Citation: Xue, Tianfan, Michael Rubinstein, Neal Wadhwa, Anat Levin, Fredo Durand, and William T. Freeman. "Refraction Wiggles for Measuring Fluid Depth and Velocity from Video." Lecture Notes in Computer Science (2014): 767–782.

As Published: http://dx.doi.org/10.1007/978-3-319-10578-9_50

Publisher: Springer-Verlag

Persistent URL: http://hdl.handle.net/1721.1/100013

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Refraction Wiggles for Measuring Fluid Depth and Velocity from Video

Tianfan Xue[1], Michael Rubinstein[2,1], Neal Wadhwa[1], Anat Levin[3], Fredo Durand[1], and William T. Freeman[1]

[1]MIT CSAIL   [2]Microsoft Research   [3]Weizmann Institute

Abstract. We present principled algorithms for measuring the velocity and 3D location of refractive fluids, such as hot air or gas, from natural videos with textured backgrounds. Our main observation is that intensity variations related to movements of refractive fluid elements, as observed by one or more video cameras, are consistent over small space-time volumes. We call these intensity variations "refraction wiggles", and use them as features for tracking and stereo fusion to recover the fluid motion and depth from video sequences. We give algorithms for 1) measuring the (2D, projected) motion of refractive fluids in monocular videos, and 2) recovering the 3D position of points on the fluid from stereo cameras. Unlike pixel intensities, wiggles can be extremely subtle and cannot be known with the same level of confidence for all pixels, depending on factors such as background texture and physical properties of the fluid. We thus carefully model uncertainty in our algorithms for robust estimation of fluid motion and depth. We show results on controlled sequences, synthetic simulations, and natural videos. Different from previous approaches for measuring refractive flow, our methods operate directly on videos captured with ordinary cameras, do not require auxiliary sensors, light sources or designed backgrounds, and can correctly detect the motion and location of refractive fluids even when they are invisible to the naked eye.

1 Introduction

Measuring and visualizing the flow of air and fluids has great importance in many areas of science and technology, such as aeronautical engineering, combustion research, and ballistics. Multiple techniques have been proposed for this purpose, such as sound tomography, Doppler LIDAR and Schlieren photography, but they either rely on complicated and expensive setups or are limited to in-lab use. Our goal is to make the process of measuring and localizing refractive fluids cheaper, more accessible, and applicable in natural settings.

In this paper, we develop passive, markerless techniques to measure the velocity and depth of air flow using natural video sequences. Our techniques are based on visual cues produced by the bending of light rays as they travel through air of differing densities. Such deflections are exploited in various air measurement techniques, described below in related work. As the air moves, small changes in the refractive properties appear as small visual distortions (motions) of the background, similar to the shimmering effect experienced when viewing objects across hot asphalt or through exhaust gases. We call such motions *"refraction wiggles"*, and show they can be tracked using regular video cameras to infer information about the velocity and depth of a refractive fluid layer.
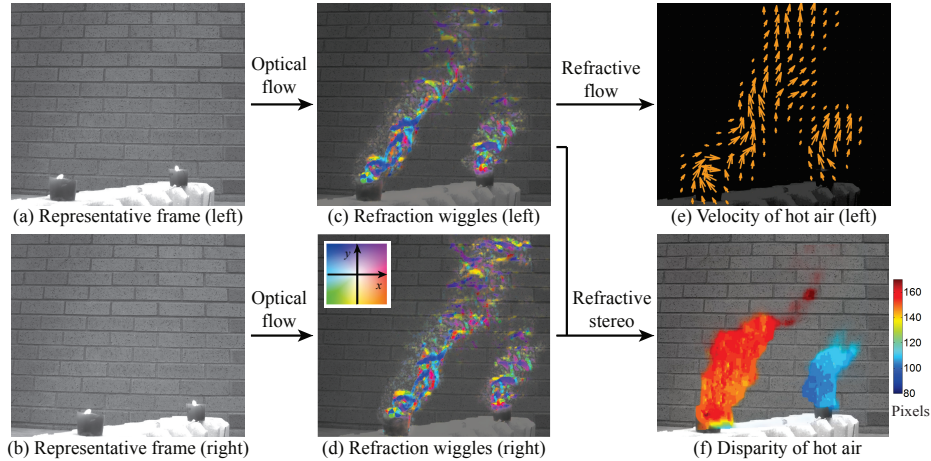
(a) Representative frame (left)	(c) Refraction wiggles (left)

Optical flow →

Refractive flow →

(e) Velocity of hot air (left)

(b) Representative frame (right)

Optical flow →

(d) Refraction wiggles (right)

Refractive stereo →

(f) Disparity of hot air

Fig. 1. Measuring the velocity and depth of imperceptible candle plumes from standard videos. The heat rising from two burning candles (a, b) cause small distortions of the background due to light rays refracting as they travel from the background to the camera passing through the hot air. Methods such as synthetic Schlieren imaging (c, d) are able to *visualize* those small disturbances and reveal the heat plume, but are unable to *measure* its actual motion. We show that, under reasonable conditions, the refraction patterns (observed motions) move coherently with the refracting fluid, allowing to accurately measure the 2D motion of the flow from a monocular video (e), and the depth of the flow from a stereo sequence (f). The full sequence and results are available in the supplementary material.

However, wiggles are related to the fluid motion in nontrivial ways, and have to be processed by appropriate algorithms.

More specifically, measuring refractive flow from video poses two main challenges. First, since the air or fluid elements are transparent, they cannot be observed directly by a camera, and the position of intensity texture features is not directly related to the 3D position of the fluid elements. We therefore cannot apply standard motion analysis and 3D reconstruction techniques directly to the intensity measurements. Our main observation in this paper is that such techniques are still applicable, but in a *different feature space*. Specifically, while *intensity* features result from a background layer and their location is not directly related to the fluid layer, *motion* features (wiggles) correspond to the 3D positions and motion of points on the transparent fluid surface. The movement of those wiggles between consecutive frames (i.e. the motion of the observed motions) is an indicator of the motion of the transparent fluid, and the disparity of these motion features between viewpoints is a good cue for the depth of the fluid surface.

Following this observation, we derive algorithms for the following two tasks: 1) tracking the movement of refractive fluids in a single video, and 2) recovering the 3D position of points on the fluid surface from stereo sequences. Both these algorithms are based on the *refractive constancy* assumption (analogous to the brightness constancy assumption of ordinary motion analysis): that intensity variations over time (the wiggles) are explained by the motion of a constant, non-uniform refractive field. This distortion is measured by computing the wiggle features in an input video, and then using those features to estimate the motion and depth of the fluid, by matching them across frames

and viewpoints. In this paper, we focus on estimating the fluid motion and depth from stationary cameras, assuming a single, thin refractive layer between the camera and the background.

