Article

Methodology for Large-Scale Camera Positioning to Enable Intelligent Self-Configuration

Yingfeng Wu, Weiwei Zhao * and Jifa Zhang

School of Mechanical and Electronic Engineering, Wuhan University of Technology, 122 Luoshi Road, Wuhan 430070, China; wyf1997@whut.edu.cn (Y.W.); zjf94131@whut.edu.cn (J.Z.)
* Correspondence: wzhao@whut.edu.cn

Abstract: The development of a self-configuring method for efficiently locating moving targets indoors could enable extraordinary advances in the control of industrial automatic production equipment. Being interactively connected, cameras that constitute a network represent a promising visual system for wireless positioning, with the ultimate goal of replacing or enhancing conventional sensors. Developing a highly efficient algorithm for collaborating cameras in the network is of particular interest. This paper presents an intelligent positioning system, which is capable of integrating visual information, obtained by large quantities of cameras, through self-configuration. The use of the extended Kalman filter predicts the position, velocity, acceleration and jerk (the third derivative of position) in the moving target. As a result, the camera-network-based visual positioning system is capable of locating a moving target with high precision: relative errors for positional parameters are all smaller than 10%; relative errors for linear velocities (v_x, v_y) are also kept to an acceptable level, i.e., lower than 20%. This presents the outstanding potential of this visual positioning system to assist in the industry of automation, including wireless intelligent control, high-precision indoor positioning, and navigation.

Keywords: large-scale positioning and navigation; intelligent self-configuration; collaborative visual network; extended Kalman filter

1. Introduction

Large-scale camera positioning systems, also known as visual positioning systems, are used for precisely locating or tracking a moving object in three-dimensional space. Compared with other positioning technologies, e.g., laser-tracking, theodolite, RFID, geomagnetic field, or indoor GPS (iGPS), visual positioning systems have received intense research interest because of advantages such as low cost, high precision, and contactless tracking functions [1–3].

Therefore, they have wide application prospects in various industrial areas, such as motion analysis, robot guidance, or aircraft manufacturing [4–7]. According to the literature, visual positioning systems can be categorized into two types, as defined by the number of cameras: single pan-tilt-zoom (PTZ) camera, and camera network [8]. Due to the superiorities of locating accuracy, field of view, and response speed, positioning systems based on camera networks are regarded as a more promising system type than those based on a single PTZ camera [9].

Positioning systems based on camera networks are able to adjust the visual scale, by choosing appropriate numbers of cameras, to meet various production requirements. Galetto et al. presented a visual positioning system called MScMS-II, which was able to track two targets in a cubic space with length of 2 m [10]. The coordinates of the targets could be calculated using data acquired from at least two cameras. Dixon et al. successfully tracked vehicles in an urban environment by setting up seven cameras [11]. A system using the interaction of 30 cameras was set up by Wei et al. for monitoring an area of 1670 m^2 [12].

Approximately 100 stationary cameras were integrated by Kuo et al. [13], enabling a high degree of collaborative working among the cameras, for campus monitoring. Jiarui Lin et al. presented a ceiling-mounted workshop measurement positioning system (C-wMPS), to compensate for many deficiencies shown by conventional metrology systems [14]. Pablo Puerto et al. designed a methodology to evaluate the measurement uncertainty of portable photogrammetry for LVM [15]. Domenico Augusto Maisano et al. presented some diagnostic tests for combinations of LVM systems that are equipped with distance and/or angular sensors [16]. Yongkang Lu et al. developed a novel nonlinear optimization method to reduce the overall calibration error of the LRS [3]. Apparently, then, the number of cameras is varied according to the corresponding application requirements. Effective management of the cameras is able to improve the locating accuracy, increase the responding speed, and minimize the energy consumption. Soro et al. developed two methods for choosing the cameras' visual area, aiming to avoid redundant data and reduce the processing time [17]. The research group of Liu et al. introduced a dynamic collaborative scheme for tracking moving targets, based on wireless camera networks [18]: sequential Monte Carlo (SMC) method, i.e., particle filtering, was utilized for calculating the probability of detecting a target and assigning the camera with the highest probability as the chief camera from the cluster. Another particle filter, based on color, was investigated by Liang et al. for switching cameras within the network [9]. They achieved a significant reduction in redundant video data, and enabled the simultaneous location of five targets, in real time, using an ordinary personal computer.

We proposed an update to the system, by software synchronization method, and formulated the solution for the synchronization of cameras. The mean value of standard deviation of eight cameras, mounted on two workstations, was 12.53 ms, the localization performance of LSVLS was enhanced [19]. We optimized the cameras' placement using a relative positioning algorithm (RPA). The result was that optimal camera placement greatly enhances the efficiency of camera placement in LSVLS, as verified with a field-winding mobile vehicle model [20].

In this work, a visual positioning system, based on a camera network consisting of 16 cameras, was built for locating a filament winder. A novel strategy for: selecting cameras; assigning a chief camera; and conducting collaborative data transfer, which enables intelligent self-configuration of the camera network; was developed and described in detail. The feasibility and reliability of the positioning system, in terms of the locating method and predicting algorithm, was experimentally demonstrated, i.e., prediction of the position and velocity of the filament winder.

2. Methodology

2.1. Positioning System

A filament winder was recognized as the moving target in this work. A mandrel with a diameter of 5 m was located at the center of the working space (see Figure 1a). The filament winder moved around the mandrel, as the default trajectory, with a constant speed. The circular motion of the filament winder was attended by a slight ascending motion of the mandrel, with the aim of avoiding overlap of the fiber during winding. As is illustrated in Figure 1a, 16 cameras were installed above the working space, directly above the circular trace of the filament winder. The visual positioning system was thereby established, with these 16 cameras, to navigate the filament winder while the winding operation took place. The location of the cameras in the visual positioning system, in this work, were designed following guidelines from the literature [19,20].
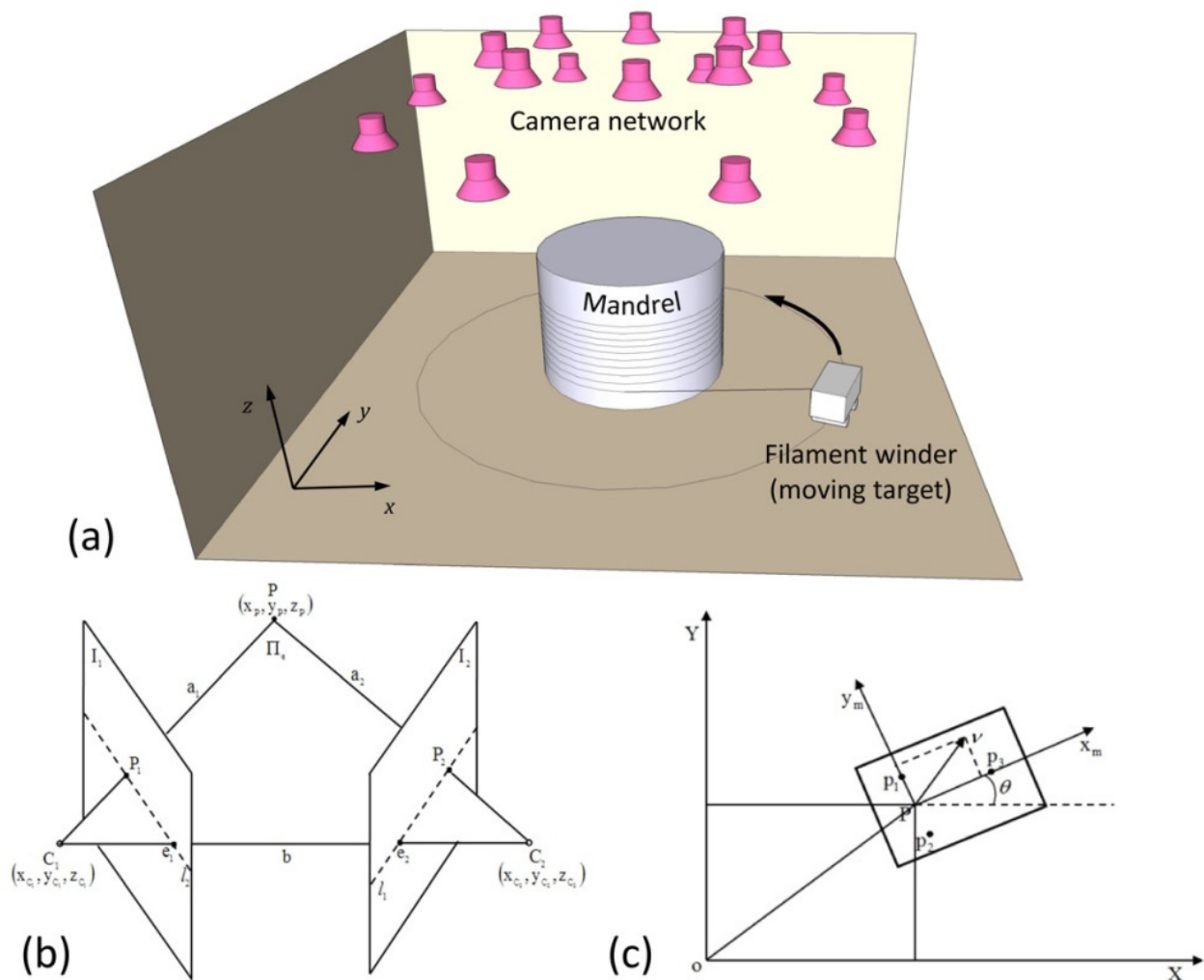
Figure 1. (**a**) Schematic diagram for the experimental setup in this work. A total of 16 cameras constitute a camera network for the visual positioning system. (**b**) Graphic illustration of the principle of stereo vision. (**c**) Method of locating the moving target (filament winder).

