

Video-based Bottleneck Detection utilizing Lagrangian Dynamics in Crowded Scenes

Maik Simon, Markus Kuechhold, Tobias Senst, Erik Bochinski, Thomas Sikora
 Communication Systems Group
 Technische Universität Berlin

simon, kuechhold, senst, bochinski, sikora@nue.tu-berlin.de

Abstract

Avoiding bottleneck situations in crowds is critical for the safety and comfort of people at large events or in public transportation. Based on the work of Lagrangian motion analysis we propose a novel video-based bottleneck-detector by identifying characteristic stowage patterns in crowd-movements captured by optical flow fields. The Lagrangian framework allows to assess complex time-dependent crowd-motion dynamics at large temporal scales near the bottleneck by two dimensional Lagrangian fields. In particular we propose long-term temporal filtered Finite Time Lyapunov Exponents (FTLE) fields that provide towards a more global segmentation of the crowd movements and allows to capture its deformations when a crowd is passing a bottleneck. Finally, these deformations are used for an automatic spatio-temporal detection of such situations. The performance of the proposed approach is shown in extensive evaluations on the existing Jülich and AGORASET datasets, that we have updated with ground truth data for spatio-temporal bottleneck analysis.

1. Introduction

The analysis of crowd movements is of importance for the safety and comfort of people in transport infrastructures. Handling crowded scenes during public events (e.g. fan parks, concerts, sport events) is a challenging task for security personnel, police and crisis management teams. Especially the occurrence of bottlenecks during an event can lead to panics due to overcrowding. An automatic bottleneck identification system can aid the operator to prevent critical situations by assessing characteristic crowd-movement patterns. The aim of this work is to identify such events in the spatial and temporal domain.

In computer vision the analysis of high density crowds is performed on macroscopic perspective [17], i.e. the crowd is assessed as a single entity. The behaviors of individuals

in a crowd are dependent on the crowd behavior [17, 9] and modelled by fluid dynamic processes [2, 23, 11, 8]. Hughes work [7] supports the assumption that crowds are a flowing continuum and proposed three main behavioral hypotheses for persons moving in a crowd. In [4], Bain and Bartolo also contemplate pedestrian flows with the help of a hydrodynamic model. Here, the flow behavior of polarized crowds was examined by considering the border movements of the crowd at the start of various marathons. To describe crowd behavior for crowd simulation Still proposed in [21] three main effects: i.) least-effort hypothesis means that people are looking for the least strenuous route ii.) lane formation implies that people walk most easily behind each other and iii.) bottleneck effect occurs at a narrowing point with a significant speed change of the crowd and represents at the same time a critical point.

The following studies investigate the influence of the bottleneck to the behaviour of the crowd before and within the narrow pass. Seyfried *et al.* [15] shows an experimental study in which the flow of unidirectional pedestrian streams through bottlenecks was evaluated. The result was a linear growth of the flow with a simultaneous increase in the width of the bottleneck and the observation of the phenomenon of lane formation within bottlenecks. Krüchten *et al.* [22] recorded a dataset under laboratory conditions, which represents persons with different age, group sizes and social group sizes in case of evacuation through a bottleneck. In the study, the social aspect of passing through a bottleneck was presented, which showed that with increasing social group strength, the flow through the bottleneck increased. The study of Sieben *et al.* [16], showed the influence of the spatial structure and the perception of the participants in comparison to physical measurements.

The data recorded in [10, 15, 16, 22] has been published at the pedestrian dynamic archive and will be denoted as Jülich dataset. Allain *et al.* [3] proposed the AGORASET for crowd behaviour analysis containing corridors, obstacles and escapes. The dataset consists of synthetic rendered images and provides a higher variation and different point

of views of the scenes.

Each scene of the two datasets contains physical bottlenecks which are not always leading to congestion situations in pedestrian movements. Bottleneck situations can also arise through situation-dependent events, whereby an occurring bottleneck is defined by the flow of movement of the persons. For this reason, the datasets were extended with ground truth data that has both spatial and temporal properties. Furthermore a new metric is presented for this spatio-temporal problem, which makes the proposed and future methods comparable.

In the work of Solmaz et al. [19], four additional crowd scene properties (blocking, lane, ring/arch, fountainhead) are detected in addition to bottleneck situations using optical flow. The method performs well for these properties, but has problems with the superposition of motion patterns.

In this work we propose a novel video-based bottleneck detector based on the evaluation of characteristic stowage patterns in crowd-movements by segmenting the crowd flow. The idea of this detector is that physical bottlenecks are related to bottlenecks in the contours of the crowd flow segments. To segment the crowd flow contour we apply long-term analysis based on the Lagrangian framework proposed in [8] and use the Finite Time Lyapunov Exponents (FTLE) field to extract motion boundaries. High ridges in the FTLE field indicate Lagrangian features that are assumed to be located at motion boundaries. In addition we propose a long-term temporal low-pass filtered FTLE to suppress unsteady local features in the Lagrangian field that are caused by heterogeneous motion of the people in the crowd and lead to erroneous crowd flow contours.

