting the amplitude envelope of the camera's transmission wave. More precisely, it uses the fact that the amplitude envelope of the ToF camera's transmission wave is constant. The envelope of the reflected signal in the case of MpI deviates from that of the MpI-free envelope. By minimizing the deviation using an adaptive filter, the depth measurement error can be corrected. The method is applied at the sensor side of the ToF camera and lays the theoretical foundation for MpI suppression in ToF cameras. The simulation and experimental results demonstrate that the proposed algorithm is able to significantly improve the depth measurement accuracy.

INDEX TERMS Adaptive filter, constant envelope, depth acquisition, multipath interference, time-of-flight.

I. INTRODUCTION

In recent times, 3D imaging has undergone tremendous growth and imaging applications have been introduced in various fields, such as the automotive, gaming, smartphone, and robotics domains. Time-of-flight (ToF) is a 3D imaging technique that determines the depth of an object in a scene by illuminating the scene with modulated light and measuring the phase shift of the reflected light. ToF imaging has been applied in areas such as material sensing, vehicle parking assistance, underwater situational awareness, robot navigation, obstacle detection and collision avoidance, etc. [1]–[6]. Moreover, with the emergence of low- and medium-cost cameras with very high resolution and frame rates, the ToF camera market is experiencing unprecedented growth. ToF imaging has finally made a leap into mass production in consumer markets, with more advanced and sophisticated cameras being manufactured. The major ToF camera manufacturers include PMD, MESA, Optrima, and Microsoft [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Yongqiang Zhao [ID].

A ToF camera operates by illuminating the scene with a modulated pulsed or continuous-wave (CW) light signal and then measuring the phase displacement between the transmitted and received signals [8], [9]. The camera emits an amplitude-modulated light in the near-infrared (NIR) range. The objects in the scene then reflect the light, and the camera optic projects the reflected light into a complementary metal-oxide semiconductor (CMOS) or a charged coupled device (CCD) pixel matrix. With the use of advanced electronics, the depth is then computed from the phase shift between the transmitted and reflected signal [10]. In reality, the light reaching the camera sensor is reflected not just directly from the target object but also from other objects and surfaces present in the vicinity. The phase shift, in this case, is computed from the superposition of the transmitted and multiple reflected light signals. The measured depth, as a result, is either more or less than the actual depth. Multipath interference (MpI) is one of the primary sources of error in ToF data acquisition and can cause significant inaccuracies in depth estimations.

Several studies have addressed the impact of MpI on ToF depth measurements and several others have proposed

MpI compensation and depth correction: For an intensive review, see Whyte *et al.* [11]. A broad categorization of the MpI suppression methods can be made according to whether the method removes the MpI from the final ToF depth image or during the signal capture at the camera sensor.

The first category of methods aims to suppress the MpI from the final depth image by either developing an MpI model or by using structured light patterns to project the scene. Jimenez *et al.* incorporated a radiometric model to predict the ToF measurements and remove MpI contamination by implementing an iterative optimization algorithm [12]. Bhandari *et al.* employed a method for the backscattering vector estimation through a custom-coded waveform instead of a standard sinusoidal waveform and used the sparse deconvolution procedure to recover the sequence of signals corresponding to MpI [13]. Gupta *et al.* proposed an approach where light with a high modulation frequency was used to illuminate the scene [14]. For high-frequency illuminations, the global radiance is initially expressed as a DC component and is then canceled out from the total illumination. Naik *et al.* used a similar approach where the scene was illuminated with a spatial checkerboard pattern. The direct-reflected light and the echoes are then acquired and separated through the decomposition of the measured radiance at the sensor [15]. The technique developed by Baek *et al.* eliminates the depth error by employing a juxtaposition of three ToF cameras set to different integration times [16]. Although the results were successful, additional hardware and precise synchronizations for multiple ToF cameras are required.

The second category of methods attempts to eliminate the MpI from the sensor side of the ToF camera during depth acquisition. Freedman *et al.* developed a method based on the use of multiple frequencies and a sparse reflection analysis to remove the MpI using the L1 optimization algorithm [17]. In [18], Agresti *et al.* presented a method for MpI correction by separating the direct and global components of the received signal using high-frequency projection patterns. Patil *et al.* proposed a technique called microToF imaging to mitigate the effect of MpI in ToF 3D sensors using the direct and global separation technique [19]. Dashpute *et al.* proposed the elimination of non-systematic errors (including MpI) by incorporating a polarizer with a ToF camera and measuring the state of the reflected light at particular polarization angles [20]. The depth error is at a minimum at the maximum intensity of the received light, which occurs at the polarization angle. Lickfold *et al.* suggested simultaneous stepping of the phase and the frequency of the CW modulation of the camera to remove the error and the offset in the depth measurement [21]. Although these are efficient techniques for MpI suppression, they are not easily implementable for the complete range of ToF devices, as they either require some assumptions regarding light characteristics or require additional custom coding and structural light projections to remove MpI.

Many researchers have previously exploited the signal envelope for ToF measurements. The detection of the
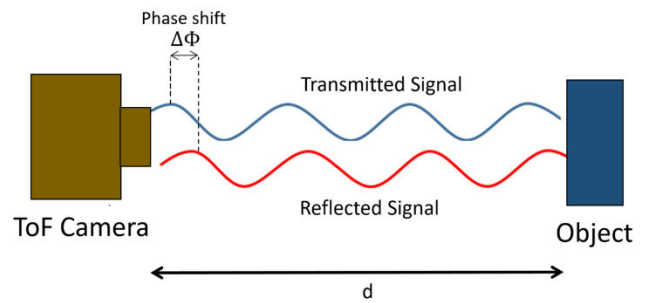

FIGURE 1. ToF camera principle.

envelope's zero crossings of the transmitted and received signals for ultrasonic ToF measurements has been used in [22]–[24]. However, unlike the zero-crossing detection or the envelope extraction used by other researchers, the method proposed in this article exploits the fact that the amplitude envelope of a sinusoidally modulated wave is a constant and that the envelope of a signal affected by MpI deviates from the constant one. The amplitude envelope of the signal, also referred to as the analytic envelope of the signal, is the magnitude of the sum of the squared in-phase and quadrature phase of the signal [25], [26]. In this article, this envelope is exploited to suppress MpI during ToF depth measurements, which happens at the sensor side of the ToF camera. The algorithm suppresses the MpI from the reflected light signal, irrespective of factors such as directivity and the reflection properties of the surrounding objects, and it does not require any specifically designed multipath model. The performance of the proposed method was validated through simulations as well as experiments with real measurements. The experimental scenarios were designed to be similar to what a ToF camera measures in the real world. The results confirmed the effectiveness of the proposed method.

