

Flow Feature Extraction Based on Entropy and Clifford Algebra

Xiaofan Liu, Wenyao Zhang^(✉), and Ning Zheng

Beijing Key Laboratory of Intelligent Information Technology,
School of Computer Science, Beijing Institute of Technology,
Beijing 100081, People's Republic of China
zhwenyao@bit.edu.cn

Abstract. Feature extraction is important to the visualization of large scale flow fields. To extract flow field features, we propose a new method that is based on Clifford algebra and information entropy theory. Given an input 3D flow field defined on uniform grids, it is firstly converted to a multi-vector field. We then compute its flow entropy field according to information theory, and choose high entropy regions to do the Clifford convolution with predefined multi-vector filter masks. Features are determined on the convolution results. With this method, we can locate, identify, and visualize a set of flow features. And test results show that our method can reduce computation time and find more features than the topology-based method.

Keywords: Flow visualization · Feature extraction · Clifford convolution · Flow field entropy

1 Introduction

Feature extraction is an important way to visualize flow fields in scientific computing. Feature visualization obtains meaningful structures from data fields, overcoming the disadvantages of traditional visualization methods. For 2D flow fields, topology analysis [1] is a traditional visualization based on feature extraction. Most methods based on topology analysis [2–4] work well in 2D feature extraction, but can not be simply extended to 3D flow fields. For example, the vortex core line reported in [5] can not be detected by topology analysis.

With such problems, methods based on Clifford algebra were developed to detect flow features [6], where multi-vector filter masks are predefined and used to convolve with the underlying flow fields to determine features. The main drawback of these methods is the heavy convolution, because it has to be applied on the whole field. Although Hitzer et al. [7] proposed the Clifford Fourier transformation to accelerate it, there is still an overhead to be cut down.

On the other hand, information theory was applied to measuring the complexity of streamlines [8]. And an information-theoretic framework was proposed and used for flow visualization [9–11]. With the information theory, flow field complexity can be measured with flow field entropy. The more complicated the areas are, the higher the entropy values are. This gives us some kinds of information about the structures in the underlying field.

With the motivation to make feature extraction easier, we propose a new method to extract features for 3D flow fields. In this method, we first compute the entropy of the flow field according to information theory, and then choose high entropy regions to do the Clifford convolution with multi-vector filter masks. Finally, features are determined on the convolution results. This method integrates information computing into flow field Clifford convolution, and increases the efficiency of feature extraction by avoiding most of unnecessary convolutions.

The remainder of this paper is organized as follows. Section 2 briefly reviews the related work. Section 3 gives our approach, including some basics about Clifford algebra and flow field entropy. Test results are included in Sect. 4. The final is the conclusions.

2 Related Work

In flow visualization, feature extraction is often used to simplify and visualize flow fields in large scale. And many methods have been proposed for flow feature extraction. Here, we only briefly review the work directly related to our method.

Based on the scalar product between two arbitrary vectors, Heiberg et al. [12] defined a scalar convolution on flow vector fields, and presented a template matching method to extract features in vector fields. It can detect most of flow features, for example, vortices, swirling flow, diverging, converging, etc. In [6], Ebling and Scheuermann proposed a pattern matching method based on Clifford convolution. This method can detect flow features, but its computation is heavy. In [7], Hitzer extended Fourier transform to Clifford algebra and got the accelerated Clifford convolution.

On the other hand, information theory was introduced for flow visualization. Chen and Jäenicke [8] presented an information-theoretic framework for flow feature analysis and visualization. Wang et al. [9] presented a block-wise technique to analyze the important aspect of time-varying data. In [10], Furuya and Itoh applied information theory to measure the complexity of existing streamlines based on the distribution of the tangent orientations. In [11], Xu et al. gave an information-theoretic framework for flow visualization, where the information of vector fields is measured by Shannon's entropy. In [13], Tao et al. proposed an information-theoretic framework, which applied dual information channels to solve streamline selection and viewpoint selection problems in flow visualization.

3 Our Approach

3.1 Overview

In this paper, we propose a feature extract method for 3D flow fields. This method shares the framework of pattern matching that is based on Clifford convolution, but includes a processing of region selection that is guided by flow field entropy. In the traditional pattern matching, convolution operator has to be applied on the whole input field. For some regions, however, there is no need to do convolution, since no meaningful features are contained in them. In order to reduce redundant overhead, we

calculate the flow field entropy firstly, and then restrict the flow field convolution to regions with high entropy values. By this means, we can identify flow patterns and features precisely and efficiently. As a whole, our method works as followed.

