Parallel Unsteady Flow Line Integral Convolution for High-Performance Dense Visualization

Zi’ang Ding†
Department of Computer Science,
Purdue University

Zhanping Liu‡
Division of Computer Science,
Kentucky State University

Yang Yu‡
State Key Lab of CAD&CG,
Zhejiang University

Wei Chen‡
State Key Lab of CAD&CG,
Zhejiang University

ABSTRACT

This paper presents an accurate parallel implementation of unsteady flow line integral convolution (UFLIC) for high-performance visualization of large time-varying flows. Our approach differs from previous implementations by using a novel value scattering-gathering mechanism to parallelize UFLIC and designing a pathline reuse strategy to reduce the computational cost of pathline integration. By exploiting the massive parallelism of modern graphical processing units (GPU), the proposed method allows for real-time dense visualization of unsteady flows with high spatial-temporal coherence.

Index Terms: I.3.3 [Computer Graphics]: Pictures / Image Generation

1 INTRODUCTION

Flow visualization plays an important role in processing, displaying, analyzing, and interpreting vector fields resulting from a wide variety of disciplines such as computational fluid dynamics simulation, oceanographic-atmospheric phenomena modeling, and electro-magnetic field analysis. There have been many geometry-based techniques [8] for visualizing flows ranging from steady to unsteady, and from planar to surface and further to volume. Accelerated with graphics libraries and/or hardware, a lot of these methods enable real-time visualization. However, inappropriate seed distribution and insufficient control over the density of geometric primitives tend to incur an incomplete view or a cluttered image.

Texture-based flow visualization techniques [8] avoid the aforementioned problem by taking image space as a large seed set, usually one seed per pixel, to synthesize a dense depiction of the vector field. Instead of explicitly rendering intermediate graphical primitives (points, lines, and polygons) at discrete sample points or along seed-dependent integral traces, the dense representation created by texture-based techniques is a powerful way of conveying both local features and global patterns of a flow. In particular, unsteady flow line integral convolution (UFLIC) proposed by Shen and Kao [15, 16] provides a novel solution for visualizing time-varying flows with high image and animation quality. The use of a time-accurate value-scattering scheme coupled with a successive texture feed-forward strategy in UFLIC mimics the behavior of real-world flows to achieve high spatial-temporal coherence. However, the high computational cost due to a huge amount of pathline integration remains a major issue [12, 11, 10].

In this paper, we propose an accurate parallel implementation of UFLIC for high-performance dense visualization of time-varying flows. This method is unique for incorporating GPU-based pathline integration and texture synthesis with pathline reuse. Our approach runs at real-time frame rates while maintaining high spatial-temporal coherence.

The remainder of this paper is organized as follows. Section 2 introduces previous work closely related to our approach. Section 3 presents our parallel UFLIC algorithm, followed by a GPU implementation. In section 4, we provide some results to demonstrate both the efficiency and accuracy of the proposed method. We conclude this paper with a brief summary and outlook on future works in section 5.

2 RELATED WORK

The past two decades have seen significant research on texture-based techniques, as indicated in a comprehensive survey [8], while our discussion herein is primarily focused on those closely related to our work. Among the texture-based family, spot noise [18] and line integral convolution (LIC) [2] draw the most attention and have far-reaching impact. In particular, LIC has been widely used because of the high-resolution realistic representation and ease of implementation. The basic idea of LIC consists in the application of a 1.5D an-isotropic low-pass filter, shifted along the streamline that is symmetrically advected from each pixel center, to white noise to perform texture convolution. This image synthesis procedure emulates what occurs when a rectangular area of (massless) fine sand is blown by strong wind, exploiting and revealing the spatial correlation in terms of intensity that exists between the pixels hit by the same streamline. LIC is well suited for visualizing steady flows and there have been many variations, optimizations, and extensions such as oriented LIC [20], enhanced LIC [13], fast LIC [17], parallel LIC [22], and volume LIC [6, 14], to name only a few. However, it cannot be directly used to handle time-varying flows due to the inability to address temporal issues. Forssell and Cohen [3] proposed to use pathlines in place of streamlines to convey temporal coherence. Unfortunately, collecting texture values from downstream pixels does not comply with the real-world observation of the behavior of a time-varying flow and therefore their attempt fails to construct temporal coherence.

