CrossMark

LETTER

Three-dimensional particle tracking velocimetry using dynamic vision sensors

D. Borer[1] · T. Delbruck[2,3] · T. Rösgen[1]

Abstract A fast-flow visualization method is presented based on tracking neutrally buoyant soap bubbles with a set of neuromorphic cameras. The "dynamic vision sensors" register only the changes in brightness with very low latency, capturing fast processes at a low data rate. The data consist of a stream of asynchronous events, each encoding the corresponding pixel position, the time instant of the event and the sign of the change in logarithmic intensity. The work uses three such synchronized cameras to perform 3D particle tracking in a medium sized wind tunnel. The data analysis relies on Kalman filters to associate the asynchronous events with individual tracers and to reconstruct the three-dimensional path and velocity based on calibrated sensor information.

1 Introduction

Flow visualization remains a valuable tool in aerodynamic testing and optimization. Recent technological developments have facilitated the quantitative analysis of the acquired image data, using for example particle image velocimetry (PIV) or particle tracking velocimetry (PTV).

Both PIV and PTV have evolved from planar into volumetric and time-resolved methods, providing four-dimensional data sets (3D space + time) of instantaneous velocity vectors, see, e.g. Elsinga et al. (2006) and Schanz et al. (2016). The rapid growth of the associated data volume creates new challenges both on the hardware side (data transfer and storage requirements) and for the postprocessing (pattern recognition, volumetric rendering, etc.).

Consequently, there is a parallel, growing interest to reduce the data volume without sacrificing the detail of the achievable results. On the processing side, algorithms such as proper orthogonal decomposition (POD) are employed to enable automatic data reduction towards significant "modes". Similarly, tracking techniques such as PTV are experiencing renewed interest due to the inherent selectivity of the tracking process and the achievable spatial (single particle) resolution.

At the level of the image sensing and acquisition, such an evolution towards efficient data generation is still outstanding. The availability of high quality CMOS sensors has enabled the cost-efficient use of high resolution, high speed cameras, but it also has driven up the bandwidth and storage requirements.

We demonstrate an alternative approach to flow sensing based on 4D particle tracking using a novel type of camera, the dynamic vision sensor (DVS). The main difference to established techniques is the working principle of these cameras, which sense and transmit only changes in brightness, thus leading to a significant data reduction already at the sensing level. The DVS was developed at the Institute of Neuroinformatics, University of Zurich in Switzerland. An in-depth description of the sensing technology is given in Lichtsteiner et al. (2008) and Posch et al. (2014).

✉ T. Rösgen
  roesgen@ethz.ch

1 Institute of Fluid Dynamics, ETH Zurich, 8092 Zurich, Switzerland

2 Institute of Neuroinformatics, University of Zurich, 8057 Zurich, Switzerland

3 ETH Zurich, 8092 Zurich, Switzerland

2 Dynamic vision sensor

The DVS128 is a 128 × 128 array of independent, asynchronous pixels (pixel size 40 μm × 40 μm, fill factor 8.1%) and operates as a temporal contrast sensor. It reacts to the change of light intensity in the observed scene. Each pixel chooses its own exposure time and responds by itself, in contrast to the standard working principle of digital camera chips where exposure and readout are globally controlled. The pixels communicate using "address event representation" (AER) circuits, in which each pixel has two addresses, one for ON-spikes and one for OFF-spikes. An event indicates a positive (ON) or negative (OFF) change in the logarithmic intensity or temporal contrast,

$$C_{\text{temporal}} = \frac{1}{I(t)}\frac{\mathrm{d}I(t)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\ln\left[I(t)\right],$$

where $I(t)$ is the photo-current produced by the photodiode that is proportional to the pixel illumination. An event is generated if the integral of the temporal contrast exceeds a given threshold since the last event from that pixel.

