

MoiréTag: Angular Measurement and Tracking with a Passive Marker

Simeng Qiu

King Abdullah University of Science
and Technology
Saudi Arabia
simeng.qiu@kaust.edu.sa

Hadi Amata

King Abdullah University of Science
and Technology
Saudi Arabia
amata.hadi@kaust.edu.sa

Wolfgang Heidrich

King Abdullah University of Science
and Technology
Saudi Arabia
wolfgang.heidrich@kaust.edu.sa

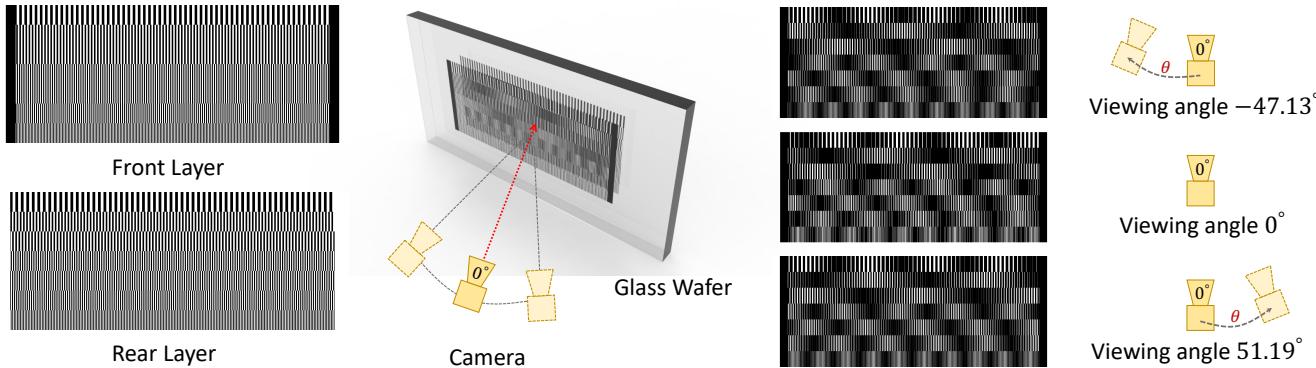


Figure 1: Angular measurement via Moiré Patterns. We encode angular information into a Moiré pattern generated by periodic binary structures printed on the two sides of a glass wafer. Angular shifts manifest themselves as a phase shift in the lowest frequency of the pattern (see rectified patches on the right). We combine multiple Moiré patterns of different base frequencies to perform accurate measurements over a wide angular range.

ABSTRACT

Stable, low-cost, and precise visual measurement of directional information has many applications in domains such as virtual and augmented reality, visual odometry, or industrial computer vision. Conventional approaches like checkerboard patterns require careful pre-calibration, and can therefore not be operated in snapshot mode. Other optical methods like autocollimators offer very high precision but require controlled environments and are hard to take outside the lab. Non-optical methods like IMUs are low cost and widely available, but suffer from high drift errors.

To overcome these challenges, we propose a novel snapshot method for angular measurement and tracking with Moiré patterns that are generated by binary structures printed on both sides of a glass plate. The Moiré effect amplifies minute angular shifts and translates them into spatial phase shifts that can be readily measured with a camera, effectively implementing an optical Vernier scale. We further extend this principle from a simple phase shift

to a chirp model, which allows for full 6D tracking as well as estimation of camera intrinsics like the field of view. Simulation and experimental results show that the proposed non-contact object tracking framework is computationally efficient and the average angular accuracy of 0.17° outperforms the state-of-the-arts.

CCS CONCEPTS

• Computing methodologies → Computational photography.

KEYWORDS

Moiré effect, Moiré Patterns, Object angular tracking, Calibration-free.

ACM Reference Format:

Simeng Qiu, Hadi Amata, and Wolfgang Heidrich. 2023. MoiréTag: Angular Measurement and Tracking with a Passive Marker. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings (SIGGRAPH'23 Conference Proceedings)*, August 6-10, 2023, Los Angeles, CA, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3588432.3591538>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SIGGRAPH'23 Conference Proceedings, August 6-10, 2023, Los Angeles, CA, USA

© 2023 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

<https://doi.org/10.1145/3588432.3591538>

1 INTRODUCTION

We introduce an optical marker to accurately measure angular and directional information. This carefully designed encoded optical element, called MoiréTag, is based on the Moiré effect [Amidror 2009; Gabrielyan 2007; Oster and Nishijima 1963], which occurs when two periodic structures interfere in a multiplicative way. Minute

changes in the relative offsets of the two structures manifest themselves as macroscopic phase shifts of the low frequencies in the Moiré pattern. In the two-layer design of MoiréTag, these offsets correspond to the parallax under different observation directions. A subtle shift of one layer over another causes apparent visual movement of the superposed Moiré patterns along different directions. The movement amplification property of the Moiré effect has been

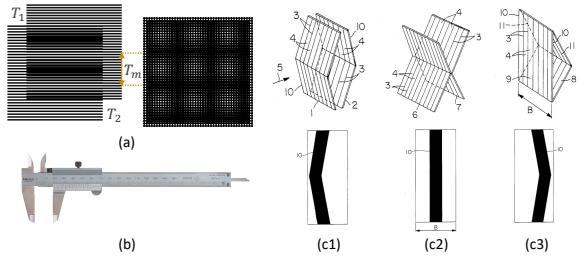


Figure 2: (a) Multiplicative combination of two periodic structures with slightly different frequency introduces a low frequency component known as the Moiré effect. The phase of this low frequency component is very sensitive to minor changes in the offset between the two periodic structures. (b) The Moiré effect is similar to the Vernier scale, which combines two scales of slightly different period to provide high accuracy measurements. (c) The Inogon [Bergkvist and Forsen 1986] is a passive marker that uses the Moiré effect to provide visual feedback on positioning in shipping lanes.

widely utilized as a measuring instrument in various applications such as microscopes, telescopes, micrometers, and Vernier scales (see Fig. 2 (b)). More recently, works like camera tracking for 3-DoF translations [Banks et al. 2019; Xiao and Zheng 2021], and camera pose estimation for 6-DoF extrinsic parameters [Ning et al. 2022] also applied the magnification attributes of Moiré patterns.

However, the Moiré effect has not been utilized for high precision angular and directional measurements, which is the main focus of this work. To the best of our knowledge, this is the first work that exhaustively utilizes the Moiré effect's coarse-to-fine frequency properties to achieve high-precision object angular measurements. In addition, we discuss extensions of the basic approach to estimate positional information as well as camera intrinsics from a single snapshot image. Specifically, we make the following contributions:

- We introduce a novel, stable, low-cost, computationally efficient, and high-accuracy object tracking method. It works out of the box without camera calibration.
- We implement the coarse-to-fine scaling structure of a *Vernier scale* that can measure the subtle motion of an object.
- Our model enables accurate angular estimation in a wide angular range and long measuring distance. A particular hyperbolic chirp model can acquire the capturing distance, the camera's intrinsic parameters, and the field of view (FOV).
- Simulation and experiment show a snapshot angular measurement accuracy of 0.17° as well as a positional accuracy of 3.04 mm .