The bottleneck location is defined by the center of a point pair, which is found by geometrical and temporal consistency constraints applied to bottleneck candidates. Bottleneck candidates are defined by defects on the contour, i.e. points on the contour with a maximum distance to the contours convex hull. To evaluate the performance of the proposed system we manually annotated a selected set of bottleneck sequences from the synthetic AGORASET and the Jülich dataset.

2. Lagrangian Measures for Bottleneck Detection

The origin of Lagrangian methods lies in the visualization and analysis of unsteady flows and has been proven to be a powerful tool for analyzing computational fluid dynamics for instance to design fluid-dynamic systems. These methods are used to describe non-linear dynamic systems that are represented by a series of time-dependent vector fields. The pioneering work by Ali and Shah [2] first showed that the Lagrangian methodology can be useful for video-based crowd segmentation. Inspired by this work

Kuhn *et al.* proposed in [8] a compact and applicable framework that implements Lagrangian concepts for video analytics. At its core, this framework is based on characterizing motion as a sequence of optical flow fields $\mathbf{v}(\mathbf{x}, t)$ to assemble a time-dependent vector field that encodes the dynamics of the video sequence in space-time of a temporal range $[t_0, t_0 + \tau]$. In this work we will follow this framework where the analysis of the optical flow fields is based on so called path lines [8]. Path lines can be interpreted as traces of massless particles advected in the flow fields. Their computation, i.e. advection, is based on the computation of the flow map $\phi_{t_0}^\tau(\mathbf{x}) = \phi(\mathbf{x}, t_0, \tau)$ which is a core aspect of Lagrangian methods. The flow map defines the mapping of all massless particles at time t_0 seeded at the position \mathbf{x}_0 to their corresponding positions after an integration time τ :

$$\phi_{t_0}^\tau : D \rightarrow D : \phi_{t_0}^\tau(\mathbf{x}_0) = \mathbf{x}(t : t_0, \mathbf{x}_0), \quad (1)$$

t_0 is the so called frame of reference denoting the basis of the projection of the path line properties. The flow map $\phi_{t_0}^\tau$ is constructed by integrating path lines in the optical flow fields over τ time steps, i.e. propagating the massless particles at position \mathbf{x}_t and time t based on the flow vector $\mathbf{v}(\mathbf{x}_t, t)$. Since the optical flow fields are discrete in space and time, trilinear interpolation is applied and the particle position is updated by:

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \tilde{\mathbf{v}}(\mathbf{x}_t, t), \quad (2)$$

where $\tilde{\mathbf{v}}$ denotes the interpolated motion vectors. It is assumed that these path lines characterize the overall dynamics, i.e. motions, and can provide quantitative information about the observed objects in the scene. Instead of considering individual trajectories only, this information can be compactly represented within so called Lagrangian fields. Examples of Lagrangian fields that have been applied for video analysis are the arc length field [8] for segmentation or the direction field for violence detection [12] or action recognition [1].

One specifically popular type of Lagrangian fields are Finite-Time Lyapunov Exponents (FTLE) which quantify the amount of separation between neighboring path lines. With respect to features in non-linear dynamic systems high ridges in the FTLE scalar field are assumed to be in close relationship with Lagrangian Coherent Structures [5] (regions of maximum change over time) that can serve as basic features to capture and quantify advanced motion patterns [6]. In the video domain FTLE fields have been successfully used in crowd segmentation [11, 20, 2], motion anomaly detection [23] and person behaviour analysis [14]. High ridges are assumed to be in close relationship with motion boundaries of physical objects with respect to small and large entities. For the task of crowd segmentation these ridges have shown to be salient and stable features. On those grounds

we want utilize FTLE fields to extract the distinct shape of crowds moving through bottlenecks. The FTLE is derived from the spatial gradients of the flow map. With

$$\nabla\phi_{t_0}^\tau(\mathbf{x}) = \frac{\partial\phi_{t_0}^\tau(\mathbf{x})}{\partial\mathbf{x}} \quad (3)$$

being the spatial gradients of the flow map and

$$\mu_i = \ln\sqrt{\lambda_i(\nabla^T\nabla)}, \quad (4)$$

the FTLE value for integration time τ is obtained as

$$\text{FTLE}^\tau(\mathbf{x}, t_0) = \frac{1}{\tau}\max\{\mu_1, \mu_2\}. \quad (5)$$

∇^T is the transposed of ∇ and $\lambda_i(\nabla^T\nabla)$ denotes the i -th eigenvalue of the symmetric matrix $\nabla^T\nabla$. In general the FTLE can be computed in forward and backward direction resulting in the description of FTLE+ and FTLE-. The forward FTLE field describes regions of repelling LCS, while features in the backward FTLE describe attracting LCS structures over the considered time scope. Only the intersections of FTLE+ and FTLE- ridge structures can segment regions of coherent movement and group invariant moving areas within the motion field.

3. Video-based Bottleneck Detection

With the proposed method we want to localize the physical bottlenecks by analyzing person flow patterns around narrow places. This localization will be based on the segmentation of the crowd flow. We have observed that physical bottlenecks result in bottlenecks in the contour of the crowd flow segments. We assume that the shape of crowd flow segments can be estimated by extracting high ridges in FTLE fields and propose a long-term analysis of the scene since the movements in that area can be very small. So called defects in the contour of the crowd flow shape allow us to detect salient points that restrict the bottleneck.

