Measuring Complexity in Lagrangian and Eulerian Flow Descriptions

H. Jänicke\textsuperscript{1} and G. Scheuermann\textsuperscript{2}

\textsuperscript{1}Computer Science Department, Swansea University, UK
\textsuperscript{2}Computer Science Department, Universität Leipzig, Germany

Abstract

Automatic detection of relevant structures in scientific data sets is still one of the big challenges in visualization. Techniques based on information theory have shown to be a promising direction to automatically highlight interesting subsets of a time-dependent data set. The methods that have been proposed so far, however, were restricted to the Eulerian view. In the Eulerian description of motion, a position fixed in space is observed over time. In fluid dynamics, however, not only the site-specific analysis of the flow is of interest, but also the temporal evolution of particles that are advected through the domain by the flow. This second description of motion is called the Lagrangian perspective. To support these two different frames of reference widely used in CFD research, we extend the notion of local statistical complexity to make them applicable to Lagrangian and Eulerian flow descriptions. Thus, coherent structures can be identified by highlighting positions that either feature unusual temporal dynamics at a fixed position or that hold a particle that experiences such dynamics while passing through the position. A new area of application is opened by LagrangianLSC, which can be applied to short pathlines running through each position in the data set, as well as to individual pathlines computed for longer time intervals. Coloring the pathline according to the local complexity helps to detect extraordinary dynamics while the particle passes through the domain. The two techniques are explained and compared using different fluid flow examples.

Categories and Subject Descriptors (according to ACM CCS): Simulation and Modeling [I.6.6]; Simulation Output Analysis—Physical Sciences and Engineering [J.2]; Mathematics and Statistics—

1. Introduction

The visualization of large unstructured time-dependent data sets is still one of the big challenges in visualization. Especially in 3D, it is often not possible to use direct visualization techniques such as isosurfacing, volume rendering or streamline visualization, as the resulting pictures easily suffer from occlusion and visual clutter. To decrease the visual complexity of the visualization, scientists commonly extract the relevant structures in their data sets and restrict the visualization to those key components. In flow visualization, such structures are for example vortices, separation and attachment lines, shear flow or critical points. A summary of feature extraction techniques in the field of flow visualization is given in [PVH’03] and [SJWS07]. Although, this reduction to the essential is very powerful, the scientist still has to make sure that all relevant structures are found. For most important structures, detection algorithms have been developed that rely on physical or topological descriptions. Using these so called feature detection algorithms, the corresponding structures can be extracted and visualized. However, this procedure has three shortcomings. First, a mathematical description of the structure’s properties is required to make it detectable by a computer. Some structures, though, are hard to define. For vortices, for example, no unique mathematical definition exists and several detection algorithms have been proposed to capture them automatically. Second, a data set commonly comprises only a subset of the known features. To make sure that they are all identified, all different detection algorithms have to be applied, which can be very time-consuming for large data sets. And third, it is not possible to detect unknown anomalies or structures in the simulation.

To overcome these problems, a new direction of extraction methods was proposed recently. Based on information theory, these new techniques introduce filters that automat-
physically extract those positions in the data set that feature an unusual evolution of values compared to the entire data set. Jänicke et al. [JWSK07, JBTS08] used local statistical complexity (LSC), a measure that locally quantifies how much information is required to predict the local future given the local past. Each position in the unsteady multivariate data set is automatically assigned a scalar value that captures this quantity. Volume rendering or isosurfacing can be used to identify the most informative regions in the data set that commonly correspond to the structures the user is interested in, as the standard behavior of the data set is usually known and extraordinary events are the interesting ones. Wang et al. [WYM08] extended the concept to block structured neighborhood structures and researched the temporal evolution of the complexity at individual positions. Woodring and Shen [WS09] analyzed temporal activity in a multiscale approach.

The advantage of such information-based measures is that they identify all statistically relevant structures at a time and can be applied to a large variety of applications. Moreover, an inherent characteristic is that they can easily handle multivariate data. However, these techniques have still the major restriction that they are limited to structured neighborhoods of positions. LSC uses conical neighborhood structures and boxes are used by Wang. A second drawback is the fact that all methods only support the Eulerian view of fluid flow and not the Lagrangian view. Both frames of reference provide different insight into the evolution of the flow and reveal either site-specific (Eulerian) or fluid-specific (Lagrangian) information about the flow.