A commercial laptop (Lenovo notebook, equipped with 1.9 GHz CPU and 2 GB memory) was used as the hardware for data-processing in the visual positioning system. Cameras used in this work, which constitute the camera network, were infrared cameras from Nintendo Co., Ltd., Kyoto, Japan (also known as Wiimote, IR camera), with 1024×768 pixels of interpolated resolution, $45° \times 30°$ field of vision, and 100 Hz sampling rate. The real-time connection between the laptop and the cameras was enabled by long-distance Bluetooth modules. The remotely controllable filament winder was wirelessly connected with the laptop, and navigated via instructions from the laptop.

2.1.1. Stereo Vision

The function attribute of the positioning of the moving target was stereo vision, which was obtained by at least one camera. Figure 1b illustrates a simplified model for establishing stereo vision on one target (point P) by two cameras (C_1 and C_2). As shown in Figure 1b, the images of the target acquired by the cameras are marked as P_1 on plane I_1, and P_2 on plane I_2, corresponding to the two cameras C_1 and C_2, respectively. Thus, the coordinates of P_1 and P_2, i.e., x_1 or x_2, are given by [21]:

$$x = CX \tag{1}$$

where X denotes the coordinates of target P, C is the projective matrix of cameras.

Therefore, the bunch of divergence lines passing through the center of the cameras and the target area can be expressed in Equation (2):

$$X(\mu) = \mu \begin{pmatrix} M^{-1}x \\ 0 \end{pmatrix} + \begin{pmatrix} -M^{-1}p4 \\ 1 \end{pmatrix} = \begin{pmatrix} M^{-1}(\mu x - p4) \\ 1 \end{pmatrix} \tag{2}$$

where, $M = C$ (1:3, 1:3), $p4 = C$ (:,4). $-M^{-1}p4$ denotes the camera center, and $M^{-1}x$ gives the direction of the divergence lines. a_1 and a_2 connect the target and the cameras, which stand for the central visual paths. By intersecting these central visual paths with the image planes (I_1 and I_2), the coordinate of the target P can, thereby, be calculated, using Ray 3D in robotics and machine vision toolboxes [22].

2.1.2. Methodology of Locating the Moving Target (the Filament Winder)

According to the experimental setup in this work, the filament winder, as the tracking target, moved around the mandrel, while the mandrel lifted slightly, with a constant speed, to complete the winding process. Thus, the circular motion of the filament winder was the primary dynamic motion, which had to be navigated. Based on conventional locating techniques reported in the literature [23,24], the coordinates (x, y, z) and orientation (θ) of the moving target in this work can be described as:

$$q = [x \; y \; z \; \theta]^T \tag{3}$$

As is illustrated in Figure 1c, three feature points on the moving target were chosen and marked as p_1, p_2, and p_3. Point P, shown in Figure 1b, is the midpoint of p_1 and p_2, having a location point at the rear of the filament winder. Point p_3 is located on the front of the filament winder. The line connecting P and p_3 is vertical to the line p_1 to p_2. The angle between the line P-p_3 and the X axis denotes the orientation θ of the moving target. Because the coordinates of p_1, p_2, and p_3 can be precisely acquired by the camera-network-based visual positioning system, the position and orientation of the moving target can be calculated by using Equations (4) and (5) as follows:

$$P = \frac{1}{2}(p_1 + p_2) \tag{4}$$

$$\theta = arctan2\left(y_{p3} - y_P, \; x_{p3} - x_P\right) \tag{5}$$

2.2. Data Processing Using Extended Kalman Filter

Kalman filter is known as a linear, discrete time, finite-dimensional time-varying system, used to evaluate estimates for the state variables and minimize the mean-square error [24]. A non-optimal approach to solving the problems with non-Gaussian functions, in the frame of linear filters, is the extended Kalman filter (EKF). The EKF implements a Kalman filter for system dynamics that results from the linearization of the original nonlinear filter dynamics around the previous state-variable estimates [25–27]. In this work, EKF was conducted to cope with the prediction of velocity, acceleration, and jerk (the third derivative of position) in the moving target.

According to the Equation (3), the velocity can be derived by taking the derivative with respect to the position q, which is given by $v_x = \dot{x}$, $v_y = \dot{y}$, $w = \dot{\theta}$. Similarly, the acceleration and jerk can be obtained by taking two or three times the derivative to the position q, respectively. Thus, the expressions for velocity, acceleration, and jerk of the moving target are listed in Equation (6):

$$\begin{cases} \dot{q} = [v_x \; v_y \; w]^T \\ \ddot{q} = [\dot{v}_x \; \dot{v}_y \; \dot{w}]^T \\ \dddot{q} = [\ddot{v}_x \; \ddot{v}_y \; \ddot{w}]^T \end{cases} \tag{6}$$

These differential operations, applied to the position, require the discretization of the moving target model when conducting the EKF calculation. The discretization of the model in terms of position, velocity, acceleration, and jerk is listed in Equation (7):

$$
\begin{cases}
x_k = x_{k-1} + Ts * v_{x(k-1)} \\
y_k = y_{k-1} + Ts * v_{y(k-1)} \\
z_k = z_{k-1} + Ts * v_{z(k-1)} \\
\theta_k = \theta_{k-1} + Ts * \dot{\theta}_{(k-1)} \\
v_{xk} = v_{x(k-1)} + Ts * \dot{v}_{x(k-1)} \\
v_{yk} = v_{y(k-1)} + Ts * \dot{v}_{y(k-1)} \\
v_{zk} = v_{z(k-1)} + Ts * \dot{v}_{z(k-1)} \\
\dot{\theta}_k = \dot{\theta}_{k-1} + Ts * \ddot{\theta}_{(k-1)} \\
\dot{v}_{xk} = \dot{v}_{x(k-1)} + Ts * \ddot{v}_{x(k-1)} \\
\dot{v}_{yk} = \dot{v}_{y(k-1)} + Ts * \ddot{v}_{y(k-1)} \\
\dot{v}_{zk} = \dot{v}_{z(k-1)} + Ts * \ddot{v}_{z(k-1)} \\
\ddot{\theta}_k = \ddot{\theta}_{k-1} + Ts * \dddot{\theta}_{(k-1)} \\
\ddot{v}_{xk} = \ddot{v}_{x(k-1)} \\
\ddot{v}_{yk} = \ddot{v}_{y(k-1)} \\
\ddot{v}_{zk} = \ddot{v}_{z(k-1)} \\
\dddot{\theta}_k = \dddot{\theta}_{(k-1)}
\end{cases}
\tag{7}
$$

