Morse–Smale decomposition of multivariate transfer function space for separably-sampled volume rendering

Natallia Kotava\textsuperscript{a, *}, Aaron Knoll\textsuperscript{b}, Hans Hagen\textsuperscript{a}

\textsuperscript{a} University of Kaiserslautern, Gottlieb-Daimler-Str., 67663 Kaiserslautern, Germany
\textsuperscript{b} Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA

\textbf{A R T I C L E   I N F O}

\textbf{Article history:}
Available online 16 April 2012

\textbf{Keywords:}
Topology
Morse theory
Volume rendering
Multidimensional transfer functions

\textbf{A B S T R A C T}

We present a topology-guided technique for improving performance of multifield volume rendering with peak finding and preintegration with 2D transfer functions. We apply Morse–Smale decomposition to segment the multidimensional transfer function domain. This segmentation helps to reduce the number of cases where sampling in transfer function space should be performed, effectively reducing the rendering cost for equivalent sampling quality. We show that the overall performance is increased depending on the topology of a transfer function.

© 2012 Elsevier B.V. All rights reserved.

\section{1. Introduction}

Direct volume rendering is a popular tool for scientific visualization of both single-variable and multivariable data. Its cost is dependent on the number of samples computed. When using a transfer function to classify volume data, the number of samples required for antialiased renderings depends on the Nyquist frequency of the scalar field, convolved by the transfer function. When either the data or transfer function contain high-frequency features, a high sampling rate is needed, hampering performance.

Principally, two techniques, preintegration and peak finding, exist for decoupling the sampling rate of the volume and transfer function. These methods allow the user to sample at the maximum of either the volume or transfer function’s frequency, as opposed to their convolved (multiplied) frequency. In the univariate case, both preintegration and peak finding can be classified by a simple 2D table lookup. For multivariate data, these processes are significantly more expensive. Multidimensional preintegration (Kraus, 2008b) requires a summed area table lookup and more sophisticated sampling along the viewing ray. Multidimensional peak finding (Knoll et al., 2010) entails chordal sampling of transfer function space between world-space sampling points. In both cases, rendering can be greatly simplified by only performing separable classification in regions that need it. Monotonic regions of transfer function space have no peaks, and can be sampled regularly in world-space with little change in quality.

Determining these monotonic regions in 1D is trivial; however in the multivariate case we can use topological methods to analyze the transfer function in the preprocessing stage. In this paper we present the application of a topology-based segmentation technique to multidimensional transfer functions.

The Morse–Smale complex of the transfer function gives us the information about regions of monotonous gradient behavior. The computationally expensive multidimensional peak finding can then be reduced to samples which fall into different MS cells. We show that depending on the topology of the transfer function this can increase the performance of the volume rendering.
rendering. In any case, the quality of the volume rendering result stays the same as in multidimensional peak finding without topological analysis.

The remainder of the paper is organized as follows. In Section 2 we give a brief overview of the work done in the areas of volume rendering, Morse–Smale complex computation and topology-based image segmentation. Section 3 introduces the reader to the basics of volume rendering with peak finding, Morse theory and Morse–Smale image segmentation. Section 4 describes the method we use for the optimization of the multidimensional peak finding technique. The main results are presented in Section 5. Section 6 discusses the advantages and drawbacks of the presented method and gives directions for future work.

2. Related work

Direct volume rendering (DVR) as a visualization modality was first introduced by Levoy (1988). While slicing (Cullip and Neumann, 1994) and splatting (Westover, 1990) remain feasible implementations, ray casting is largely the state-of-the-art rendering algorithm for both GPU (Krüger and Westermann, 2003) and CPU hardware (Knittel, 2000).

Preintegration (Röttger et al., 2000; Engel et al., 2001) employs separable integration of world and transfer function space, by precomputing the summed irradiance in the 1D transfer function domain. The irradiance for a ray segment can then be queried in a 2D lookup table. Peak finding (Knoll et al., 2009) notes that separable integration is most useful when either the frequency of the volume or transfer function is high. By sampling directly at peaks in the transfer function space, this modality effectively combines direct volume rendering and isosurface ray casting, and allows for effective rendering of high-frequency features at lower sampling rates. Ament et al. (2010) show a more robust alternative to peak finding, which renders regions between isovalue intervals in a scale-invariant manner.