As both frames of reference are important to researchers, a measure to identify interesting structures in both scenarios is desirable. So far, methods for the identification of interesting structures based on information theory are limited to the Eulerian perspective, detecting positions in a transient flow that feature an extraordinary evolution. In this paper, we will propose a technique to identify coherent structures that can be applied in both the Eulerian and the Lagrangian frame of reference and allows for the comparison of both perspectives. Therefore, we will adopt the notions of local statistical complexity to particle data, which supports the Lagrangian perspective. This new notion of LSC will be called LagrangianLSC. LagrangianLSC can be applied in two directions: First, we will show how the LagrangianLSC can be used to identify positions in the data set where the associated pathline features an extraordinary behavior and second, we will investigate the evolution of complexity of individual particles over time. Hence, with the new method proposed in this paper, interesting structures in an unsteady data set can be identified using the Eulerian, the Lagrangian and a mixed perspective based on the same underlying theory, which makes the different frames of reference directly comparable.

2. Related Work

In flow visualization, a large variety of different techniques have been proposed utilizing the two different frames of reference. Commonly, three major classes of flow visualization techniques are distinguished [PVH03]: geometry-based, texture-based, and feature-based methods. In the following we will review basic ideas in the different fields aiming at the illustration of unsteady flows in the Lagrangian frame of reference. We picked several example techniques in each class to illustrate the fundamental ideas and refer the interested reader to the survey papers cited in each section for more references on related work.

Geometry-based techniques first extract geometric objects from the flow and display them in a second step. Lagrangian approaches are the depiction of integral structures such as pathlines and pathsurfaces [GKT08, vFWTS08].

In texture-based approaches [LHD03], the structure of the flow is captured by a texture such as a colormap or a
white noise texture being manipulated to resemble the flow. These methods are commonly applied to two-dimensional data or surfaces in 3D. Colormaps and animations based on them are by nature Eulerian. Integral techniques combine techniques from the two perspectives. Pathlines are commonly used to advect properties of individual pathlines (Lagrangian) on a uniformly sampled texture (Eulerian). Different techniques have been applied to adapt the noise texture to the flow, such as local smoothing [vW91], line integral convolution [SK98, JEH02], or anisotropic diffusion [BPR01].

The broadest field in flow visualization is concerned with the identification and depiction of flow features [PVH+03]. The main branches rely on techniques from image processing and vector field topology, or identify features based on physical characteristics. The Lagrangian perspective is mainly used when applying physical characteristics and seldomly in image processing or topological approaches. The first approaches in physical feature definition aim at the identification of positions with certain value combinations and the tracking of connected components over time [SSZC94]. These methods did not take the propagation of particles into account, but applied spatial matching techniques which support the Eulerian view. In later algorithms, these fundamental physical characteristics were evaluated on a particle basis. Examples using this metaphor are pathline predicates which extract characteristics for particles in a certain time frame [SGSM07, STH+07, STW+08]. A combination of both perspectives is used in the computation of the finite time Lyapunov exponent (FTLE) to identify positions where local pathlines exhibit a strong divergent or convergent behavior [GGTH07, SP07].

3. Eulerian vs. Lagrangian Description of Fluid Flow

As we have seen earlier, two different ways to describe a fluid’s motion exist. Either the trajectories of specific fluid particles are observed (Lagrangian representation), or the fluid velocity at fixed positions is investigated (Eulerian representation) [Bat67, Pri06] (cf. Figure 1). When investigating a river’s current, one can either stand on the river bank and watch the water flow, or go rafting and observe the current while being advected by it. In the Eulerian perspective, the fluid velocity is observed at locations that are fixed in space, while observing a fluid flow from a Lagrangian perspective means that specific, identifiable fluid material volumes that are carried about with the flow are tracked.

Depending on this frame of reference, different questions concerning the evolution of an unsteady flow can be answered. The Lagrangian coordinate system is well suited to observe the advection of different quantities. If, for example, a pollutant is inserted into the river, we could easily track its spreading over time by investigating the tracks of the individual particles. By following a cluster of particles injected into the flow, we can observe how the flow deforms and rotates the fluid. On the contrary, if fluid properties such as pressure, wind or density at a certain site, e.g., along the top of the wing of an airplane or in the center of a city, are of interest, the Eulerian frame of reference is more suited. Likewise, it is easier to get a dense coverage of the flow domain in the Eulerian perspective. Particles are not traced randomly through the domain and sampled at the required location, but are specifically computed at fixed positions. Hence, they can be directly used to provide an overview over the entire flow at a certain time-step. Expressed in simplified terms, Eulerian methods provide a site-specific and Lagrangian approaches a fluid-specific representation of the unsteady flow.