In prior to the EKF calculation, measured input $\widetilde{y}_k$, and the variable to be solved in terms of state vector x_k, have been pre-described and given by:

$$
x_k = \left[\hat{x} \ \hat{y} \ \hat{z} \ \hat{\theta} \ \hat{v}_x \ \hat{v}_y \ \hat{v}_z \ \hat{\dot{\theta}} \right]
\tag{8}
$$

$$
\widetilde{y}_k = [p_1 \ p_2 \ p_3]
\tag{9}
$$

When conducting the EKF calculation, two Jacobians (*F* and *H*) are required for the predictor-corrector algorithm. These two Jacobians are defined by:

$$
F = \begin{bmatrix}
I_{4\times4} & Ts \cdot I_{4\times4} & 0_{4\times4} & 0_{4\times4} \\
0_{4\times4} & 0_{4\times4} & Ts \cdot I_{4\times4} & 0_{4\times4} \\
0_{4\times4} & 0_{4\times4} & 0_{4\times4} & Ts \cdot I_{4\times4} \\
0_{4\times4} & 0_{4\times4} & 0_{4\times4} & I_{4\times4}
\end{bmatrix}
\tag{10}
$$

$$
H = \begin{bmatrix} I_{4\times4} & 0_{4\times12} \end{bmatrix}
\tag{11}
$$

In the extended-Kalman filter, mathematically, the predictor step is given by:

$$
\hat{x}_k^- = f(\hat{x}_{k-1}, u_k, k)
$$

$$
P_k^- = F_{k-1} P_{k-1} F_{k-1}^T + Q_k
$$

$$
K_k = P_k^- K_k^T \left(H_k P_k^- H_k^T + R_k \right)^{-1}
\tag{12}
$$

$$
\hat{x}_k = \hat{x}_k^- + K_k \left(\widetilde{y}_k - h(\hat{x}_k^-, u_k, k) \right)
$$

$$
P_k = (I - K_k H_k) P_k^-
$$

In the above equations P_k is an estimate of the covariance of the measurement error and K_k is called the Kalman gain. After both the prediction and correction steps have been performed, then $\hat{x}_k$ is the current estimate of the states and $\hat{y}_k$ can be calculated directly from it.

3. Experimental Approaches

3.1. Calibration of Cameras

According to the fundamental research of our group [19,20], a relative algorithm, based on integer linear programming, was proposed. This algorithm was utilized in this work to optimize the placement of the cameras, in terms of determining the optimal position for the cameras individually, with the aim of saving processing time, reducing the computational consumption, and guaranteeing the reliability of the positioning system in controlling the filament winder. The algorithm was inserted and programmed in the Multi-Camera Self-Calibration Toolbox [28] for practical initiation of the cameras in this work. As illustrated in Figure 2a, the position of the cameras in the working space was quantitatively assigned using three parameters: distance of camera *C* to the origin of the reference coordinate *O* (*r*); the angle between the *X* axis and the projection line *OC'* (*φ*); and the angle between the *Z* axis and line *OC* (*θ*). Hence, the calibration parameters for cameras, i.e., position and orientation, were precisely determined and imported into the visual positioning system, thereby achieving the optimal initial boundary conditions. The optimum organization of the 16 cameras is demonstrated in Figure 2b. Because of the circular default trajectory of the filament winder, the cameras were organized as two circles around the center of the mandrel. Each circle contained 8 cameras, in order to guarantee that at least two cameras were able to locate the filament winder at any point in the designated trajectory.
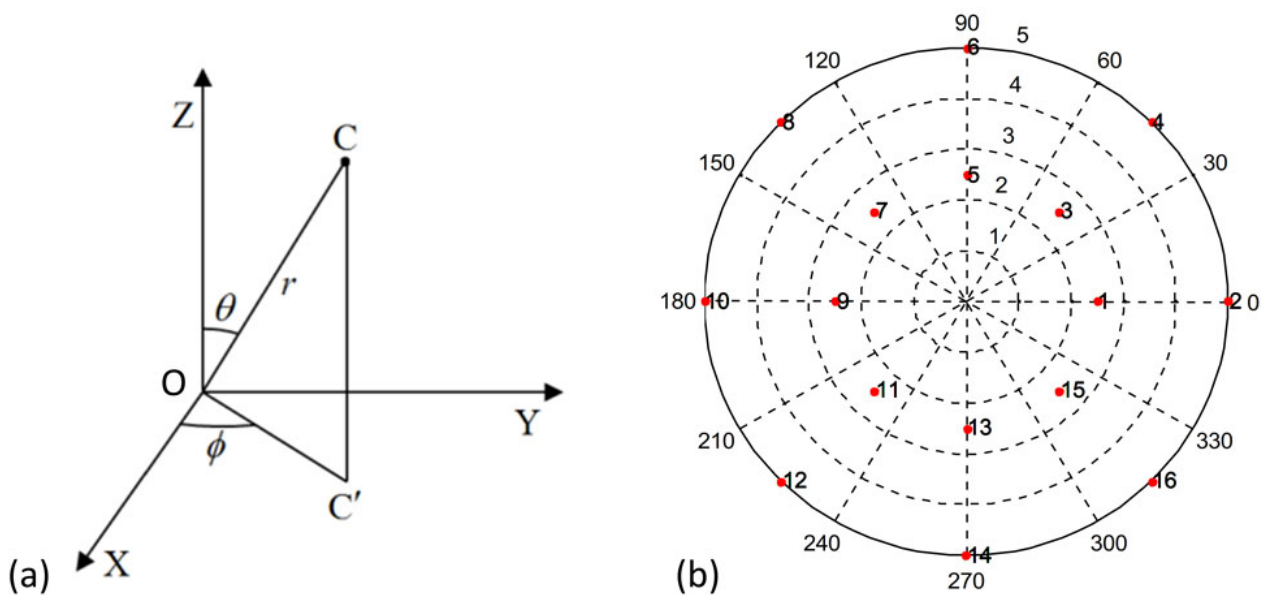

Figure 2. (**a**) The location of Camera *C* can be described by three parameters *r*, *φ*, *θ* in three-dimensional space. (**b**) The optimal arrangement of the 16 cameras for the current investigation. Cameras were arranged as two circles surrounding the center of the mandrel, with even distribution. Each circle consisted of 8 cameras.