2 RELATED WORK

Optical tracking: Checker boards [Zhang 2000] or fiducial markers such as ArUco [Garrido-Jurado et al. 2014], ARTag [Fiala 2005], STag [Benligiray et al. 2019], AprilTag [2016], or ChromaTag [DeGol et al. 2017] are commonly used for positional tracking of objects or cameras in a variety of settings, including robotics [Song et al. 2017], visual odometry [Harmat et al. 2015], or AR/VR systems. These approaches achieve good accuracy in object position, however rotational or orientation information can be quite poor and depends heavily on the careful calibration of the camera intrinsics, as well as a good distribution of the fiducials over the full field of view. These are undesirable properties for a number of reasons. First, the need for calibration complicates work outside the lab due to changes of the calibration with focus or zoom. Second, good coverage of the field of view reduces the ability to track small objects at a larger distance. Other optical methods exist for angular measurements, including, for example, autocollimators [Chen et al. 2017; Geckeler et al. 2012]. These offer superior sensitivity and precision, however, they are even more difficult to deploy outside controlled environments. In this work, we aim to achieve high angular precision while maintaining the simplicity of marker-based tracking. Our approach can be extended to position tracking and simultaneous estimation of intrinsics from a single image.

Deep-Learning-based approaches: Deep-Learning-based [Liu et al. 2020; Morency et al. 2008; Parameshwara et al. 2022; Ruiz et al. 2018] methods tend to be popular due to the fast-growing training networks. These methods strongly depend on scene structure for 3D motion recovery and require huge training datasets with complex training frameworks. It is also difficult for such approaches to measure subtle motion with high accuracy, and to deliver a generalized model. There are many deep learning methods and developed algorithms for object tracking in AR and VR areas such as Microsoft HoloLens [Microsoft. 2018] and Oculus Quest 2 [Technologies. 2020] apply Simultaneous Localization and Mapping (SLAM) [Campos et al. 2021; Saputra et al. 2018; Younes et al. 2017] and visual odometry [Scaramuzza and Fraundorfer 2011; Yang et al. 2022]. They recover the camera pose by applying the camera-captured images. But the performance by using SLAM and visual odometry depends on the environment's lighting condition, static scene structure, and scene texture [Eger Passos and Jung 2020; Hübner et al. 2020] to extract apparent motion, and the results are unreliable [Saputra et al. 2018] in some scenarios.

Moiré patterns for position and angle measurement: An early example of utilizing the Moiré effect for angular measurement is the Inogon [Lars A. Bergkvist 1983] (see Figure 2 (c)), as passive visual indicator for steering ships in narrow shipping lanes. Hideyuki et al. [Tanaka et al. 2012, 2014] propose to apply Moiré patterns for solving orientation accuracy in frontal observation and pose ambiguity.

More recent work explores two layers of line gratings for 3D position measurement [Banks et al. 2019], however, the approach has a limited working volume, the Moiré patterns are very large compared to the working volume, and the camera has to be fronto-parallel to the two planes. Some of these restrictions have later been relaxed [Xiao and Zheng 2021] to allow for tilted camera

angles. In comparison to these methods, our method offers thin markers with a small volume that works with a large standoff distance and primarily targets angular measurements. In very recent work, Ning et al. [2022] propose to use the Moiré pattern resulting from re-photographing a monitor with a digital camera for 6 DoF pose estimation. This approach requires careful calibration of the specific camera and display combination. With a reported angular error of multiple degrees and a positional error of multiple cm, their accuracy is substantially lower than our results using an uncalibrated camera. For completeness, we would like to mention that the Moiré effect has also found various other applications in imaging and graphics, including for security purposes such as a watermark-like technique [Cheng et al. 2021], a secure QR-based communication method [Pan et al. 2019], and face-spoofing detection [Garcia and de Queiroz 2015]. Another Moiré implementation is to minimize the Moiré fringes to improve the quality of a recaptured image from the display monitor [He et al. 2020; Sun et al. 2018; Yu et al. 2022].

3 OVERVIEW

The operating principle of MoiréTag is based on the Moiré effect, which is illustrated in Fig. 3: when two square waves with slightly different frequencies superimpose in a multiplicative fashion, the resulting product wave consists of a mixture of different frequencies, including a low base frequency whose phase magnifies and encodes the relative offset between the two signals.

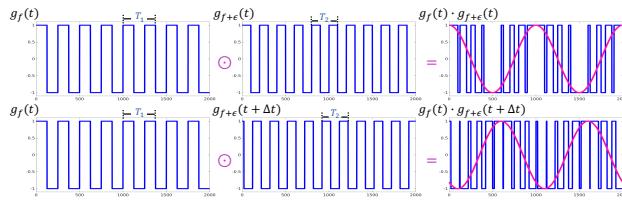


Figure 3: Left: Square wave $g_f(t)$ with period T_1 . Middle: $g_{f+e}(t)$ with period T_2 (top) and $g_{f+e}(t + \Delta t)$ with a time shift Δt (bottom). Right: A slight time shift in one wave can result in a significantly base frequency shift (bottom pink) compared to the one without any time shift (top pink).

In MoiréTag, we generate this Moiré effect using two periodic stripe patterns printed onto the two sides of a transparent glass slide, as depicted in Fig. 1. The parallax corresponding to different observation directions therefore causes an offset between the two stripe patterns, which gets encoded and magnified into the Moiré effect, and can therefore be measured with a camera.

Since a single Moiré pattern only can only encode a limited range of offsets, we propose a coarse-to-fine approach that employs multiple frequencies (Sec. 4), enabling us to achieve both a wide measuring range and high angular accuracy. Although we have derived an analytical model for the mapping individual phase measurements to angles, fusing the information provided by individual frequency patterns is a complex task. This complexity arises from changes in the reliability of each phase measurement with angle, camera standoff distance, camera resolution, motion blur, and noise

in different levels (examples are shown in Fig. 13). The effects of all these hidden parameters are difficult to model analytically. Therefore we opt instead to train a simple Multi-Layer Perceptron (MLP) to handle the data fusion between the individual raw phase measurements (Sec. 5). Finally, we extend the basic Moiré framework to a hyperbolic chirp model for full 6-DoF measurement as well as camera intrinsic calculation (Sec. 6).

3.1 Mapping Between Angles and Spatial Offsets

The physical realization of MoiréTag is to print the two-layer designs from Fig. 7 onto the two sides of a thin glass wafer. We need to account for refraction on the glass-air interface when mapping the measured raw Moiré phase to physical angles. In our prototype, the glass wafer has a thickness of $d = 510 \mu\text{m}$, and the glass's refractive index is $n_1 \approx 1.46$. Considering Snell's law [Shirley 1951], the relationship between the camera viewing angle θ and pixel offset x relative to the superposed two Moiré layers is $x = d \tan(\arcsin(\frac{\sin \theta}{n_1}))$ (see Fig. 4).