The second challenge in measuring refractive flow from video is that the distortion of the background caused by the refraction is typically very small (on the order of 0.1 pixels) and therefore hard to detect. The motion features have to be extracted carefully to overcome inherent noise in the video, and to properly deal with regions in the background that are not sufficiently textured, in which the extracted motions are less reliable. To address these issues, we develop probabilistic refractive flow and stereo algorithms that maintain estimates of the uncertainty in the optical flow, the refractive flow, and the fluid depth.

The proposed algorithms have several advantages over previous work: (1) a simple setup that can be used outdoors or indoors, (2) they can be used to visualize and measure air flow and 3D location directly from regular videos, and (3) they maintain estimates of uncertainty. To our knowledge, we are the first to provide a complete pipeline that measures the motions and reconstructs the 3D location of refractive flow directly from videos taken in natural settings.

All the videos and results presented in this paper, together with additional supplementary material, are available on: http://people.csail.mit.edu/tfxue/proj/fluidflow/.

2   Related Work

Techniques to visualize and measure fluid flow can be divided into two categories: those that introduce tracers (dye, smoke or particles) into the fluid, and those that detect the refraction of light rays through the fluid, where variations in index of refraction serve as tracers.

In tracer-based methods, the fluid motion is measured by tracking particles introduced into the fluid, a technique called particle image velocimetry (PIV). Traditional PIV algorithms are based on correlating the particles between image patches [1]. Recently, optical flow algorithms were used with PIV images [19, 20], and different regularization terms were proposed to adapt the optical flow methods to track fluids rather than solid objects [13].

In tracer-free methods, Schlieren photography is a technique to visualize fluid flow that exploits changes in the refractive index of a fluid. It works by using a carefully aligned optical setup to amplify deflections of light rays due to refraction [21, 22, 24, 25]. To measure the velocity of the fluid, researchers have proposed Schlieren PIV [14, 5], in which the motion of a fluid is recovered by tracking vortices in Schlieren photographs using PIV correlation techniques. These methods still require the optical setup for Schlieren photography, which can be expensive and hard to deploy outside a lab.

The initial stage of our approach is most similar to a technique called Background Oriented Schlieren (BOS, a.k.a Synthetic Schlieren) [6, 8, 10, 12, 15, 18]. The optical setup in Schlieren photography is replaced by optical flow calculations on a video of a fluid in front of an in-focus textured background. The refraction due to the fluid motion is recovered by computing optical flow between each frame of the video and an undistorted reference frame. Most previous BOS techniques focus on visualizing, not measuring, the fluid flow, and produce visualizations of the flow similar to the results shown in Fig. 1(c,d). Atcheson et al. use BOS tomography to recover the volumetric 3D

shape of air flow [6]. However, their technique requires a camera array that covers $180°$ of the air flow of the interest, making the whole system difficult to use outdoors. To the best of our knowledge, ours is the simplest camera-based system–only requiring stereo video–that can measure the motion and 3D location of air flow.

While artificial backgrounds are often used when performing BOS [6], Hargather et al. [12] showed that natural backgrounds, such as a forest, are sufficiently textured for BOS, allowing for the visualization of large-scale air flows around planes and boats. To address the fact that the background must be in focus, Wetzstein et al. [28] introduced light field BOS imaging in which the textured background is replaced by a light field probe. However, this comes at the cost of having to build a light field probe as large as the flow of interest.

Another limitation of the aforementioned BOS algorithms is that they require a reference frame, which has no air flow in front of it. To avoid having to capture such a reference frame, Raffel et al. proposed background-oriented stereoscopic Schlieren (BOSS) [17] , where images captured by one camera serve as reference frames for the other camera. It is important to note that BOSS uses stereo setup for a different purpose than our proposed refractive stereo algorithm: the acquisition of a background image, not depth. BOSS uses stereo to achieve a reference-frame-free capture while we use it for depth recovery. Moreover, an important weakness of all of these BOS algorithms is that they require a background that has a strong texture. While texture also helps our algorithms, the probabilistic framework we propose also allows for more general backgrounds.

Complementary to the visualization of BOS, several methods have been developed to recover quantitative information about a scene from refraction. Several authors have shown that it is possible to recover the shape of the surface of an air-water interface by filming a textured background underneath it [9, 16, 30]. Alterman et al. [4] proposed to recover refraction location and strength from multiple cameras using tomography. Tian et al. [27] showed that atmospheric turbulence provides a depth cue as larger deflections due to refraction typically correspond to greater depths. Wetzstein et al. [29] proposed to recover the shape of a refractive solid object from light field distortions. Alterman et al. [2, 3] showed that it is possible to locate a moving object through refractive atmospheric turbulence or a water interface.

3   Refraction Wiggles and Refractive Constancy

Gradients in the refractive properties of air (such as temperature and shape) introduce visual distortions of the background, and changes in the refractive properties over time show up as minute motions in videos. In general, several factors may introduce such changes in refractive properties. For example, the refractive object may be stationary but change in shape or temperature. In this paper, however, we attribute those changes to non-uniform refractive flow elements *moving* over some short time interval $[t, t + \Delta t]$. We assume that for a small enough Δt, a refractive object maintains its shape and temperature, such that the observed motions in the video are caused mostly by the motion of the object (and the object having a non-uniform surface, thus introducing variation in the refractive layer). While this is an assumption, we found it to hold well in practice as long as the video frame-rate is high enough, depending on the velocity of the flow being captured (more details in the experimental section below). In this section,
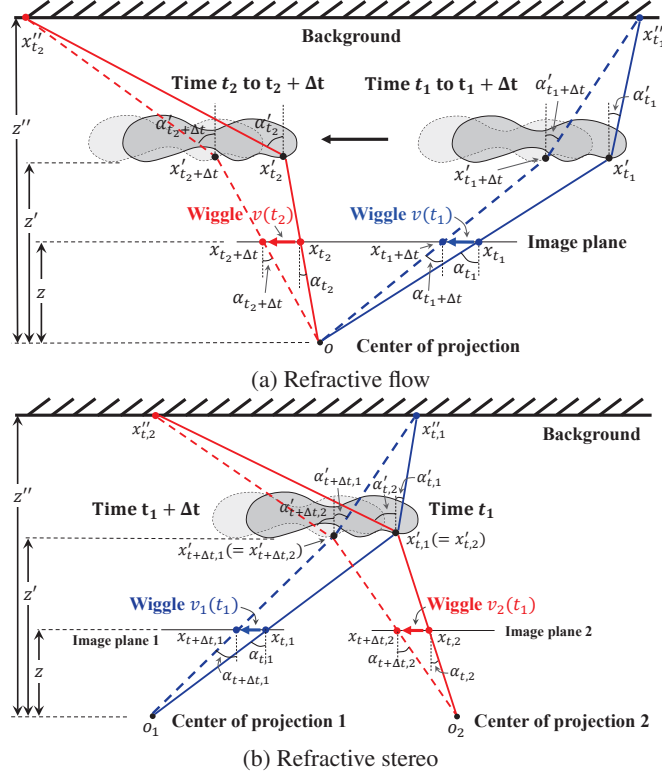
(a) Refractive flow