Due to the logarithmic sensitivity, the sensor reacts similarly to local scene reflectance changes independent of the global illumination. This results in a sensor dynamic range of 120 dB. The pixels have a latency as low as 15 μs in good lighting conditions. Compared to standard cameras the data volume produced is significantly reduced and can be processed by standard hardware in real-time despite the high event rate. The DVS output is transferred to a PC as an asynchronous stream of events, each consisting of the pixel address, the time at which the event was registered and the polarity (ON/OFF). The DVS has a time stamp clock rate of 1 MHz, resulting in a maximum temporal resolution of 1 μs. In this study, three DVS cameras were used for volumetric imaging, synchronized such that the sensors' time stamps refer to the same master clock.

An early example of particle tracking with a single DVS is given in Drazen et al. (2011) who focused on 2D tracking with ground truth provided by conventional PTV imaging. The particles in that application were fully resolved with a diameter of about 10 pixels whereas in the current setup, the HFSB are under-resolved such that they appear as streaks with a width of 1 pixel. This necessitates the development of a new tracking algorithm which can handle particle identification, tracking and the 3D reconstruction in a "featureless" environment.

3 Experimental setup

All experiments were conducted in the wind tunnel of the Institute of Fluid Dynamics at ETH Zurich, a low speed facility with a test section area of 2.1 × 3 m^2 and a length of 4 m, capable of a maximum speed of 60 m/s.

The comparatively large test volume necessitates flow tracers large enough to be discernible by the limited resolution imaging system. Here, millimeter-sized helium-filled soap bubbles (HFSB) are employed since they offer low inertia, neutral buoyancy and good visibility. The strong specular reflection on the HFSB surface ideally creates a pixel-sized, localized response in the DVS.

A special HFSB generator was built with the capability to produce individual bubbles on-demand at low rates of up to 10 bubbles/s. For routine applications, this rate can be increased, but the low and consistent bubble seeding density provided a reproducible and steady environment in the test phase. The soap and helium flows in the bubble generator are actively controlled, generating fixed sized bubbles with a lifetime of tens of seconds.

The illumination is provided by 6 white light LED lamp pairs (2 × 50 W each), labelled A–F in Fig. 1. The number and position of the light sources affects the visibility of the bubbles in front of the black wind tunnel walls. Volumetric illumination is achieved by distributing the light sources around the test section. Direct illumination of the background against which the particles are viewed should be avoided to suppress false events.

4 Calibration and measurement procedure

For a 3D recording and reconstruction of the bubble paths, a configuration with three synchronized DVS and wide-angle lenses (Schneider Kreuznach Cinegon f/1.4, 8 mm focal length) was used. The intrinsic parameters (focal length, principal points and lens distortion) and the extrinsic calibration values (position and orientation) for each DVS camera were determined following the approach introduced in Tsai (1986). A panel with 13 small, flashing LEDs arranged in a
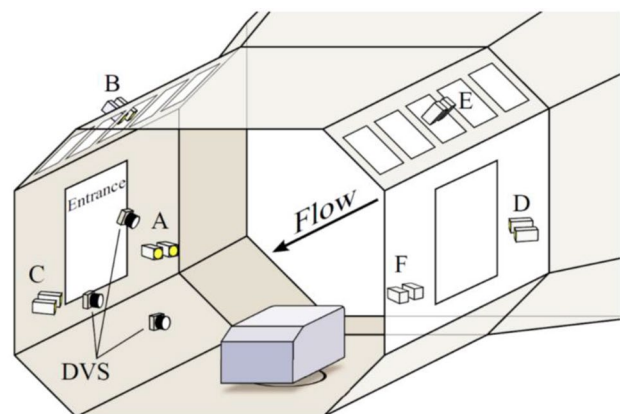

Fig. 1 Test environment (**A–F** LED lighting)

predefined pattern was used as calibration target to provide dynamic visibility for the DVS. The calibration software extracts the LED positions for each camera and calculates the optimal parameter set by minimizing the reprojection errors. Table 1 shows the achievable calibration accuracy for a typical configuration.