Our approach will be composed of three major parts: the long-term temporal filtered FTLE fields, the crowd flow contour segmentation and crowd flow contour analysis.

Long-Term Temporal Filtered FTLE: For the calculation of the crowd flow shape, a long-term analysis of the scene is of major interest, since the movements in near the area of the bottleneck can be very small. Figure 1 (top) shows the FTLE+ field of a rather short integration time $\tau = 15$. The ridges related to the crowd margin are relative weak. To assess the crowd margin while the walking speed of the people is low a large integration interval $\tau > 100$ has to be used, which requires a high computational effort. In addition, due to the heterogeneous movement of the people in the crowd each walking person causes a ridge which becomes stronger for large values of τ . However in contrast

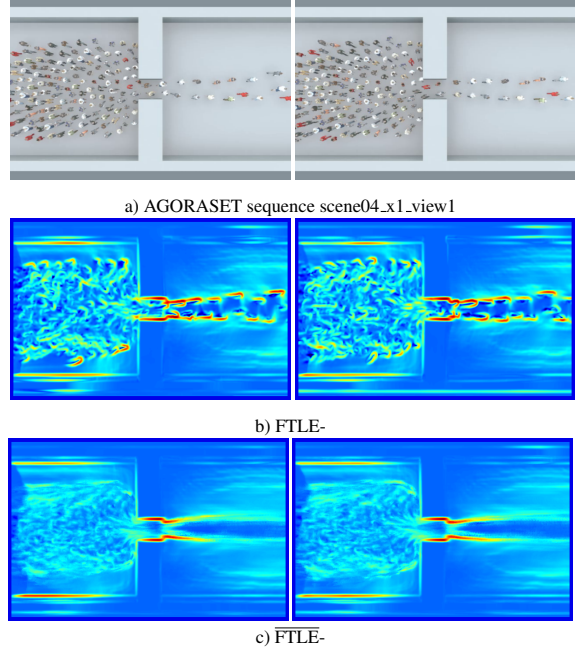


Figure 1: Comparison of the FTLE fields estimated in backward direction for $\tau = 15$ containing unsteady and long-term temporal filtered FTLE ($\overline{\text{FTLE}}^-$). In contrast to the $\overline{\text{FTLE}}^-$ the FTLE- contains more locally unsteady structures.

to ridges caused by a physical bottleneck and crowd margins, these ridge structures are not consistent for the frame of reference at different times.

To cope with the requirement of this long-term surveillance we propose to skip frames, which simply allows to increase the walking speed of the pedestrian. To remove local adverse structures caused by individuals and enhance the global ridge structure of the crowd we propose the long-term temporal filtered FTLE. This allows to use relative small integration intervals and reduces the computational effort while maintaining the global separation lines.

The long-term temporal filtered FTLE can be estimated as follows. At first we subsample the given image sequence I_t with the factor Δt , where $I_t \in \{I_0, I_{\Delta t}, \dots, I_{n \cdot \Delta t}\}$ with $n \in \mathbb{N}$ and compute the optical flow fields in forward and backward direction:

$$\begin{aligned} \mathbf{v}_t^+ &\in \{\mathbf{v}_0(I_0, I_{\Delta t}), \dots, \mathbf{v}_{n \cdot \Delta t}(I_{n \cdot \Delta t}, I_{(n+1) \cdot \Delta t})\} \\ \mathbf{v}_t^- &\in \{\mathbf{v}_0(I_{\Delta t}, I_0), \dots, \mathbf{v}_{n \cdot \Delta t}(I_{(n+1) \cdot \Delta t}, I_{n \cdot \Delta t})\}. \end{aligned} \quad (6)$$

In a next step we compute the $\text{FTLE}_+^\tau(\mathbf{x}, t_0)$ from the optical flow fields $\{\mathbf{v}_{t_0}^+, \dots, \mathbf{v}_{t_0+\tau \cdot \Delta t}^+\}$ and $\text{FTLE}_-^\tau(\mathbf{x}, t_0)$ respective from $\{\mathbf{v}_{t_0}^-, \dots, \mathbf{v}_{t_0-\tau \cdot \Delta t}^-\}$ for the reference frame at $t_0 \in \{0, \dots, n \cdot \Delta t\}$. Please note that the real integration time is $\tau \cdot \Delta t$.