Multidimensional classification was first applied in visualization of MRI data (Laidlaw, 1995). Subsequent work in medical visualization showed how 2D classification of isovalue and inverse gradient magnitude could be useful in identifying features and suppressing noise (Kniss et al., 2002; Kindlmann, 2002). 3D classification using curvature was demonstrated by Kindlmann et al. (2003). For more general multifield volume rendering, Kniss et al. (2003) used Gaussian basis functions and special analytical integration to reduce artifacts. However, constructing such transfer functions can be non-intuitive and restricts classification. Schulze and Rice (2004) demonstrate how simple blending operations can be used towards effective multifield volume visualization.

Separable sampling is more complicated in the case of multidimensional classification. Kraus (2008a) shows how summed area tables can be used to approximate area integrals in 2D transfer function space. By pairing this with beam tracing, 1D preintegration can be extended to 2D transfer functions. However, this approach would not trivially extend to higher-dimensional classification.

Knoll et al. (2010) propose sampling directly in transfer function space along curve approximations of the image of the ray. This approach can effectively extend both peak finding and preintegration to arbitrary-dimensional transfer functions with no precomputation or storage. Unlike 1D peak finding, the method must perform the sampling query along every segment, which is expensive. Since high frequencies appear more readily in high-dimensional classification, separable sampling is even more useful for multidimensional classification than for 1D transfer functions. However, performance could be significantly improved by only performing transfer-space sampling where necessary.

Morse theory (Milnor, 1963; Matsumoto, 2001) has been proven to be a powerful tool for the description of the topology of a manifold. Providing an intuitive connection between the topology and geometry, the theory has found applications in scientific visualization in recent decades. Because of the discrete nature of the data, concepts from smooth Morse theory had to be transferred to a discrete setting, leading to the development of the piecewise-linear (PL) Morse theory (Edelsbrunner et al., 2003) and the discrete Morse theory (Forman, 2001). Based on the Forman’s theory, several techniques for the computation of the discrete gradient fields have been proposed (Lewiner et al., 2004; King et al., 2005; Robins et al., 2011). The Morse–Smale (MS) complex is a well-known and useful topological structure. However, the computation and hierarchical representation of it is not a trivial task. The concept of quasi-MS complexes for PL 2-manifolds and the method for their computation and simplification were first discussed in Edelsbrunner et al. (2001). Later Bremer and Pascucci (2007) addressed the computational issues of the initial algorithm and proposed data structures to optimize its performance. Reininghaus et al. (2010) used a graph theoretical formulation of the discrete Morse theory to extract the extremal structure in 2D. Algorithms for computing the Morse–Smale complex of volumetric data were proposed in Gyulassy et al. (2008), Günther et al. (2011).

3. Background

3.1. Volume rendering

As employed in visualization, volume rendering is a discrete approximation of the emission component of the emission–absorption model in the radiative transport equation (Levoy, 1988), where irradiance on a ray segment bounded by \([a, b]\) is given by:
\[ I(a, b) = \int_a^b \rho_E(f(s))\rho_\alpha(f(s))e^{- \int_0^s \rho_\alpha(f(t))\,dt}\,ds \]  

(1)

where \( \rho_E \) is the emissive term or color and \( \rho_\alpha \) is opacity given by the transfer function. The ray is parameterized as \( \tilde{R}(t) = \tilde{O} + t \tilde{D} \), and \( f(\tilde{R}(t)) \) is the scalar field function. Eq. (1) is computed discretely as a Riemann sum

\[ e^{- \int_0^t \rho_\alpha(f(t))\,dt} \approx \prod_{i=0}^n e^{- \Delta t \rho_\alpha(f(i \Delta t))} = \prod_{i=0}^n (1 - \alpha_i) \]  

(2)

where \( \Delta t \) is the uniform sampling step, \( n = (s - a)/\Delta t \), and

\[ \alpha_i = 1 - e^{- \Delta t \rho_\alpha(f(i \Delta t))} \]  

(3)