The performance is sub-pixel accurate considering the reprojection error. The accuracy in the target space is limited due to the moderate pixel count of the sensors, but is sufficient to localize the millimeter-sized HFSBs.

The actual measurements record the data streams from the three cameras, while the bubble-seeded flow is observed. The bubble generator is manually traversed upstream of the test article to cover the measurement volume. At present the recorded DVS event data are processed offline following the measurement runs, but the algorithmic complexity is low enough to enable real-time processing at a later development stage.

Table 1 Calibration performance for DVS cameras

	Camera 1	Camera 2	Camera 3
# Of target views	124	171	250
Mean reconstruction error (mm)	1.28	1.02	0.94
Mean reprojection error (pixel)	0.21	0.17	0.17

5 Data processing

The data received from the cameras are available as a series of time-tagged spatial events. The processing of this sequential data stream must consider the fact that no synchronicity or direct correlation of the events between the cameras can be assumed.

5.1 Pre-filter step

The raw data are relatively noisy due to shot noise and other types of noise generated in the sensors as described in Lichtsteiner et al. (2008). A first filtering step is implemented for each DVS separately. Valid events must consist of an ON/OFF pair ("pair filter") and should not occur isolated in space–time ("neighbor filter"). A regular HFSB signature exhibits multiple events per pixel, often extending over a few milliseconds after the bubble's passage. This behaviour impairs a correct determination of the bubble's position in time and space, and all event pairs of the same pixel following an initial event during a defined time interval are rejected ("burst filter"). Figure 2 shows the effect of the different filters for linear HFSB motion observed by a DVS, depicted for better visibility in one spatial and the temporal direction only.
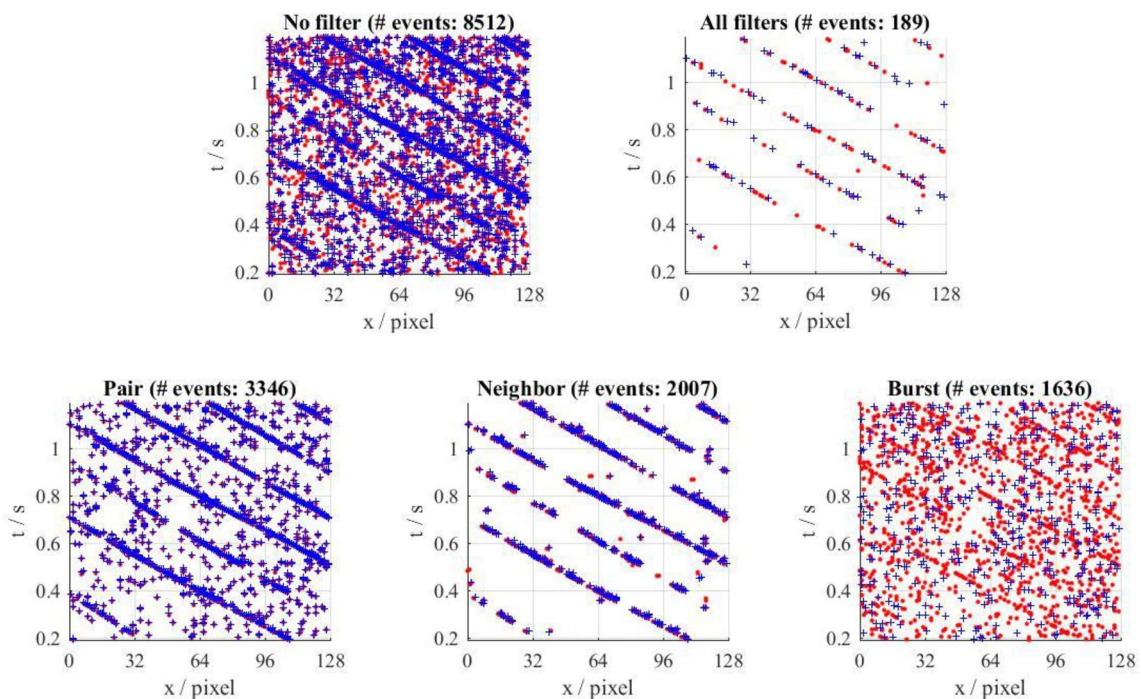