To formalize these ideas, consider a fluid flow within a three-dimensional domain \( \mathbb{R}^3 \). We want to observe the evolution of this flow over time using the two different descriptions of motion. Figure 1 illustrates the setting in two dimensions.

The more natural way to analyze a flow is the Lagrangian perspective, where particles or small flow parcels are traced over time. In a practical experiment, one could, for example, inject dye into the transient flow of water or colored smoke into an airstream. The color particles would follow the flow along pathlines and give a Lagrangian description of the motion. Now assume, we track a fluid parcel located at position \( \mathbf{q} = (\alpha, \beta, \gamma) \), where \( \alpha, \beta, \) and \( \gamma \) denote the x-, y-, and z-coordinate respectively. Determining the position of parcels at all later times, \( t_i, i \in [0, \ldots, n] \), we obtain the parcel trajectory, also called pathline, \( \mathbf{q}_t = \mathbf{q}(\mathbf{q}_t, t) \). The trajectory \( \mathbf{q}_t \) of specific fluid parcels is a dependent variable in a Lagrangian description, while the initial position \( \mathbf{q}_t \) and time \( t \) are the independent variables.

The velocity of a parcel \( V_L \) is given by the partial derivative with respect to time,

\[
V_L(\mathbf{q}_t, t) = \frac{\partial \mathbf{q}_t}{\partial t},
\]  

where \( V_L \) is the Lagrangian velocity.

As we have seen earlier, there are applications where the observations are required for a fixed position. The velocity sampled at this position is termed Eulerian velocity, \( V_E \), and is the velocity of the fluid parcel that is instantaneously present at the given position. Thus the Eulerian velocity is defined by

\[
V_E(\mathbf{x}, t)|_{\mathbf{x}_0 = \mathbf{q}_t(\mathbf{q}_t, t)} = V_L(\mathbf{q}_t, t),
\]  

where \( \mathbf{x} \) is fixed and the \( \mathbf{x}_0 \) on the left and right hand-side are the same initial position. In simplified terms, the velocity on the left hand-side is the one of the fluid parcel that happens to be at the given position at that instant of time. In the Eulerian notion, \( V_E \) is a dependent variable, while the position \( \mathbf{x} \) and time \( t \) are independent ones.

Although equation 2 seems not very impressive, it lays the foundations for our further observations: For a given fluid flow there is a unique fluid velocity that can be sampled in two quite different ways; Either a fluid parcel is followed over time by investigating the tracks of the individual particles, or the fluid parcel is fixed and the Eulerian view is used in the computation of the finite time derivative of the velocity.
(Lagrangian) or the velocity is sampled at a fixed position at different time steps (Eulerian). These two sampling strategies can be used to determine a position's past and future, which are necessary to make statements about the fluid's local dynamics. In the following, we will show how these different pasts and futures of a position are used to identify positions and structures with extraordinary local temporal dynamics either in a site- or a fluid-specific notion, and how they can be compared to one another.

4. Quantification of Complexity

To quantify the "extraordinariness" of such a temporal track, we need a measure. In information theory, complexity measures are used to fulfill this task. One of the fundamental goals of information theory is to quantify the complexity of a data stream, as its states by how much the data can be compressed. A large variety of complexity measures can be found in the literature. Common measures originating from the analysis of strings of data are Shannon entropy [Sha48] and algorithmic information [BP97]. Shannon entropy is a measure of the uncertainty associated with a random variable, whereas the algorithmic information is roughly speaking the length of the shortest program capable of generating a certain string. Both measures have in common that they are measures of randomness. In complex systems however, randomness is commonly not considered to be complex. Likewise, Hogg and Huberman [HH85] state that complexity is small for completely ordered and completely disordered patterns and reaches a maximum in-between.

A different approach was taken by Grassberger [Gra86], who defined complexity as the minimal information that would have to be stored for optimal predictions. Based on this idea, statistical complexity (SC) [CY89] was introduced, where the complexity of a system is given by the amount of information needed to specify its causal states, i.e., its classes of identical behavior. In order to analyze random fields, a point-by-point version was formulated by Shalizi [Sha05] called local statistical complexity (LSC), which was extended to data coming from scientific simulations by Jänicke et al. [JWSK07, JBTSS08].

4.1. LSC in the Different Settings

Local statistical complexity is an information theoretic measure that extracts those regions in an unsteady field that feature extraordinary behavior compared to the dynamics in the entire time-dependent data set. The idea is to measure how much information from the local past is required to predict the dynamics in the local future at a certain position. If the dynamics of a configuration match the average behavior in the data set, only little information is required as this is the behavior we would have expected. On the contrary if something unusual happens, more information is required as it is harder to predict the unexpected.