In contrast with LIC that employs an image-space oriented Eulerian-based value-gathering scheme, unsteady flow line integral convolution (UFLIC) adopts an object-space oriented Langrangian-based value-scattering mechanism to achieve high spatial coherence as well as strong temporal coherence. Specifically, scattering the footprint (or property/texture value) of a particle to its downstream locations along a pathline over several time steps not only correlates a considerable number of intra-frame pixels to create crispy images with accentuated flow streaks, but also correlates sufficient interframe pixels to produce a smooth animation. In addition, UFLIC uses a texture feed-forward strategy by which each output frame, after noise-jittered high-pass filtering, is taken as the input texture for synthesizing the next frame. In this way, an even tighter correlation can be established between any two consecutive frames to enhance temporal coherence. Compared to texture blending techniques such as LEA [7], IBFV [19], UFAC [21], and ISA [9] that incorporate
single-step flow line advection with an exponentially-decaying low-pass filter, UFLIC demonstrates its advantage in maintaining high spatial-temporal coherence. One drawback is that the large computational cost restricts its applicability to interactive visualization of large flows.

The low computational performance of UFLIC is primarily due to the intensive process of advecting long flow lines that typically needs over 100 steps of integration for each pathline. In fact, there is a huge amount of redundancy in pathline integration between different particles during the value scattering process. In other words, different particles may leave their footprints or scatter their property values to a large number of common downstream pixels. Motivated by this observation, Liu and Moorhead proposed an accelerated UFLIC (AUFLIC) algorithm [11, 12] that runs one order-of-magnitude faster than UFLIC with the same image and animation quality. The essence of AUFLIC is to reuse pathlines both within and between value scattering processes, wherever possible and whenever possible, respectively. The ultimate goal is that in any value scattering process, one and only one seed (particle) is made available within each pixel, either explicitly placed from scratch at the center at the exact beginning of an integer time step or implicitly extracted from an existing pathline at an arbitrary position and/or at a fractional time, to move downstream as it scatters its texture value to its receiver pixels.

Li et al [10] presented a GPU-based implementation of UFLIC (GPUFLIC) that exploits the programmability of commodity GPUs to achieve interactive dense visualization of unsteady flows. To deal with the restriction that the single-instruction-multiple-data (SIMD) design of GPU imposes on data structures, they proposed an on-the-fly depositing scheme to obviate the need for a ring-bucket value accumulator as adopted in UFLIC. The interval between any two consecutive time steps is sub-divided into multiple time slots, of which each invokes a simplified, single-step value scattering process to generate a sub-frame. A sub-frame is obtained by performing convolution on the texture values that the upstream particles scatter within the current time slot. On a SIMD computing architecture, this time sub-division strategy effectively lengthens pathlines and accordingly increases the actual window of a convolution kernel. As the output animation is created with only the sub-frames corresponding to the original integer time steps (instead of the shorter time slots), GPUFLIC essentially emulates multi-step pathline integration and hence maintains higher spatial-temporal coherence than by the aforementioned single-step texture advection techniques [7, 19, 21, 9].

3 Our Method

Sharing the same idea of pathline reusing [11, 12] as well as GPU-based implementation to parallelize the pathline generation and value scattering [10], here we present our highly parallel unsteady flow line integral convolution algorithm. In our approach, the entire pipeline, particle management, value scattering and depositing, post-processing is fully CUDA-accelerated and achieves real time frame rates for each test dataset. Figure 1 shows the overview of our pipeline, and each step of our implementation will be discussed in this section.