Fig. 2 Pre-filter performance (red: ON events, blue: OFF events); top row, left: raw data, right: all filters applied; bottom row, left-to-right: individual filters applied

5.2 Kalman track filters

The bubble events that pass the pre-filter stage (providing HFSB positions in space and time) are fed into a series of Kalman filters (see, e.g. Reid 1979) to further improve the data consistency and to construct continuous bubble tracks.

A 3D extended Kalman filter computes the evolution of the bubble tracks in the 3D measurement space. For this, each track is described by a 6 component "state vector" $\mathbf{x} = [X, Y, Z, U, V, W]^T$ which contains the current 3D bubble position $[X, Y, Z]^T$—the "head" of the track—and its velocity $[U, V, W]^T$ at that instant. The evolution of the state vector is described by a linear model, the "process" equations,

$$\mathbf{x}(t_{k+1}) = \mathbf{F}(t_k)\mathbf{x}(t_k) + \mathbf{v}(t_k); \qquad \mathbf{F}(t_k) = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

where a term $\mathbf{v}(t_k)$ is included to account for "process noise" with an assumed known covariance matrix $\mathbf{R}$. While the state vector includes the track velocities, only the positions are measured and provide external updates. The 3D bubble coordinates are linked to the 3 camera images ($m = 1, 2, 3$) with their corresponding 2D spatial event (pixel) positions $[x_p^{(m)}, y_p^{(m)}]^T$ through the observation equations $\mathbf{z}^{(m)} = [x_p^{(m)}, y_p^{(m)}]^T = \mathbf{h}^{(m)}(\mathbf{x})$. The relations $\mathbf{h}^{(m)}(\mathbf{x})$ are nonlinear due to the projective transform and the camera distortion models involved. This leads to the "measurement" equations for each camera,

$$\mathbf{z}^{(m)}(t_k) = \mathbf{h}^{(m)}(\mathbf{x}(t_k), t_k) + \mathbf{w}(t_k) \qquad (m = 1, 2, 3), \tag{2}$$

where again a "measurement noise" contribution $\mathbf{w}(t_k)$ with assumed known covariance $\mathbf{Q}$ is included. The process and measurement noise terms must be modelled. Validation tests

determined that the measurement noise can be expressed as a stationary, zero-mean, normal distributed process with a variance of 1 pixel^2. The process noise estimate cannot be predicted in a universal fashion and must be adjusted dependent on the experiment situation. The Kalman filter solution of the process and measurement equations for an optimal estimate of the state vector leads to a set of linear equations which are solved for each new event/data sample (Eq. (3)). It estimates not only the state vector (here: the 3D positions and velocities along a track) but also the noise covariance of that estimate. The time step in the update can remain variable, allowing for a strictly incremental processing of the time-stamped data stream.

$$\underbrace{\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1}}_{\text{predicted state vector}} \qquad \underbrace{\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k}_{\text{predicted state covariance}}$$

$$\underbrace{\mathbf{y}_k = \mathbf{z}_k^{(m)} - \mathbf{h}^{(m)}(\hat{\mathbf{x}}_{k|k-1})}_{\text{measurement residual vector}} \qquad \underbrace{\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k}_{\text{residual covariance}} \qquad \underbrace{\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}}_{\text{"Kalman gain"}}$$

$$\underbrace{\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \mathbf{y}_k}_{\text{updated state vector}} \qquad \underbrace{\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{P}_{k|k-1}}_{\text{updated state covariance}}$$

$$\underbrace{\mathbf{H}_k = \left.\frac{\partial \mathbf{h}^{(m)}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k|k-1}}}_{\text{observation matrix}} \tag{3}$$