The local statistical complexity at a certain position $p$ in the field is defined as the mutual information between the corresponding configuration’s past ($\ell^p$) and its causal state ($\epsilon(\ell^p)$):

$$LSC(p) = I[\epsilon(\ell^p); \ell^p]$$

(3)

Causal states are stochastic spatio-temporal patterns that group positions whose pasts induce the same distribution over possible futures. Mutual information is a measure from information theory, which tells how much information one random variable contains about another one:

$$I[A; B] = \sum_{a \in A,b \in B} P(a,b) \log_2 \frac{P(a,b)}{P(a)P(b)}$$

(4)

where $P(a)$ is the probability that the random variable $A$ takes the value $a$ and $P(a,b)$ is the corresponding joint probability of variables $A$ and $B$. Using this definition, the local statistical complexity of a configuration tells how much information from the past is required to identify its causal state. If one knows the causal state, the dynamics in the future are clear as well. Hence, if only little information is required to predict the future, something usual is going on. On the contrary, if the position is assigned a high LSC value, a lot of information is required to identify the causal state and the local dynamics are extraordinary compared to what is happening in the rest of the data set. To put it in a nutshell, LSC assigns each position at each time step in the transient field a scalar value that is the larger the more unusual the local dynamics are.

Figure 2 depicts a one-dimensional flow (y-axis) that changes over time (x-axis). The double cone structure in the upper part labeled EulerianLSC (Cone) illustrates the spatio-temporal pattern of a sample causal state. The position to be investigated is marked by a red dot in the right
cone structure. All positions in the past that might influence this position are comprised in the left cone. Positions possibly influenced by the red position are summarized in the right hand-side cone. If we decrease the cone angle to zero, we no longer have a double cone, but two cylinders extending in the past and future (cf. Fig. 2, EulerianLSC (Line)). If applied in this setting, the underlying idea of LSC changes. We no longer question how much information from the data set’s entire past going through position \( p \) is required to predict the future dynamics of the positions being possibly influenced by \( p \). Instead, the question is how much information from \( p \)’s past is required to predict the future evolution of values at this position. Here we are back to the parcel idea used in the Eulerian frame of reference, where the evolution of a parcel fixed in space is monitored over time. If this concept is changed to the Lagrangian setting, we do not keep the parcel fixed in space, but allow for advection by the flow. Here, we keep the particle to be tracked fixed and the corresponding form of LSC is called LagrangianLSC (cf. Fig. 2, blue structure).

The theoretical details of EulerianLSC (Cone) have been detailed in [Sha03] and [JWSK07]. In the following, we will show that LSC can also be applied to the parcel setting. We will use the particle-based Lagrangian view to detail the ideas. However, the principles are also valid for the Eulerian setting if the moving particle is substituted by a parcel fixed in space and the appropriate notions of past and future.

### 4.2. LagrangianLSC

Like with EulerianLSC, the fundamental question that LagrangianLSC answers is “How much information from the past is required to predict the dynamics in the future?” This quantity is computed for each position in the data set using its (local) past and future (cf. Fig. 2). To describe a particles past and future, we first have to compute the pathline going through the current position. The past is resolved by computing the Lagrangian velocity along the particle trace at the required discrete time steps. If a past depth of 3 is used, for example, the velocity associated with the current particle is evaluated at time steps \( t - 1 \), \( t - 2 \), and \( t - 3 \). The current position belongs to the future as well as the values in successive time steps. Now that we have computed the past and future of a particle, we can evaluate how likely such an evolution is using LagrangianLSC.

Analog to EulerianLSC, LagrangianLSC is the amount of information the past contains about the future. To evaluate this property, we use mutual information [CT91]:

\[
I(\text{Past}; \text{Future}) = \sum_{p \in \text{Past}, f \in \text{Future}} P(p, f) \log \frac{P(p, f)}{P(p)P(f)},
\]

where \( \text{Past} \) and \( \text{Future} \) are random variables, whose sample spaces comprise all the pasts and future particle traces that occur in the data set. \( P(x) \) is the probability that the random variable \( X \) takes the value \( x \). And \( P(x, y) \) is the joint probability of events \( x \) and \( y \). Mutual information can also be rewritten in terms of entropy. Entropy \( H(X) \) [CT91] is a measure of the average uncertainty in a random variable \( X \) and defined as

\[
H(X) = - \sum_{x \in X} P(x) \log P(x).
\]

Now mutual information can be rewritten as

\[
I(\text{Past}; \text{Future}) = H(\text{Future}) - H(\text{Future}|\text{Past}).
\]

Thus, we have the uncertainty about the current particle’s future \( H(\text{Future}) \) and subtract the uncertainty that remains after we know what happened in the past \( H(\text{Future}|\text{Past}) \). The difference is the amount of information that we must have learned from the past about the future. Hence, the mutual information between past and future can be thought of as the information that the past contains about the future.