3.1 Particle Management

At the beginning of each frame $f$, new particles are released at the center of every pixel in a domain of size $M \times N$, where $M$ and $N$ are the spatial resolution of the output image. Also an image value $\alpha = Tex^{f-1}$, where $Tex^{f-1}$ is the image generated by the pervious frame $f - 1$ ($Tex^0$ is the input white noise texture), is associated to each particle. This image value $\alpha$ remains unchanged during the entire life span of the particle.

During any frame the system has to maintain up to $K = M \times N \times L$ active particles, where $L$ is the global particle life span measured in number of frames. After reaching the steady state, at each frame there are $M \times N$ particles reach the life span, hence the memory is recycled to be used by the new released particles.

Similar to the "on-the-fly depositing" method used in GPUFLIC, at every frame we only advect particles from their previous position to the current time instead of tracing them through their entire life span. The segment of a pathline in this time interval, named pathlet, is used to scatter the image value $\alpha$ stored in particle to pixels in current frame.

3.2 Pathline Reuse

After new particles released, all active particles in the system need to be advected forwardly by the unsteady flow. However the advection can be accelerated by parallelizing the computation on GPU [10], the redundancy in pathline integration [1] can be eliminated by a pathline reuse strategy [11, 12].

The strategy is based on an assumption that a pathlet traced by particle $i$ at frame $f$ can be reused by any particle $j$ as long as $i$ and $j$ are in the same pixel at the beginning of this frame. Though those particles may diverge later, reusing pathlet has a very limited influence on the final result which will be discussed at section 4.

Figure 2 illustrates the pathline reuse strategy. Two pathlines $\Theta$(red) and $\Psi$green traced by particle $\theta$ and $\phi$ respectively, come into the same pixel at the beginning of frame $f$, the dot curves indicate the pathlets by continuing tracing $\theta$ and $\phi$ in this frame. Based on the above assumption each dot curve can be reused by the other. Furthermore, the system will release a new particle $\Psi$(yellow) at the center of that pixel. Its pathlet $\Psi$ can also be reused by $\theta$ and $\phi$. Thus, at each frame, only the new released $M \times N$ particles need to be advected to generate the new pathlets. All other active particles reuse those pathlets during the value scattering and gathering step.
3.3 Pathline Reuse based Value Scattering and Depositing

At every frame, the image value \( \alpha \) associated with each particle needs to be scattered over all pixels covered by its pathlet. With our pathline reuse strategy, all particles at the same pixel will share a newly created pathlet and therefore their associated image values can be scattered together by the same pathlet. In contrast, GPU-UFLIC needs to scatter the value separately because all particles are advected individually.

When the system releases new particles, each of them is stored based on the initial starting position. In order to perform the value scattering and depositing by using the benefit of the pathline reuse strategy, each particle has to be fetched based on its current position. This can be achieved by creating a mapping from particle’s current position to its initial position.

The mapping process contains following three operations:

1. **Pixel Particle Counting** Each pixel counts the number of particles that located in it at the starting time of current frame.

2. **Pixel Offset Computing** A prefix sum operation applied on the result of previous operation, and generates the offset for each pixel in the mapping buffer.

3. **Particle to Pixel Mapping** Each particle maps itself to the pixel by the current position and stores it to the mapping buffer based on the offset computed in the previous operation.

Figure 3 illustrates the result after the mapping process. Each pixel has a unique offset which point to the memory location in the mapping buffer and the number of particles inside it, while the mapping buffer contains the actual memory address of each particle.