(b) Refractive stereo

Fig. 2. Refractive distortions (wiggles) in a single view (a) and multiple views (b). A single, thin refractive layer is moving between one or more video cameras and a background. As the refractive fluid moves between time t_i (solid lines) and time $t_i + \Delta t$ (dashed lines), changes in the refractive patterns move points on the background (shown in blue and red) to different positions on the image plane, generating the observed *"wiggles"* (red and blue arrows). The direction of the wiggles on the image plane can be arbitrary, but they are *consistent* over short time durations and between close viewpoints as the fluid moves (see text). By tracking the wiggles over time we can recover the projected 2D fluid motion (a), and by stereo-fusing the wiggles between different views, we can recover the fluid's depth (b). **Note:** as discussed in the text, wiggle constancy holds if the refraction, the motion of the object and the baseline between the cameras are small. In these illustrations we exaggerated all these quantities for clarity.

we establish the relation between those observed motions in one or more cameras and the motion and depth of refractive fluids in a visual scene.

To understand the basic setup, consider Fig. 2. A video camera, or multiple video cameras, are observing a static background through a refractive fluid layer (such as hot air). In this paper, we assume that the cameras are stationary, and that a single, thin and moving refractive layer exists between the camera and the background. We use the notation $x_{t,j}$ for points on the j'th camera sensor plane at time t, $x'_{t,j}$ for points on the (locally planar) fluid, and $x''_{t,j}$ for background points, where t denotes the time index. The depths of these planes are denoted by z, z', z'', respectively. We denote the camera centers by o_j, and denote by $\alpha_{t,j}, \alpha'_{t,j}$ the angles between the optical axis to the ray

from the j'th camera center to points on the image plane and background, respectively. For brevity, we will omit the subscript j in the case of a single camera.

Now consider the stereo setup in Fig. 2(b). An undistorted ray from the camera center o_j to points $x_{t,j}, x'_{t,j}$ on the j'th camera plane has angle (assuming small $\alpha_{t,j}$, such that $\tan \alpha_{t,j} \approx \alpha_{t,j}$)

$$\alpha_{t,j} = (x_{t,j} - o_j)/z = (x'_{t,j} - o_j)/z'. \tag{1}$$

This ray is distorted as it passes through the fluid, and transfers into angle $\alpha'_{t,j}$. The exact distortion is determined by Snell's law. As is common in the optics literature, if the difference between the incident and refraction angles is small, we can use first order paraxial approximations, which imply that Snell's refraction law is effectively an additive constant,

$$\alpha'_{t,j} \approx \alpha_{t,j} + \Delta\alpha_t, \tag{2}$$

where the angle difference $\Delta\alpha_t$ depends only on the local geometric and refractive properties of the fluid around the intersection point x'_t, and is independent of the incoming ray direction $\alpha'_{t,j}$, and in particular independent of the location of the observing camera o_j. The distorted ray then hits the background surface at the point

$$x''_{t,j} = x'_{t,j} + \zeta \cdot \alpha'_{t,j} = x'_{t,j} + \zeta\,(\alpha_{t,j} + \Delta\alpha_t) \tag{3}$$

where $\zeta = z'' - z'$ is the distance between the fluid object and the background.

At a successive time step $t + \Delta t$ the fluid moves (dashed gray blob in Fig. 2(b)). We assume this motion introduces an observed motion of the projection of the background point $x''_{t,j}$ on the image plane, from point $x_{t,j}$ to point $x_{t+\Delta t,j}$. We call that motion a *"refraction wiggle"*, or *"wiggle"* for short. The geometry of the fluid can be complex and predicting the exact path of the light ray mathematically is not straightforward. Nevertheless, We define as $x'_{t+\Delta t,j}$ the point at which a ray connecting $x''_{t,j}$ to the camera center o_j intersects and refracts at the fluid layer.

Let us now fix a point $\bar{x}'_t = x'_{t,j}$ on the fluid surface, and refer to the image $x_{t,j}$ of that same physical point in all cameras. The rays from each camera to that point have different angles $\alpha'_{t,j}$, and as a result they intersect the background layer at different points $x''_{t,j}$. Thus, the texture intensity observed by each camera at the points $x_{t,j}$ can be arbitrarily different. Therefore, we cannot match intensities between the cameras if we wish to match features on the fluid layer and recover their depth. However, we observe that the wiggles correspond to features on the fluid layer rather than on the background, which allows us to stereo-fuse them to estimate the fluid depth. For this we need to show that despite the fact that, without loss of generality, $x''_{t,1} \neq x''_{t,2}$, the two points are refracted at time $t + \Delta t$ via the same fluid point $x'_{t+\Delta t,1} = x'_{t+\Delta t,2}$. To see that, we use Eq. (3) to express

$$x'_{t+\Delta t,j} - x'_{t,j} = -\zeta(\alpha_{t+\Delta t,j} - \alpha_{t,j}) - \zeta(\Delta\alpha_{t+\Delta t} - \Delta\alpha_t), \tag{4}$$

and from Eq. (1), we have

$$\alpha_{t+\Delta t,j} - \alpha_{t,j} = (x'_{t+\Delta t,j} - x'_{t,j})/z'. \tag{5}$$

Plugging Eq. (5) in Eq. (4) we thus have

$$x'_{t+\Delta t,j} - x'_{t,j} = c \cdot (\Delta\alpha_{t+\Delta t} - \Delta\alpha_t), \tag{6}$$

where $c = -(z'' - z')z'/z''$. Therefore, since the terms on the RHS of Eq. (6) are all camera-independent (under our setup and assumptions), if $x'_{t,1} = x'_{t,2} = \bar{x}'_t$, we conclude that $x'_{t+\Delta t,1} = x'_{t+\Delta t,2}$ and the wiggles in both cameras are equal.

This *refraction constancy* can be shown similarly for the monocular case (Fig. 2(a)). That is, if the fluid object is moving at constant velocity over a short spatio-temporal window, then the observed refraction wiggles move coherently with the fluid object between consecutive frames. This stands in contrast to the fact that the observed motions themselves (the wiggles) are *unrelated* to the actual fluid motion, and in fact can point in opposite directions (Fig. 1, Fig. 4). We refer the interested reader to our supplementary material for the derivation, as well as for a more detailed version of the derivation we gave here. Also note that in Eq. (1) and Eq. (3) we assumed the viewing angle α is small to simplify the derivation, however refraction constancy is not restricted by the viewing angle (not to be confused with $\Delta\alpha$, which does need to be small). In the supplementary material we give the derivation without making this assumption.