The direct computation of LagrangianLSC would require us to evaluate the probabilities \( P(p, f) \), \( P(p) \) and \( P(f) \), which would be very time consuming and error prone, as commonly too few samples exist to correctly evaluate the joint probability. To ease the computation we apply a combination step as used for EulerianLSC, namely the identification of causal states. A causal state, \( CS \), is the set of all pasts that induce the same future, i.e.,

\[
CS(\phi) = \{ \text{past} | P(\text{future}|\text{past}) = P(\text{future} | \phi) \}.
\]

The causal states are the equivalence classes induced by the equivalence relation \( \sim_c \) that defines two histories as being equivalent if and only if they have the same conditional distribution of futures. What still has to be shown, is that \( \sim_c \) is a minimal sufficient statistic. Sufficiency means, that the statistic retains all the information present in the original data.

**Definition 1:** \( T(X) \) is sufficient relative to \( \{ f_0(x) \} \) if and only if \( I(\theta; X) = I(\theta; T(X)) \). [CT91]

As shown in [Kul68], this is the same as requiring that \( P(\text{future}|\text{past}) = P(\text{future}|\text{CausalState}(\text{past})) \). This is guaranteed by the definition of the equivalence relation \( \sim_c \) and hence, the mapping from pasts to causal states is a sufficient statistic, i.e., it retains all the information present in the original data.

Additionally, it can be shown that the mapping to causal states is a minimal sufficient statistic, i.e., it is a function of every other sufficient statistic, and that it is unique. As the mapping to causal states fulfills both criteria, uniqueness and minimal sufficiency, each causal state comprises the least amount of information from the past that is required to still be able to predict the future and we achieve highest lossless compression rates. As the statistic is unique there is only one such mapping (except for relabeling).

So far we have shown that the mapping to causal states retains all the information, provides maximal compression and is unique. What remains to do, is to compute the actual...
amount of information that is required to find a positions causal state, which is a measure of the complexity of the local evolution as detailed before.

First, let us replace the future in the earlier considerations by the causal state formulation. This is motivated by the fact that if a past’s causal state is known, its distribution over possible futures is clear. Thus, it is sufficient to compute how much information is required to determine the causal state to which a past corresponds to and we can replace Equation 7 by

$$I(CS(\text{Past}) ; \text{Past}) = H(CS(\text{Past})) - H(CS(\text{Past})|\text{Past})$$

(9)

This equation is evaluated for each past $p$ that occurs in the field separately and the random variable $CS(\text{Past})$ in Equation 10 has the distribution

$$P(CS(\text{Past}) = s) = \begin{cases} 1 & \text{if } s = CS(p) \\ 0 & \text{otherwise} \end{cases}$$

(11)

Using Equation 10 as complexity measure, LagrangianLSC is given by

$$\text{LagrangianLSC}(\text{past}) = - \log P(CS(\text{past})).$$

(12)

This simple formula is the measure used to compute the complexity of a particle trace discriminating between normal and unusual dynamics. All we have to do to evaluate it, is to compute the particle’s causal state, derive the probability of the causal state and take the negative logarithm of it. The result is a scalar value telling whether the particle experiences unusual dynamics or not.

### 4.3. Computation of LagrangianLSC

In order to compute the LagrangianLSC at each position in the data set, we first have to identify the causal states of the system. Afterwards, one sweep through all positions is required to assign each position the corresponding causal state and compute the probability of each causal state. In a final step each position is assigned the LagrangianLSC of its corresponding causal state. The resulting complexity field features high values in interesting regions and low ones in regions with an ordinary temporal evolution.

For the identification of pasts and futures, pathlines have to be integrated starting from each position in the time step to be investigated. The backwards integrated part of the pathline is used to determine the values of the position’s past (cf. fig. 2) and the forward integrated pathline for the evaluation of values in the future. Pasts and futures are stored in two separate vectors and are transmitted to the causal state computation.

For the computation of the causal states we use an algorithm similar to the one proposed in [JBTS08], which consists of the following steps:

- **Discretization**: Compute the past- and future-vector at each position and store the discretized vectors in two trees.
- **Density-driven Voronoi Tessellation**: Partition the high-dimensional data space using a Voronoi tessellation that takes the underlying distribution of the data vectors into account.
- **Clustering**: Cluster those past vectors that have similar distributions over future vectors.
- **Ids**: Assign each such cluster a unique id.