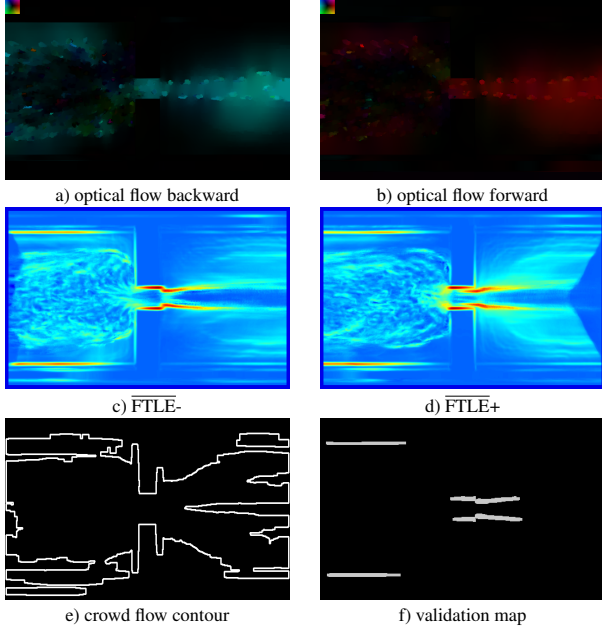


Figure 2: Illustration of crowd flow contour segmentation. Long-term filtered FTLE fields (c,d) are computed based on RLOF [13] optical flow (a,b) and lead to the basis of segmentation map (e) and the validation map (f). In (e) the combination of forward and backward ridges in the FTLE fields leads to the crowd flow segment. Only the dominant ridges with high values that occurs in both FTLE fields are in (f).

For a given set of τ_s consecutive FTLE fields we apply the temporal low-pass filter using median:

$$\overline{\text{FTLE}}^\tau(\mathbf{x}, t_0) = \text{median}_{n \in [0, \tau_s - 1]} \text{FTLE}^\tau(\mathbf{x}, t_0 - n \cdot \Delta t). \quad (7)$$

Figure 1 gives an example of the temporal filtered FTLE. It can be shown that in contrast to the FTLE the $\overline{\text{FTLE}}$ fields are less affected by unsteady, temporal local, ridge structures caused by individual motions and contain features that are steady on a global temporal scale.

Crowd Flow Contour Segmentation: Figure 2 shows exemplary the extraction of the salient motion boundary contour caused by the crowd flow. We extract ridges with high and low FTLE values. The low ridge contour will be the basis to generate possible bottleneck candidates while the high ridge contour will be used to evaluate the bottleneck candidates. Both are computed by the binarization of the temporal filtered FTLE fields based on the two thresholds σ_{low} and σ_{high} :

$$M_{(low/high)}^{(+)} = \left\{ \mathbf{x} \mid \overline{\text{FTLE}}_{(+)}^\tau(\mathbf{x}) > \sigma_{(low/high)} \right\}. \quad (8)$$

After dilating all four binary maps to close gaps in the contours a segmentation map $M_{seg}(\mathbf{x}, t_0) = M_{low}^-(\mathbf{x}, t_0) \vee$

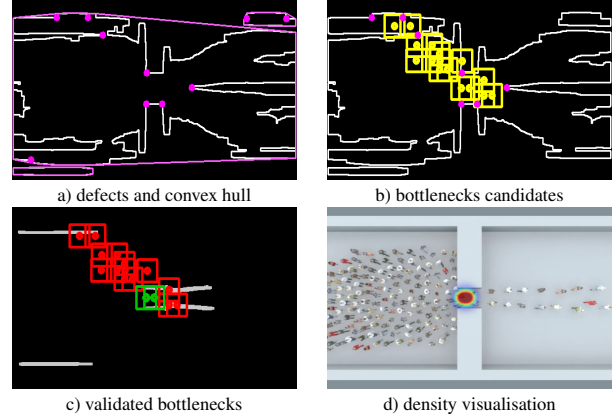


Figure 3: Illustration of the crowd flow contour analysis for bottleneck detection. Bottleneck candidates are selected by comparing the contour with convex hull (a). Selected candidates (b) are filtered by geometrical filter and the validation map (c). The visual result can be seen in (d).

$M_{low}^+(\mathbf{x}, t_0)$ is computed by combining the forward and backward low ridge maps. The segmentation map contains the contour of the crowd flow and can be prone to oversegmentation and artifacts. A second validation map $M_{val}(\mathbf{x}, t_0) = M_{high}^-(\mathbf{x}, t_0) \wedge M_{high}^+(\mathbf{x}, t_0)$ will be computed by the overlap of the forward and backward high ridge maps that contains the most stable ridges of the Lagrangian fields. The ridges of this map relate to the barriers of the physical bottleneck. Unless this map can not contain the complete crowd contour it contains ridges that are at the bottleneck location with a high probability.

Crowd Flow Contour Analysis: An example of the crowd flow contour is shown in Figure 3(a). The bottleneck candidates are a set of point tuples $\hat{C}_{t_0} = \{(\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1)_{t_0}^0, \dots, (\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1)_{t_0}^m\}$ that are located at indentations of the contour (purple dots). The candidates can be located by computing so called defects. Defects are points computed by evaluating the distance between the contour at its convex hull [18]. Bottleneck candidates are filtered by two geometrical constraints between the point pair $(\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1)_{t_0}^m$:

i) The relation d_c/l_s between euclidean distance $d_c = \|(\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_1)\|$ and the crowd flow segment contour length l_s has to be below a given threshold σ_s (see Figure 3(b)). ii) The relation between the distance of the points on the contour and the euclidean distance has to be greater than $2 \cdot d_c$ to remove point pairs that are likely to be not on opposite location of the contour. Finally, the points are projected on the validation map $M_{val}(\mathbf{x}, t_0)$. If the point, as shown in Figure 3(c), contains at least two different ridges within a region of size $\sigma_r \times \sigma_r$, which are not on the same contour, the center point is selected as a bottleneck detection $c_{t_0}^m = (\mathbf{x}_0, \mathbf{x}_1)_{t_0}^m$.