![Figure 3: Illustration of the result after the mapping process.](image)

The process of value scattering and depositing begins with tracing a seed \( s(x, y) \) placed in the center of each pixel \( p(x, y) \). In each step, a seed \( s(x, y) \) is advected by solving an ordinary differential equation on the input unsteady vector field. We employ higher-order numerical integration methods such as 4th order Runge-Kutta to compute the new position \( POS_k \) from the previous one \( POS_{k-1} \). All particles located at the pixel \( p(x, y) \) corresponding to that seed \( s(x, y) \), which could be fetched by the mapping buffer, scatter their associated image values by depositing them to the pixels on the image plane covered by the line segment \( POS_{k-1}POS_k \). We use one buffer \( R_v \) to store the sum of weighted values \( \sum V \times w \), and another one \( R_w \) to store the accumulated weights \( \sum w \). In our implementation, each weight is determined by two factors, namely, the life time \( \tau \), and the remaining life time \( \tau \). This memory is divided into \( L \) layers, each layer is further divided into \( M \times N \) blocks which store the information of one particle. At any frame \( f \), all particles in the layer \( f \%L \) (\( f \) mod \( L \)) will reach their life span, and reuse by the new released particles. Please refer to Figure 4 as the layout of the global memory for particles.

![Figure 4: The layout of the global memory for particles.](image)

3.4 Post-Processing

Mentioned in UFLIC [15, 16], the LIC method is really a low-pass filtering process on the input image. Simply using the current result as the input of next frame would lead to an over-blurred image. Postprocessing is required to maintain a satisfying contrast of the result throughout the animation.

The first step of the postprocessing is applying a high-pass filter to enhance the image contrast. Similar to the UFLIC, the following Laplacian operator is applied to the current result:

\[
\begin{bmatrix}
-1 & -1 & -1 \\
-1 & 9 & -1 \\
-1 & -1 & -1
\end{bmatrix}
\begin{bmatrix}
0 \\
1 \\
0
\end{bmatrix}
= \begin{bmatrix}
1 \\
-8 \\
1
\end{bmatrix}
\]

(1)

In order to prevent the unnecessary high frequencies introduced by the high pass filter, we apply a second step, named noise jittering, to the output from the high-pass filter. Similar to the GPU-UFLIC, to address the frozen patterns in steady regions, this step replaces the lower seven bits of the output by a periodic noise.

3.5 CUDA Implementation

The above presented algorithm can be fully parallelized on the massive parallel architecture such as CUDA (Compute Unified Device Architecture) [10]. Although most parts of the CUDA-based codes are identical to C code on CPU, there are still some details need to be presented in our implementation.

3.5.1 Memory

As mentioned before, at any frame the number of total particles in the system is bounded by \( K = M \times N \times L \). Thus, a pre-allocated memory on the global memory in CUDA is used to store the information of particles, namely, the current position \( POS \), the image value \( \alpha \), and the remaining life time \( \tau \). This memory is divided into \( L \) layers, each layer is further divided into \( M \times N \) blocks which store the information of one particle. At any frame \( f \), all particles in the layer \( f \%L \) (\( f \) mod \( L \)) will reach their life span, and reuse by the new released particles. Please refer to Figure 4 as the layout of the global memory for particles.

The input 2D unsteady vector field is stored as a 3D texture memory in CUDA, which is cached and no extra cost on trilinear interpolation. Each slice on the z axis corresponds to a time step in the 2D unsteady vector field. With this configuration, the spatial-temporal coordinate \((x,y,t)\), where t is the global time, is transformed to the texture coordinate \((u,v,w)\) in our implementation to access the flow velocity at any position of any time.
3.5.2 Concurrency

A race condition arises when two or more threads in CUDA attempt to access the same global memory concurrently and at least one access is a write operator. In our algorithm the value scattering and depositing step will trigger the race condition, for example, two threads attempt to update the value stored in the same memory address of $B_V$ at the same time. Atomic operators supported by CUDA such as, atomicAdd, are used in order to avoid unexpected results caused by the race condition.

3.5.3 Bandwidth

To maximally optimize the use of the bus bandwidth, our algorithm is implemented entirely on CUDA. At the initial state of the program, the input unsteady vector field and white noise image are transferred from the host memory to the video memory. After that, no data transfer is needed. The result of each frame is displayed effectively by using the pixel buffer object feature supported by CUDA.