In the following section, the imaging process at the sensor of the ToF camera is briefly explained. Then, the algorithm proposed for the suppression of echoes is presented. After that, the simulations and the experiments based on the proposed method are discussed at the end of the paper.

II. ToF CAMERA PRINCIPLE
A CW ToF camera emits a temporally modulated light signal, usually in the NIR range, employing LEDs as the light source to illuminate the scene. The light reflected by the scene reaches the camera sensor as a temporally varying signal with an attenuated amplitude and a phase shift. The depth of the object is estimated by obtaining the phase displacement between the transmitted and received signal. The modulated transmitted signal is expressed as

$$s_t(t) = a_t \cos(\omega t - \phi_i) \qquad (1)$$

where a_t is the amplitude of the transmitted light signal, ω represents the transmitted frequency, and ϕ_i is the initial phase of the signal. Without MpI, the reflected light reaching

the receiver is given by

$$s_r(t) = b_0 + a_r \cos(\omega t - \phi) \tag{2}$$

where b_0 represents the ambient light, a_r is the amplitude of the reflected light, and ϕ is the phase shift of the reflected light. The depth d is calculated from the phase shift $\Delta\phi$, i.e., $(\phi_i - \phi)$ of the reflected light by the following relation:

$$d = \frac{c}{4\pi f}\Delta\phi \tag{3}$$

where c is the speed of light, and f is the frequency of the transmitted light signal. To retrieve the phase shift of the received signal, synchronous demodulation of the received light signal is carried out. In this method, also commonly known as the "four-bucket" sampling technique, the received signal is sampled at four specifically chosen instances, usually at 0, $\pi/2$, π, and $3\pi/2$ for every period. The phase shift is then calculated by the following formula [27]:

$$\Delta\phi = \arctan\left(\frac{c\left(\frac{3\pi}{2}\right) - c\left(\frac{\pi}{2}\right)}{c(\pi) - c(0)}\right) \tag{4}$$

The depth of the object is further estimated by the relation given in (3). A ToF-based depth estimation assumes that the light reaching each pixel is reflected from a single point in the scene without inter-reflections. The light, however, may undergo multiple reflections from other reflecting objects, mirror-like and semi-transparent materials, or from a complicated geometry of walls and corners in the vicinity. A number of multiple reflected signals then combine at each pixel. This distorts the depth calculations and results in incorrect depth results. Usually, the received light is considered as a superposition of two components: the direct reflection from the target object and the indirect component, which includes all the other reflections from the surrounding objects and surfaces. The received signal $s_r(t)$ can be mathematically represented as a sum of the direct-reflected and the N superimposing multipath reflections:

$$s_r(t) = b_0 + a_0 \cos(\omega t - \phi_0) + \sum_{k=1}^{N} a_k \cos(\omega t - \phi_k) \tag{5}$$

where k represents the number of superimposing multipath reflections. Most of the MpI is comprised of the light signals that undergo up to three reflections from the scene. Further higher-order reflections suffer from large attenuation and they decay exponentially due to scattering [13].

The ToF depth measurement also encounters many other practical issues, which are categorized into two broad categories: systematic and non-systematic errors. The systematic errors include the integration time (IT) error, amplitude ambiguity, temperature error, depth distortion error, and the built-in pixel error, and the non-systematic errors include MpI, background illumination, scattering noise, and motion blur [28], [29]. Modern ToF cameras use techniques such as the suppression of background illumination [30] or employ special background-light immune sensors to suppress the

effect of ambient light in the measurements [31], [32]. Thus, the term b_o in the above equation is neglected in the calculations in this article. Depth measurements in ToF imaging are also affected by motion blur; i.e., the error occurring when the objects move. In [33], the ToF raw data are considered as a noisy time series, and the error between raw frames due to transverse motion is reduced by the application of statistical processing using a Kalman filter. This work, however, is only concerned with MpI and its suppression, and the other errors are excluded from the scope of this article.

III. MpI SUPPRESSION EXPLOITING THE CW CHARACTERISTICS OF LIGHT

The method proposed in this article exploits the amplitude envelope of the sinusoidal waveform to mitigate against the effect of MpI from the incident light signal. As already mentioned in Section I, the amplitude envelope of a signal is given by

$$E^2(t) = s^2(t) + H(s(t))^2 \tag{6}$$

where $H(s(t))$ is the 90° phase shift or the quadrature of the signal. This method employs the fact that the amplitude envelope of a sinusoidal signal is a constant, whereas the envelope of the signal with MpI deviates from the constant one. Consider an example where the received signal is comprised of one direct-reflected and two echo signals. The received signal is given by the equation

$$s_r(t) = a_0 \cos(\omega t - \phi_0) + a_1 \cos(\omega t - \phi_1) + a_2 \cos(\omega t - \phi_2) \tag{7}$$

where a_0, a_1, and a_2 are the amplitudes of the direct-reflected signal and the echoes, respectively, and ϕ_0, ϕ_1, and ϕ_2 are the respective phase shifts. The light signal incident on the camera sensor without any MpI is given by

$$s_r(t) = a_o \cos(\omega t - \phi_0) \tag{8}$$

and the envelope of the above signal is

$$(a_0 \cos(\omega t - \phi_0))^2 + (a_0 \sin(\omega t - \phi_0))^2 = a_0^2 \tag{9}$$

Similarly, the envelope of the light signal with multiple echoes is given by

$$\left(\begin{array}{c} a_0 \cos(\omega t - \phi_0) + a_1 \cos(\omega t - \phi_1) \\ + a_2 \cos(\omega t - \phi_2) \end{array}\right)^2$$
$$+ \left(\begin{array}{c} a_0 \sin(\omega t - \phi_0) + a_1 \sin(\omega t - \phi_1) \\ + a_2 \sin(\omega t - \phi_2) \end{array}\right)^2 \tag{10}$$

which, on further simplification, gives

$$a_0^2 + a_1^2 + a_2^2 + 2a_1 a_2 \cos(\phi_0 - \phi_1)$$
$$+ 2a_0 a_2 \cos(\phi_0 - \phi_2) + 2a_1 a_2 \cos(\phi_1 - \phi_2) \tag{11}$$