The practical implication of the observations we made in this section is that **if we wish to match the projection of a point on the fluid object across frames or viewpoints, we can match the observed wiggle features**. That is, while the position of *intensity texture* features is unrelated to the fluid 3D structures, the position of *wiggle* features respect features on the fluid surface, and can serve as an input feature for optical flow and stereo matching algorithms. In the following sections we will use this to derive optical flow and stereo algorithms for tracking and localizing refractive fluids.

4   Refractive Flow

The goal of fluid motion estimation is to recover the projected 2D velocity, $\mathbf{u}(x,y,t)$, of a refractive fluid object from an observed image sequence, $I(x,y,t)$ (Fig. 2(a)). As discussed in the previous section, the wiggle features $\mathbf{v}(x,y,t)$, not the image intensities $I(x,y,t)$, move with the refractive object. Thus, there are two steps in estimating the fluid's motion: 1) computing the wiggle features $\mathbf{v}(x,y,t)$ from an input image sequence $I(x,y,t)$, and 2) estimating the fluid motion $\mathbf{u}(x,y,t)$ from the wiggle features $\mathbf{v}(x,y,t)$. We will now discuss each of these steps in turn.

Computing wiggle features. We use optical flow to compute the wiggles $\mathbf{v}(x,y,t)$ from an input image sequence. Recall the *brightness constancy* assumption in optical flow is that any changes in pixel intensities, I, are assumed to be caused by a translation $\mathbf{v} = (v_x, v_y)$ over spatial horizontal or vertical positions, x and y, of the image intensities, where v_x and v_y are the x and y components of the velocity, respectively. That is,

$$I(x,y,t+dt) = I(x - v_x dt, y - v_y dt, t). \tag{7}$$

Based on this brightness constancy equation, a traditional way to calculate the motion vector v is to minimize the following optical flow equation [7]:

$$\tilde{\mathbf{v}} = \arg\min_{\mathbf{u}} \sum_{\mathbf{x}} \alpha_1 \left\| \frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} \right\|^2 + \alpha_2 \left(\left\| \frac{\partial \mathbf{v}}{\partial x} \right\|^2 + \left\| \frac{\partial \mathbf{v}}{\partial y} \right\|^2 \right), \tag{8}$$

where α_1 and α_2 are weights for the data and smoothness terms, respectively.

(a) Source
(*candle*)

(b) Wiggles, visualized in two ways   (c) Refractive flow (d) Prob. refractive flow
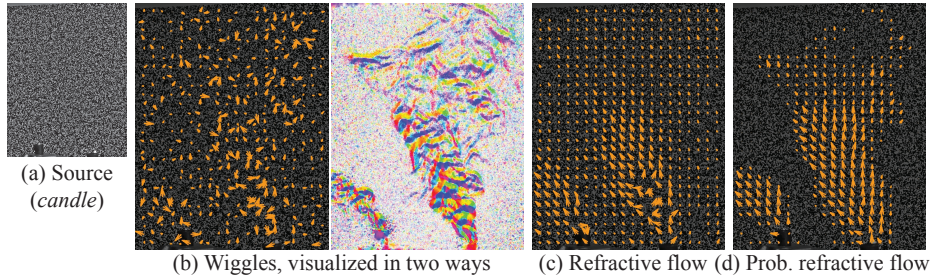
Fig. 3. Example result of our refractive flow algorithm (single view) on a sequence of burning candles. Wiggle features (b) are extracted from the input video (a). Notice how the directions of the wiggles (observed motions) are arbitrary and inconsistent with the air flow direction (the right visualization in (b) uses the same wiggle color coding as in Fig. 1). (c) and (d) show refractive flows calculated by two algorithms, refractive flow and probabilistic refractive flow, respectively.

Estimating fluid motion. Let u_x and u_y be the x and y components of the fluid's velocity as seen in the image. Following Section 3, we thus define the *refractive constancy* equation for a single view sequence as

$$\mathbf{v}(x, y, t + \Delta t) = \mathbf{v}(x - u_x \Delta t, y - u_y \Delta t, t). \tag{9}$$

Notice that refractive constancy has the exact same form as brightness constancy (Eq. 7), except that the features are the wiggles, $\mathbf{v}$, rather than the image intensities, I. This implies that running an optical flow algorithm on the wiggle features $\mathbf{v}$ (i.e. the motion of the motion features), will yield the fluid motion $\mathbf{u}$.

Formally, we calculate the fluid motion $\mathbf{u}$ by minimizing the following equation:

$$\tilde{\mathbf{u}} = \arg\min_{\mathbf{u}} \sum_{\mathbf{x}} \beta_1 \left\| \frac{\partial \tilde{\mathbf{v}}}{\partial x} u_x + \frac{\partial \tilde{\mathbf{v}}}{\partial y} u_y + \frac{\partial \tilde{\mathbf{v}}}{\partial t} \right\|^2 + \beta_2 \left(\left\| \frac{\partial \mathbf{u}}{\partial x} \right\|^2 + \left\| \frac{\partial \mathbf{u}}{\partial y} \right\|^2 \right). \tag{10}$$

This is similar to the Horn-Schunck optical flow formulation, except that we use L_2-norm for regularization, as opposed to robust penalty functions such as L_1-norm traditionally used by optical flow methods. This is because fluid objects, especially hot air or gas, do not have clear and sharp boundaries like solid objects. We use a multi-scale iterative algorithm to solve Eq. 10, as is common in the optical flow literature.

Fig. 3 demonstrates a result of this refractive flow algorithm, when applied to a video of burning candles. First, wiggle features (Fig. 3(b)) are extracted from the input video (Fig. 3(a)). Since wiggle features move coherently with the air flow, the algorithm correctly estimates the motion of the thermal plume rising from the candle. Notice that the observed motions (wiggles) have arbitrary directions, yet the estimated refractive flow is much more coherent.

Such processing is very sensitive to noise, however, as can be seen in Fig. 3(b) (the noise is more obvious in an enlarged image). The problem is even more severe for less textured backgrounds. This motivates the probabilistic formulation, which we will now describe.

4.1   Probabilistic Refractive Flow

We seek to estimate both the refractive flow, and its uncertainty. Consider a background that is smooth in the x direction and textured in the y direction. Due to the aperture

problem [11], the flow in the x direction may be dominated by noise, while the optical flow in the y direction can be clean. Knowing the uncertainty in the flow allows uncertain estimates to be down-weighted, increasing the robustness of the algorithm.