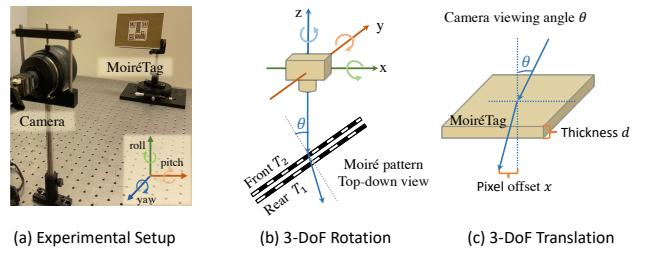


Figure 4: (a) We measure the viewing angle by fixing the camera and rotating the stage per degree. (b) The relationship between viewing angles and the horizontal Moiré pattern in MoiréTag. (c) The correlation between the translation x and the viewing angle θ .

4 ANGLE MEASUREMENT VIA MOIRÉTAG

In the following, we derive the mathematical model for the phase magnification and encoding due to the Moiré effect. Since we have previously established a straightforward mapping between angles and spatial offsets. In the following we choose to make all derivations with respect to spatial offsets instead of angles for simplicity of notation.

4.1 Phase Estimation from Moiré Patterns

The binary patterns consist of an even area proportion of black-and-transparent linear stripes, hence having constant frequencies. In 1D, each of the two layers is modeled as a square wave $g_f(t)$. The Fourier series expansion of a square wave $g_f(t)$ is

$$g_f(t) = \frac{1}{2} + \sum_{k=1}^{\infty} \left[\frac{2}{k\pi} \sin(k2\pi ft) \right]. \quad (1)$$

The product of two square waves $g_f(t)$ and $g_{f+e}(t + \Delta t)$ with frequencies f and $f + e$ generates a superposed binary Moiré pattern, where the offset Δt can be geometrically mapped to the parallax

angle between the two layers. The Fourier expansion of the observed Moiré pattern is then given as

$$\begin{aligned}
 g_s(t) &= g_f(t) \cdot g_{f+\epsilon}(t + \Delta t) \\
 &= \left(\frac{1}{2} + \sum_{k=1}^{\infty} \left[\frac{2}{k\pi} \sin(2\pi k f t) \right] \right) \cdot \left(\frac{1}{2} + \sum_{l=1}^{\infty} \left[\frac{2}{l\pi} \sin(2\pi l(f + \epsilon)(t + \Delta t)) \right] \right) \\
 &= \frac{1}{4} + \sum_{k=1}^{\infty} \left[\frac{1}{k\pi} \sin(2\pi k f t) \right] + \sum_{l=1}^{\infty} \left[\frac{1}{l\pi} \sin(2\pi l(f + \epsilon)(t + \Delta t)) \right] \\
 &\quad - \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \eta \cos(2\pi [t(k+l)f + l\epsilon + l(f+\epsilon)\Delta t]) \\
 &\quad + \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \eta \cos(2\pi [t(k-l)f - l\epsilon - l(f+\epsilon)\Delta t]).
 \end{aligned} \tag{2}$$

where η is a scaling factor $\frac{2}{kl\pi^2}$, and we have made use of the trigonometric identity $\sin \alpha \cdot \sin \gamma = \frac{1}{2} [\cos(\alpha - \gamma) - \cos(\alpha + \gamma)]$.

We can see from the last term of the equation that the lowest frequency component of the Moiré pattern corresponds to the *frequency difference* ϵ . Many high frequency terms are also present, but they are either suppressed through limited camera resolution, e.g. at large standoff distances from the pattern, or they can be filtered out post capture. Moreover, the phase offset $\phi = 2\pi l(f + \epsilon)\Delta t$ in the low frequency term encodes and magnifies the shifted geometric alignment of the grating patterns. The observable (macroscopic) phase shift of the lowest frequency therefore encodes the microscopic offset Δt of the two layers due to parallax. The Moiré design utilizes this effect to achieve high precision angular measurements.

4.2 Phase Unwrapping Design

Unfortunately, the Moiré magnification effect also implies that the measurement range is quite limited due to phase wrapping. This is similar to the phase wrapping problem encountered in time of flight 3D cameras. The solution to increase the measurement range is to combine different Moiré patterns with a variety of base frequencies f and frequency differences ϵ , and then to fuse the phase information obtained from these individual measurements. This approach is also inspired by the coarse-to-fine approach of classical Vernier scales.

In the following we describe how to select these frequency pairs so that both a large angular range is achieved, and the contrast in the individual Moiré patterns is sufficiently high for detection from a large variety of observation distances.

Contrast maximization. Since we would like Moiré to work over a large range of distances as well as angles, one design criterion is to preserve sufficient contrast in the lowest frequency over a wide range of magnification levels. The exact working volume will depend on factors such as camera resolution and zoom, however we can optimize the design without considering these factors by analyzing the contrast of the Moiré pattern under different down-sampling factors. Within a local area, the maximum intensity of the pattern is 0.5, which occurs when the lines of the two layers align perfectly. On the other hand, if the lines of one layer align with the gaps in the other layer, the local average intensity reaches its minimum of 0. To manually optimize the frequencies, we calculate the average area intensity of the superposed Moiré pattern by pixels

and plot the interference intensity variation of each pixel (see the first row in Fig. 5) as translating the camera location horizontally from the left (negative degrees) to the middle and the right (positive degrees). Each colored line in Fig. 5 represents a specific pixel, here 14 pixels in total with a maximum intensity change within a full angular range (from -75° to 75°). Almost all of the pixels cover a wide intensity range, which means that the grated Moiré pattern changes obviously under the projection of a moving camera. Then we repeat the previous experiments until finding the appropriate Moiré pairs. When the camera is located at $(0, 0, z_w)$, the camera's rotation matrix is identity and the camera angle is zero.

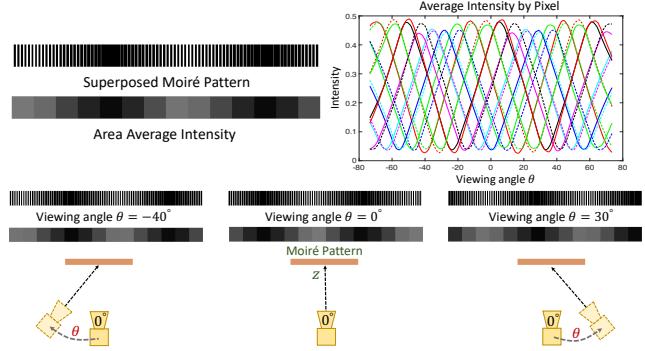


Figure 5: High pixel contrast and pattern movement. The grated Moiré pattern shows the adjacent pixels generate different intensity values. The intensity display of each pixel (different color) on the right shows that the pixel can generate a wide intensity range across a large camera viewing angles. The bottom Moiré pair shows apparent pattern variation.