For more details on the individual steps see [JBTS08].

In a final step, the LagrangianLSC for each causal state has to be evaluated. Using formula 12, the LagrangianLSC of a causal state $cs$ is obtained by computing the negative logarithm of the probability estimate of $cs$. This procedure is done for each causal state and resulting values are assigned to the corresponding positions in the data set.

The computation of LagrangianLSC along pathlines is analog, with complexity values not being assigned to grid positions but to particle positions along the pathline.

### 5. Results

In the following examples, we will analyze the structures identified by the different measures of complexity. As sample data sets we chose the simulation of a swirling flow, of the flow through a draft tube and the flow around a delta wing. The swirling flow will be used to compare the three measures EulerianLSC (Cone), EulerianLSC (Line) and LagrangianLSC in an unsteady setting with stable features. The flow through the draft tube is more turbulent and interesting spatial structures quickly change their spatial position. Using the third sample data set, we will follow a different analysis technique based on LagrangianLSC. We will no longer concentrate on the entire domain, but follow selected particles and identify sections when these particles feature extraordinary behavior.

#### 5.1. Swirling Flow

The development of a recirculation zone in a swirling flow is investigated by numerical simulation. This type of flow is relevant to several applications where residence time is important to enable mixing and chemical reactions.

The unsteady flow in a swirling jet is simulated with an accurate finite-difference method. The Navier-Stokes equations for an incompressible, Newtonian fluid are set up in cylindrical coordinates assuming axi-symmetry in terms of
streamfunction and azimuthal vorticity. All equations are dimensionless containing the Reynolds number \( Re \) and the swirl number \( S \) as defined by Billant et al. [BCH99]

\[
Re \equiv \frac{v_c(0,z_0)D}{\nu} \quad S \equiv \frac{2v_0(R/2,z_0)}{v_c(0,z_0)}
\]

where \( z_0 = 0.4D, D = 2R \) is the nozzle diameter and \( \nu \) the kinematic viscosity, as dimensionless parameters.

The flow domain is the meridional plane \( D = \{(r,z) : 0 \leq r \leq R, 0 \leq z \leq L\} \) with \( R = 5D, L = 8D \) and \( D \) denoting the nozzle diameter at the entrance boundary. The flow domain is mapped onto the unit rectangle which is discretized with constant spacing. The mapping is separable and allows to a limited extent crowding of grid points in regions of interest. The present simulation uses \( n_r = 91 \) and \( n_z = 175 \) grid points in radial and axial directions. The boundary conditions are of Dirichlet type at the entrance section and the outer boundary and at the exit convective conditions are imposed for the azimuthal vorticity. The initial conditions are stagnant flow and the entrance conditions are smoothly ramped up to their asymptotic values within four time units.

Figure 3(a) shows a line integral convolution (LIC) of time-step 4 out of 10. The relevant features in this data set are the conical shear region where the flow enters at the bottom and ringlike vortex-structures further upstream. In the LIC image these features are rather difficult to discover, as no clue concerning relevance or the strength of different phenomena is given. Overlaying the LIC image with a colormap of norm of velocity (fig. 3(b)) gives a better impression of the structure of the data. The conical shear region is the structure on the bottom in full blue. The plane in which the ringlike vortex corelines lie are orthogonal to the LIC plane. Each ring pierces the LIC plane twice and these four positions intersecting the plane are surrounded by light blue in the norm of velocity map (fig. 3(b)).

Figures 3(d), 3(e), and 3(f) show complexity fields computed using EulerianLSC (Cone), EulerianLSC (Line) and LagrangianLSC respectively. The structures marked with highest complexity are essentially the same, revealing the shear flow structure and the vortices. Differences occur at the vortex centers, at the interior of the jetstream and at the upper parts, where the flow leaves the domain. EulerianLSC (Cone) marks large areas of the flow and the complexity slowly decreases at the boundaries. This happens as the cones consider a large area of influence that gradually moves out of the relevant structures at the boundaries. This more extended spatial pattern has its strengths when it comes to the analysis of structures that are in itself not very interesting such as vortex center, where the velocity is zero, but are relevant due to the surrounding flow. LSC is perfectly

\[ \text{© 2009 The Author(s)} \]
\[ \text{Journal compilation © 2009 The Eurographics Association and Blackwell Publishing Ltd.} \]
Figure 4: Draft tube: (a) Isosurface in the $\lambda_2$ field with iso value $-1000$ displaying two major and several minor vortices. (b) A volume rendering of the EulerianLSC (Line) field highlights the vortices. (c) An isosurface in the LagrangianLSC field with a moderate isovalue shows that the particle dynamics at the boundary are very peculiar. (d) With a high isovalue, LagrangianLSC reveals the corelines of the two major vortices and the areas, where it interacts with the boundary.