The tracking scheme using the 3D Kalman filter is depicted in Fig. 3. The pre-filtered DVS events from all cameras are sorted into a single, time-monotonic sequence. Instead of looking for direct coincidence between different events from different cameras, each event is processed separately. A first test checks whether a new event can be associated with one or more of the 3D tracks projected back onto the corresponding sensor planes (blue). "Association" means here that a new event location lies within a finite-sized ellipsoid around the 2D track head (Bar-Shalom and Fortmann 1988; Reid 1979). Each successful association leads to an update component for the measurement vector and an update of the 3D track (red).

Fig. 3 Event tracking in 3D using an extended Kalman filter and 3 camera inputs
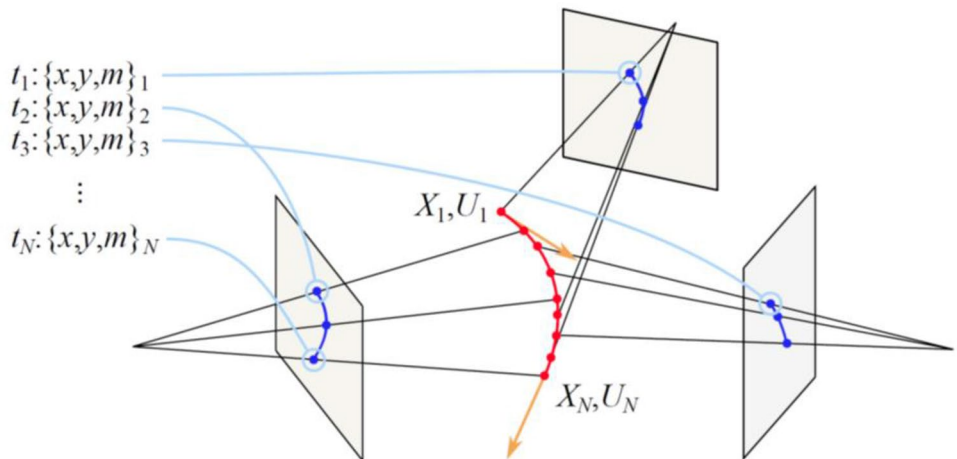
This approach has the advantage that partial information (i.e. events that cannot be associated with all projected tracks simultaneously) can contribute to the 3D filter individually by updating the respective projected displacement component only.

If no association can be found in any of the projected 2D tracks, a new 3D track segment may be initiated. No direct estimation of the new track's 3D root point can be derived from a single event, and reference is made to a separate collection of 2D tracks that are being built up for each camera. These tracks are concurrently computed using 2D Kalman filters and can directly associate an event to its measured position. The filters are implemented in a fashion similar to the 3D filter, but are based on 2D state and measurement vectors. If an event is found to match at least two tracks in different cameras based on the epipolar constraint (Hartley and Zisserman 2003), the 3D position in target space can be triangulated to initiate a new, validated 3D track. To assure temporal coincidence of the candidate 2D tracks under consideration, the positions are interpolated to the same time instant using the continuous trajectory descriptions available from the 2D Kalman filters.

6 Results

The first example presents the flow around the rear face of an Ahmed body (length 1044 mm, height 288 mm), a generic shape often studied as a reference configuration in automotive aerodynamics tests.