To find the variance of the optical flow, let us reformulate Eq. (8) as a posterior distribution:

$$P(\mathbf{v}|I) = \exp\left(-\sum_{\mathbf{x}} \alpha_1 \left\|\frac{\partial I}{\partial x}v_x + \frac{\partial I}{\partial y}v_y + \frac{\partial I}{\partial t}\right\|^2 + \alpha_2 \left\|\frac{\partial \mathbf{v}}{\partial x}\right\|^2 + \alpha_2 \left\|\frac{\partial \mathbf{v}}{\partial y}\right\|^2\right). \quad (11)$$

Here, $P(\mathbf{v}|I)$ is a Gaussian distribution, and the mean of $P(\mathbf{v}|I)$ is equal to the solution of the original optical flow equation (8). With this formulation, we can also calculate the variance of the optical flow (the wiggle features). Please refer to the supplementary material for the detailed calculation. Let $\tilde{\mathbf{v}}$ and Σ_v be the mean and covariance, respectively, of the wiggle features computed from Eq. (11).

Then, with the variance of the wiggle features, we can reweight the fluid flow equation as follows:

$$\tilde{\mathbf{u}} = \arg\min_{\mathbf{u}} \sum_{\mathbf{x}} \beta_1 \left\|\frac{\partial \tilde{\mathbf{v}}}{\partial x}u_x + \frac{\partial \tilde{\mathbf{v}}}{\partial y}u_y + \frac{\partial \tilde{\mathbf{v}}}{\partial t}\right\|_{\Sigma_v}^2 + \beta_2 \left(\left\|\frac{\partial \mathbf{u}}{\partial x}\right\|^2 + \left\|\frac{\partial \mathbf{u}}{\partial y}\right\|^2\right) + \beta_3\|\mathbf{u}\|^2, \quad (12)$$

where $\|\mathbf{x}\|_{\Sigma}^2 = \mathbf{x}^{\mathsf{T}}\Sigma^{-1}\mathbf{x}$ is the squared Mahalanobis distance. In this formulation, the data term is reweighted by the variance of the optical flow to robustly estimate the fluid motion: wiggle features with less certainty, such as motions measured in regions of low-contrast, or of flat or one-dimensional structure, will have lower weight in the fluid flow equation. To increase the robustness, we also penalize the magnitude of $\mathbf{u}$ to avoid estimating spurious large flows. Including the uncertainty information leads to more accurate estimation of the fluid motion, as shown in Fig. 3(d).

In practice, calculating the covariance matrix precisely for each pixel is computationally intractable, as we need to compute the marginal probability distribution for each optical flow vector. To avoid this calculation, we concatenate all the optical flow vectors into a single vector and compute its covariance. See the supplementary material for the details. Also, notice that the fluid flow equation (12) still has a quadratic form, so we can model the posterior distribution of the fluid flow $\mathbf{u}$ as a Gaussian distribution, and compute its variance. This variance serves as a confidence measure in the estimated fluid motion.

5   Refractive Stereo

The goal of fluid depth estimation is to recover the depth $z'(x, y)$ of a refractive fluid object from a stereo sequence $I_L(x, y, t)$ and $I_R(x, y, t)$ (Fig. 2(b)). Following Section 3, we can find the depth of the fluid object by stereo-matching the refraction wiggles from the left and right views. Therefore, we first use the algorithm discussed in the previous section to calculate the mean and variance of the optical flows in the left and right views, $\mathbf{v}_L \sim N(\tilde{\mathbf{v}}_L, \Sigma_L)$ and $\mathbf{v}_R \sim N(\tilde{\mathbf{v}}_R, \Sigma_R)$, respectively. We then use a discrete Markov Random Field (MRF), commonly used in the stereo literature [23], to regularize the depth estimates.

Formally, let x_L and x_R be the projection of a point on the fluid object onto the left and right image plane, respectively, and define disparity as $d = x_L - x_R$. We first solve the disparity map by minimizing the objective function

$$\tilde{\mathbf{d}} = \min_{\mathbf{d}} \sum_{x,y} f(\mathbf{v}_R(x,y), \mathbf{v}_L(x + \mathbf{d}(x,y), y)) + \alpha \left(\left\| \frac{\partial \mathbf{d}}{\partial x} \right\|^2 + \left\| \frac{\partial \mathbf{d}}{\partial y} \right\|^2 \right), \quad (13)$$

where $f(\mathbf{v}_R, \mathbf{v}_L)$ is the data term based on the observed wiggles $\mathbf{v}_R$ and $\mathbf{v}_L$, and the last two terms regularize the disparity field. We found that using the L_2-norm for regularization generates better results overall, better explaining the fuzzy boundaries of fluid refractive objects (similar to what we observed for estimating the optical flow in Section 4).

As with the refractive flow, we weigh the data term by the variance of the optical flow to make the depth estimation robust to points in a scene where the extracted wiggles are not as reliable. To achieve this, we define the data term, $f(\mathbf{v}_R, \mathbf{v}_L)$, as the log of the covariance between the two optical flows from the left and right views,

$$f(\tilde{\mathbf{v}}_R, \tilde{\mathbf{v}}_L) = \log cov(\mathbf{v}_R, \mathbf{v}_L) = \frac{1}{2} \log |\Sigma_L + \Sigma_R| + \frac{1}{2} \|\overline{\mathbf{v}}_R - \overline{\mathbf{v}}_L\|^2_{\Sigma_L + \Sigma_R}, \quad (14)$$

where $\|\overline{\mathbf{v}}_R - \overline{\mathbf{v}}_L\|^2_{\Sigma_L + \Sigma_R} = (\overline{\mathbf{v}}_R - \overline{\mathbf{v}}_L)^{\intercal}(\Sigma_L + \Sigma_R)^{-1}(\overline{\mathbf{v}}_R - \overline{\mathbf{v}}_L)$. This data term will assign a higher penalty to inconsistencies in the wiggle matching where the wiggles are more reliable, and a lower penalty where the wiggles are less reliable (typically where the background is less textured and the optical flow is noisy). The choice of the log of the covariance as the data term is not arbitrary. It is the log of the conditional marginal distribution of $\mathbf{v}_L$ and $\mathbf{v}_R$, given that $\mathbf{v}_L$ and $\mathbf{v}_R$ match. See the supplementary material for a detailed derivation.

With calibrated cameras, we can then compute the depth map, $z'(x, y)$, from the disparity map, $d(x, y)$.

6   Experiments

We show several examples of measuring and localizing hot air radiating from several heat sources, such as people, fans, and heating vents. **Please refer to the supplementary materials for the full sequences and results.**

All the videos were recorded in raw format to avoid compression artifacts. To deal with small camera motions or background motions, we subtracted the mean flow for each frame from the optical flow result. For each sequence we captured, we first applied a temporal Gaussian blur to the input sequence to increase SNR. For fast-moving flow, we recorded high-speed videos using a Phantom v10 high-speed camera. In some of the indoor high-speed sequences that required additional lighting, we used a temporal band-stop filter to remove intensity variations from the lighting due to AC power.