Comparing Eq. (9) and Eq. (11), it can be seen that the envelope of the received light with echoes deviates from that of the direct-reflected signal. By minimizing this deviation, the MpI from the received signal can be suppressed.
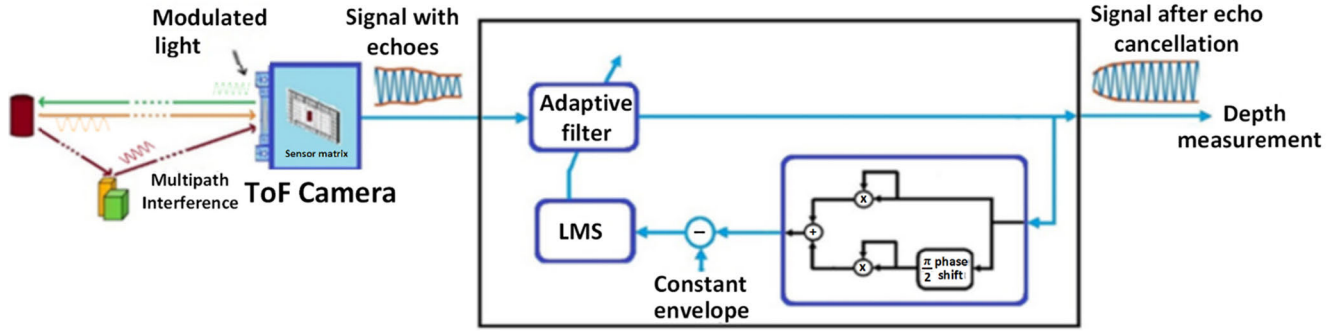
FIGURE 2. Workflow of the proposed algorithm.

The proposed method suppresses the MpI from the received signal by incorporating the least mean square (LMS) algorithm. The algorithm updates the adaptive filter coefficients and consequently restores the received signal in a form whose envelope matches that of the signal without any MpI. The workflow of the proposed MpI suppression algorithm is described in Fig. 2. The light received at the sensor, distorted due to the MpI, is passed through an adaptive filter whose output is then routed into two 90° phase-shifted signals, which are squared and added to each other. The 90° phase shift is achieved through the Hilbert transformation. The output of the filter with weight vector $\overrightarrow{W} = [W_0, W_1, \ldots, W_L]^T$ and the input signal vector $\overrightarrow{X} = [X_1, X_2, \ldots, X_M]^T$ is given by $\overrightarrow{Y} = \overrightarrow{W}^T \cdot \overrightarrow{X}$. Here $(\cdot)^T$ is the transpose of a vector, L is the number of filter coefficients, and M is the size of the input signal. Next, the resulting signal is compared with a constant envelope. The instantaneous error at a time instance k is given by

$$e(k) = d(k) - y_{env}(k) \qquad (12)$$

where $d(k)$ is the constant envelope and y_{env} is the envelope of the filter output signal given by

$$y_{env} = y_c^2 + y_s^2 \qquad (13)$$
$$= W^2 \cdot X_c^2 + W^2 \cdot X_s^2 = W^2 \cdot X_{env} \qquad (14)$$

The LMS algorithm is a stochastic gradient descent method, i.e., the update is based on the error at the current instance. The cost function is formulated as the instantaneous mean squared error (MSE) between the envelope of the received signal and the constant, given by

$$J(k) = \frac{1}{4}e^2(k) \, (15) \qquad (15)$$
$$= \frac{1}{4}\left(d(k) - W^2 \cdot X_{env}(k)\right)^2 \qquad (16)$$

The gradient of the MSE with respect to the weights at each iteration is given by

$$\Delta(k) = \frac{\partial}{\partial W}(J(k)) \qquad (17)$$

After further simplifications, the gradient becomes

$$\Delta(k) = y_{filt}(k) \cdot e(k) \qquad (18)$$

where the term $y_{filt}(k) = W^T \cdot X_{env}(k)$ represents the filter output of the signal envelope. For a particular choice of step-size μ, the update of the weights, according to the steepest descent adaptive algorithm, can be written as

$$\overrightarrow{W}(k+1) = \overrightarrow{W}(k) - \mu \cdot y_{filt}(k) \cdot e(k) \qquad (19)$$

The filter coefficients are updated at each subsequent iteration of the algorithm until the cost function gradually reaches a minimum. The filter output converges in a way that its envelope is as close to the constant envelope as possible. This implies that the echo components in the received signal (Eq. 11), which deviate the envelope from the constant, are suppressed. By suppressing the echoes from the received signal, the influence of MpI on the ToF depth measurement is minimized.

The filter's coefficients are adapted such that the echoes are effectively subtracted from the main signal during equalization. Consider an example where the reflected light consists of one direct-reflected and one echo signal whose amplitude is 0.25 mV. The echo arrives at the sensor with a delay of three samples, as shown by the received signal in Fig. 3.

In the course of the equalization, the filter coefficients adapt to their optimum values so that the echo is effectively subtracted from the signal. Each coefficient is adjusted accordingly to compensate for any deficiency that remains. As the echo may be received with a certain delay after the main signal, the first tap of the filter need not be the main tap. In the simulations it was observed that the filter weight $\overrightarrow{W}[3]$ (corresponding to the echo at n = 3) converges to the final value of ~ -0.25 (Fig. 4) and thus subtracts the echo from the main signal.

IV. SIMULATIONS AND RESULTS

The efficiency of the proposed method was investigated by simulating ToF operation under different scenarios. The simulations were carried out using MATLAB. The transmitted signal was modeled with an amplitude $a_0 = 0.5$ mV and
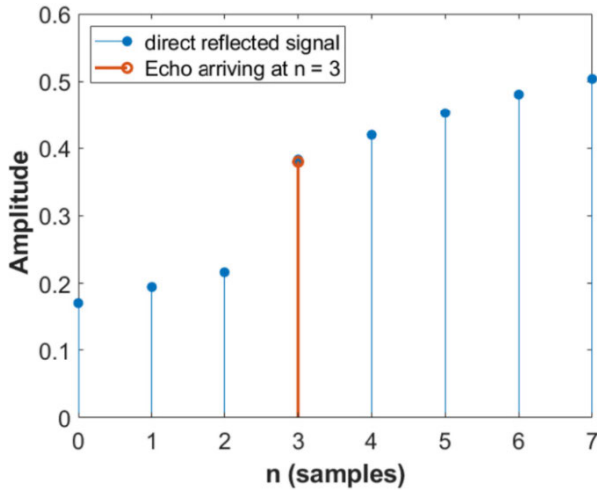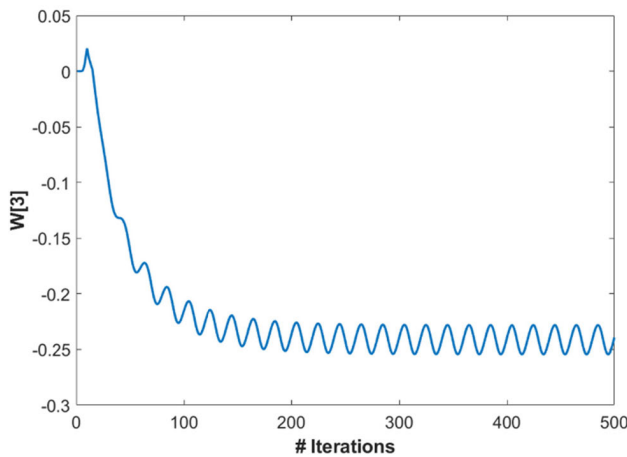
FIGURE 3. Received signal and echo at n = 3.