3.2. Algorithm for Automatically Selecting Cameras

The task for the positioning system, following the calibration of cameras, was assigning the cameras that could detect and locate the target, initially. It required the system to be able to conduct intelligent self-configuration, in terms of automatically identifying the appropriate camera(s), according to the location of the target, while it was moving. The criterion of activating the camera was that the target appeared in the field of vision of the camera. Figure 3a schematically illustrates a target (*P*), six cameras and their corresponding field of visions. If the current location of the target was not in the camera's field of vision, the camera was in its standby mode, such as the cameras numbered 1, 2, 15, and 16 in Figure 3a. If the current location of the target was in the rectangular field of vision, the

camera was activated and in its operative mode. For example, the two cameras, numbered 13 and 14, shown in Figure 3a:
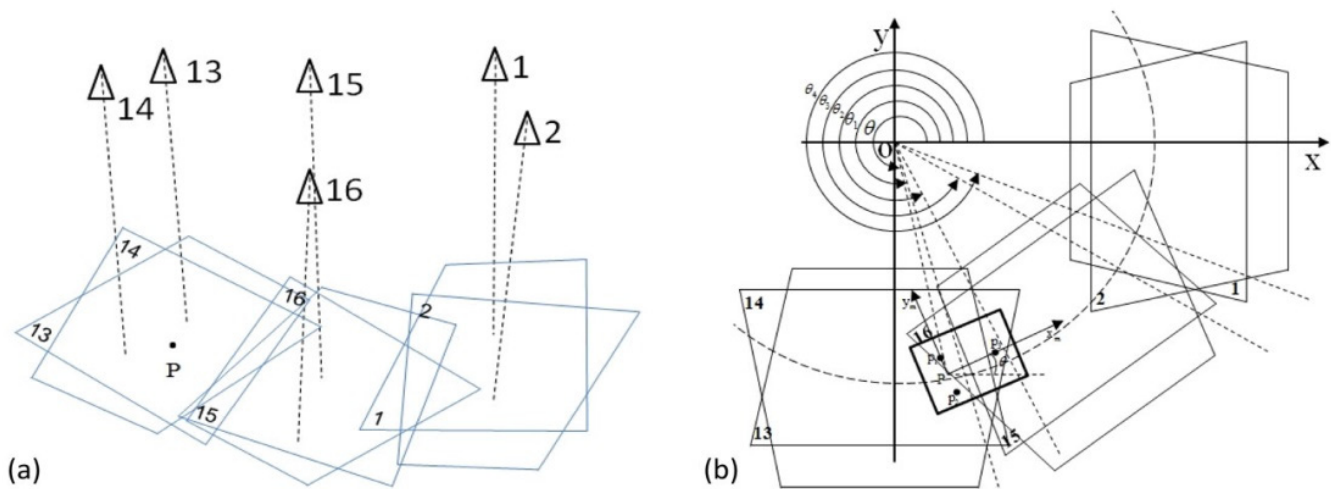

Figure 3. (**a**) Schematic diagram of the rectangular field of visions, which are projected on the horizontal plane (*XY* plane), corresponding to the cameras with numbers 1, 2, 13, 14, 15, 16. (**b**) The moment of moving target (*p*) leaving the rectangular fields of vision of cameras 13 and 14, and entering into the fields of vision of cameras 15 and 16. The parameter *θ*, the angle between *X* axis and the line *OP*, is the numerical criterion for selecting appropriate camera(s) according to the location of the moving target.

Practically, the field of vision projected on the work plane (*XY* plane) was a rectangular area, because of the geometry of the CCD (charge-coupled device)-sensors matrix of cameras. This rectangular visual region was demarcated by four feature points, i.e., four pixels on the corners of the CCD matrix. Their coordinates were set to (0, 0), (0, 768), (1024, 768), and (1024, 0).

Aiming to ensure that every point in the default trajectory for the moving target was surveilled by at least two cameras, overlap of the rectangular visual regions was set up during switching cameras, when the target was moving around the boundaries of the visual areas. Figure 3b illustrates the situation whereby the target (*P*) moved from the region covered by cameras 13 and 14 to the region covered by cameras 15 and 16. In this transitive approach, it took a little time to switch cameras and process data, requiring the region to be visually covered by at least four cameras, including the two current cameras and two cameras which were to be activated next. To precisely determine the appropriate moment to proceed with the camera switching approach, the parameter *θ*, the angle between *X* axis and the line *OP* (see Figure 3b), was adopted in this work, and acted as the numerical criterion for selecting appropriate camera(s), according to the location of the moving target. To carry out such camera-switching processes in this work, an EKF predictor was utilized. It was able to identify the camera number within the visual region from which the target was about to leave, or into which the target was about to enter. The detailed structure of the EKF predictor was built up in Matlab/Simulink. A model was performed in Matlab, according to velocity (v_x, v_y) predicted by EKF, as in Figure 4:
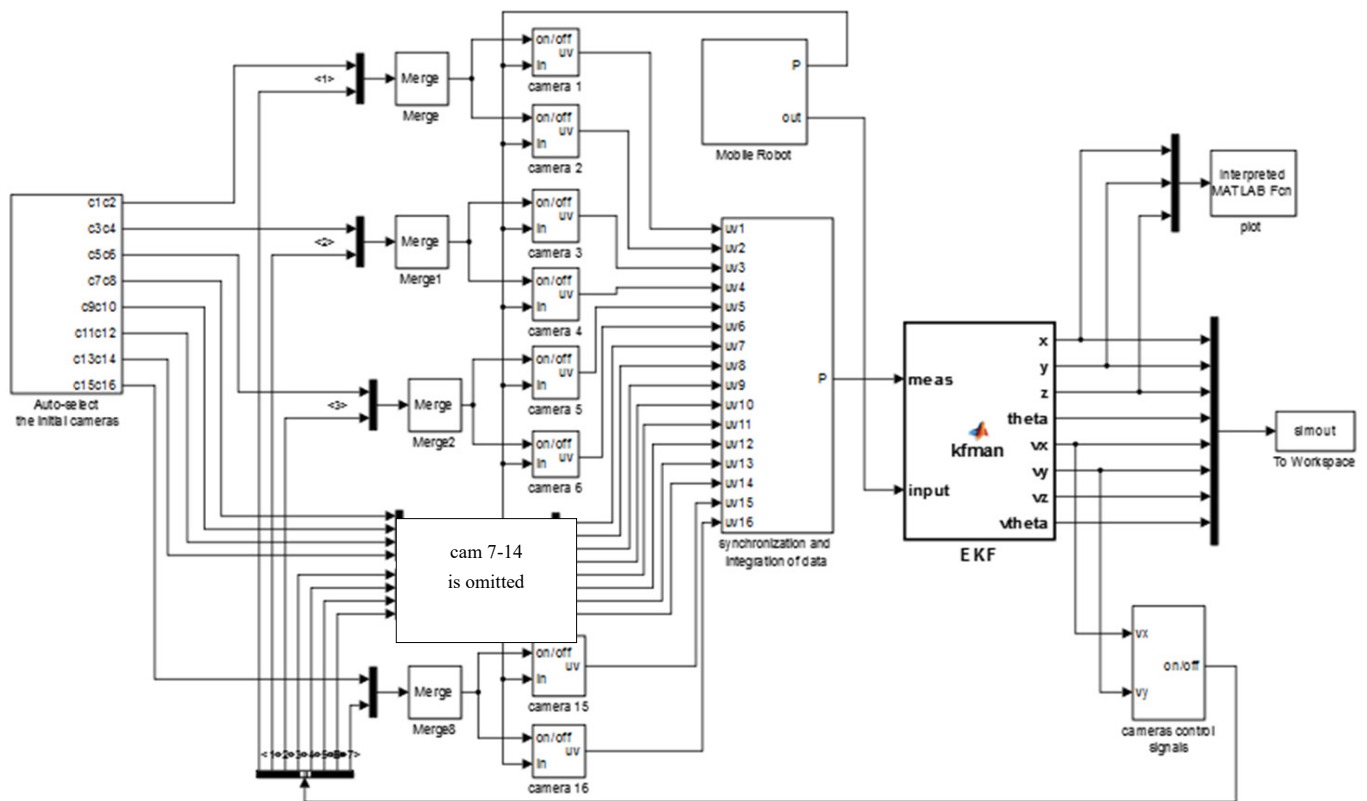
Figure 4. A total of 16 cameras provide the coordinates, in pixels, to subsequent modules. Synchronization and integration of data integrates the datum from cameras, calculates the coordinates of p_1, p_2, p_3, and calculates the position and orientation of the mobile robot (P). The EKF estimates the position and velocity ($x, y, z, \theta, v_x, v_y, v_z, \dot{\theta}$) of the mobile robot. According to v_x, v_y, the camera control signals model the output signals, to let cameras work (or sleep).