6.1   Refractive Flow

Qualitative Results. In Fig. 4 we show several results of refractive flow analysis from a single camera.

We first tested our algorithm in a controlled setting, using a textured background. In *hand*, we took a 30fps video of a person's hand after he held a cup of hot water. The algorithm was able to recover heat radiating upward from the hand. In *hairdryers*, we took a 1000fps high-speed video of two hairdryers placed in opposite directions (the
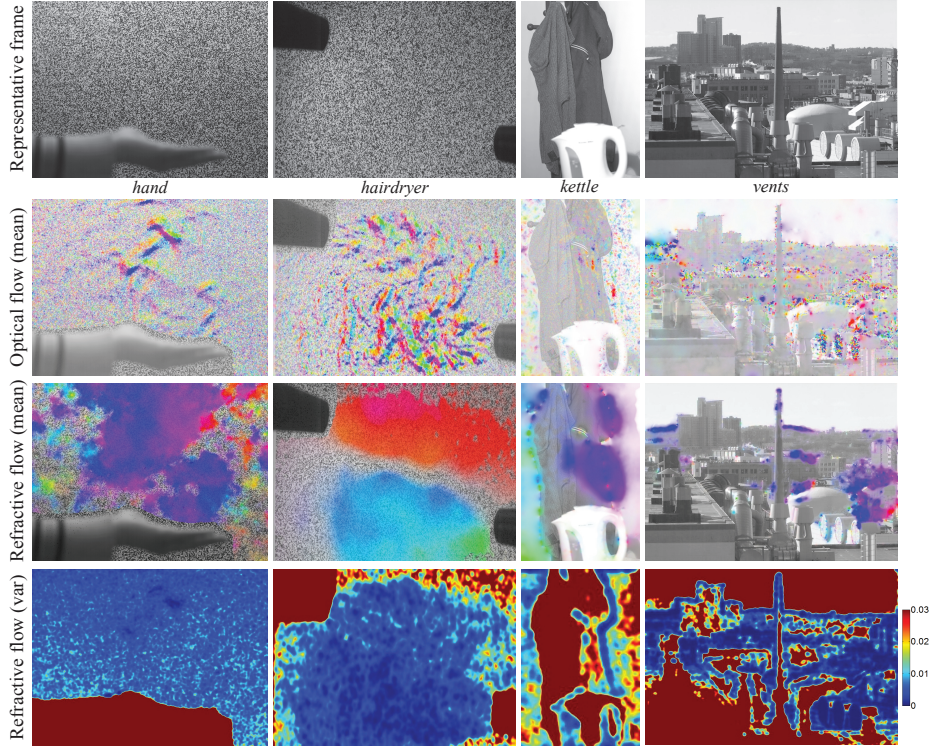
Fig. 4. Refractive flow results. First row: sample frames from the input videos. Second row: the mean of the optical flow for the same representative frames (using the colormap shown in Fig. 1), overlayed on the input images. Third row: the mean of the refractive flow, weighted by the variance. Fourth row: the variance of the estimated refractive flow (the square root of the determinant of the covariance matrix for each pixel). For this visualization, variance values above 0.03 were clipped. The full videos are available in the accompanying material.

two dark shapes in the top left and bottom right are the front ends of the hairdryers), and our algorithm detected two opposite streams of hot air flows.

kettle and *vents* demonstrate the result on videos with more natural backgrounds. In *vents* (700fps), the background is very challenging for traditional Background Oriented Schlieren algorithms (BOS), as some parts of the background are very smooth or contain edges in one direction, such as the sky, the top of the dome, and the boundary of the buildings or chimneys. BOS algorithms rely on the motion calculated from input videos, similar to the wiggle features shown in the second row of Fig. 4, which is very noisy due to the insufficiently textured background. In contrast, the fluid flow (third row of Fig. 4) clearly shows several flows of hot air coming out from heating vents. By modeling the variance of wiggles in the probabilistic refractive flow (bottom row of Fig. 4), most of noise in the motion is suppressed.

Quantitative Evaluation. To quantitatively evaluate the fluid velocity recovered by the proposed algorithm, we tested it on simulated sequences with precise ground truth. We generated a set of realistic simulations of dynamic refracting fluid using Stable
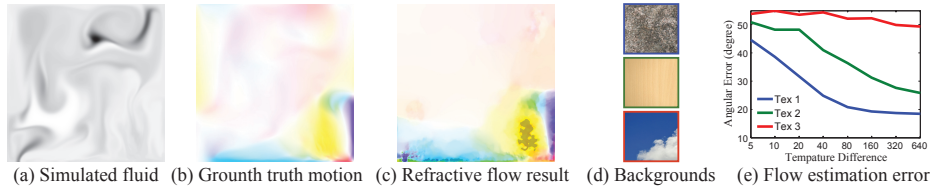
(a) Simulated fluid (b) Grounth truth motion (c) Refractive flow result (d) Backgrounds (e) Flow estimation error

Fig. 5. Quantitative evaluation of refractive flow using synthetic sequences. Simulated fluid density (a) and velocity field (b) were generated by Stable Fluids [26], a physics-based fluid flow simulation technique, and rendered on top of three different textures (d). The recovered velocity field from one of the simulations in which the fluid was at $320°$ Celsius (c) is similar to the ground truth velocity field (b). Quantitative evaluation is given in (e). As expected, larger temperature-related index of refraction differences between the fluid and the background give better flow estimates. The error also increases for backgrounds that do not contain much texture.
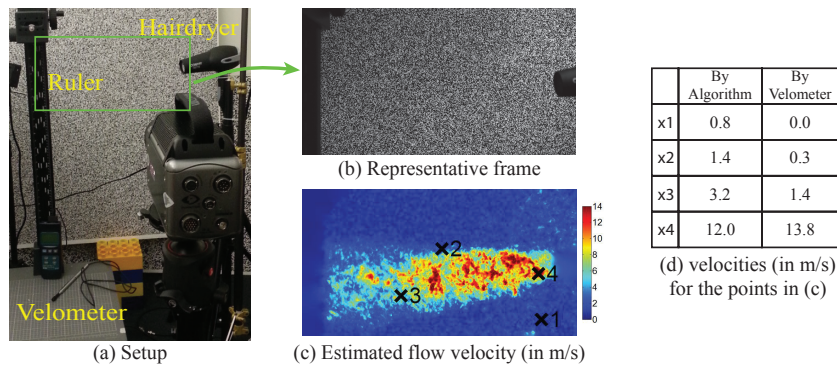

(b) Representative frame

(a) Setup   (c) Estimated flow velocity (in m/s)

	By Algorithm	By Velometer
x1	0.8	0.0
x2	1.4	0.3
x3	3.2	1.4
x4	12.0	13.8

(d) velocities (in m/s) for the points in (c)

Fig. 6. Quantitative evaluation of refractive flow using a controlled experiment. (a) The experiment setup. (b) A representative frame from the captured video. (c) The mean velocity of the hot air blown by the hairdryer, as computed by our algorithm, in m/s. (d) Numerical comparison of our estimated velocities with velocities measured using a velometer, for the four points marked $x_1 - x_4$ in (c).