FIGURE 4. Convergence of the filter coefficient.

a frequency of 10 MHz, which is commonly used in many ToF cameras. An object at a distance of 1 m from the ToF camera was considered for the simulations. The received light without any echo gives the actual distance of the object. The phase shift of the reflected light for a distance of 1 m can be derived from the known formula in Eq. 3. To determine the effectiveness of the algorithm, a number of different scenarios, with the received light consisting of one, two, and three echoes, were considered separately. It is known from [13] that the amplitude of an echo signal attenuates after each bounce. Thus, for a case with two echoes, their respective amplitudes were set to be $a_1 = a_0/10$ and $a_2 = a_0/100$, assuming one echo reaches the sensor after more bounces than the other one. The simulated received light is then cross-correlated with the transmitted signal, and with the "four-bucket" sampling, the phase shift, and consequently, the depth is computed. The depth in this case was estimated to be 1.09 m, which does not match the ground truth, resulting in a depth error of 0.09 m. The received signal was then equalized by employing the proposed adaptive filter.

A major challenge in designing an adaptive filter is the compromise between the computational complexity and convergence accuracy. An infinite impulse response (IIR) filter structure, in comparison with a finite impulse response (FIR) structure, involves a fixed number of computations with respect to memory locations and mathematical operations, thus providing the additional advantage of reduced computational complexity [34]. The rate of convergence of the MSE depends on a number of factors, such as the filter initialization, filter length, and the step size. The selection of the step size regulates the convergence speed and stability of the adaptation. There are many methods that propose dynamically updating the step size during adaptation to improve the speed and accuracy of the convergence: A detailed compilation of the methods and the formulas can be found in the paper by Hwang and Li [35]. Although there is no analytical definition that characterizes the convergence and the factors influencing it, there are, however, qualitative results that can help in choosing the parameters to attain the desired convergence. The speed of convergence increases with the step size; however, an upper bound on the step size remains for a stable convergence given by

$$0 < \mu < \frac{1}{(L+1)\,(signal\ power)} \quad (20)$$

where L is the filter length [36]. The length of the filter has an inverse relation with the speed of the convergence as it influences the "condition number" for the cross-correlation of the signal and the filter, and consequently reduces the upper bound on the step size for the MSE to converge [36]. There are several algorithms, such as fractional tap-length LMS (FT-LMS), variable-tap variable-step LMS (VTVS LMS), and variable-length LMS (VL-LMS) to help choose an optimum filter length L for a given step-size μ [37].

Taking the various factors into consideration, the equalization of the simulated received signal was carried out with an IIR filter with five coefficients (initialized to zero) and a step size of 0.01. It was observed that the error in the depth estimation had been minimized. The depth after the equalization was computed to be 1.001 m, which is very close to the actual depth, and the error is reduced to 0.001 m.

TABLE 1. Depth correction for different echo signals.

Number of echoes	Measured depth (m)	Error before (in %)	Corrected depth (m)	Error after (in %)
0 (ground truth)	1	0	1	0
1	1.088	8.8	1.002	0.2
2	1.090	9	1.001	0.1
3	1.098	9.8	1.002	0.2

Table 1 lists the simulation results for the extended reflection model by considering the reflected signal corrupted by a different number of echoes. The measured and the corrected depths for each case and the corresponding errors before and after the echo cancelation are shown in the table. The error

in each case is minimized, and the depth after correction is much closer to the actual depth.

For further simulations, it was considered that the reflected signal reaches the sensor with an additive white Gaussian noise (AWGN) with a 30-dB signal-to-noise ratio (SNR). Table 2 lists the depth measurements before and after the echo correction for different numbers of echo signals. The data from Table 2 show that the MpI has been suppressed from the received signal in every case, and the depth computed after the equalization is very close to the ground truth.

TABLE 2. Depth correction in a noisy medium.

Number of echoes	Measured depth (m)	Error before (in %)	Corrected depth (m)	Error after (in %)
0 (ground truth)	1.063	6.3	1.041	4.1
1	1.151	15.1	1.073	7.3
2	0.874	13.6	1.013	1.3
3	0.827	17.3	1.075	7.5

More simulations were carried out by assuming that the object was placed at different distances from the camera. It was observed that the MpI suppression and the subsequent depth correction were equally good, even when the object was quite far from the camera.

TABLE 3. Depth correction for an object at different distances.

Actual depth (m)	Measured depth (m)	Error before (in %)	Corrected depth (m)	Error after (in %)
1	1.11	11	1.001	0.1
2	1.92	4	1.999	0.05
5	4.88	2.4	4.99	0.2
10	10.058	8.5	10.002	0.02
25	24.89	0.44	24.908	0.3

The results in Table 3 show the comparison of the depth of the object before and after the correction for different distances from the camera. It can be seen from the data that for each case, the corrected distance closely matches the ground truth.