Phase unwrapping. Phase ϕ in the low-frequency term can be taken as a unique symbol to describe the camera's angular position. The distribution of phases from a single Moiré pair has repetitive properties in a wide angular range and the frequency of the repetitiveness depends on the periods of two Moiré layers. A Moiré pair with high periods has a large phase range, and with low periods has a small but repetitive range. We propose to design seven different Moiré pairs with phase repetitiveness from coarse to fine and frequencies from low to high (Fig. 6). Hence, each camera viewing angle corresponds to a unique phase combination. Our coarse-to-fine mutual correction design allows phase correction between each periodic pair (Fig. 1) and guarantees the camera's angular accuracy.

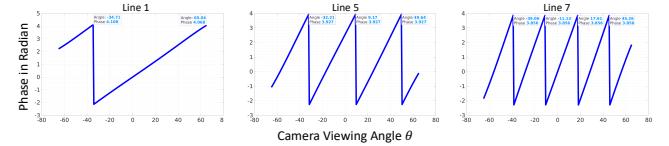


Figure 6: Phase distribution of each Moiré pair. The phase value of each edge point elaborates the repetitive character and range of Moiré pair. The corresponding angular range is from -65° to 65° (see Fig. 2 in supplementary).

4.3 MoiréTag Layout

Finally, we combine two identical groups of seven Moiré patterns into a marker layout to measure both horizontal and vertical angles (Fig. 7). The Moiré markers furthermore contain four corner markers in the style of QR codes, which allow for the automatic detection and approximate localization of the tags in images. Then we introduce small bowtie markers that allow for subpixel-accurate localization of the Moiré patterns for accurate phase estimation. The bowties define a homography that is used to resample each Moiré group into a rectified configuration before further processing. The front and back layer of the MoiréTag are shown together with the superimposed Moiré effect in Fig. 7. The design fabrication is discussed in Sec. 7.

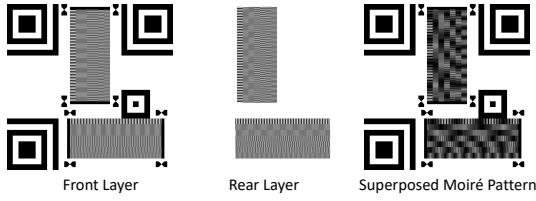


Figure 7: MoiréTag layout with double-side encoding.

5 INFORMATION FUSION BY MLP

5.1 Information Fusion

We previously discussed in Sec. 3 that combining information from individual frequency patterns is a complicated task and involves analyzing complex dependencies such as capturing distance, which is difficult to do analytically. To address this, we chose to use a small Multi Layer Perceptron (MLP) neural network (Fig. 8) to handle the data fusion between the individual raw phase measurements. The input size of MLP is 7. The input layer, two hidden layers, and the output layer are 90, 50, 15, and 1 neurons each. Each layer is followed by a ReLU activation function.

Given a camera image, we first extract rectified versions of the groups of Moiré patterns. In the rectified image, the spatial frequency of each pattern is independent of the observation conditions, and we can obtain an initial raw phase estimate for each frequency by fitting a sinusoid of the appropriate frequency to Eq. 2. The resulting 7 raw phase measurements are processed by the MLP to calculate the angle. The neural network is trained on a dataset that is obtained through a combination of simulation and physical data collection. In particular, it is difficult to set up an accurate 3D rotation system for data collection. We simulate different rotation angles, display the resulting MoiréTag images on a monitor, and re-image those scenes with a camera mounted on a translation stage for accurate measurements under a large range of distances. We believe that this process combined the best of both a relatively simple physical setup with real camera characteristics and the flexibility of simulation.

For each simulated angular pose we capture 40 images covering a wide distance range. The complete input training dataset consists of extracted phases from $203 \times 40 = 8210$ captured images. Each extracted Moiré pattern of the captured images contains 7 phases, which correspond to a known angular pose. We use the

Adam optimizer to train the MLP network with a learning rate of 5×10^{-3} and batch size of 2800 for 2500 epochs. Fig. 2 in the supplement validates the stability of the proposed model by plotting all the phases from the training dataset. The accuracy regards to the viewing angle is 0.17° .

5.2 Validation with Real Data

The orthogonal design of two Moiré patterns allows us to measure both roll and pitch. Since the two patterns are identical, we analyze only roll measurement during the experiments. In real-scene experiments, it is hard to qualitatively analyze the precision of the training model by moving the camera toward the MoiréTag. Hence, we fix the position of the camera while rotating the rotation stage at an arbitrary location to capture the MoiréTag and extract the phases of the horizontal pattern for the first angular θ_0 prediction. Then we fine-tune the rotation stage by a constant $\Delta\theta$ and snapshot the target along a wide angular range to collect the experiment dataset. We compare the difference between θ_0 and $\theta_0 + n\Delta\theta$ among the whole dataset to calculate the root mean square error. A quantitative evaluation of the training model's overall accuracy is performed by capturing several groups of angular images ranging from 1 m to 4 m (see Fig. 11 and Fig. 1 in supplementary). By varying measuring distance and viewing angle, we obtain an error distribution (see Fig. 12, middle). Fig. 11 shows one of the viewing angle estimation results from 65.20° to -65.28° . The purpose of doing these large-range experiments is to validate the stability of the training model. The RMSE of the above wide angular range is 0.1719 (Fig. 12, left), which proves that the model can cover a wide measurement range with high credibility. The overall angular accuracy from multiple real experiments is approximately equal to 0.17° , which is close to the training results. We demonstrate the precision of MoiréTag by minute changes in angle rather than drastic motions. In addition, we fix the target location and rotate the camera to acquire a dataset to show the practicability of the proposed approach (see the supplemental video).

6 HYPERBOLIC CHIRP MODEL FOR 6 DOF TRACKING

Hyperbolic chirp. The phase shift analysis presented above assumes that the MoiréTag is observed with an orthogonal camera, so that the front and back layer have the same relative scale. This is an accurate model for determining directional information at larger standoff distances. However, if the marker is large compared to the camera distance, the two layers are scaled differently, and especially at oblique angles, perspective foreshortening results in spatially varying frequencies of the line patterns from each layer.