Fluids [26], a physics-based fluid flow simulation technique, resulting in fluid densities and (2D) ground truth velocities at each pixel over time (Fig. 5(a-b)). We then used the simulated densities to render refraction patterns over several types of background textures with varying spatial smoothness (Fig. 5(d)). To render realistic refractions, we assumed the simulated flow is at a constant distance between the camera and background, with thickness depending linearly on its density (given by the simulation). For a given scene temperature (we used $20°$ Celsius) and a temperature of the fluid, we can compute exactly the amount of refraction at every point in the image. We then apply our algorithm to the refraction sequences. The angular error of the recovered fluid motion at different temperature is shown in (e). All the sequences and results are available in the supplemental material.

To further demonstrate that the magnitude of the motion computed by the refractive flow algorithm is correct, we also performed a controlled experiment shown in Fig. 6. We use a velocimeter to measure the velocity of hot air blown from a hairdryer, and compare it with the velocities extracted by the algorithm. To convert the velocity on the image plane to the velocity in the real world, we set the hairdryer parallel to the imaging

plane and put a ruler on the same plane of the hairdryer. We measured the velocity of the flow at four different locations as shown in Fig. 6(c). Although our estimated flow does not match the velocimeter flow exactly, it is highly correlated, and we believe the agreement is accurate to within the experimental uncertainties.

6.2   Refractive Stereo

Qualitative Results.   We captured several stereo sequences using our own stereo rig, comprised of two Point Grey Grasshopper 3 cameras recording at 50fps. The two cameras are synchronized via Genlock and global shutter is used to avoid temporal misalignment. All videos were captured in 16-bit grayscale.

Several results are shown in Fig. 7. The third row in Fig. 7 shows the disparity map of air flow as estimated by our refractive stereo algorithm, and the forth row shows a 3D reconstruction of the scene according to the estimated depths of the solid and refractive objects. For the 3D reconstructions, we first used a standard stereo method to reconstruct the (solid) objects and the background, and then rendered fluid pixels according to their depth as estimated by our refractive stereo algorithm, colored by their disparities (and weighted by the confidence of the disparity). In *candle*, two plumes of hot airs at different depth are recovered. In *lamps*, three lights were positioned at different distances from the camera: the left one was the closest to the camera and the middle one was the furthest. The disparities of three plumes recovered by the algorithm match the locations of the lights. In *monitor*, the algorithm recovers the location of hot air radiating from the center top of a monitor. We intentionally tilted the monitor such that its right side was closer to the camera to introduce variation in the depth of the air flow. The algorithm successfully detects this gradual change in disparities, as shown in the right column of Fig. 7.

Quantitative Evaluation.   We compared the recovered depths of the refractive fluid layers with that of the heat sources generating them, as computed using a standard stereo algorithm (since the depth of the actual heat sources, being solid objects, can be estimated well using existing stereo techniques). More specifically, we picked a region on the heat source and picked another region of hot air right above the heat source (second row in Fig. 7), and compared the average disparities in these two regions. The recovered depth map of the (refractive) hot air matched well the recovered depth map of the (solid) heat sources, with an average error of less than a few pixels (bottom row in Fig. 7). We show more evaluations of our refractive stereo algorithm using synthetic experiments in the supplementary material.

7   Conclusion

We proposed novel methods for measuring the motion and 3D location of refractive fluids directly from videos. The algorithms are based on wiggle features, which corresponds to the minute distortions of the background caused by changes in refraction as refractive fluids move. We showed that wiggles are locally constant within short time spans and across close-by viewpoints. We then used this observation to propose a refractive flow algorithm that measures the motion of refractive fluids by tracking wiggle features across frames, and a refractive stereo algorithm that measures the depth
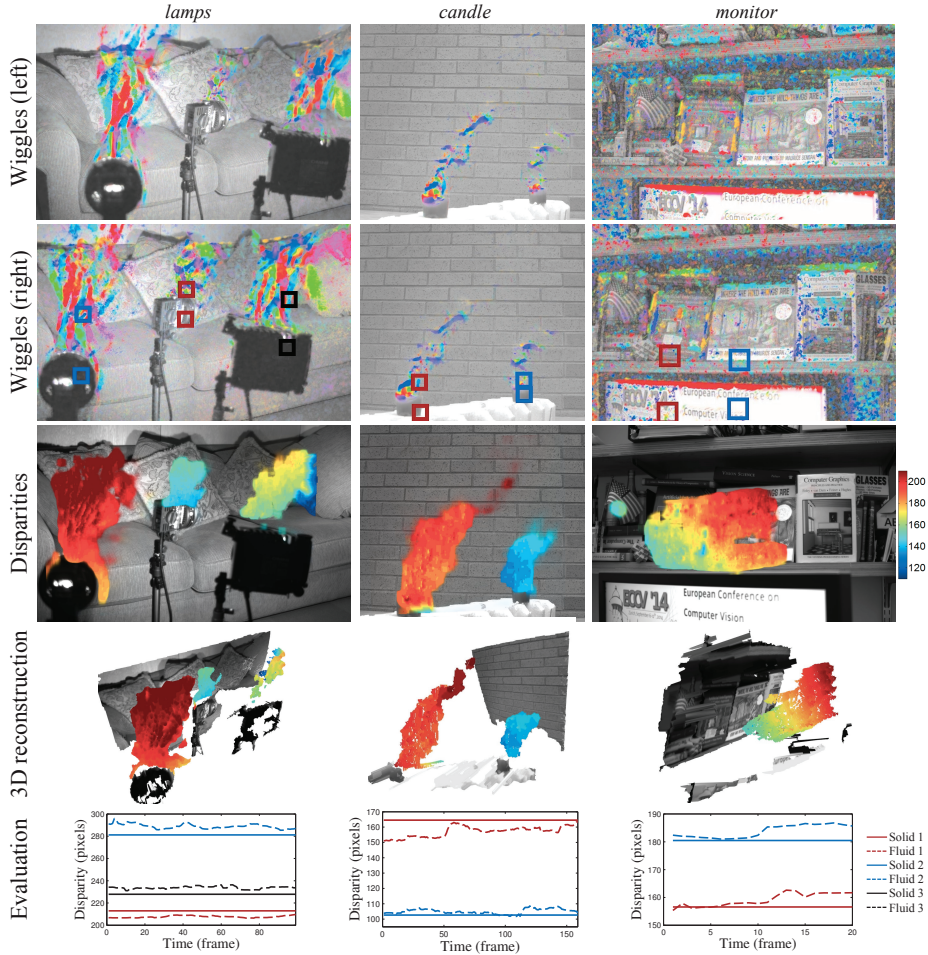
Fig. 7. Refractive stereo results. First and second rows: representative frames from the input videos. The observed wiggles are overlayed on the input frames using the same color coding as in Fig. 1. Third row: the estimated disparity maps (weighted by the confidence of the disparity) of the fluid object. Forth row: 3D reconstruction of the scene, where standard stereo is used for solid objects (and the background), and refractive stereo is used for air flows (the depth was scaled for this visualization). Bottom row: Comparison of our estimated refractive flow disparities, with the disparities of the (solid) heat sources that generated them as computed with a standard stereo algorithm, for the points marked as rectangles on the frames in the second row.