V. EXPERIMENTS AND RESULTS

In addition to the simulations, experiments with real ToF measurements were also performed to validate the efficiency of the proposed method in real-world scenarios. The different experimental scenarios involved objects of different sizes and material being placed strategically at different distances from the ToF camera. The experimental setup consisted of a laser diode as the light source that emitted sinusoidally modulated light at a frequency of 100 kHz with 30 mW power and 0.6 mV peak amplitude. The light is reflected from a target object placed at a certain distance from the light source. A photodetector placed adjacent to the emitter receives the reflected light. The reflected light signal is then passed through a low-noise amplifier and recorded by
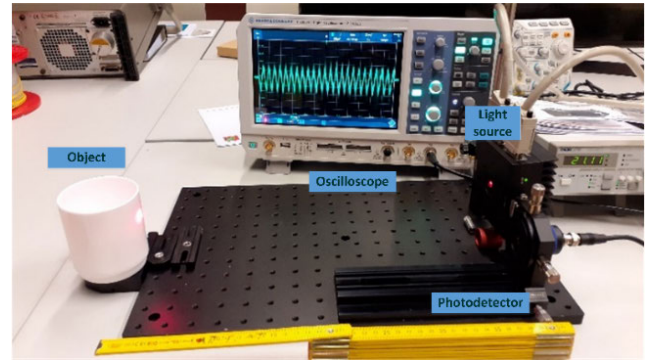

FIGURE 5. Experimental setup.

an oscilloscope. The acquired data are processed with a program in MATLAB, which computes the depth of the object and carries out the correction algorithm. The experimental setup is shown in Fig. 5.

The experiment was carried out in a completely dark room with a non-reflective screen placed behind the target object to avoid as much background illumination as possible. The experimental setup employs a fixed modulation and demodulation frequency, which cannot be increased in order to measure a larger depth range. The algorithm is scalable to any frequency in the NIR range used by commercial ToF cameras. To test the scalability of the method, measurement data from a camera sensor are essential, and unfortunately, access to raw data from commercial ToF cameras is not possible at present. This experimental setup was only used to validate the method proposed in this article.

A qualitative analysis of the performance of the MpI suppression algorithm is shown in Fig. 6. Plot (a) in Fig. 6 shows the received light signal reflected by a coffee mug placed at a distance of 400 mm from the ToF camera. The reflected light is affected by MpI and noise. Next to it, plot (b) shows the corrected output of the filter, i.e., the signal after MpI suppression, and plot (c) shows the convergence of the error during the adaptation process. Later in this section, a quantitative analysis of the algorithm is performed on different objects, such as a metal wire, a pencil lead, and an eraser placed at distances of 150 mm, 250 mm, and 350 mm away from the camera. The selection of the different objects was made by considering different materials and their reflective properties. The pencil lead exhibits high diffusion and low reflectivity, the metal wire has high reflectivity and exhibits specular reflections, and the eraser features Lambertian reflections. Since the experimental setup is very susceptible to noise, multiple measurements were taken and averaged to provide a more general representation of the analysis. For the depth measurement, four samples corresponding to the correlation function were captured at 0, $\pi/2$, π, and $3\pi/2$ for every period over multiple periods so that the measured depth was averaged out. The measured depths before and after correction for the different objects in different scenarios were compiled in the form of a table. Table 4 shows the depth estimations
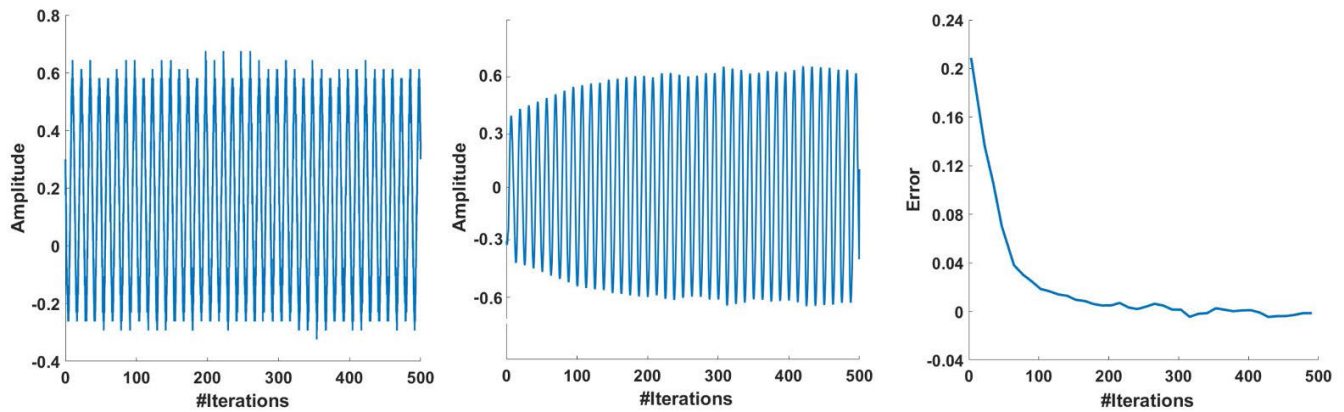
FIGURE 6. a) Attenuated and phase-shifted received light signal. b) Corrected output signal. c) Convergence of error during adaptation.

TABLE 4. Depth correction for a lead pencil.

Actual depth (mm)	Measured depth (mm)	Error before (in %)	Corrected depth (mm)	Error after (in %)
150	108.3	27.8	139.6	6.93
250	218	12.8	249.2	0.32
350	367.6	5.02	359.6	2.73

TABLE 5. Depth correction for a metal wire.

Actual depth (mm)	Measured depth (mm)	Error before (in %)	Corrected depth (mm)	Error after (in %)
150	162.9	8.6	160	6.66
250	206.7	17.3	242.3	3.08
350	355.0	1.42	348.8	0.34

TABLE 6. Depth correction for an eraser.

Actual depth (mm)	Measured depth (mm)	Error before (in %)	Corrected depth (mm)	Error after (in %)
150	201	34	172	14.6
250	296.8	18.4	240.1	3.96
350	376	7.42	352.2	0.62

before and after the MpI correction for a lead pencil placed at different distances from the camera. Tables 5 and 6 show similar comparisons for metal wire and an eraser.

During the experiments, it was observed that different objects consisting of different materials tended to be estimated at different depths when placed at the same distance from the camera. Besides, different lighting conditions and the presence of a non-reflective background also affect the measurements. A thorough analysis of the factors affecting the depth measurement of an object has been done by He *et al.* [29]. In addition, other typical systematic and non-systematic errors were not taken into consideration during the measurements in the experiments.

As a ToF camera is essentially employed to capture depth images of moving objects, the performance of the proposed method was also tested for objects in motion. For this scenario, the experiments were designed such that an object initially placed at a distance of 500 mm from the camera moved away from the camera continuously. The camera captured the depth of the object at each frame. A typical ToF camera runs at a rate of 30–40 frames per second [7]. As the object continuously changes its position relative to the camera and the surroundings, the amplitude and the phase of the direct-reflected light, as well as those of the echoes, also change. The algorithm must be able to minimize the error before the next frame. The accepted duration of convergence is around 25 milliseconds. It was observed that the algorithm was able to converge the error to a minimum within the timeframe. Table 7 lists the three time instances at which the object was captured at different depths. A metallic rail with micrometer precision was used to place the object at different precise positions from the camera (for ground truth). It was observed that the depth of the object was successfully corrected for each frame.