As a motivating example, consider Fig. 9(a), which shows an example from our experiments, where the marker has been rotated horizontally relative to the camera by 49.88° . The orange line in Fig. 9 (b) represents a top view of the horizontal Moiré pattern, in which perspective foreshortening causes a spatial variation in the frequency, in this case decreasing from the left to the right. A frequency change with time is known as a chirp effect. If we extract the horizontal Moiré pattern without rectifying the horizontal (“roll”) rotation (see Fig. 10 (b)), the frequencies of the superposed Moiré pattern become inconsistent. We show in the following that

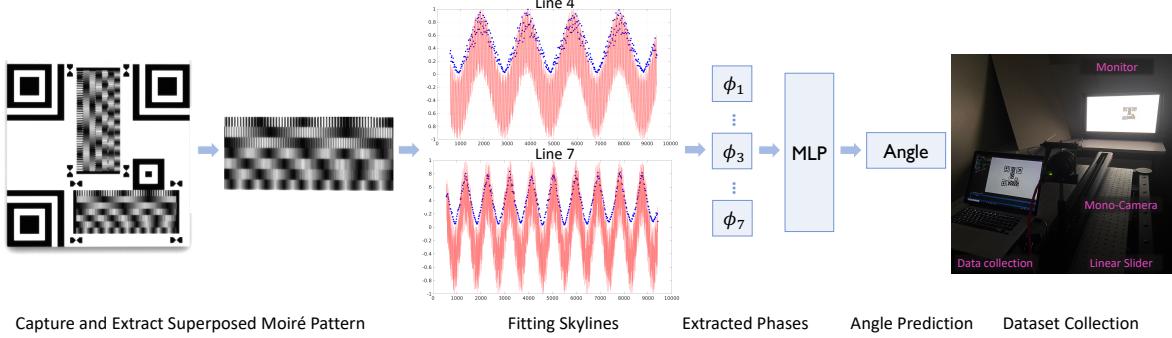


Figure 8: Information fusion framework. From raw camera images we extract rectified Moiré patterns to which we fit the analytical model in order to obtain raw phase measurements. These raw measurements are fused by an MLP to produce the final measured angle. The MLP is trained on data obtained through a combination of a physical setup and simulation.

this perspective effect can be modeled as a chirp with a hyperbolic change in frequency. The hyperbolic chirp model can be used to obtain positional and angular information as well as camera intrinsics.

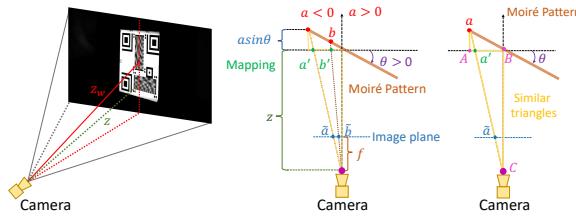


Figure 9: Hyperbolic chirp model (see supplementary).

To derive this model, we consider how the Moiré frequencies from Eq. 2 are mapped from world space to image space under a perspective transformation. The top view of our model in Fig. 9 (b) shows the setup between the camera and the horizontal Moiré pattern (orange color), which is extracted without rectifying the “roll” rotation. According to the pinhole camera projection, two coordinate points a and b (negative when sitting on the left side of the mask) on the Moiré pattern are mapped onto a horizontal plane, which is parallel to the camera image plane. Then a' and b' are projected to the camera image plane at \tilde{a} and \tilde{b} . Based on the geometry alignment, we utilize two similar triangles (yellow lines in Fig. 9 (c)) $Aa'a$ and $Ba'C$ to derive the equation $\frac{a \sin \theta}{z} = \frac{a \cos \theta - a'}{a}$ (see Fig. 9 (c)). z is the distance between the camera and the middle of the mask. θ represents the viewing angle. The period of one of the lines in the superposed Moiré pattern is $T_m = b - a$. Then T_m is mapped to the horizontal plane to get $T'_m = b' - a'$. We assume the z_w -axis of the camera is aligned with the middle of the mask, so the rotation matrix of the camera remains identity. Hence, the superposed Moiré pattern’s period in the image plane is $\tilde{T}_m = f_l \frac{T'_m}{z}$, where f_l is the focal length. From the above equations, a full representation of the period \tilde{T}_m on the image plane is

$$\begin{aligned} \tilde{T}_m &= \frac{f_l T_m z \cos \theta}{\sin^2 \theta a^2 + (2 \sin \theta z + T_m \sin^2 \theta) a + T_m \sin \theta z + z^2}, \\ &= \frac{f_l T_m \cos \theta}{T_m \sin \theta + z}, \quad (\text{if } a = 0). \end{aligned} \quad (3)$$

Eq. 3 shows that the frequency ($1/\tilde{T}_m$) of the grated Moiré pattern with respect to point a is a hyperbolic curve. From the low-frequency term, the hyperbolic chirp model as a function of coordinate point a is expressed as

$$x(a) = \cos\left(\frac{2\pi a}{\tilde{T}_m} + \phi_0\right). \quad (4)$$

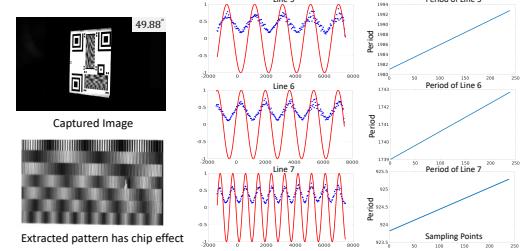


Figure 10: Example of Hyperbolic chirp (see supplementary).

6.1 Validation of Hyperbolic Chirp Model

To illustrate the accuracy of the hyperbolic chirp model, we use the example from Fig. 10 to show the frequency change of each Moiré pair. We attach the MoiréTag at the middle of a rotation stage and turn the stage clockwise at a relatively large angle 49.88° . Then we capture the target with a monochrome camera. After capturing, we extract the horizontal Moiré pattern and fit each Moiré pair’s skyline (one-third area average of a Moiré pair) to the hyperbolic period function Eq. 4 by considering the period \tilde{T}_m in Eq. 4. ϕ_0 in Eq. 4 is from the previous angular fitting. We plot the fitting curves and the frequency change of each Moiré pair in Fig. 10 (c) and (d). The red curve for each skyline has decreasing frequencies from the left to the right. The designed grated periods of Line 5, 6, and 7 are 1984, 1740, and 924. The above periods can be found from the period axis in Fig. 10 (d), and each period has a corresponding sampling point, which is sitting on the rectified horizontal pattern in Fig. 10 (b). We can accurately estimate the capturing origin with the aid of seven changing frequencies. Therefore, the capturing distance z and the focal length $f_l = \frac{\tilde{T}_m(T_m \sin \theta + z)}{T_m \cos \theta}$ ($a = 0$) can be

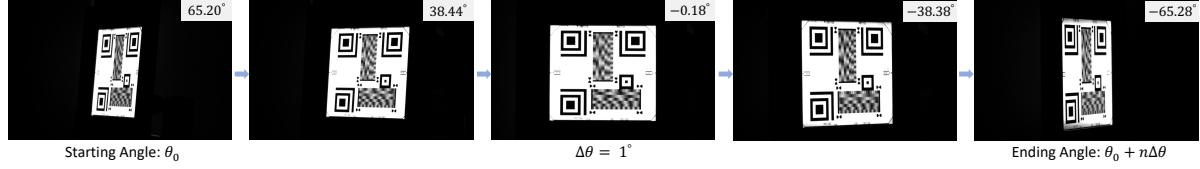


Figure 11: We put the MoiréTag at an arbitrary position and capture the first image. Then we feed the extracted phases of the bottom Moiré pattern from the first image to our pre-trained MLP model in order to predict the angular $\theta_0 = 65.20^\circ$. After obtaining the first angle, we carefully rotate the rotation stage by $\Delta\theta = 1^\circ$ to capture 132 MoiréTag images. We estimate the angle of each image by applying the MLP model and compare the difference with θ_0 .