In the EKF model, the noise of the input and estimated output were pre-set, as follows:
R = diag([0.005^2 0.005^2 0.005^2 0.0005^2]);
Q = diag([0.01 0.01 0.01 0.001 0.01 0.01 0.01 0.001 0.01 0.01 0.01 0.001 0.01 0.01 0.01 0.001]);
The algorithm code for automatically selecting the cameras' output-controlling signals is listed below:
if theta<theta1
then c13,c14 working, the other sleeping
if theta >= theta1&theta <theta2
then c13,c14,c15,c16 working, the other sleeping
if theta >= theta2&theta <theta3
then c15,c16 working, c13,c14 and the other sleeping
if theta >= theta3&theta <theta4
then c15,c16, c1,c2 working, the other sleeping
. . . . . . etc.

3.3. Self-Configuration for Camera Network

The visual positioning system developed in this work is capable of tracking, as well as controlling, the moving target, thereby enabling intelligent navigation of the target. The procedures for understanding the internal logical steps are illustrated in Figure 5, as a flow chart. The procedures can be divided into three modules, i.e., system initialization, intelligent self-configuration of the camera network, and a visual servo system for tracking and controlling the moving target.
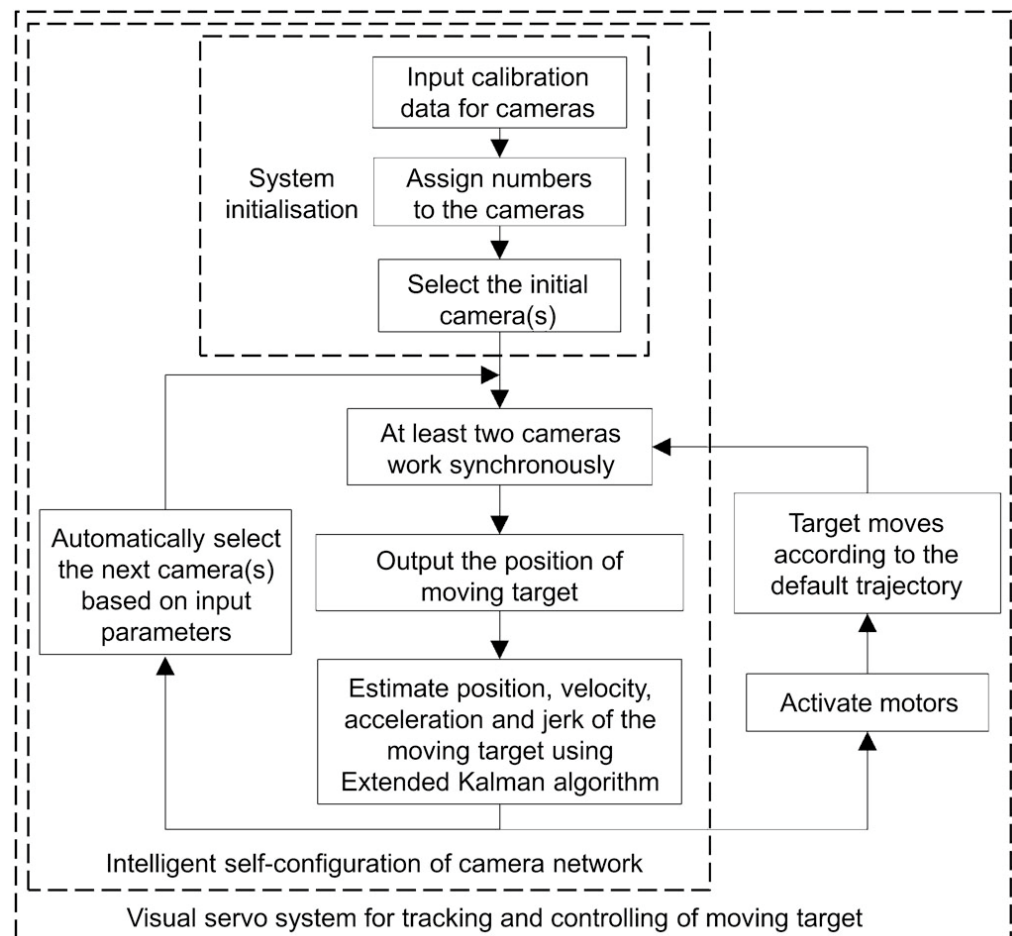
Figure 5. Flow chart of the logical procedures enabling intelligent navigation of the target.

The system initialization involves three tasks: calibrating the cameras, individually assigning numbers to the cameras, and determining the activation of the cameras, based on the initial position of the target. The second module, which is the core of the visual positioning system, enables the intelligent self-configuration. The function of this module is primarily based on the application of the extended Kalman Filter (EKF) algorithm. The procedure for automatically selecting cameras in this module can be summarized by the following step-by-step instructions: (1) first, cameras provide the acquired visual information about the moving target, in terms of pixels, to the subsequent modules; (2) synchronization and integration of data from cameras are then conducted, to calculate the coordinates of point p_1, p_2, and p_3, thereby locating the current position and orientation of the moving target, i.e., the filament winder, in this work; (3) the EKF operation is then used to estimate the velocity, acceleration, and jerk of the moving target; (4) predictions for the route of the moving target can then be achieved, based on the information given for the velocity, acceleration, and jerk, thereby determining the status of the cameras, from one of three action states: be activated, remain unchanged, or get deactivated in the next step. This critical signal is output from the EKF and represents the feedback loop for controlling the cameras. The process loop then starts again from step 1.

Following the EKF calculation, the output signals for both selecting the cameras and controlling the moving target can be obtained from the system. Such output controlling signals consisting of current velocity (speed and direction), acceleration and steering angle, ensure that the system has full control of the moving target and is capable of navigating the target so that it moves exactly following the designated route.

4. Results

4.1. Experimental Observations

The trajectory of the fibers was digitalized and displayed simultaneously with the experimental winding process, which was done by the filament winder. The digitalized trajectory is illustrated in Figure 6a. The trajectory is observed as a helical three-dimensional curve with uniform diameter of 4 m. The distance between two circles in the trajectory is entirely dependent on the rising speed of the mandrel when the filament winder moves along the circular route. The dynamic coefficients of the filament winder, in terms of position and velocity, were obtained as experimental observation data, in real time. Acceleration and jerk of the filament winder were then worked out and output, by the module of the EKF predictor, as the numerical calculated data.
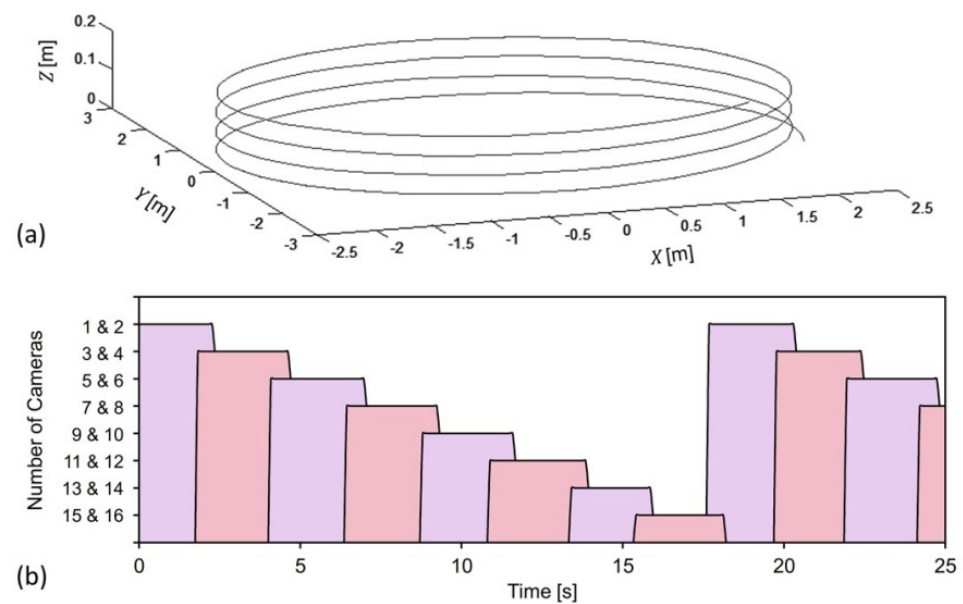

Figure 6. (**a**) The trajectory of fibers on the mandrel, which is constructed by the filament winder in this work. (**b**) Experimentally obtained time schedule for cameras in their alternate operation mode.