...able to capture vortex corelines as it takes neighboring values into account. Hence, the unusual pattern is a stagnant center surrounded by stronger current. The two line-based measure, EulerianLSC (Line) and LagrangianLSC are not able to capture such structures, but nevertheless highlight the vortices as the surrounding flow is very unusual. The more crisp boundaries of the last two measures are however, well suited to detect sharp boundaries as the conical shear region. Here the boundaries do not get washed out as with EulerianLSC (Cone) but precisely mark the outer and inner boundary.

As the features in this data set hardly change their position, the results of EulerianLSC (Line) and LagrangianLSC look very similar. However, differences occur at the top of the images where the flow leaves the domain. Here the flow splits, one part being dragged into the vortex and the other one leaving the domain. From a cell-based point of view, the evolution in this area is perfectly normal. However, when changing the focus to the particles, we see that this area becomes more interesting and the dynamics of particles passing this part of the domain are quite unusual.

5.2. Draft Tube

The second data set represents the draft tube of a Francis turbine as illustrated in Figure 4. The water enters the turbine from top and acts on the runner, causing it to spin. The runner is not illustrated in the image and sits horizontally on top of the circular inlet. After passing the runner, the flow enters the draft tube, where it is decelerated. Thereby the kinetic energy is converted into static pressure. The flow leaves the tube through the rectangular regions in the lower left part. The data set consists of 980,000 positions and 258 time-steps.

An interesting feature of the data set is the formation of vortices. The $\lambda_2$-criterion [JH95] is a standard technique in vortex detection. Visualizing an isosurface with value $0.01$ covers almost the entire data set with different intricate structures and barely anything is visible. Therefore we chose $\lambda_2 < -1000$, which resulted in the structures illustrated in Figure 4(a). The visualization of this data set is rather difficult, as structures near the boundaries occlude the inner part. The $\lambda_2$ criterion extracts two major entwined vortices that run through the entire tube. Over the course of the simulations, the vortices rotate through the draft tube and form a rather unstable set of features. Several minor vortices in the lower part of the tube accompany the two large vortices and follow their motion. When circulating in the draft tube, the vortices heavily interact with the boundary surface of the tube.

Figures 4(b) and 4(c) depict the complexity fields for Eulerian- and LagrangianLSC. While EulerianLSC assigns rather uniform values to cells being traversed by the vortices with values slowly decreasing at the boundaries, LagrangianLSC features a higher gradient and more precise boundaries. EulerianLSC covers almost the entire domain, while LagrangianLSC clearly focuses on the vortices and their interaction with the boundaries. This can be explained by the way the two different kinds of LSC are computed. While the features are followed by LagrangianLSC over time providing a more focused computation, EulerianLSC computes the complexity of the cell for a time interval where the feature is present only for a subinterval.
5.3. Delta Wing

The Delta Wing data set represents the airflow around a delta wing at low speeds with an increasing angle of attack. Multiple vortex structures form on top of the wing due to the rolling-up of the viscous shear layers that separate from the upper surface. These formations of three vortices can be observed on either side of the wing (Fig. 5). With increasing angle of attack the intensity of the primary vortices (vortices nearest the symmetry axis) increases until in time-step 700 a vortex breakdown occurs (bubbles at the end of the vortices). The analysis of vortex breakdown is highly interesting, as it is one of the limiting factors of extreme flight manoeuvre. The grid consists of approximately 3.1 million positions.

Figures 5(b) and 6(b) depict the evolution of pathlines started on a line in front of the frontal tip of the wing. Pathlines started in the center of this line form the major vortices. As we go further away from the symmetry axis, the pathlines first pass over the major vortex, once twine around them and then evolve into the minor vortices at the boundaries of the wing. The pathlines in figures 5(b) and 6(b) are color coded using norm of velocity and pressure respectively.