Figure 4 presents the raw data in the form of selected individual tracks of HFSBs. In total, 1080 tracks were evaluated. The local velocity magnitude is color coded to enhance the visibility of the wake flow. The bubble tracks extend over a length of up to 1 m, confirming the large-scale
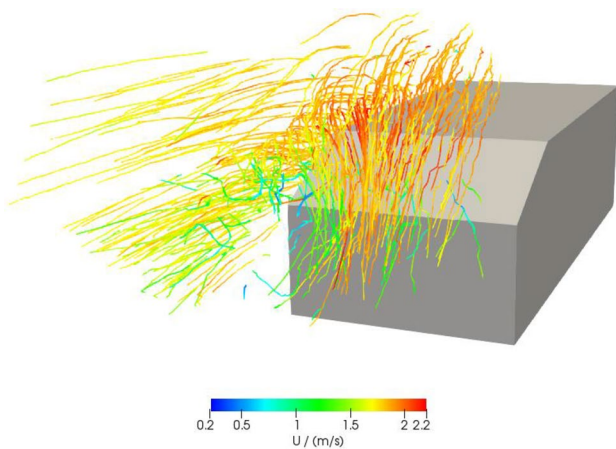
tracking capability of the DVS system. Using a suitable post-processing step (here: "natural neighbor" interpolation in MATLAB®) the Lagrangian track data can be converted into gridded information, as exemplified in Fig. 5.

The second example is chosen to further demonstrate the potential of Lagrangian sampling of large flow volumes. Here, the complex 3D flow field developing behind an air jet (diameter 120 mm) emerging from an orifice into an external cross-flow is studied. A subset of the recorded 3D bubble tracks is shown in Fig. 6, selected to visualize the interaction region between the cross-flow and the jet.

The full set of 8648 HFSB tracks (after filtering) can be used in the interpolation onto a 3D grid. This allows the reconstruction of streamlines and the velocity field in arbitrary planes, as shown in Fig. 7.
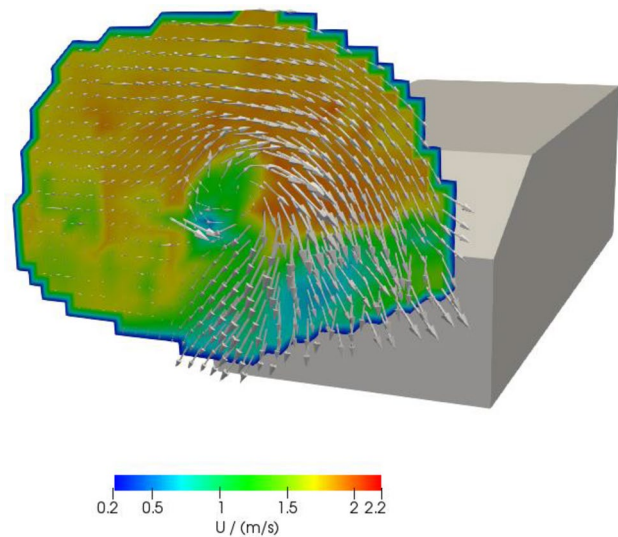

Fig. 5 Flow field in a cut-plane 10 mm behind the rear face of an Ahmed body (color encodes velocity magnitude)

Fig. 4 Flow pattern in the wake of an Ahmed body; selection of HFSB tracks, color coded with local velocity magnitude
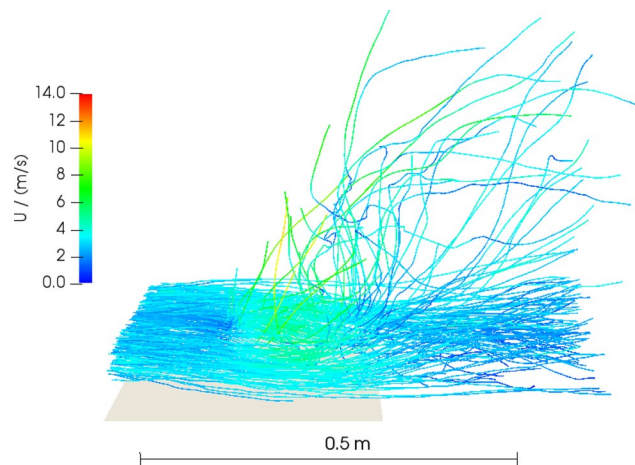

Fig. 6 Selected bubble tracks around a jet in crossflow

In both examples, a stationary flow field has to be assumed in the processing because the complete set of bubble tracks is recorded sequentially by manually guiding the HFSB generator through the upstream flow. This drawback is partially offset by the possibility of injecting bubbles only into "interesting" parts of the flow, reducing the measurement time, data volume and required overall HFSB seeding density.