of the fluids by matching wiggle features across different views. We further proposed a probabilistic formulation to improve the robustness of the algorithms by modeling the variance of the observed wiggle signals. This allows the algorithms to better handle noise and more natural backgrounds. We demonstrated results on realistic, physics-based fluid simulations, and recovered flow patterns from controlled and natural monocular and stereo sequences. These results provide promising evidence that refractive fluids can be analyzed in natural settings, which can make fluid flow measurement cheaper and more accessible.

References

1. Adrian, R.J., Westerweel, J.: Particle image velocimetry. Cambridge University Press (2011)
2. Alterman, M., Schechner, Y.Y., Perona, P., Shamir, J.: Detecting motion through dynamic refraction. IEEE Trans. Pattern Anal. Mach. Intell. 35(1), 245–251 (2013)
3. Alterman, M., Schechner, Y.Y., Swirski, Y.: Triangulation in random refractive distortions. In: Computational Photography (ICCP), 2013 IEEE International Conference on. pp. 1–10. IEEE (2013)
4. Alterman, M., Schechner, Y.Y., Vo, M., G, S.N.: Passive tomography of turbulence strength. In: Proc. of European Conference on Computer Vision (ECCV). IEEE (2014)
5. Arnaud, E., Mémin, E., Sosa, R., Artana, G.: A fluid motion estimator for schlieren image velocimetry. In: Proc. of European Conference on Computer Vision (ECCV). pp. 198–210. IEEE (2006)
6. Atcheson, B., Ihrke, I., Heidrich, W., Tevs, A., Bradley, D., Magnor, M., Seidel, H.P.: Time-resolved 3D capture of non-stationary gas flows. ACM Trans. Graph. 27(5), 132:1–132:9 (Dec 2008)
7. Baker, S., Scharstein, D., Lewis, J., Roth, S., Black, M., Szeliski, R.: A database and evaluation methodology for optical flow. International Journal of Computer Vision 92(1), 1–31 (2011)
8. Dalziel, S.B., Hughes, G.O., Sutherland, B.R.: Whole-field density measurements by synthetic schlieren. Experiments in Fluids 28(4), 322–335 (2000)
9. Ding, Y., Li, F., Ji, Y., Yu, J.: Dynamic fluid surface acquisition using a camera array. In: Computer Vision (ICCV), 2011 IEEE International Conference on. pp. 2478–2485 (2011)
10. Elsinga, G., van Oudheusden, B., Scarano, F., Watt, D.: Assessment and application of quantitative schlieren methods: Calibrated color schlieren and background oriented schlieren. Experiments in Fluids 36(2), 309–325 (2004)
11. Fleet, D., Weiss, Y.: Optical flow estimation. In: Handbook of Mathematical Models in Computer Vision, pp. 237–257. Springer US (2006)
12. Hargather, M.J., Settles, G.S.: Natural-background-oriented schlieren imaging. Experiments in Fluids 48(1), 59–68 (2010)
13. Has, P., Herzet, C., Mmin, E., Heitz, D., Mininni, P.D.: Bayesian estimation of turbulent motion. IEEE Trans. Pattern Anal. Mach. Intell. 35(6), 1343–1356 (June 2013)
14. Jonassen, D.R., Settles, G.S., Tronosky, M.D.: Schlieren "PIV" for turbulent flows. Optics and Lasers in Engineering 44(3-4), 190–207 (2006)
15. Meier, G.: Computerized background-oriented schlieren. Experiments in Fluids 33(1), 181–187 (2002)
16. Morris, N.J., Kutulakos, K.N.: Dynamic refraction stereo. In: Computer Vision (ICCV), 2005 IEEE International Conference on. vol. 2, pp. 1573–1580 (2005)
17. Raffel, M., Tung, C., Richard, H., Yu, Y., Meier, G.: Background oriented stereoscopic schlieren (BOSS) for full scale helicopter vortex characterization. In: Proc. of International Symposium on Flow Visualization (2000)
18. Richard, H., Raffel, M.: Principle and applications of the background oriented schlieren (BOS) method. Measurement Science and Technology 12(9), 1576 (2001)

19. Ruhnau, P., Kohlberger, T., Schnrr, C., Nobach, H.: Variational optical flow estimation for particle image velocimetry. Experiments in Fluids 38(1), 21–32 (2005)
20. Ruhnau, P., Schnrr, C.: Optical stokes flow estimation: an imaging-based control approach. Experiments in Fluids 42(1), 61–78 (2007)
21. Ruhnau, P., Stahl, A., Schnrr, C.: On-line variational estimation of dynamical fluid flows with physics-based spatio-temporal regularization. In: Pattern Recognition, Lecture Notes in Computer Science, vol. 4174, pp. 444–454. Springer Berlin Heidelberg (2006)
22. Schardin, H.: Die schlierenverfahren und ihre anwendungen. In: Ergebnisse der exakten Naturwissenschaften, pp. 303–439. Springer (1942)
23. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision 47(1-3), 7–42 (2002)
24. Settles, G.S.: Schlieren and shadowgraph techniques: visualizing phenomena in transparent media, vol. 2. Springer Berlin (2001)
25. Settles, G.S.: The penn state full-scale schlieren system. In: Proc. of International Symposium on Flow Visualization (2004)
26. Stam, J.: Stable fluids. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques. pp. 121–128. SIGGRAPH '99, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA (1999)
27. Tian, Y., Narasimhan, S., Vannevel, A.: Depth from optical turbulence. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. pp. 246–253. IEEE (June 2012)
28. Wetzstein, G., Raskar, R., Heidrich, W.: Hand-held schlieren photography with light field probes. In: Computational Photography (ICCP), 2011 IEEE International Conference on. pp. 1–8 (April 2011)
29. Wetzstein, G., Roodnick, D., Heidrich, W., Raskar, R.: Refractive shape from light field distortion. In: Computer Vision (ICCV), 2011 IEEE International Conference on. pp. 1180–1186 (2011)
30. Ye, J., Ji, Y., Li, F., Yu, J.: Angular domain reconstruction of dynamic 3d fluid surfaces. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. pp. 310–317 (2012)