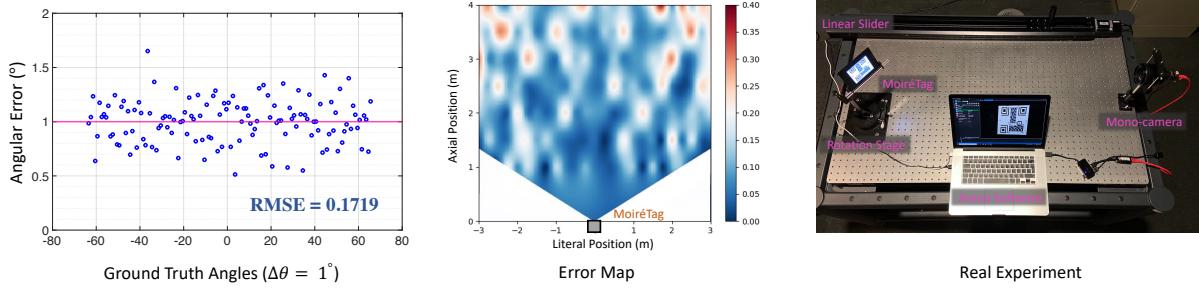


Figure 12: Left: The accuracy of each captured image $\theta_0 + n\Delta\theta$ is measured by calculating the RMSE error based on the gradient of the slope equal to 1, with an angle variation $\Delta\theta$ of 1° . So points in the second image should be close to 1° . The overall RMSE error illustrates the accuracy of the experimental results. Middle: The location of the MoiréTag is shown at the center bottom, with an error distribution obtained by varying measuring distance and viewing angle. Right: All equipment used during the experiment is marked in the image.

estimated from Eq. 4 and the capturing origin. Then the field of view can be derived as $FOV = 2 \times \arctan \frac{\text{sensor size}}{2f_i}$.

To evaluate the overall accuracy of the chirp model, we fix the monochromatic camera on a tilted linear slider and adjust the distance relative to the fixed MoiréTag. We first put the camera closer to the marker and then move the camera backward within a certain distance to collect the captured MoiréTag for each point. Then we move the camera forward at the same path along the slider. Fig. 4 in supplementary shows the distance and focal length accuracy.

7 EXPERIMENTS

Prototype fabrication, cameras, and setup: To achieve a compact measuring instrument, we fabricated the prototype MoiréTag on both sides of a 4 inch fused silica wafer using state-of-the-art photolithography techniques. The mask patterns are etched into chromium layers to achieve high absorption. However since chromium is highly reflective, a layer of 100 nm of silicon dioxide SiO_2 is first deposited by PECD (plasma-enhanced chemical vapor deposition) on both sides wafer. The SiO_2 film works as a light absorber to reduce interreflection between the chromium layers. The vertical and horizontal Moiré patterns are aligned using the front-to-back alignment technique available on the contact aligner (*EVG6200*). The training data was collected with a monochrome camera from Lucid Lab that contains a Sony IMX392 CMOS sensor. The pixel size of the camera is $3.45\text{ }\mu m$ and the sensor size is 7.9 mm . In our other experiments, we also used different cameras including the built-in camera of a DJI Mavic 3 UAV without retraining the network. The

experimental setups used a 180 cm translation stage from Zaber, and the real validation experiment a rotation stage from Thorlabs.

8 COMPARISONS OF OBJECT TRACKING

Next, we compare our work to marker-based object tracking methods. Specifically, we chose the classical checker board pattern [Zhang 2000], the relatively recent fiducial marker system ArUco [Garrido-Jurado et al. 2014], and the state-of-the-art Moiré-guided method [Ning et al. 2022]. In real experiments, with rotation and translation stages, we measure the translations and rotations separately. We test each method at different capture distances ($\Delta D = 1\text{ m}$) to better reflect the overall accuracy. The distance range is from 1.0 to 4.0 meters. All of the tracking methods produce errors to some extent. Our method behaves more stable in the experiment and the estimation accuracy is high. To measure translation accuracy, we fix the target pose and put the camera on a translation stage to adjust the distance and position relative to the target. For the computational cost, the checkerboard is fast but generates large errors if the intrinsics are estimated from the same image, as in our approach. ArUco requires recovering the camera’s intrinsic parameters before tracking. MoiréBoard from [Xiao and Zheng 2021] state that they can achieve good positional accuracy. However, it is limited in its ability to measure angles. Additionally, Moiré pattern requires adjustments to accommodate varying camera distances measuring range. The comparison results show that our proposed method is competitive with the state-of-the-art snapshot position tracking, and offers an order of magnitude improvement in the quality of angle measurements (see Table 1).

Table 1: Comparison of RMS angular and positional errors produced by various competing methods. In real experiments. The positional accuracy was averaged over a distance range of 1 – 4m.

Methods	Angular Error	Position Error
Checkerboard [Zhang 2000]	6.23°	18.65 mm
ArUco [Garrido-Jurado et al. 2014]	2.79°	16.32 mm
MoiréPose [Ning et al. 2022]	1.66°	7.50 mm
MoiréTag (ours)	0.17°	3.04 mm

9 APPLICATION SCENARIOS

We show three experimental results relative to MoiréTag to illustrate the application value of the proposed angular measurement system.

Camera angular retrieval. In the first experiment, we use the MoiréTag to retrieve the camera angle on an iPhone 11. We move the camera from approximately the middle of the fixed MoiréTag to the left. Then we move upward at a certain distance and the right. After that, we move the phone in the opposite direction along the approximately same path. Fig. 14 (a) shows the angular movement. The middle of the angular figure illustrates the direction reversal.

Object tracking. In the second and third experiments, we manage to track the angular position of a flying drone with the aid of MoiréTag. The drone we use is DJI Mavic 3, which has Mavic 3 4/3 CMOS Hasselblad camera. The experiment details are shown in Fig. 14 (b) and (c) (see figures only pages and the corresponding supplement video). The results show that we can retrieve the angular variation for both indoor and outdoor flight tests.

10 CONCLUSION

We present a novel vision-based marker MoiréTag that incorporates the use of Moiré patterns for high-accuracy object angular measurement. A mutual correction permutation of the Moiré patterns makes the estimation more robust and extends its working volume. MoiréTag is a scalable design that can cover large capturing distances and a wide range of viewing angles with high sensitivity. Moreover, MoiréTag has a thin form factor, making it applicable in virtually all scenarios where 2D markers are used today, including AR/VR tracking, autonomous vehicle tracking, visual odometry for UAVs, etc. The unique design of 6-DoF object tracking provides a snapshot procedure without calibration or complex environmental control compared with other state-of-the-art methods. Unlike other object tracking methods, we derive a hyperbolic chirp model to acquire the camera intrinsic parameters, capturing distance and field of view. We hope our proof of concept can motivate further research for high-precision angular measurement.