- (1) Initially, the input flow field is normalized and converted to a multi-vector field.
- (2) Multi-vector filter masks are prepared for different flow patterns and features.
- (3) Calculate the flow field entropy according to the information theory of flow field, and get the entropy field of the input field.
- (4) Find out all local maxima in the entropy field, and put them into the candidates of potential feature points.
- (5) For each candidate points, do the pattern matching using predefined multi-vector filter masks in the framework of Clifford convolution. If the neighborhood region of the candidate point matches with one mask, it is then recognized as a feature with the corresponding pattern.
- (6) All flow patterns and features obtained in previous step are visualized with streamlines.

To help the understanding of our method, some concepts and processes are further described in the following subsections.

3.2 Vector Field Clifford Convolution

Convolution operator is usually applied to scalar fields. If X is a scalar field and Y is a scalar-value filter, the convolution between X and Y is defined as:

$$S_n(r) = \int \int \int \langle X_n(\xi), Y(r - \xi) \rangle d\xi \quad (1)$$

Ebling and Scheuermann [6] extended the convolution operator to vector field by using the geometric product of multi-vectors given in the Clifford algebra framework [14]. In the Clifford algebra, there are 8 basis blades, $\{1, e_1, e_2, e_3, e_1e_2, e_1e_3, e_2e_3, e_1e_2e_3\}$, which form an eight-dimensional algebra space whose bases are in fact 3D vectors. The general elements in this algebra are called multi-vectors. A multi-vector of an n -dimensional Clifford algebra is a linear combination with real coefficients of the 2^n basis blades. The rule of multiplication operator is defined by the following equations:

$$1e_j = e_j, \quad j = 1, 2, 3 \quad (2)$$

$$e_j e_j = 1, \quad j = 1, 2, 3 \quad (3)$$

$$e_j e_k = -e_k e_j, \quad j, k = 1, 2, 3, j \neq k \quad (4)$$

$$(e_1 e_2 e_3)^2 = e_1 e_2 e_3 e_1 e_2 e_3 = e_1 e_1 e_2 e_3 e_2 e_3 = -e_2 e_2 e_3 e_3 = -1 \quad (5)$$

The multiplication of two vectors \mathbf{a} and \mathbf{b} can be written as:

$$\begin{aligned}
 \mathbf{ab} &= (a_1e_1 + a_2e_2 + a_3e_3)(b_1e_1 + b_2e_2 + b_3e_3) \\
 &= (a_1b_1 + a_2b_2 + a_3b_3) + (a_1b_2 - a_2b_1)e_1e_2 \\
 &\quad + (a_2b_3 - a_3b_2)e_2e_3 + (a_3b_1 - a_1b_3)e_3e_1 \\
 &= \langle \mathbf{a}, \mathbf{b} \rangle + \mathbf{a} \wedge \mathbf{b}
 \end{aligned}
 \tag{6}$$

The Clifford multiplication is in fact a convolution of a point in the vector field with a 1×1 filter mask. Therefore, if X is a 3D multi-vector filter mask and Y is a 3D multi-vector field, then the discrete convolution between X and Y can be defined as:

$$C(j, k, l) = \sum_{p=-m}^m \sum_{q=-m}^m \sum_{r=-m}^m X(p, q, r)Y(j - p, k - q, l - r) \tag{7}$$

where j, k, l, p, q, r are coordinates of grid nodes, m is the dimension of the grid of the filter mask and $(j, k, l), (p, q, r)$ are grid nodes.

Given a filter mask with fixed direction, it will not match all the kind of the corresponding features in different directions. So we have to rotate the filter mask in a great deal of directions, which will cost too much computation time. As we can see in Eq. (6), however, $\langle \mathbf{a}, \mathbf{b} \rangle$ is the inner product, and $\mathbf{a} \wedge \mathbf{b}$ is the outer product of the two vectors. So their multiplication gives us the sine and cosine of the angle between two vectors as follow:

$$\langle \mathbf{ab} \rangle_0 = \langle \mathbf{a}, \mathbf{b} \rangle = \|\mathbf{a}\| \|\mathbf{b}\| \cos \alpha \tag{8}$$

$$\|\langle \mathbf{ab} \rangle_2\| = \|\mathbf{a} \wedge \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin \alpha \tag{9}$$

where α is the angle between \mathbf{a} and \mathbf{b} . Therefore, when we get the discrete convolution between X and Y as shown in Eq. (7), we can estimate the angle between X and Y from the following equations:

$$\langle C(j, k, l) \rangle_0 \approx \eta \cos \alpha \tag{10}$$

$$\langle C(j, k, l) \rangle_2 \approx \eta \sin \alpha \tag{11}$$

where $\eta = \sum_{p,q,r} |X(p, q, r)| \leq \sum_{p,q,r} 1$, $|X(p, q, r)|$ is the magnitude of vector $X(p, q, r)$. With the estimated angle, we can directly rotate the filter mask to the desired direction, and do the Clifford convolution to detect features, avoiding checking all possible directions.