As introduced in the methodology section, to guarantee the locating accuracy in the moving target, the filament winder, at any position in the designated route, should be covered by at least two cameras. According to the arrangement of the cameras and their corresponding assigned numbers, as shown in Figure 2b, the cameras can be grouped with those nearest. For instance, cameras no.1 and no. 2 are always in synchronized working status, because these two cameras are in the same radius of the circular working area (see Figure 2b).

The time schedule of the 16 cameras, when positioning the filament winder, is displayed in Figure 6b. Such a time schedule, which represents the intelligent self-configuration of the camera network, is entirely dependent upon the speed and the designated route of the filament winder. As is displayed in Figure 6b, grouped cameras are activated and deactivated in turn. Cameras no. 1 and 2 were first activated, and deactivated in the time of 2.5 s, approximately. Before the deactivation of cameras no. 1 and 2, cameras no. 3 and 4 were activated in the time of 2 s, approximately. Evidently, there were four cameras in their working status during the time of 2 to 2.5 s. This corresponds to the transitive approach of the cameras, i.e., the camera switching progress, and indicates that the filament winder is concurrently in the overlap area of the rectangular visual regions of cameras no. 1, 2, 3, and 4, during the time of 2~2.5 s. Because the speed of the filament winder is constant, the length of the activated time for all groups of cameras is the same, at the value of approximately 3 s. After the activation of cameras no. 15 and 16, cameras no. 1 and 2 were activated again, at the time of 17.5 s. This suggests that the filament winder had already completed a full circular trajectory, and retuned back to the place where the

designated route started. Thus, the period of the movement can be read as taking17.94 s for the filament winder to go all the way around the mandrel. Duplications of such moving periods, round and round, enable continuous winding of the fibers on the mandrel.

Figure 7 compares real data, from actual observation (black solid lines), with measured data, estimated from the visual positioning system (purple dashed lines), in terms of positional parameters (x, y, θ) and kinetic parameters (v_x, v_y, w). Displacement of the moving exhibits target, on X axis or Y axis, a simple harmonic behavior, with respect to time, because of the circular designated route for the filament winder. As observed in Figure 7a–c, the purple dashed lines and the black solid lines are identical. This qualitatively indicates that the visual positioning system developed in this work can precisely locate the moving target. However, for the linear velocities (v_x, v_y), there is a slight offset between the curves, which can be seen in Figure 7d,e. The purple dashed lines appear slightly ahead of the black solid lines for the whole period of time. Such behavior prediction is primarily owing to the application of the extended Kalman filter (EKF) in the cameras' self-configuration stage. Specifically, the measured velocity from the EKF predictor, at every point, is numerically calculated by linear prediction, based on the tendency of the former two points, thereby inheriting the mathematical characteristics of the trend before the current point. Such technique was also used in the literature [29] for plotting the tangent modulus from the stress-strain correlations. Thus, the estimated results calculated by the visual positioning system are theoretically ahead of the actual observed results, which are displayed in Figure 7d,e. The angular speed of the moving target in terms of w has been depicted in Figure 7f. Significant divergence can be observed between the real data and the measured data. The reason for such disparity is similar to the cause of the offset between the real data and measured data for the parameters of linear velocities (v_x, v_y). Because the prediction of the current angular speed inherits the tendency of the former angular speed, controlling signals from the output of the navigating system lead to an abrupt increase or decrease in the angular speed. As a consequence, overshoot of the predictive curve (see fluctuations of purple curve in Figure 7f), which comprises estimated data points, is likely to occur.
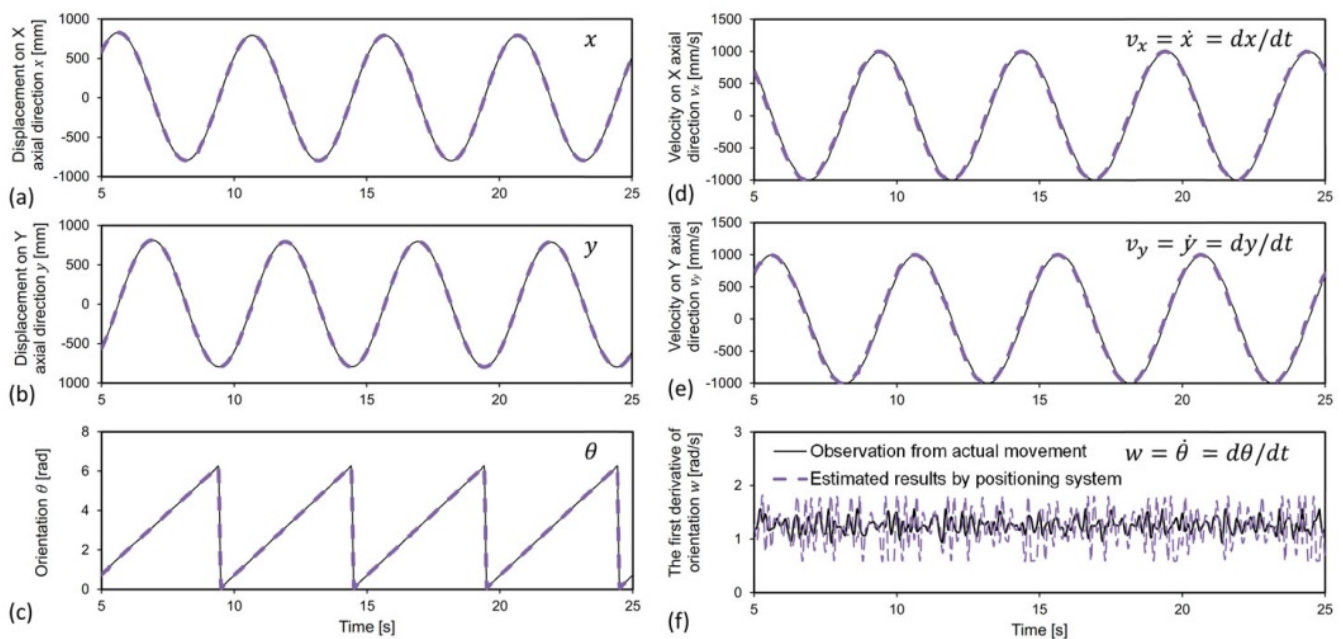

Figure 7. Comparison of parameters from actual observation (black solid lines) and that estimated by the visual positioning system (purple dashed lines), including the positional parameters x, y, θ, and the kinetic parameters v_x, v_y, and w, which are defined in Equation (6).