ACKNOWLEDGMENTS

This work was supported by KAUST baseline funding. We would like to thank Abdelrahman Amr Morsi and Mohamed Shalaby for their valuable assistance with the drone experiment. Additionally, we would like to thank the Siggraph reviewers for their insightful comments.

REFERENCES

- Isaac Amidror. 2009. *The Theory of the Moiré Phenomenon: Volume I: Periodic Layers*. Vol. 38. Springer Science & Business Media.
- Samuel Banks, Richard Green, and Jun Junghyun. 2019. Use of Moiré Patterns in Camera Position Estimation. In *2019 International Conference on Image and Vision Computing New Zealand (IVCNZ)*. IEEE, 1–7.
- Burak Benligiray, Cihan Topal, and Cuneyt Akinlar. 2019. STag: A stable fiducial marker system. *Image and Vision Computing* 89 (2019), 158–169.
- Lars A Bergkvist and Ivan Forseen. 1986. Leading mark indicator. <https://patents.google.com/patent/US4629325A/en>. US Patent 4,629,325.
- Carlos Campos, Richard Elvira, Juan J Gómez Rodríguez, José MM Montiel, and Juan D Tardós. 2021. Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam. *IEEE Transactions on Robotics* 37, 6 (2021), 1874–1890.
- Yuan-Liu Chen, Yuki Shimizu, Jun Tamada, Yukitoshi Kudo, Shuhei Madokoro, Kazuki Nakamura, and Wei Gao. 2017. Optical frequency domain angle measurement in a femtosecond laser autocollimator. *Optics Express* 25, 14 (2017), 16725–16738.
- Yushi Cheng, Xiaoyu Ji, Lixu Wang, Qi Pang, Yi-Chao Chen, and Wenyan Xu. 2021. {mID}: Tracing screen photos via {Moiré} patterns. In *30th USENIX Security Symposium (USENIX Security 21)*, 2969–2986.
- Joseph DeGol, Timothy Bretl, and Derek Hoiem. 2017. Chromatag: A colored marker and fast detection algorithm. In *Proceedings of the IEEE International Conference on Computer Vision*. 1472–1481.
- Daniel Eger Passos and Bernhard Jung. 2020. Measuring the accuracy of inside-out tracking in XR devices using a high-precision robotic arm. In *International Conference on Human-Computer Interaction*. Springer, 19–26.
- Mark Fiala. 2005. ARTag, a fiducial marker system using digital techniques. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 2. IEEE, 590–596.
- Emin Gabrielyan. 2007. The basics of line moiré patterns and optical speedup. *arXiv preprint physics/0703098* (2007).
- Diogo Caetano Garcia and Ricardo L de Queiroz. 2015. Face-spoofing 2D-detection based on Moiré-pattern analysis. *IEEE transactions on information forensics and security* 10, 4 (2015), 778–786.
- Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco José Madrid-Cuevas, and Manuel Jesús Marín-Jiménez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47, 6 (2014), 2280–2292.
- Ralf D Geckeler, Oliver Kranz, Andreas Just, and Michael Krause. 2012. A novel approach for extending autocollimator calibration from plane to spatial angles. *Advanced optical technologies* 1, 6 (2012), 427–439.
- Adam Harmat, Michael Trentini, and Inna Sharf. 2015. Multi-camera tracking and mapping for unmanned aerial vehicles in unstructured environments. *Journal of Intelligent & Robotic Systems* 78, 2 (2015), 291–317.
- Bin He, Ce Wang, Boxin Shi, and Ling-Yu Duan. 2020. FHDe 2 Net: Full high definition demoiré network. In *European Conference on Computer Vision*. Springer, 713–729.
- Patrick Hübner, Kate Clintworth, Qingyi Liu, Martin Weinmann, and Sven Wursthorn. 2020. Evaluation of HoloLens tracking and depth sensing for indoor mapping applications. *Sensors* 20, 4 (2020), 1021.
- Sidensjö, Lars A. Bergkvist, Själevad; Ivan Forseen. US patent 4,629,325A, 1983. LEADING MARK INDICATOR. <https://patentimages.storage.googleapis.com/46/23/fb/0098457a5f9483/US4629325.pdf>.
- Jianwei Liu, Jinsong Han, Feng Lin, and Kui Ren. 2020. Adversary helps: Gradient-based device-free domain-independent gesture recognition. *arXiv preprint arXiv:2004.03961* (2020).
- Microsoft. 2018. Hololens. <https://www.microsoft.com/en-us/hololens/hardware>.
- Louis-Philippe Morency, Jacob Whitehill, and Javier Movellan. 2008. Generalized adaptive view-based appearance model: Integrated framework for monocular head pose estimation. In *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*. IEEE, 1–8.
- Jingyi Ning, Lei Xie, Yi Li, Yingying Chen, Yanling Bu, Baoliu Ye, and Sanglu Lu. 2022. MoiréPose: ultra high precision camera-to-screen pose estimation based on Moiré pattern. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 106–119.
- Gerald Öster and Yasunori Nishijima. 1963. Moiré patterns. *Scientific American* 208, 5 (1963), 54–63.
- Hao Pan, Yi-Chao Chen, Lanqing Yang, Guangtao Xue, Chuang-Wen You, and Xiaoyu Ji. 2019. mqrcode: Secure qr code using nonlinearity of spatial frequency in light. In *The 25th Annual International Conference on Mobile Computing and Networking*. 1–18.
- Chethan M Parameshwara, Gokul Hari, Cornelia Fermüller, Nitin J Sanket, and Yiannis Aloimonos. 2022. DiffPoseNet: Direct Differentiable Camera Pose Estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 6845–6854.
- Nataniel Ruiz, Eunji Chong, and James M Rehg. 2018. Fine-grained head pose estimation without keypoints. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2074–2083.