The detection of candidates can be affected by small changes of the ridge segmentation which can result in a

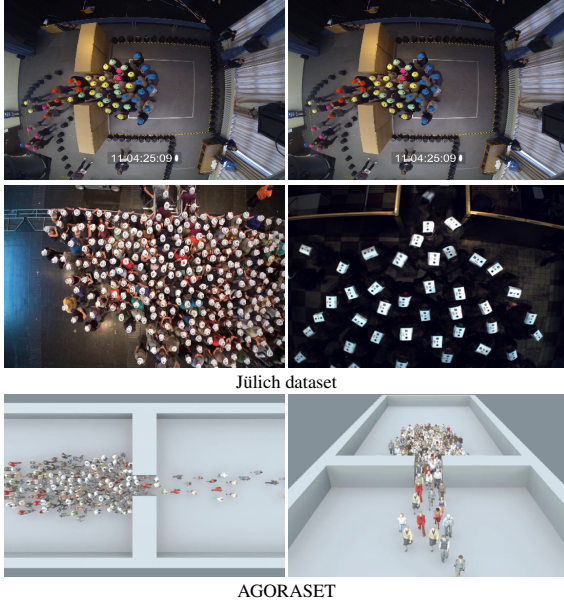


Figure 4: Example of the Jülich dataset containing bottlenecks and the synthetic AGORASET.

missed detection. To remain consistent over time the detection points $(\mathbf{x}_0, \mathbf{x}_1)_{t_0}$ will be propagated to the next frame. The detection $c_{t_0}^m$ will only be accepted if it has been detected along $\sigma_o \cdot \Delta t$ frames within the same radius σ_r .

4. Evaluation

In this section, we assess the performance of the Lagrangian-based bottleneck detection approach. To evaluate the results we introduce an appropriate metric for this novel problem and provide supplemented ground truth for the existing datasets Jülich and AGORASET which are applicable to the new use case. The ground truth as well as the evaluation script are publicly available for future work¹.

The evaluation will be based on 76 sequences from the Jülich dataset and four from the AGORASET² [3] showing crowds passing bottleneck scenarios. An example of the used sequences can be found in Figure 4. AGORASET is a synthetic rendered dataset. It contains different viewing angles as well as a high variation of the peoples density and movement characteristics under constant environmental conditions. The Jülich dataset is a composition of datasets related to [10, 15, 16, 22] and published via the pedestrian dynamics data archive³. Different bottleneck sizes were examined as well as different social aspects and their consequences under laboratory conditions. This has resulted in a broad field of data in which no bottleneck is available

¹<https://github.com/simonmaik/bottleneck-detection-benchmark>

²<https://www.sites.univ-rennes2.fr/costel/corpetti/agoraset/Site/Scenes.html>

³<http://ped.fz-juelich.de/da/>

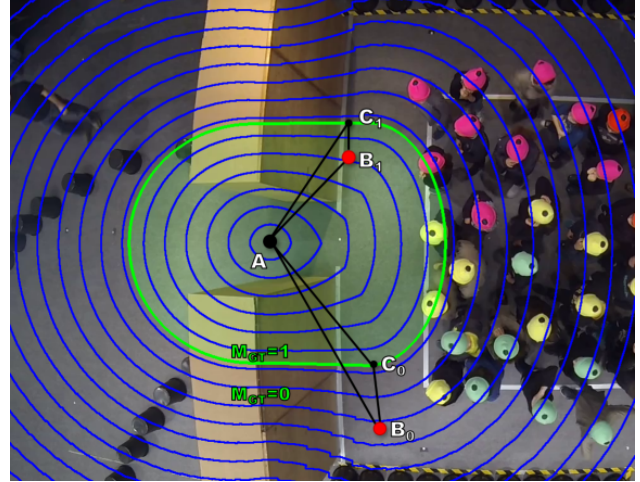


Figure 5: Visualization of the localization error ϵ_d estimation, the ground truth bottleneck detection (A) and the ground truth bottleneck mask M_{GT} (framed by green line). ϵ_d is computed by the relation between the distances between the ground truth detection (A), the estimated detection's (B_0, B_1) and the nearest point to the bottleneck mask (C_0, C_1). Each blue line shows the increasing isometric ϵ_d .

for longer periods of time or in which different motion sequences repeatedly occur due to constrictions of varying sizes. The presented algorithm detects a bottleneck both temporally and spatially. The temporal characteristics of the event were determined under two essential aspects: i) Pedestrians cross the bottleneck and ii) the individuals of the crowd try to take the shortest route, which depends on the density of the crowd's dependent speed. The last mentioned aspect is based on hypotheses describing a crowd, more details can be found in the work of Hughes [7]. The characterization is necessary, since a narrowing of individual persons in our understanding does not represent a bottleneck. The central point of the constriction was determined and carefully annotated after the subjective evaluation by scientific staff. In order to measure the accuracy while taking the distortion, camera angle and scaling/height into account, a binary ground truth map M_{GT} is created around the determined point. Whereby M_{GT} only exists at the t points in time when a bottleneck exists according to the above time definition. In our evaluation, the detection of a bottleneck is then treated as a frame-wise binary classification problem with: True positives TP (mask hit), false positives FP (detection outside mask), true negatives TN (no detection outside mask), false negatives FN (no detection inside mask). The accuracy is defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (9)$$