TABLE 7. Depth correction for a moving object.

Actual depth (mm)	Measured depth (mm)	Error before (in %)	Corrected depth (mm)	Error after (in %)
500	447.9	10.42	517.2	3.44
600	556.2	7.3	614.1	2.35
700	629.6	10.05	711.1	1.58

The filter coefficients reach their optimum values after the convergence, and the optimum coefficients are then used as the initial coefficients for the next optimization. Fig. 7 shows the convergence of the error for the consecutive time instances. It takes around 150 iterations for the algorithm to converge the error to a minimum.

Another scenario was considered where the target object remained in the same position, and another object suddenly appeared near it. Although the actual distance of the object in
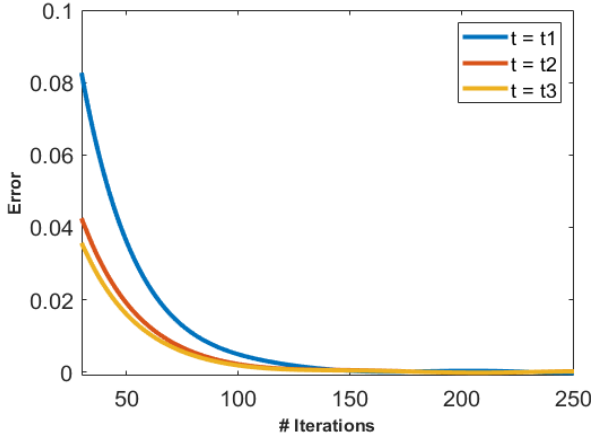
FIGURE 7. Error convergence for consecutive instances.

this case remained the same (500 mm), the echo component reaching the sensor changed suddenly. Consequently, due to the abrupt change in the received light, the measured depth was affected. It was observed that the algorithm was able to carry out the echo cancelation well inside the time period. The depths measured before and after the correction for both frames are shown in Table 8 below.

TABLE 8. Depth correction for a sudden change in scene.

ToF field of view	Measured depth (mm)	Error before (in %)	Corrected depth (mm)	Error after (in %)
One object	539.1	7.8	507.2	1.44
Two objects	554.5	10.9	507.7	1.54

Other Adaptive Algorithms: An appropriate selection of the step size such that the filter's coefficients do not diverge is a big challenge in filter design. If a small μ is chosen, the convergence will be accurate but very slow, and for a larger μ the weights might diverge. To overcome this difficulty, several methods have been developed over the years. One example is the normalized LMS (NLMS) algorithm, which computes the dynamic step size as

$$\mu(n) = \frac{\tilde{\mu}}{\delta + L \cdot \hat{P}_x(n)} \quad (21)$$

where the adaptive step-size $\mu(n)$ is normalized by filter length L and the signal power $\hat{P}_x$, and δ is a small constant to avoid division by zero. The NLMS algorithm optimizes the speed of convergence of the error and guarantees the convergence regardless of the statistics of the input signal [38]. The block-LMS (BLMS) algorithm converts the input signal into blocks that are applied to the adaptive filter one block at a time. The adaptation of the filter weights occurs on a block-by-block basis, as compared to the conventional LMS. The BLMS algorithm reduces the computational complexity and hence the runtime of the algorithm; however, this improvement generally does not infer faster convergence. The convergence rate of the BLMS is improved by combining

it with the NLMS algorithm and computing the adaptive step size at each iteration [38]. The updated equation for the tap-weight vector for the block-normalized LMS (BN-LMS) becomes

$$\overrightarrow{W}(k+1) = \overrightarrow{W}(k) - \mu(k) \cdot \Phi(k) \quad (22)$$

where $\mu(k)$ is the adaptive step size calculated from the NLMS formula and $\Phi(k)$ is the cross-correlation given as

$$\Phi(k) = \sum_{i=0}^{L_B - 1} X(kL_B + i) \cdot e(kL_B + i)$$

$$k : index\ of\ the\ block$$
$$L_B : Block\ size \quad (23)$$

One of the alternatives to the LMS algorithm is the recursive least squares (RLS) algorithm, which provides an excellent convergence rate but introduces increased complexity and computational cost. While the LMS filter adapts the coefficients until the difference between the desired and the actual signal is minimized, the RLS algorithm recursively finds the filter coefficients that minimize the weighted linear least squares. The signal paths and the adaptation are the same for both the LMS and the RLS filters [39].

The computational complexity of an algorithm is based on the total floating-point operations (additions and multiplications) involved in its implementation on a processor. For L filter coefficients, the complexity of the LMS (as well as the NLMS and BN-LMS) algorithm is of the order $O(L)$, while that of the RLS algorithm is of the order $O(L^2)$ [34]. The performance of the above-mentioned algorithms compared with the LMS algorithm on the basis of the root mean square error (RMSE) and the computation time is compared in Table 9 below. To compare the runtime of the algorithms, the code (implemented in MATLAB) was run on a Windows 10 computer equipped with an intel core i5 (7th Gen) processor with four cores, 8 GB of RAM, and a clock speed of 2.80 GHz. It can be seen from the table that the BN-LMS algorithm is much faster than the conventional LMS algorithm is, and the RLS algorithm, though improving the convergence, is much more time-consuming than the LMS algorithm is.

TABLE 9. Comparison of the different adaptive algorithms.

Parameter	LMS	NLMS	BN-LMS	RLS
# Iterations	150	150	150	150
Step size	0.01	0.01	0.01	λ= 0.999 δ = 0.5
RMSE	0.0054	0.0050	0.0036	0.0045
Computation Time (msec)	15.91	19.60	13.21	25.16

Comparison: The method proposed in this article successfully corrects the error in the depth measurement without

any physical supplementation to the ToF camera and without using multiple frequencies. The depth correction performance of the proposed method was compared with recent depth correction methods. The comparisons, based on the depth correction for an object at a distance of about 500 mm from the camera, are listed in Table 10 below.

TABLE 10. Comparison of depth correction with other methods.