- Muhamad Risqi U Saputra, Andrew Markham, and Niki Trigoni. 2018. Visual SLAM and structure from motion in dynamic environments: A survey. *ACM Computing Surveys (CSUR)* 51, 2 (2018), 1–36.
- Davide Scaramuzza and Friedrich Fraundorfer. 2011. Visual odometry [tutorial]. *IEEE robotics & automation magazine* 18, 4 (2011), 80–92.
- John W Shirley. 1951. An early experimental determination of Snell's law. *American Journal of Physics* 19, 9 (1951), 507–508.
- Kai-Tai Song, Cheng-Hei Wu, and Sin-Yi Jiang. 2017. CAD-based pose estimation design for random bin picking using a RGB-D camera. *Journal of Intelligent & Robotic Systems* 87, 3 (2017), 455–470.
- Yujing Sun, Yizhou Yu, and Wenping Wang. 2018. Moiré photo restoration using multi-resolution convolutional neural networks. *IEEE Transactions on Image Processing* 27, 8 (2018), 4160–4172.
- Hideyuki Tanaka, Yasushi Sumi, and Yoshio Matsumoto. 2012. Application of moiré patterns to AR markers for high-accuracy pose estimation.
- Hideyuki Tanaka, Yasushi Sumi, and Yoshio Matsumoto. 2014. A solution to pose ambiguity of visual markers using moiré patterns. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 3129–3134.
- LLC. Facebook Technologies. 2020. Oculus Quest 2. <https://www.oculus.com/quest-2/>.
- John Wang and Edwin Olson. 2016. AprilTag 2: Efficient and robust fiducial detection. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 4193–4198.
- Chang Xiao and Changxi Zheng. 2021. MoiréBoard: A Stable, Accurate and Low-cost Camera Tracking Method. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. 881–893.
- Binchao Yang, Xinying Xu, Jinchang Ren, Lan Cheng, Lei Guo, and Zhe Zhang. 2022. SAM-Net: Semantic probabilistic and attention mechanisms of dynamic objects for self-supervised depth and camera pose estimation in visual odometry applications. *Pattern Recognition Letters* 153 (2022), 126–135.
- Georges Younes, Daniel Asmar, Elie Shammas, and John Zelek. 2017. Keyframe-based monocular SLAM: design, survey, and future directions. *Robotics and Autonomous Systems* 98 (2017), 67–88.
- Xin Yu, Peng Dai, Wenbo Li, Lan Ma, Jiajun Shen, Jia Li, and Xiaojuan Qi. 2022. Towards Efficient and Scale-Robust Ultra-High-Definition Image Demoiréing. *arXiv preprint arXiv:2207.09935* (2022).
- Zhengyou Zhang. 2000. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence* 22, 11 (2000), 1330–1334.

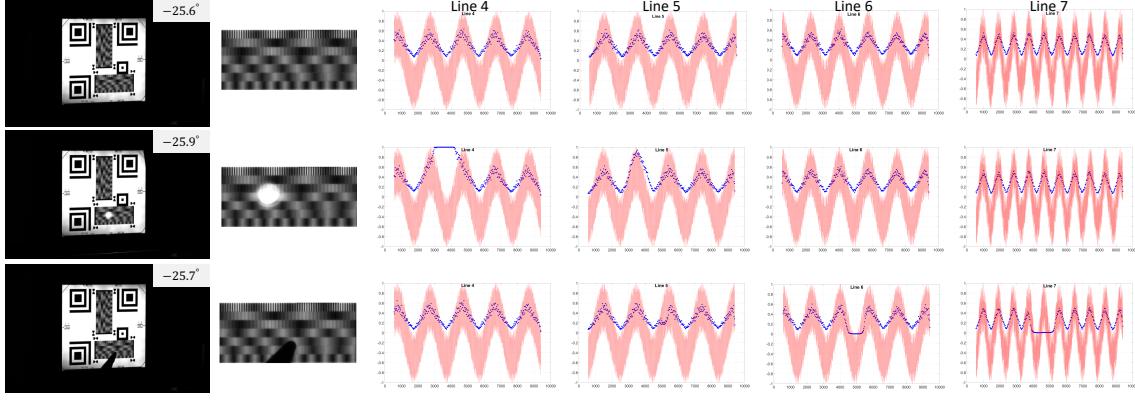


Figure 13: We present an example where a portion of the MoiréTag was partially obscured by either a flashlight or an obstacle. We demonstrate that our method can still accurately measure the viewing angle in comparison to images captured without obstruction (the first row). The images on the right show the extracted pattern relative to the corresponding extracted Moiré patterns in the middle. Even when parts of the Moiré pattern are obscured, the mutual correction design of the MoiréTag enables accurate angle prediction.

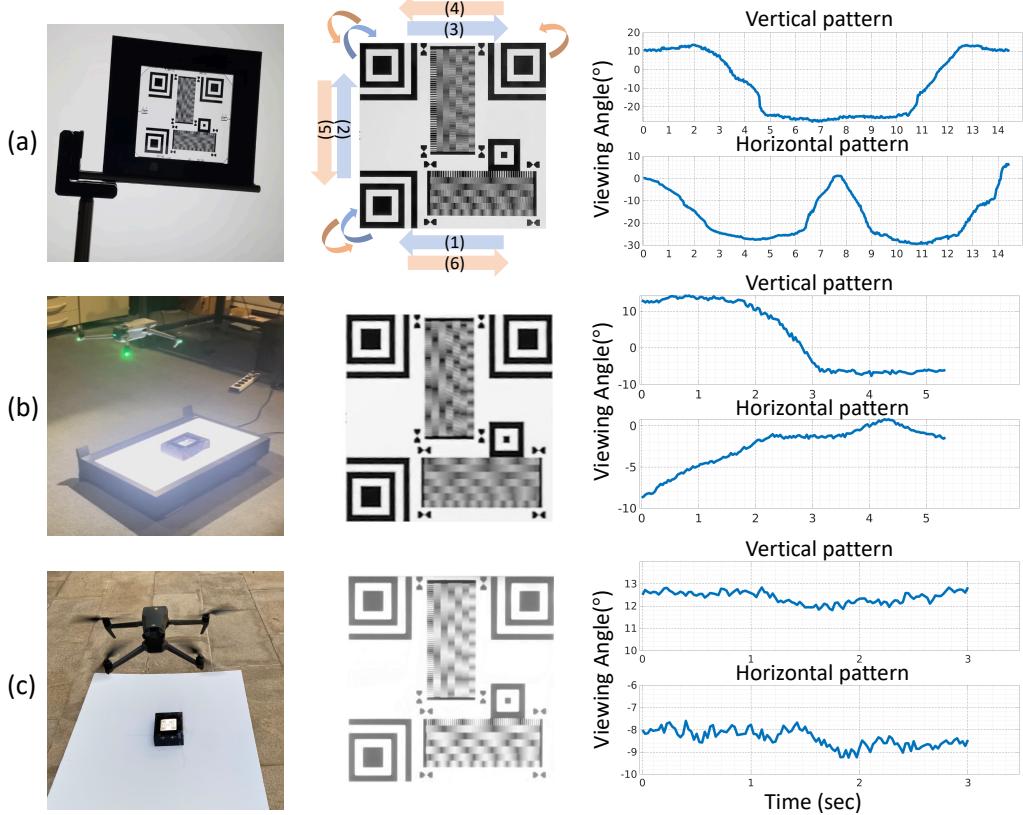


Figure 14: (a) Camera angular retrieval. The left image shows the rectified MoiréTag captured by the camera, with colored arrows indicating the direction of movement of the phone camera. The starting point of the path is labeled as (1). The total duration of the captured video is 14 sec. (b) Indoor application. We track the angular movement of a drone and present the experimental setup on the left. In the middle, we show the rectified MoiréTag captured by the camera and the corresponding angles for both directions. The video has a duration of approximately 5 sec. (c) Outdoor application. The drone's angular movement is tracked in the experiment, with the video lasting around 3 sec. The videos are included in the supplement.