Thus, irradiance is given by the summation:

\[ I \approx \sum_{i=0}^n \tilde{\rho}_E(i) \prod_{j=0}^{i-1} (1 - \alpha_j), \quad \tilde{\rho}_E(i) \approx \rho_\alpha(f(i \Delta t))\rho_E(f(i \Delta t)) \]

We refer to as separably sampled techniques that use a separate discretization of the transfer function \( \rho \) in integrating \( I \). Preintegration (Engel et al., 2001) approximates \( \rho \) as its own Riemann sum between two samples \( f_0 = f(t_0) \) and \( f_1 = f(t_1) \), and assumes that \( f \) varies linearly between these points. Then, the emissive term is estimated as:

\[ \alpha_i = 1 - e^{- \int_0^t \rho_\alpha((1 - \omega)f_0 + \omega f_1)\,d\omega} \]  

(4)

Peak finding (Knoll et al., 2009) assumes that the frequency of the transfer function is high where one wishes to sample separably, and that over a segment \( I \) is best approximated by the supremum:

\[ \alpha_i \approx 1 - e^{\left(\sup_{f \in [f_0, f_1]} \rho_\alpha(f)\right)} \]  

(5)

Samples at peaks are integrated in a scale-invariant (Kraus, 2005) manner, assuming that an isosurface is always sampled regardless of world-space discretization \( \Delta t \). In 1D peak finding (Knoll et al., 2009), all other samples (without peaks) employed postclassification with scale-variant integration. In multidimensional peak finding (Knoll et al., 2010), there is no mechanism to determine which samples are at peaks, and all samples are treated as potential peaks in a scale-invariant integration. In this paper, we maintain scale-invariant integration, noting that when \( f \) is monotonic over a segment, a peak exists at one of the endpoints.

Peak finding and preintegration both provide higher-quality volume rendering at lower sampling cost. Preintegration is preferable when the volume and transfer function have moderately high frequencies, but more or less bounded derivative (i.e. they are Lipschitz). Though it exhibits more discontinuity, peak finding is better at rendering data with extremely high frequencies in the transfer function or volume domain. As shown in Knoll et al. (2010), one can sample in transfer-space and integrate using a Riemann sum, effectively achieving the preintegrated modality using the multidimensional peak finding sampling algorithm.

### 3.2. Morse theory

In this section we describe the basic concepts from Morse theory. For more detailed information we refer to Milnor (1963), Matsumoto (2001).

Morse theory characterizes the homotopy type (under a more precise observation even the diffeomorphic type) of an \( n \)-dimensional Riemannian manifold through the behavior of a differentiable function \( f \) in its critical points, assuming that \( f \) has only non-degenerate critical points.

Let \( f \) be a smooth real-valued function on a manifold \( M \). A point \( p \in M \) is called a critical point of \( f \) if the differential of \( f \) vanishes at this point.

A critical point is called non-degenerate if the Hessian matrix (the matrix of second partial derivatives) of \( f \) is non-singular at this point. Otherwise a critical point is called degenerate.

The index of a non-degenerate critical point \( p \) of \( f \) is the dimension of the largest subspace of the tangent space to \( M \) at \( p \) on which the Hessian is negative definite. In other words, index is the number of negative eigenvalues of the Hessian.

A smooth map \( f : M \to \mathbb{R} \) is a Morse function if none of its critical points are degenerate and no two critical points have the same function value.

**Lemma 3.1** (Morse lemma). Let \( p \) be a non-degenerate critical point of a function \( f \). Then there is a local coordinate system \((y_1, \ldots, y_n)\) in a neighborhood \( U \) of \( p \) with \( y_i(p) = 0 \) for all \( i \) such that the identity \( f = f(p) - (y_1)^2 - (y_2)^2 - \cdots - (y_k)^2 + (y_{k+1})^2 + \cdots + (y_n)^2 \) holds throughout \( U \).
The Morse lemma states that non-degenerate critical points are isolated. \( \lambda \) in the above equation represents the index of a critical point \( p \).

In two dimensions there are three types of non-degenerate critical points: minima, saddles and maxima. In a three-dimensional case the saddles can be of two kinds: 1-saddles and 2-saddles.