In order to evaluate the detection spatially, an additional

Test set	Accuracy
All sequences	0.70
Bottleneck and social groups	0.73
Entrance 1	0.83
AGORASET	0.87

Table 1: The Accuracy values for the considered sets shown in Figure 8 are listed for localization error values $\epsilon_d = 1$.

isometric score ϵ_d , called localization error, is mapped to simulate the dilatation/contraction of M_{GT} :

$$\epsilon_d = \begin{cases} \frac{\overline{AB}}{AB+BC} & , if M_{GT}(c_{t_0}^m) = 1 \\ \frac{\overline{BC}}{AC} + 1 & , else. \end{cases} \quad (10)$$

Point A is the annotated centre of the bottleneck ground truth M_{GT} . The respective detected points $c_{t_1}^m$ are marked with B_1 (inside the ground truth mask M_{GT}) and B_0 (outside), and C_0, C_1 are indicating the nearest corresponding points on the contour of M_{GT} . The localization error ϵ_d is defined within M_{GT} between $[0, 1]$ and outside M_{GT} between $(1, \infty]$. Figure 5 shows that ϵ_d reaches the value of 0 if a detected point lies within the smallest assumed isometric contour of M_{GT} . If the detected point lies on the green contour, ϵ_d becomes equal to 1. The results of the evaluation are shown in Figure 8 and shows the achieved accuracy in dependence of the localization error ϵ_d . The higher the accuracy value, the better the events were detected. The smaller the localization error, the more accurately the event was detected spatially.

The accuracy increases in all results with increasing distance from the optimal point, i.e. with increasing ϵ_d . The arrangement in Figure 8 from left to right shows different parameter tests. The Accuracy for a Localization Error value of $\epsilon_d = 1$ with the best settings are listed in Table 1. Figure 8 (top) shows the evaluation results for all 80 sequences. The results for the entire test dataset show that the proposed method performs well within the ground truth ($\epsilon_d = 1$) with an accuracy value of 0.7. The neighborhood of the optimal point of the bottleneck ($\epsilon_d = 0$) could be reached with an accuracy value of 0.4. Visual results can be seen in Figure 6. Furthermore, it becomes clear that the results for a small integration time τ (Figure 8 (top, left)) are in accordance with the method, because a small τ also means a lower calculation effort for the calculated path lines. Seen over all sequences, the buffer size parameter τ_s seems to have the smallest effect. This is due to the length of the AGORASET sequences. The larger the buffer, the more stable the results can be. However, by averaging the FTLE values with the help of the median, the method becomes sluggish, so that the buffer size also presents itself as a limitation of the system. The result of the partial sequence Entrance_1 (Figure 8) shows a different effect regarding to the buffer size. For the largest buffer value, the best result is obtained

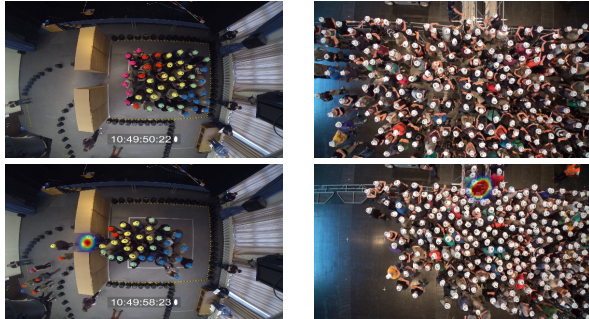


Figure 6: Detection results of the method for the sequences *Bottleneck and social groups 01_02* (left) and *Entrance 1, entry without guiding barriers (semicircle setup)* (right). At the beginning of the sequences there is no bottleneck (top). In the later part of the sequence the crowd starts running respectively the gate is opened.

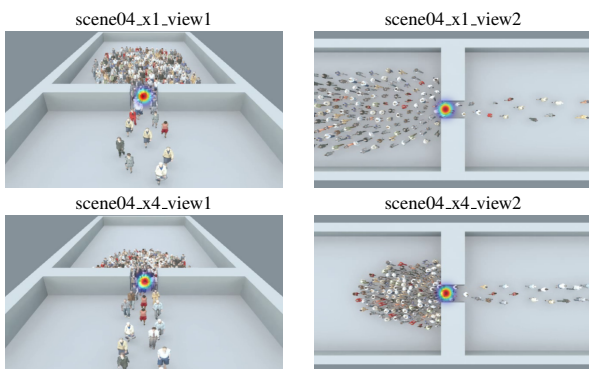


Figure 7: Detection result of the method for the same time of the escape sequences of AGORASET, at different viewing angles and same parameters.

here. The sequence is very long and only at the end of the sequence an entry is opened, which represents the bottleneck. A large buffer has more frames of reference, so that many small movements of the group can be caught, which can lead to errors with smaller buffers. In the evaluation of the radius σ_r it turns out that a smaller value $\sigma_r = 30$ over all sequences achieves the best result. This is due to the number of detected bottlenecks. A large radius for the ROI can also enclose unrelated ridges in the validation map. Certainly there are also sequences in the test dataset which have very large bottlenecks related to the image content. The filter fails because the ridges in the validation map cannot be included at all.