Method	Error Correction	
Dashpute [20]	Error reduction of 30% (1.3 mm) at a distance of 500 mm	External element (polarizer) is added to the camera. Selection of polarization angle critical.
LiWu Li [40]	Error reduced to under 2% at a distance of 500 mm	Requires capture of multiple (25) depth pictures for DLL.
Islam [41]	Error reduced to around 4% at a distance of 250 mm	Only simulated results. Practical implementation may lead to more errors.
Naik [15]	Depth error reduced to about 5 mm at a distance of 25 mm (~20%) and about 13 mm at a distance of 110 mm	External projection required. Precise calibration necessary.
This work	Error reduced to about 2 mm (~3%) at 500 mm and to about 4% (avg.) at 250 mm	No external projections or multiple captures required.

VI. SUMMARY AND CONCLUSION

In this article, the authors presented a new approach to suppress the MpI from the received light signal in a ToF camera. The proposed method operates directly at the sensor side of the camera and can be employed with any ToF camera. Simulations demonstrated that the algorithm is able to correct the MpI-induced depth error under different scenarios. Furthermore, the algorithm was employed on a ToF camera setup, and experiments with real objects under real conditions were performed. To replicate the results in a commercial ToF camera, raw data from the camera sensor are required. Unfortunately, access to the raw data from the sensor interface of ToF cameras is not possible with commercially available ToF cameras. Future work will involve a more comprehensive implementation of the method suitable for commercially available real-time ToF cameras and the adaptations required in the physical structure of the cameras. The simulations, as well as the experiments, validate the method proposed in this article, although for the latter, there is still much room for improvement, which will definitely be overcome in future work.

REFERENCES

[1] M. H. Conde, "A material-sensing time-of-flight camera," *IEEE Sensors Lett.*, vol. 4, no. 7, Jul. 2020, Art. no. 7002804, doi: 10.1109/LSENS.2020.3005042.

[2] L. P. Pelaez, M. E. V. Recalde, E. D. M. Munoz, J. M. Larrauri, J. M. P. Rastelli, N. Druml, and B. Hillbrand, "Car parking assistance based on Time-or-Flight camera," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 1753–1759, doi: 10.1109/IVS.2019.8813820.

[3] K. V. Mack, W. D. Jemison, L. K. Rumbaugh, D. W. Illig, and M. K. Banavar, "Time-of-Flight (ToF) cameras for underwater situational awareness," in *Proc. OCEANS MTS/IEEE SEATTLE*, Oct. 2019, pp. 1–5, doi: 10.23919/OCEANS40490.2019.8962695.

[4] H. Yu, J. Zhu, Y. Wang, W. Jia, M. Sun, and Y. Tang, "Obstacle classification and 3D measurement in unstructured environments based on ToF cameras," *Sensors*, vol. 14, no. 6, pp. 10753–10782, Jun. 2014.

[5] A. M. Pinto, P. Costa, A. P. Moreira, L. F. Rocha, G. Veiga, and E. Moreira, "Evaluation of depth sensors for robotic applications," in *Proc. IEEE Int. Conf. Auto. Robot Syst. Competitions*, Apr. 2015, pp. 139–143.

[6] M. Kraft, M. Nowicki, R. Penne, A. Schmidt, and P. Skrzypczyáski, "Efficient RGB–D data processing for feature–based self–localization of mobile robots," *Int. J. Appl. Math. Comput. Sci.*, vol. 26, no. 1, pp. 63–79, Mar. 2016.

[7] Y. He and S. Chen, "Recent advances in 3D data acquisition and processing by Time-of-Flight camera," *IEEE Access*, vol. 7, pp. 12495–12510, 2019, doi: 10.1109/ACCESS.2019.2891693.

[8] H. Sarbolandi, M. Plack, and A. Kolb, "Pulse based time-of-flight range sensing," *Sensors*, vol. 18, no. 6, p. 1679, May 2018, doi: 10.3390/s18061679.

[9] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sensors J.*, vol. 11, no. 9, pp. 1917–1926, Sep. 2011, doi: 10.1109/JSEN.2010.2101060.

[10] C. Dal Mutto, P. Zanuttigh, and G. M. Cortelazzo, *Time-of-Flight Cameras and Microsoft Kinect*. Boston, MA, USA: Springer, 2012.

[11] R. Whyte, L. Streeter, M. J. Cree, and A. A. Dorrington, "Review of methods for resolving multipath interference in time-of-flight range cameras," in *Proc. IEEE SENSORS*, Valencia, Spain, Nov. 2014, pp. 629–632, doi: 10.1109/ICSENS.2014.6985077.

[12] D. Jimenez, D. Pizarro, M. Mazo, and S. Palazuelos, "Modelling and correction of multipath interference in time of flight cameras," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Providence, RI, USA, Jun. 2012, pp. 893–900, doi: 10.1109/CVPR.2012.6247763.

[13] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar, "Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization," *Opt. Lett.*, vol. 39, no. 6, p. 1705, Mar. 2014, doi: 10.1364/OL.39.001705.

[14] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin, "Phasor imaging," *ACM Trans. Graph.*, vol. 34, no. 5, pp. 1–18, Nov. 2015, doi: 10.1145/2735702.

[15] N. Naik, A. Kadambi, C. Rhemann, S. Izadi, R. Raskar, and S. B. Kang, "A light transport model for mitigating multipath interference in Time-of-flight sensors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 73–81, doi: 10.1109/CVPR.2015.7298602.

[16] E.-T. Baek, H.-J. Yang, S.-H. Kim, G. Lee, and H. Jeong, "Distance error correction in Time-of-Flight cameras using asynchronous integration time," *Sensors*, vol. 20, no. 4, p. 1156, Feb. 2020, doi: 10.3390/s20041156.

[17] D. Freedman, Y. Smolin, E. Krupka, I. Leichter, and M. Schmidt, "SRA: Fast removal of general multipath for ToF sensors," in *Computer Vision*, vol. 8689, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham, Switzerland: Springer, 2014, pp. 234–249.

[18] G. Agresti and P. Zanuttigh, "Combination of Spatially-Modulated ToF and Structured Light for MPI-Free Depth Estimation," in *Computer Vision*, vol. 11129, L. Leal-Taixé and S. Roth, Eds. Cham, Switzerland: Springer, 2019, pp. 355–371.

[19] S. S. Patil and V. S. Inamdar, "Multipath mitigation in time of flight 3D sensor based on direct and global seperation technique," in *Proc. Int. Conf. Comput., Analytics Secur. Trends (CAST)*, Pune, India, Dec. 2016, pp. 632–636, doi: 10.1109/CAST.2016.7915044.