3.3 Flow Field Information Entropy

In our method, we introduce information entropy theory to reduce the calculation of Clifford convolution. Information theory provides a theoretical framework to quantitatively measure the information contained in a distribution of data values. In flow fields, flow directions around flow features are often in a high degree of variation, and then have high entropy values.

Given a random variable X with a sequence of possible values $\{x_1, x_2, \dots, x_n\}$, if the probability for X to have the value x_i is $p(x_i)$, then the information content for the random variable X can be computed using Shannon's entropy:

$$H(x) = - \sum_{x_i} p(x_i) \log_2 p(x_i) \quad (12)$$

Xu et al. [11] took Shannon's entropy to compute the information of a vector field, where a map is created to approximate the probability mass function $p(x)$. For 3D vector fields in our consideration, a unit sphere is firstly divided into 642 uniform bins by the icosahedrons subdivision method [15]. Then a map is built from vectors to bins. With the map, the probability of the vectors in the bin x_i can be defined as:

$$p(x_i) = \frac{C(x_i)}{\sum_{i=1}^n C(x_i)} \quad (13)$$

where $C(x_i)$ is the number of vectors in the bin x_i . With the probabilities calculated here, we can get the flow field entropy according to the Eq. (12).

In practice, the flow field entropy is calculated in local regions. Given a grid point, a local neighborhood around the point, whose sizes are $3 \times 3 \times 3$, is used to build the map and get the estimated probability. By this means, we can get a scalar field which is called entropy field. The value in the entropy field shows the degree of vector variation in its neighborhood. Such a 2D example is given in Fig. 1, where regions with higher degree of variation have higher entropy values, and they are considered to contain more information than others. The entropy $H(x)$ is convex and reaches its maximum when $p(x_i)$ is equal for all x_i . We use this property of entropy field to reduce the overhead of Clifford convolution.

3.4 Multi-vector Filter Masks

Multi-vector filter masks can be regarded as very small vector fields, each of which contains a kind of flow patterns that is defined by the vectors. With predefined filter masks, we can detect the desired features through pattern matching. For 3D flow fields, interesting features [16] include as vortices, critical points, convergence, divergence, etc. We can design a representative filter masks for each of them. Two example masks are shown in Fig. 2. Due to the length limitation, other masks are not included in this paper.

4 Test Results

To evaluate the performance of our feature extraction method, we compared it with the topology-based method [17] where feature points are located by flow field interpolation.

Firstly, the flow field shown in Fig. 3(a) was used to test the method. This is a vortex field with a core line. Due to the fact that no critical point exists in this field, the

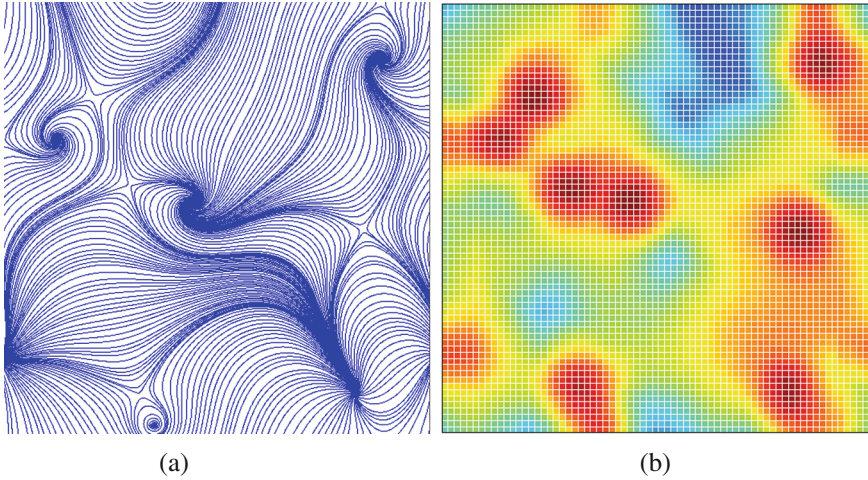


Fig. 1. Example of entropy field for a vector field (a) is the vector field visualized by streamlines, and (b) is its the entropy field, where warmer colors indicate higher entropy values.