Figure 7 shows the result of the AGORASET escape sequences for the same point in time from two different perspectives. The outcome emphasizes that the presented procedure can act independently of the point of view.

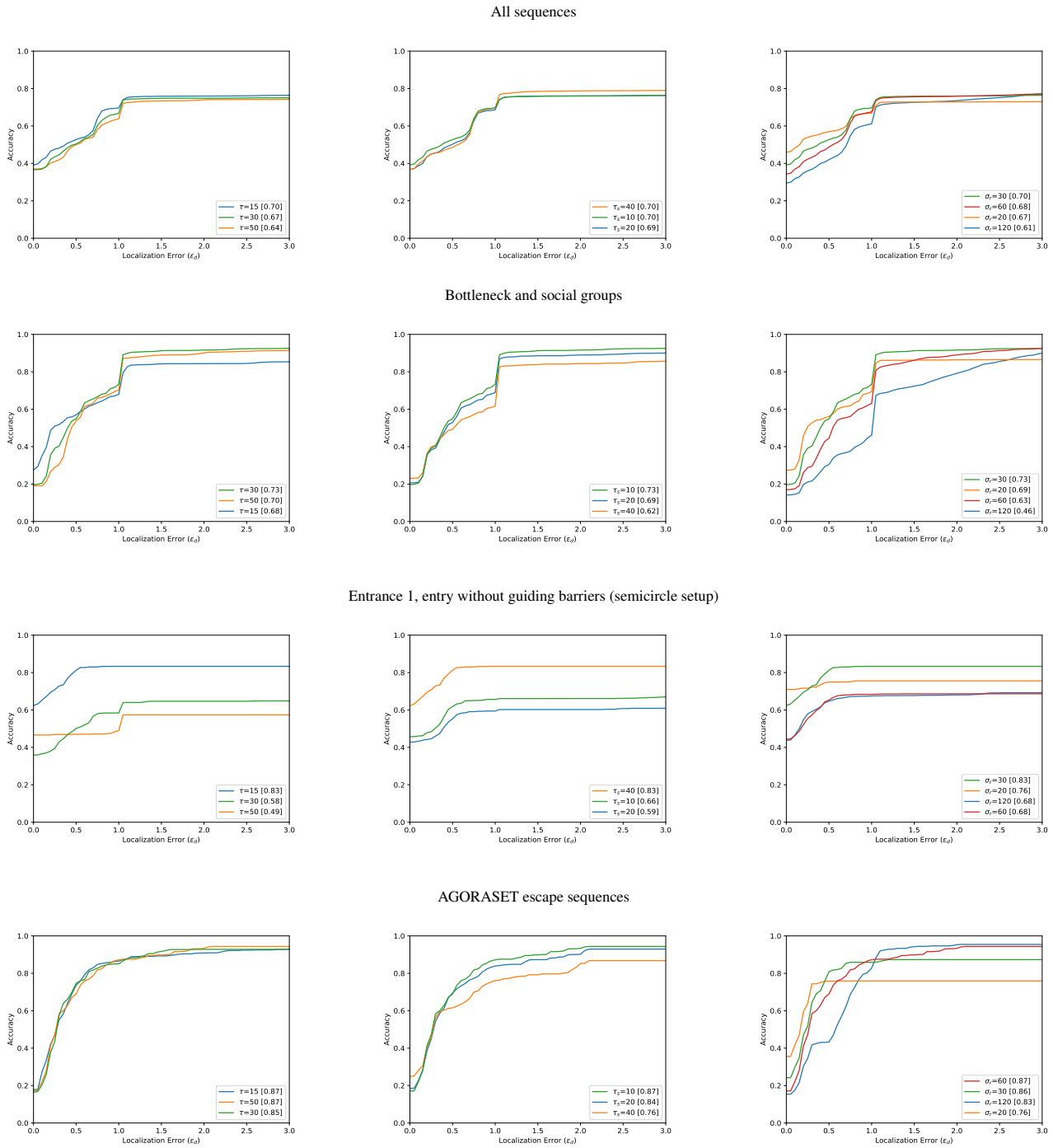


Figure 8: The figure shows the results of the bottleneck detector related to the accuracy value and the localization error. The arrangement from left to right shows the results of the parameter tests: integration time τ (left), buffer size τ_s (middle) and radius σ_r (right). The plots at the top show the results for all sequences, below are the results for the sequences of the subsets *Bottleneck and social groups* and the single sequence *Entrance 1, entry without guiding barriers (semicircle setup)*. The results for the four escape sequences of the AGORASET dataset are shown at the bottom of this figure. The order within the legends arranges the results after the accuracy at the position $\epsilon_d = 1$. The accuracy value at this position is given in brackets.