[20] A. Dashpute, C. Anand, and M. Sarkar, "Depth resolution enhancement in time-of-flight cameras using polarization state of the reflected light," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 1, pp. 160–168, Jan. 2019, doi: 10.1109/TIM.2018.2838819.

[21] C. A. Lickfold, L. Streeter, M. J. Cree, and J. B. Scott, "Simultaneous phase and frequency stepping in time-of-flight range imaging," in *Proc. IEEE SENSORS*, Oct. 2019, pp. 1–4, doi: 10.1109/SENSORS43011.2019.8956800.

[22] X. Wang and Z. Tang, "Optional optimization algorithms for time-of-flight system," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 10, pp. 3326–3333, Oct. 2011.

[23] C. Cai and P. P. L. Regtien, "Accurate digital time-of-flight measurement using self-interference," *IEEE Trans. Instrum. Meas.*, vol. 42, no. 6, pp. 990–994, Dec. 1993, doi: 10.1109/19.245651.

[24] Z. Lu, "Estimating time-of-flight of multi-superimposed ultrasonic echo signal through envelope," in *Proc. Int. Conf. Comput. Intell. Commun. Netw.*, Nov. 2014, pp. 300–303, doi: 10.1109/CICN.2014.74.

[25] Y. Yang, "A signal theoretic approach for envelope analysis of real-valued signals," *IEEE Access*, vol. 5, pp. 5623–5630, 2017, doi: 10.1109/ACCESS.2017.2688467.

[26] S. O. Rice, "Envelopes of narrow-band signals," *Proc. IEEE*, vol. 70, no. 7, pp. 692–699, Jul. 1982.

[27] R. Lange, "3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. Siegen, Siegen, Germany, 2000.

[28] M. Hansard, S. Lee, O. Choi, and R. Horaud, *Time-of-Flight Cameras: Principles, Methods and Applications*. London, U.K.: Springer, 2013.

[29] Y. He, B. Liang, Y. Zou, J. He, and J. Yang, "Depth errors analysis and correction for Time-of-Flight (ToF) cameras," *Sensors*, vol. 17, no. 1, p. 92, Jan. 2017, doi: 10.3390/s17010092.

[30] M. H. Conde, K. Hartmann, and O. Loffeld, "Adaptive high dynamic range for Time-of-Flight cameras," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 7, pp. 1885–1906, Jul. 2015.

[31] G. Zach and H. Zimmermann, "A 2×32 range-finding sensor array with pixel-inherent suppression of ambient light up to 120 klx," *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Dec. 2009, pp. 352–353.

[32] M. Davidovic, M. Hofbauer, K. Schneider-Hornstein, and H. Zimmermann, "High dynamic range background light suppression for a ToF distance measurement sensor in 180nm CMOS," *IEEE SENSORS*, Oct. 2011, pp. 359–362.

[33] L. Streeter, "Time-of-flight range image measurement in the presence of transverse motion using the Kalman filter," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 7, pp. 1573–1578, Jul. 2018.

[34] V. Madisetti and D. B. Williams, *The Digital Signal Processing Handbook*. Boca Raton, FL, USA: CRC Press, 1998.

[35] J.-K. Hwang and Y.-P. Li, "LMS algorithm with an adaptive neural network cost function," *IEEE Sensors Lett.*, vol. 4, no. 7, Oct. 2009, Art. no. 7002804.

[36] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1985.

[37] Y. Gu, K. Tang, and H. Cui, "LMS algorithm with gradient descent filter length," *IEEE Signal Process. Lett.*, vol. 11, no. 3, pp. 305–307, Mar. 2004, doi: 10.1109/LSP.2003.822892.

[38] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[39] P. S. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. Norwell, MA, USA: Kluwer, 2002.

[40] L. Li, H. Liu, Y. Xu, and Y. Zheng, "Measurement linearity and accuracy optimization for Time-of-Flight range imaging cameras," in *Proc. IEEE 4th Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Jun. 2020, pp. 520–524, doi: 10.1109/ITNEC48623.2020.9084901.

[41] A. Islam, M. A. Hossain, and Y. Min Jang, "Interference mitigation technique for Time-of-Flight (ToF) camera," in *Proc. 8th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2016, pp. 134–139, doi: 10.1109/ICUFN.2016.7537001.

AYUSHYA PARE was born in Bhopal, India, in 1989. He received the bachelor's degree in electronics and communication engineering from Rajiv Gandhi Technical University, in 2011, and the M.Sc. degree in electrical and electronics engineering from the University of Duisburg-Essen, Germany, in 2017. He has worked as a System Engineer with Tata Consultancy Services, Pune, India, from 2011 to 2014. Since 2017, he has been working as a Scientific Assistant with the Institute of Measurement and Sensor Technology, University of Applied Sciences Ruhr West, Germany. His research interests include ToF imaging, signal processing, and deep learning.

SHUFANG ZHANG (Member, IEEE) received the B.S. degree in mechanical engineering from the Taiyuan University of Technology, Taiyuan, China, in 1996, and the M.S. and Ph.D. degrees in mechanical and electronic engineering as well as signal and information processing major from Tianjin University, Tianjin, China, in 2004 and 2007, respectively. She joined the School of Electrical and Information Engineering, Tianjin University, where she is currently an Associate Professor. Her research interests include digital image and video processing, deep learning, and video quality assessment.

ZHICHUN LEI was born in Qian'an, China, in 1964. He received the B.S. degree in communication technology and the M.S. degree in communication and electronic system from Tianjin University, China, in 1986 and 1989, respectively, and the Ph.D. degree in electronic engineering and information technology from the University of Dortmund, Germany, in 1998. From 1989 to 1993, he was a Scientific Assistant with the Department of Electronic Engineering, Tianjin University. From 1994 to 1998, he was a Scientific Assistant with the Circuits and Systems Laboratory, Chair for Communication Technology, University of Dortmund. From 1998 to 2001, he was a Scientific Assistant with the Philips Research Laboratory, Aachen, Germany. From 2001 to 2011, he was an Engineer and a Senior Engineer with the Sony European Technology Centre, Stuttgart, Germany. Since 2011, he has been a Professor with the Institute of Measurement Engineering and Sensor Technology, University of Applied Sciences Ruhr West, Germany. Since 2014, he has held a 2nd Professorship at the School of Microelectronics, Tianjin University. His research interests include high-dynamic range and wide color gamut.

• • •