7 Discussion

The experiments demonstrate the capabilities of DVS cameras for particle tracking in a comparatively large 3D measurement volume. Limitations at present are the low resolution of the DVS and a reduced responsivity at higher speeds. This can be explained by considering the refractory period
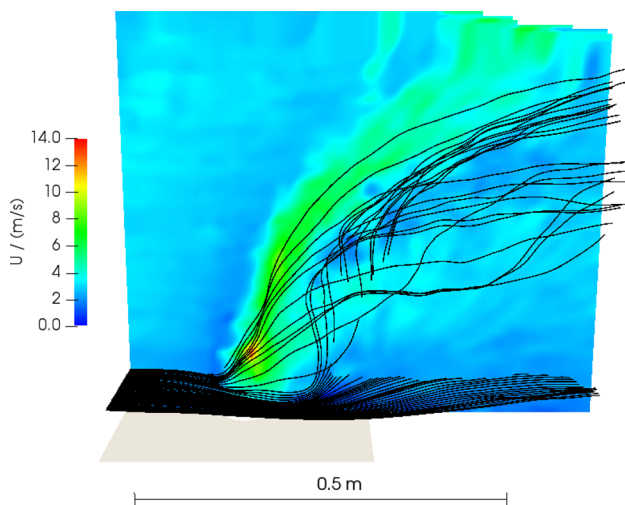

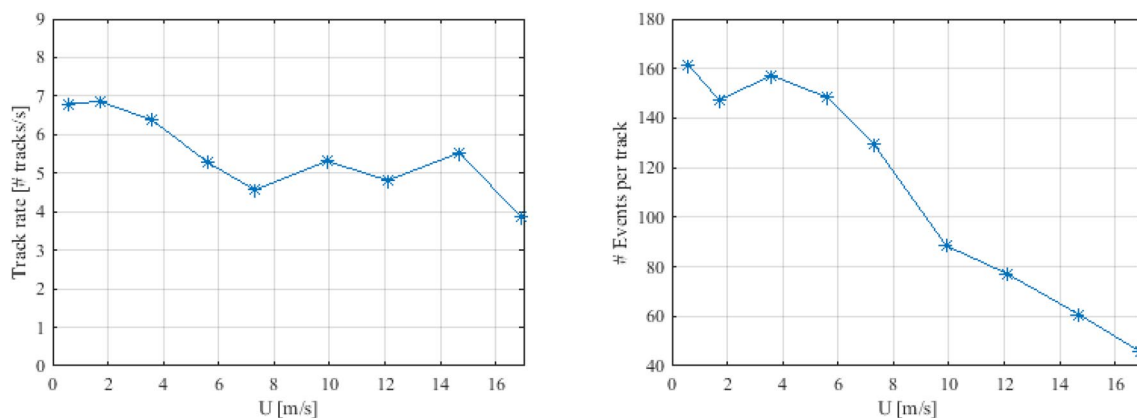**Fig. 7** Reconstructed streamlines and velocity magnitude in the center plane of a jet in crossflow

that is built into the DVS pixels to limit their firing rate and the low pass behavior with a bandwidth on the order of 1 kHz (see Lichtsteiner et al. 2008). Once the passage time of a HFSB across a pixel falls below a time constant $\tau_0 \approx 1$ ms the sensor will compute a reduced contrast and start missing events. Note that this time constant limit refers only to the responsivity of an individual pixel. The rapid passage of a bubble across multiple pixels can be tracked easily because neighboring pixels can produce distinguishable events at the sensor's clock resolution of 1 microsecond. The pixel-level detectability problems may be alleviated using the latest DVS camera (DAVIS346). Preliminary tests which an intermediate DVS type (DAVIS240C, 240×180 pixel, fill factor 22%) have confirmed an improved bubble detectability.