4.2. Error Analysis

To quantitatively evaluate the fitness of the real data (actual observation) and the measured data (results estimated by the positioning system), variation of the relative error, in percentage, with time, was calculated and collected in Figure 8, with respect to the positional parameters (x, y, θ) and kinetic parameters (v_x, v_y, w). In general, the error for positional parameters (x, y, θ) corresponding to the Figure 8a–c are all within 10%. The averaged errors are calculated as 3.65%, 2.65%, and 1.31%, for the positional parameters x, y, and θ, respectively. These values indicate the high precision of the visual positioning system in obtaining the positional information for the moving target. The averaged errors of the kinetic parameters v_x, v_y, and w are calculated as 18.96%, 13.27%, and 29.13%, respectively. These values are slightly larger than that for positional parameters. It may be limited by the resolution of the cameras, or the blurry object boundaries resulting from residual images obtained by the cameras. These residual images or trajectories are very likely to occur when the object moves fast.
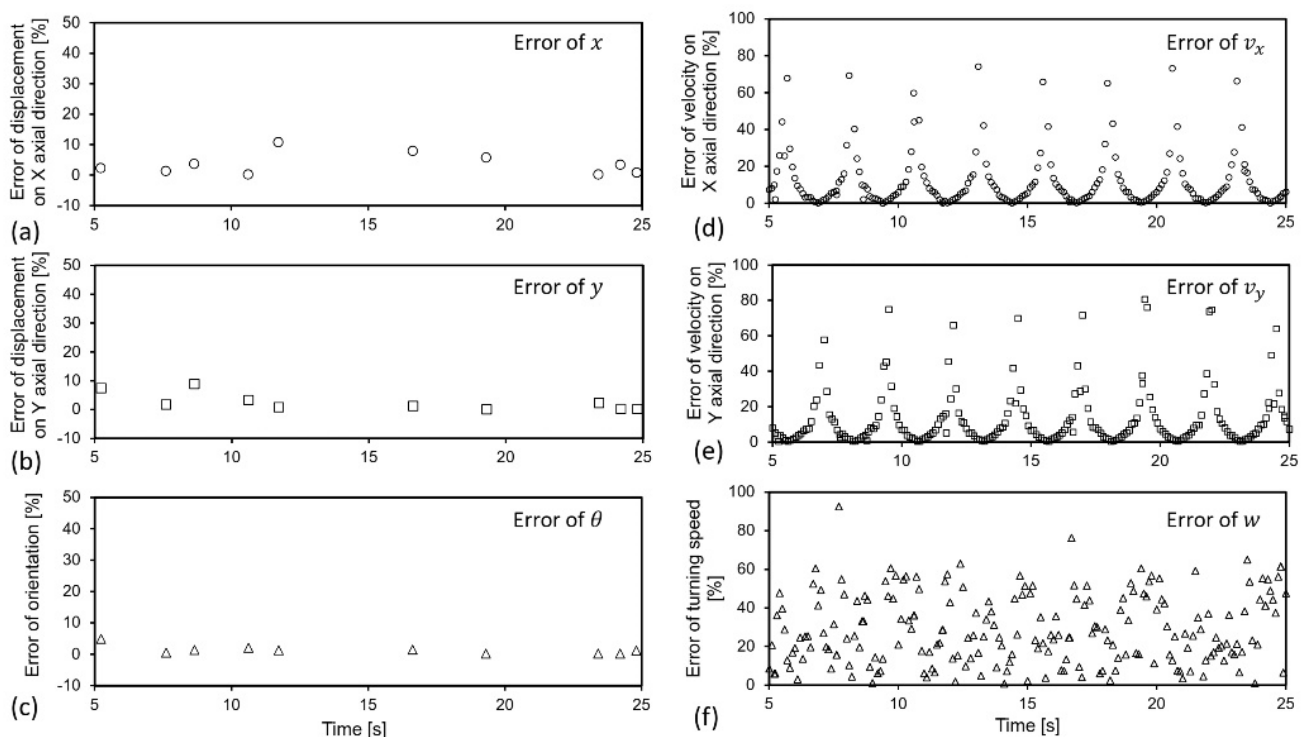

Figure 8. Relative errors plotted versus time for positional parameters (**a**–**c**) and kinetic parameters (**d**–**f**). Errors of positional parameters are kept within 10%, which demonstrates the high-precision locating functions of the visual positioning system.

Further quantitative evaluations on the fitness between the real data and the measured data have been conducted using the technique of the one-way analysis of variance (ANOVA test). The ANOVA test is often used for judging the fitness between two groups of data [30]. If the calculated value F is smaller (or larger) than the critical value, the ANOVA test confidently indicates that the two groups of data are statistically equal (or not statistically equal). In this work, an ANOVA test with a significance level of $\alpha = 0.01$ was performed. Table 1 lists the results from the ANOVA test on both positional and kinetic parameters. It can be observed that obtained values F are smaller than the critical value, except the data group for angular speed, w. Therefore, it shows 99% confidence $(1-\alpha)$ that the real data and the measured data for the positional parameters (x, y, θ) and the linear velocities (v_x, v_y) are statistically equal. Acceptable fitness of real data and measured data has been achieved for these parameters. For the angular speed w, the ANOVA test demonstrates that the real data and the measured data are not equal, which implies that the result, in terms of angular speed, obtained from the visual positional system, is not statistically accurate.

Table 1. Quantitative evaluation of the fitness of real data and measured data, in terms of positional parameters (x, y, θ) and kinetic parameters (v_x, v_y, w).

	x	y	θ	v_x	v_y	w
F *	2.31×10^{-6}	1.69×10^{-6}	2.16×10^{-4}	5.23×10^{-6}	4.14×10^{-4}	9.72
F^{CV} *	6.69	6.69	6.69	6.69	6.69	6.69
Interpretation	$F < F^{CV}$	$F < F^{CV}$	$F < F^{CV}$	$F < F^{CV}$	$F < F^{CV}$	$F > F^{CV}$

* F and F^{CV} are the obtained value and critical value, respectively, from one-way analysis of variance (ANOVA) test.

5. Conclusions

An intelligent, self-configurable visual positioning system based on a camera-network was developed and investigated in this work. The present paper introduces this visual positioning system in relation to a specific industrial application, i.e., tracking the movement of a filament winder. The methodology of locating the filament winder has been elaborated, including the algorithms for enabling calibration of cameras, the automatic selection of cameras, and self-configuration for the camera network. The technique of extended Kalman filter (EKF) was involved in the data processing, aiming to numerically estimate the kinetic coefficients (i.e., velocity, acceleration, and jerk) of the filament winder on the basis of experimentally obtained positioning.

The moving trajectory, time schedule of automatically switching cameras, and time-domain chart of positional parameters (x, y, and θ) and kinetic parameters (v_x, v_y, and w) have been experimentally recorded and illustrated. General error analysis and an ANOVA test have been carried out to determine the quality of positioning of the system. Relative errors for positional parameters are all smaller than 10%; relative errors for linear velocities (v_x, v_y) are also kept to an acceptable level, i.e., lower than 20%. The result from the ANOVA test verifies such a conclusion, which suggests that the real data and the measured data for the positional parameters (x, y, θ) and the linear velocities (v_x, v_y) are statistically equal, with 99% confidence.

In conclusion, these results demonstrate that the camera-network-based visual positioning system is capable of locating a moving target with high precision, and predicting the linear velocities of the moving target with acceptable accuracy. It presents the outstanding potential of this visual positioning system to assist in the industry of automation, including wireless intelligent control, high-precision indoor positioning, and navigation.