Although the color coding of the pathlines for the original quantities norm of velocity and pressures looks quite different, the assigned complexity values are very similar. Figure 5(a) depicts the evolution of LagrangianLSC along the different pathlines. Highest complexity values are assigned to the particles swirling around the corelines of the two major vortices, while particles close to it feature only medium complexity. Another extraordinary evolution is faced by particles that come close to the different vortices. The light yellow pathline parts at the front of the wing belong to particles the come close to the vortices but do not yet merge with them. First these particles pass below the major vortex, then above the minor vortex, before they finally become part of them. In the close-up on the left hand-side of figure 5(a), we see similar effects happening further downstream. The small yellow subsections of the pathlines correspond to regions where the particle trace hits a minor vortex.

The close-up on the right hand-side of figure 5(a) shows how particles from one of the major vortices evolve into the recirculating bubbles. The complexity gradually increases in front of the phenomenon and is highest right in front and in the frontal halt of it. After exiting the recirculating bubble, the local dynamics of the particles are quickly back to normal.

Figure 6(a) depicts the same scenarios with pressure being the investigated variable along the pathlines. Although at first sight both Lagrangian complexity illustrations look very similar, several small differences can be spotted. Looking at the close-ups, we see that the evolution of the pressure...
values is not that peculiar when the particles hit the vortices. We can see an increase in complexity, when they pass under the major vortices, but the values along the minor vortices are very homogenous. Similar findings can be made for the recirculating bubbles. While the local dynamics in the velocity field become quite unusual at the end of the recirculating bubble, they remain very stable when looking at the evolution of pressure. As soon as particles from the major vortices meet the recirculating bubble, their complexity remains almost the same throughout the remainder of the vortex structure.

6. Timings

All timings are given in the format minutes:seconds. The tabular provides information about the following quantities: Data-set gives the name of the data-set as used in the Results section. The capital letter in brackets indicates the type of region of interest: EC = Eulerian (Cone), EL = Eulerian (Line) and L = Lagrangian. nPos is the number of positions each time-step of the data-set contains. nTS gives the number of time-steps that were used for the computation. The first summand gives the number of evaluated time-steps, the second one the additional time-steps at the end and beginning of the time-series. ROI states the parameter for the region of interests that were used. The first number indicates the past depth, the second one the future depth (past/future). System information: operating system - Linux, language - C++, processor - 1 CPU of an AMD Opteron Quad-Core (2,110 MHz), RAM - 31.5 GB.

<table>
<thead>
<tr>
<th>Data-set</th>
<th>nPos</th>
<th>nTS</th>
<th>ROI</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swirling (EL)</td>
<td>33,425</td>
<td>1+5</td>
<td>3/3</td>
<td>0.2 sec</td>
</tr>
<tr>
<td>Swirling (EC)</td>
<td>33,425</td>
<td>1+2</td>
<td>1/2</td>
<td>0:02</td>
</tr>
<tr>
<td>Draft tube (EL)</td>
<td>0.98 Mio</td>
<td>1+5</td>
<td>3/3</td>
<td>0:56</td>
</tr>
<tr>
<td>Draft tube (EL)</td>
<td>0.98 Mio</td>
<td>251+5</td>
<td>3/3</td>
<td>3h 8min</td>
</tr>
<tr>
<td>Draft tube (L)</td>
<td>0.98 Mio</td>
<td>1+9</td>
<td>5/5</td>
<td>3:24</td>
</tr>
<tr>
<td>Delta wing (L)</td>
<td>3.1 Mio</td>
<td>1+5</td>
<td>3/3</td>
<td>8:04</td>
</tr>
</tbody>
</table>

7. Conclusions

In this paper we introduced two complexity measures that identify coherent structures in Lagrangian and Eulerian flows and are directly comparable. We applied the measures to two different flow scenarios, one with features that hardly move over time and a second one with very unsteady features. In general, the results of the two complexity measures applied to different frames of reference are similar. This is due to the effect that interesting structures are depicted no matter if they pass through a particular position or are followed by pathlines. Differences can be seen with respect to the local extent of the extracted features. While EulerianLSC produces rather soft boundaries as if the complexity field was filtered by a Gaussian kernel, the boundaries in the Lagrangian complexity field are more crisp. Using LagrangianLSC we were able to find a few additional features that originate from more salient events such as pathlines coming close to vortices.

A new direction of application is provided by LagrangianLSC as it can be directly applied to particle traces and highlight periods during which a particle faces unusual dynamics. We applied this technique to the flow around a delta wing and could identify new interesting structures that had not been captured by standard features descriptions.

An interesting direction for future work, is a more intensive research on LagrangianLSC with a special focus on the application to particle traces. A further direction is the extension of the method to other areas of application, such as meteorological data or biological applications in order to give automatic hints where and when to look for interesting phenomena in time-dependent data sets.

References