The sequential filtering approach provides a robust basis to process the data. As shown in the headers of Fig. 2 the event-related filters lead to a reduction by a factor of ~ 40 in the number of events to be processed. The resulting number of new tracks being established with the Kalman filters can be compared against the known rate of bubble generation (10 Hz). Ideally, each released bubble should generate a new track. In practice, "false" tracks will increase the number, "missed" tracks decrease it. Furthermore, "broken" tracks may lead to an artificial increase in the number of new tracks. Figure 8 shows this performance measure for the case of the flow across the Ahmed body. In the same figure the number of events per track is shown, a value that should be comparable with the number pixels in a linear track on the sensor surface (~ 128). While the measured track generation rate proves nearly independent of flow velocity, the event-per-track count shows a decrease with velocity due to an increasing number of missed events (see above).

In achieving these performance figures a proper choice of the parameters for the filter statistics, especially the noise models, remains important.

Fig. 8 Tracking statistics: new tracks/second (left), events/track (right)

The tracking system is not designed to provide dense velocity fields from a limited number of observed image frames as would be the case for PIV and PTV recordings. Rather, the aim is to create a cumulative, real-time visualization and analysis of bubble tracks in a steady flow, establishing a streamline field with overlaid velocity information. An upper limit for the number of simultaneous tracks being detected and processed will depend on several factors such as flow velocity, image magnification, or the scene dependent event noise floor. While these factors are difficult to quantify in a universal fashion, a rough estimate for the tolerable/achievable seeding density can be based on the pixel dead-time described above. Assuming a flow velocity of 10 m/s and a camera magnification of, say, 0.01, the passage time of a bubble is (pixel size/magnification/velocity) = 40 µm/0.01/10 m/s = 400 µs. The pixel dead-time of 1 ms discussed above then corresponds to a travel distance of 2.5 pixels and thus implies a minimum bubble spacing of the same order.

8 Conclusion

The measurement system represents an inexpensive and fast flow visualization method suitable for large testing environments. In comparison with other whole field imaging techniques (e.g. tomographic PIV and PTV) the DVS approach significantly reduces the data volume to be processed and appears amenable to real time processing. This may prove relevant in commercial aerodynamics testing where the speed and selectivity of the data acquisition process often represents a decisive performance factor.

References

Bar-Shalom Y, Fortmann TE (1988) Tracking and data association. Mathematics in science and engineering, vol 179. Academic Press, Cambridge

Drazen D, Lichtsteiner P, Häfliger P, Delbrück T, Jensen A (2011) Toward real-time particle tracking using an event-based dynamic vision sensor. Exp Fluids 51(5):1465

Elsinga GE, Scarano F, Wieneke B, van Oudheusden BW (2006) Tomographic particle image velocimetry. Exp Fluids 41:933–947

Hartley R, Zisserman A (2003) Multiple view geometry in computer vision. Cambridge University Press, Cambridge

Lichtsteiner P, Posch C, Delbruck T (2008) A 128 × 128 120 dB 15 µs latency asynchronous temporal contrast vision sensor. IEEE J Solid-State Circuits 43(2):566–576

Posch C, Serrano-Gotarredona T, Linares-Barranco B, Delbruck T (2014) Retinomorphic event-based vision sensors: bioinspired cameras with spiking output. Proc IEEE 102(10):1470–1484

Reid D (1979) An algorithm for tracking multiple targets. IEEE Trans Autom Control 24(6):843–854

Schanz D, Schröder A, Gesemann S (2016) Shake-The-Box: lagrangian particle tracking at high particle image densities. Exp Fluids 57:70

Tsai RY (1986) An efficient and accurate camera calibration technique for recognition. In: Proceedings of IEEE CV PR'86. pp 364–374