4 Result and Discussion

We implemented our algorithm described in section 3 using C++ and CUDA on a Windows PC. The platform we used in our experiment has a hex-core i7 processor at 3.2GHz with 12GB of RAM and an NVIDIA GTX590 CUDA-Enabled graphics card (dual GPUs, 1536MB video RAM per GPU).

Four time-varying vector fields, including PSI(figure 5), double gyre(figure 6), wind flow(figure 7), and convection(figure 8) were tested in our experiment. The life span used in each test case is listed in table 1. Both PSI and Double Gyre are standard synthetic datasets, while wind flow and convection are results from the numerical simulation of certain physical phenomena. The color coding in figure 5 and figure 7 is based on the magnitude of local velocity, while in figure 8 it highlights the temperature. FTLE color coding is used in figure 6. For more information about the FTLE color coding, please refer to [4].

Figure 5: PSI dataset.

Table 1 gives the performance of each dataset by our implementation. In all cases we are able to achieve real-time frame rates at different image resolutions. We have implemented GPUFLIC algorithm and make the performance comparison in the same hardware platform. Table 1 indicates our algorithm outperforms it because of the pathline reuse strategy. It is apparent in figure 9 that the performance of GPUFLIC decreases faster and runs less stable than our approach with the increasing of the life span.

Figure 9: Performance comparison in FPS with different life spans. All results are tested on PSI dataset with resolution of $512 \times 512$.

In order to evaluate the influence of pathline reuse, we made a side by side comparison between our approach and GPUFLIC. The first row in figure 10 shows the result of our approach on PSI dataset without color coding, and the second row is the result of GPUFLIC at corresponding frames. Our approach is able to capture the same high quality crisp line patterns matching those generated by GPUFLIC.

5 Conclusion and Future Work

We have presented a GPU-based parallel algorithm that synthesizes LIC-like textures at real-time frame rates for dense visualization of time-varying 2D flows. It is built on UFLIC to not only convey very high spatial coherence in individual frames but also exhibit very strong temporal coherence in the resulting animation. Our algorithm runs faster than AUFLIC by taking advantage of GPU parallelism while it differs from GPUFLIC due to a pathline reuse mechanism as well as a relatively efficient on-the-fly value scattering strategy. The results show that the proposed approach has a great potential, in both visualization quality and computational performance, to bring UFLIC to practical applications.

As for future work, we would like to extend our algorithm for curved surface and volumetric flows. We are also interested in exploring the use of an adaptive step size integration method in pathline advection.

Acknowledgements

The authors wish to thank Prof. Xavier Tricoche of Purdue University for his helpful suggestions and comments. This work was supported by Major Program of National Natural Science Foundation of China (61232012) and National Natural Science Foundation of China (61422211).

References


Table 1: Performance in average FPS.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Resolution</th>
<th>Time steps</th>
<th>Life span</th>
<th>GPUFLIC (FPS)</th>
<th>Ours (FPS)</th>
<th>Speedup Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSI</td>
<td>512 × 512</td>
<td>101</td>
<td>4</td>
<td>19.5</td>
<td>86.6</td>
<td>4.4</td>
</tr>
<tr>
<td>PSI</td>
<td>1024 × 1024</td>
<td>101</td>
<td>4</td>
<td>7.0</td>
<td>31.9</td>
<td>4.6</td>
</tr>
<tr>
<td>Double Gyre</td>
<td>1024 × 512</td>
<td>100</td>
<td>4</td>
<td>14.1</td>
<td>59.3</td>
<td>4.2</td>
</tr>
<tr>
<td>Wind Flow</td>
<td>1024 × 512</td>
<td>41</td>
<td>4</td>
<td>17.4</td>
<td>65.4</td>
<td>3.7</td>
</tr>
<tr>
<td>Convection</td>
<td>1024 × 512</td>
<td>200</td>
<td>4</td>
<td>11.5</td>
<td>48.2</td>
<td>4.2</td>
</tr>
</tbody>
</table>

Figure 8: Convection dataset.

Figure 10: A side-by-side comparison between our approach and GPUFLIC on PSI dataset.