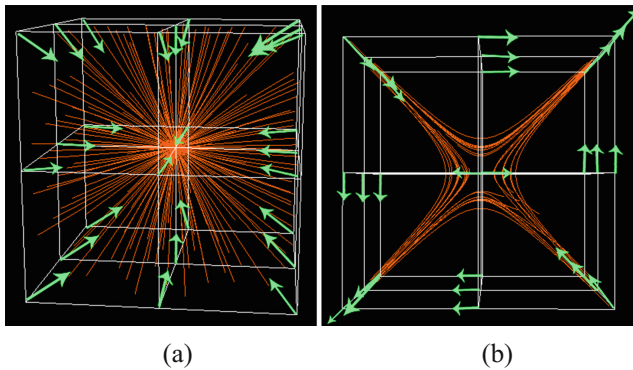


Fig. 2. Examples of multi-vector filter masks (a) is for sink point, (b) is for saddle point.

topology-based method can not find out the core line, which is the important feature of this field. For our method, it found the feature point as shown in Fig. 3(b). When a few streamlines associated with the feature point were rendered, we got the results in Fig. 3(c), where we can notice a straight line through the feature point. This line is the vortex core line.

For general flow fields, both our method and the topological method can find the same features, but our method is general faster than the later. For example, for the flow field in Fig. 4(a), it cost 926 ms for the topological method to find out the feature point shown in Fig. 4(b), while it was only 208 ms for our method. The flow pattern around the feature point was clearly disclosed by streamlines in Fig. 4(c).

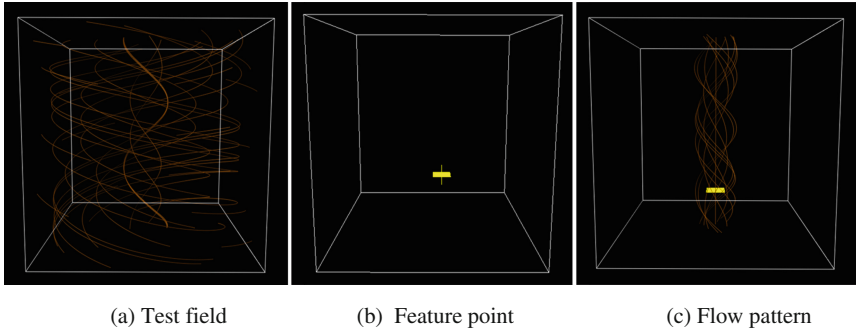


Fig. 3. Test results for the flow field with a vortex core line (a) is the flow field shown by streamlines, (b) is the feature point found by our method, and (c) is the flow pattern around the feature point, where the vortex core line is disclosed.

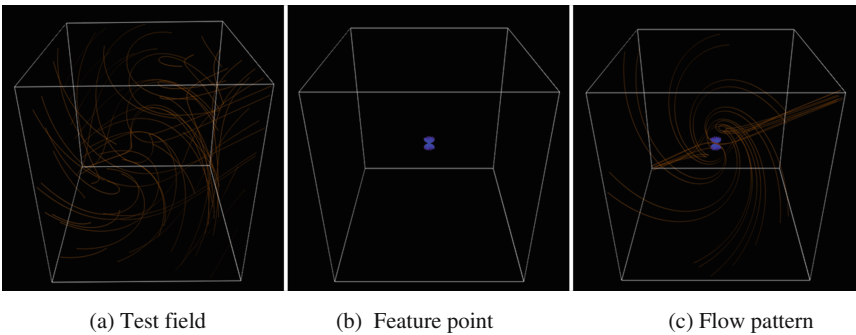


Fig. 4. Test results for a simple flow field (a) is the original field, (b) is the feature point in the field, and (c) is the flow pattern around the feature point.

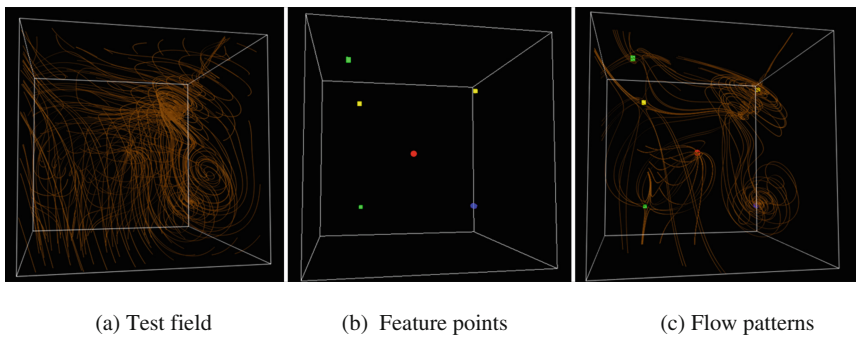


Fig. 5. Test results for the complex flow field (a) is the original field, (b) is the feature points, and (c) is the flow patterns built with feature points and a few streamlines.