References

1. Bae, H.J.; Choi, L. Large-Scale Indoor Positioning using Geomagnetic Field with Deep Neural Networks. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–6. [CrossRef]
2. Chang, L.; Wang, J.; Meng, H.; Chen, X.; Fang, D.; Tang, Z.; Wang, Z. Towards Large-Scale RFID Positioning: A Low-cost, High-precision Solution Based on Compressive Sensing. In Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom), Athens, Greece, 19–23 March 2018; pp. 1–10. [CrossRef]
3. Lu, Y.; Liu, W.; Zhang, Y.; Xing, H.; Li, J.; Liu, S.; Zhang, L. An Accurate Calibration Method of Large-Scale Reference System. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 6957–6967. [CrossRef]

4. Maisano, D.A.; Jamshidi, J.; Franceschini, F.; Maropoulos, P.G.; Mastrogiacomo, L.; Mileham, A.R.; Owen, G.W. A comparison of two distributed large-volume measurement systems: The mobile spatial co-ordinate measuring system and the indoor global positioning system. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2009**, *223*, 511–521. [CrossRef]

5. Liu, Z.G.; Xu, Y.Z.; Liu, Z.Z.; Wu, J.W. A large scale 3D positioning method based on a network of rotating laser automatic theodolites. In Proceedings of the IEEE International Conference on Information and Automation (ICIA), Harbin, China, 20–23 June 2010; pp. 513–518.

6. Cuypers, W.; van Gestel, N.; Voet, A.; Kruth, J.-P.; Mingneau, J.; VanGestel, P.B.N.; Voet, A.; Kruth, J.-P.; Mingneau, J.; Bleys, P. Optical measurement techniques for mobile and large-scale dimensional metrology. *Opt. Lasers Eng.* **2009**, *47*, 292–300. [CrossRef]

7. Zhang, L.; Zhang, Z.; Siegwart, R.; Chung, J.J. Distributed PDOP Coverage Control: Providing Large-Scale Positioning Service Using a Multi-Robot System. *IEEE Robot. Autom. Lett.* **2021**, *6*, 2217–2224. [CrossRef]

8. Yen, H.H. Novel visual sensor deployment algorithm in PTZ wireless visual sensor networks. In Proceedings of the Asia Pacific Conference on Wireless and Mobile, Bali, Indonesia, 28–30 August 2014; pp. 214–218.

9. Liang, X.F.; Sumi, Y.; Kim, B.K.; Do, H.M.; Kim, Y.S.; Tomizawa, T.; Ohara, K.; Tanikawa, T.; Ohba, K. A large planar camera array for multiple automated guided vehicles localization. In Proceedings of the International Conference on Advanced Intelligent Mechatronics, Xi'an, China, 2–5 July 2008; pp. 608–613.

10. Galetto, M.; Mastrogiacomo, L.; Pralio, B. MScMS-II an innovative IR-based indoor coordinate measuring system for large-scale metrology applications. *Int. J. Adv. Manuf. Technol.* **2011**, *52*, 291–302. [CrossRef]

11. Dixon, M.; Jacobs, N.; Pless, R. An Efficient System for Vehicle Tracking in Multi-Camera Networks. In Proceedings of the Third ACM/IEEE International Conference on Distributed Smart Cameras, Como, Italy, 30 August–2 September 2009; pp. 1–8.

12. Wei, G.; Petrushin, V.; Gershman, A. Multiple-camera people localization in a cluttered environment. In Proceedings of the 5th International Workshop on Multimedia Data Mining, Brighton, UK, 1–4 November 2004; pp. 9–12.

13. Kuo, T.; Ni, Z.; De Leo, C.; Manjunath, B.S. Design and Implementation of a Wide Area, Large-Scale Camera Network. In Proceedings of the Computer Vision and Pattern Recognition Workshops, San Francisco, CA, USA, 13–18 June 2010; pp. 25–32.

14. Lin, J.; Chen, J.; Yang, L.; Ren, Y.; Wang, Z.; Keogh, P.; Zhu, J. Design and development of a ceiling-mounted workshop measurement positioning system for large-scale metrology. *Opt. Lasers Eng.* **2020**, *124*, 105814. [CrossRef]

15. Puerto, P.; Heißelmann, D.; Müller, S.; Mendikute, A. Methodology to Evaluate the Performance of Portable Photogrammetry for Large-Volume Metrology. *Metrology* **2022**, *2*, 320–334. [CrossRef]

16. Maisano, D.A.; Mastrogiacomo, L. Cooperative diagnostics for combinations of large volume metrology systems. *Int. J. Manuf. Res.* **2019**, *14*, 15–42. [CrossRef]

17. Soro, S.; Heinzelman, W. Camera Selection in Visual Sensor Networks. In Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, London, UK, 5–7 September 2007; pp. 81–86.

18. Liu, L.; Zhang, X.; Ma, H.D. Dynamic Node Collaboration for Mobile Target Tracking in Wireless Camera Sensor Networks. In Proceedings of the INFOCOM, Rio de Janeiro, Brazil, 24 April 2009; pp. 1188–1196.

19. Wu, Y.F.; Li, G.Y.; Yan, H. Optimal Camera Placement of Large Scale Volume Localization System for Mobile Robot. In Proceedings of the ICMSE, Shanghai, China, 14–16 April 2014; pp. 1390–1395.

20. Tian, H.F.; Yan, H.; Wu, Y.F. Relative Position Algorithm for Optimal Camera Placement of Large Scale Volume Localization System. In Proceedings of the ICFMD, Hong Kong, China, September 2014; pp. 1442–1446.

21. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2004; Part I.

22. Corke, P. *Robotics Vision & Contro*, 1st ed.; Springer: Berlin/Heidelberg, Germany; GmbH & Co. K: Berlin, Germany, 2011; Chapter 2.

23. Kim, G.W.; Nam, K.T.; Lee, S.M. Visual Servoing of a Wheeled Mobile Robot using Unconstrained Optimization with a Ceiling Mounted Camera. In Proceedings of the 16th IEEE International Conference on Robot and Human interactive Communication, Jeju, Korea, 8–12 August 2007; pp. 212–217.

24. Dutkiewicz, P.; Kiełczewski, M.; Kozłowski, K. Vision localization system for mobile robot with velocities and acceleration estimator. *Bull. Pol. Acad. Sci. Tech. Sci.* **2010**, *58*, 29–41. [CrossRef]

25. Wan, E. Sigma-Point Filters: An Overview with Applications to Integrated Navigation and Vision Assisted Control. In Proceedings of the 2006 IEEE Nonlinear Statistical Signal Processing Workshop, Cambridge, UK, 13–15 September 2006; pp. 201–202, ISBN 978-1-4244-0579-4. [CrossRef]

26. Pelka, M.; Hellbruck, H. Introduction discussion and evaluation of recursive Bayesian filters for linear and nonlinear filtering problems in indoor localization. In Proceedings of the 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Alcala de Henares, Spain, 4–7 October 2016; pp. 1–8.

27. Haifeng, Z.; Huabo, W.; Min, W.; Huizhu, Z.; Yu, L. A Vision Based UAV Pesticide Mission System. In Proceedings of the 2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC), Xiamen, China, 10–12 August 2018; pp. 1–6. [CrossRef]

28. Svoboda, T.; Martinec, D.; Pajdla, T. A convenient multi-camera self-calibration for virtual environments. *Teleoperators Virtual Environ.* **2005**, *14*, 407–422. [CrossRef]

29. Zhao, W.; Shi, Z.; Chen, X.; Yang, G.; Lenardi, C.; Liu, C. Microstructural and Mechanical Characteristics of PHEMA-based Nanofibre-reinforced Hydrogel under Compression. *Compos. Part B Eng.* **2015**, *76*, 292–299. [CrossRef]

30. Zhao, W.; Li, X.; Gao, S.; Feng, Y.; Huang, J. Understanding Mechanical Characteristics of Cellulose Nanocrystals Reinforced PHEMA Nanocomposite Hydrogel: In Aqueous Cyclic Test. *Cellulose* **2017**, *24*, 2095–2110. [CrossRef]