5. Conclusion

In this paper we presented a novel video-based bottleneck detection method for crowded scenes based on the evaluation of characteristic stowage patterns in crowd-movements. We utilized the proposed long-term temporal filtered Finite Time Lyapunov Exponents (FTLE) fields for a global segmentation of the crowd flow, which enables to extract its deformations. Furthermore we showed that high ridges in the FTLE field indicate Lagrangian features that are assumed to be located at bottlenecks.

Ground truth data was generated for 80 tested sequences, which were evaluated in dependence of the localization error. The results show that the method can detect bottleneck events spatially and temporally well for both natural and synthetic data. Our method is independent from camera angle and distortion, but is currently limited in the width of the bottleneck due to a fixed size of the region of interest. For future work, an adaptive adjustment of the search area is planned, whereby the current restriction will be solved.

6. Acknowledgements

The research leading to these results has received funding BMBF-VIP+ under grant agreement number 03VP01940 (SiGroViD).

References

- [1] E. Acar, T. Senst, A. Kuhn, I. Keller, H. Theisel, S. Albayrak, and T. Sikora. Human action recognition using lagrangian descriptors. In *IEEE Workshop on Multimedia Signal Processing*, pages 360–365, 2012.
- [2] S. Ali and M. Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6, 2007.
- [3] P. Allain, N. Courty, and T. Corpetti. AGORASET: a dataset for crowd video analysis. In *International Workshop on Pattern Recognition and Crowd Analysis*, pages 1–6, 2012.
- [4] N. Bain and D. Bartolo. Dynamic response and hydrodynamics of polarized crowds. *Science*, 363(6422):46–49, 2019.
- [5] G. Haller. A variational theory of hyperbolic Lagrangian Coherent Structures. *Physica D: Nonlinear Phenomena*, 240(7):574 – 598, 2011.
- [6] G. Haller. Lagrangian coherent structures. *Annual Review of Fluid Mechanics*, 47(1):137–162, 2015.
- [7] R. L. Hughes. The flow of human crowds. *Annual Review of Fluid Mechanics*, 35(1):169–182, 2003.
- [8] A. Kuhn, T. Senst, I. Keller, T. Sikora, and H. Theisel. A lagrangian framework for video analytics. In *IEEE International Workshop on Multimedia Signal Processing*, pages 387–392, 2012.
- [9] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, and S. Yan. Crowded scene analysis: A survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(3):367–386, 2015.
- [10] W. Liao, A. Seyfried, J. Zhang, M. Boltes, X. Zheng, and Y. Zhao. Experimental study on pedestrian flow through wide bottleneck. *Transportation Research Procedia*, 2:26 – 33, 2014.
- [11] B. E. Moore, S. Ali, R. Mehran, and M. Shah. Visual Crowd Surveillance Through a Hydrodynamics Lens. *Communications of the ACM*, 54(12):64–73, 2011.
- [12] T. Senst, V. Eiselein, A. Kuhn, and T. Sikora. Crowd Violence Detection Using Global Motion-Compensated Lagrangian Features and Scale-Sensitive Video-Level Representation. *IEEE Transactions on Information Forensics and Security*, 12(12):2945–2956, 2017.
- [13] T. Senst, J. Geistert, and T. Sikora. Robust local optical flow: Long-range motions and varying illuminations. In *2016 IEEE International Conference on Image Processing*, pages 4478–4482, 2016.
- [14] T. Senst, A. Kuhn, H. Theisel, and T. Sikora. Detecting people carrying objects utilizing lagrangian dynamics. In *International Conference on Advanced Video and Signal-Based Surveillance*, pages 398–403, 2012.
- [15] A. Seyfried, O. Passon, B. Steffen, M. Boltes, T. Rupprecht, and W. Klingsch. New Insights into Pedestrian Flow Through Bottlenecks. *Transportation Science*, 43(3):395–406, 2009.
- [16] A. Sieben, J. Schumann, and A. Seyfried. Collective phenomena in crowds - Where pedestrian dynamics need social psychology. *PLOS ONE*, 12(6):1–19, 2017.
- [17] J. C. Silveira Jacques Junior, S. R. Musse, and C. R. Jung. Crowd Analysis Using Computer Vision Techniques. *IEEE Signal Processing Magazine*, 27(5):66–77, 2010.
- [18] J. Sklansky. Finding the Convex Hull of a Simple Polygon. *Pattern Recognition Letters*, 1(2):79–83, 1982.
- [19] B. Solmaz, B. E. Moore, and M. Shah. Identifying behaviors in crowd scenes using stability analysis for dynamical systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10):2064–2070, 2012.
- [20] U. Soori and M. R. Arshad. Underwater crowd flow detection using Lagrangian dynamics. In *International Conference Underwater System Technology*, pages 359–364, 2008.
- [21] G. K. Still. Crowd dynamics. *PhD thesis, University of Warwick*, July 2000.
- [22] C. von Krüchten, F. Müller, A. Svachiy, O. Wohak, and A. Schadschneider. Empirical Study of the Influence of Social Groups in Evacuation Scenarios. In V. L. Knoop and W. Daamen, editors, *Traffic and Granular Flow '15*, pages 65–72, 2016.
- [23] S. Wu, B. E. Moore, and M. Shah. Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes. In *Conference on Computer Vision and Pattern Recognition*, pages 2054–2060, 2010.