Figure 5 gives another test results for a complex flow field. This field was generally visualized by streamlines in Fig. 5(a), where it is hard to capture the flow pattern. For this field, our method got the feature points shown in Fig. 5(b), and further built the flow patterns in Fig. 5(c) with a few streamlines.

5 Conclusions

We present a new method in this paper to extract features in 3D flow fields. The main advantage of our method is that it can extract various types of flow features in a unified and efficiency way. Our method is based on the Clifford algebra convolution, but it reduces the cost of convolution by the computing of flow field entropy. Test results show that our method improves the extraction of flow features. Currently our method is only suitable for flow fields defined on rectangular grids. It would be extended to arbitrary grids in future.

Acknowledgments. This work was partially supported by the grants from the Natural Science Foundation of China (Nos.: 11472049, 11272066).

References

1. Gyulassy, A., Natarajan, V.: Topology-based simplification for feature extraction from 3D scalar fields. In: Visualization, 2005 VIS 2005, pp. 535–542. IEEE, Oct 2005
2. Zhang, W., Deng, J.: Topology-driven streamline seeding for 2D vector field visualization. In: IEEE International Conference on Systems, Man and Cybernetics, 2009, SMC 2009, pp. 4901–4905, Oct 2009
3. Reininghaus, J., Lowen, C., Hotz, I.: Fast combinatorial vector field topology. *IEEE Trans. Vis. Comput. Graph.* **17**(10), 1433–1443 (2011)
4. Otto, M., Germer, T., Hege, H.C., Theisel, H.: Uncertain 2D vector field topology. *Comput. Graph. Forum* **29**(2), 347–356 (2010)
5. Tricoche, X., Garth, C., Kindlmann, G., Deines, E., Scheuermann, G., Ruetten, M., Hansen, C.: Visualization of intricate flow structures for vortex breakdown analysis. In: Proceedings of the Conference on Visualization 2004, VIS 2004, pp. 187–194. IEEE Computer Society, Washington, DC (2004)
6. Ebling, J., Scheuermann, G.: Clifford convolution and pattern matching on vector fields. In: Visualization, 2003, VIS 2003, pp. 193–200. IEEE, Oct 2003
7. Hitzer, E., Mawardi, B.: Clifford fourier transform on multivector fields and uncertainty principles for dimensions $n = 2 \pmod{4}$ and $n = 3 \pmod{4}$. *Adv. Appl. Clifford Algebras* **18** (3–4), 715–736 (2008)
8. Chen, M., Jäenicke, H.: An information-theoretic framework for visualization. *IEEE Trans. Vis. Comput. Graph.* **16**(6), 1206–1215 (2010)
9. Wang, C., Yu, H., Ma, K.L.: Importance-driven time-varying data visualization. *IEEE Trans. Vis. Comput. Graph.* **14**(6), 1547–1554 (2008)
10. Furuya, S., Itoh, T.: A streamline selection technique for integrated scalar and vector visualization. In: Vis Š08: IEEE Visualization Poster Session, vol. 2, no. 4 (2008)

11. Xu, L., Lee, T.Y., Shen, H.W.: An information-theoretic framework for flow visualization. *IEEE Trans. Vis. Comput. Graph.* **16**(6), 1216–1224 (2010)
12. Heiberg, E., Ebbers, T., Wigstrom, L., Karlsson, M.: Three-dimensional flow characterization using vector pattern matching. *IEEE Trans. Vis. Comput. Graph.* **9**(3), 313–319 (2003)
13. Tao, J., Ma, J., Wang, C., Shene, C.K.: A unified approach to streamline selection and viewpoint selection for 3D flow visualization. *IEEE Trans. Vis. Comput. Graph.* **19**(3), 393–406 (2013)
14. Fernandez, V., Moya, A., Rodrigues, W.: Euclidean clifford algebra. *Adv. Appl. Clifford Algebras* **11**(3), 1–21 (2001)
15. Shen, L., Ford, J., Makedon, F., Saykin, A.: A surface-based approach for classification of 3D neuroanatomic structures. *Intell. Data Anal.* **8**(6), 519–542 (2004)
16. Ebling, J., Scheuermann, G.: Template matching on vector fields using clifford algebra. In: *International Conference on the Applications of Computer Science and Mathematics in Architecture and Civil Engineering (IKM 2006)* (2006)
17. Wang, Y., Zhang, W., Ning, J.: Streamline-based visualization of 3D explosion fields. In: *2011 Seventh International Conference on Computational Intelligence and Security (CIS)*, pp. 1224–1228, Dec 2011