An integral line of \( f \) is a maximal path in \( M \) whose tangent vectors agree with the gradient of \( f \) at every point. Integral lines connect critical points. The set of integral lines sharing a common origin/destination is called an ascending/descending manifold. Given a Morse function \( f \), the descending manifolds of its critical points constitute the Morse complex of this function.

We call a function Morse–Smale if its ascending and descending manifolds intersect only transversally.

The Morse–Smale complex is a fundamental topological construct that partitions the domain of a real-valued function into regions having uniform flow behavior, namely the regions, where integral lines share common origin and destination. Given a Morse–Smale function, its Morse–Smale complex is the intersection of its ascending and descending manifolds. The connected components of the intersection form the Morse cells.

### 3.3. Forman’s discrete Morse theory

Discrete Morse theory was first described by Forman (2001). It is a combinatorial extension of smooth Morse theory to cell complexes.

Let \( B^d = \{ x \in \mathbb{R}^d : |x| \leq 1 \} \) be a closed unit ball in a \( d \)-dimensional Euclidean space. A \( d \)-cell \( \sigma \) is a topological space which is homeomorphic to \( B^d \). Let \( \alpha \) and \( \beta \) be two cells. \( \alpha \) is a face of \( \beta \) or \( \alpha < \beta \) if \( \alpha \) is a proper subset of \( \beta \). In this case \( \beta \) is a co-face of \( \alpha \). If the dimensions of \( \alpha \) and \( \beta \) differ only by 1, they are called facet and co-facet of each other respectively.

A cell complex is obtained by attaching cells of higher dimensions to cells of lower dimensions along their boundaries (see Fig. 1).

Let \( K \) be a cell complex. A discrete Morse function \( f : K \rightarrow \mathbb{R} \) is a function which assigns scalar values to all cells in \( K \), such that for every cell \( \alpha \):

1. the number of co-faces of \( \alpha \) with function value lower than \( f(\alpha) \) is at most one;
2. the number of faces of \( \alpha \) with function value higher than \( f(\alpha) \) is at most one.

A cell is critical if all of its co-faces have a function value higher than \( f(\alpha) \) and all of its faces have a function value lower than \( f(\alpha) \).

A discrete vector field is a collection of pairs of cells \( \{ \alpha(i)^{(d)} < \beta(i)^{(d+1)} \} \) such that each cell is in at most one pair. Each pair represents a discrete vector, where the vector arrow points from \( \alpha(i)^{(d)} \) to \( \beta(i)^{(d+1)} \). A gradient vector in a discrete Morse function \( f \) is a pair of cells \( \{ \alpha (d) < \beta(d+1) \} \) for which \( f(\alpha^{(d)}) > f(\beta^{(d+1)}) \). According to the definition of the discrete Morse function each cell can have at most one facet with higher function value, this is why the discrete gradient is uniquely defined for each cell. Critical cells according to their definition cannot be paired. This simulates the vanishing gradient for these cells.

A V-path is a sequence of cells:

\[
\alpha_0^{(d)}, \ \beta_0^{(d+1)}, \ \alpha_1^{(d)}, \ \beta_1^{(d+1)}, \ \alpha_2^{(d)}, \ \ldots, \ \beta_{r+1}^{(d+1)}, \ \alpha_r^{(d)}
\]

where \( \{ \alpha(i)^{(d)} < \beta(i)^{(d+1)} \} \in V \), and \( \{ \beta(i)^{(d+1)} > \alpha(i+1)^{(d)} \} \). This path is closed if \( \alpha_0 = \alpha_{r+1} \). V-paths in a discrete gradient field are monotonic and contain no closed V-paths.

To define a discrete Morse–Smale complex we have to explain the combinatorial equivalents of ascending and descending manifolds. The following definitions are due to Gulyassy (2008).

---

**Fig. 1.** A cell complex. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)
The boundary operator maps a cell to its facets and induces an orientation on the facets. The discrete tangent operator \( T(\beta) \) maps the cell at the tail of a gradient arrow to the head of the gradient arrow. The discrete flow operator \( \Phi \) is defined as \( \Phi(\alpha) = \alpha + \beta T(\alpha) + T(\partial \alpha). \) The discrete descending manifold \( D_{\alpha} \) of a critical cell \( \alpha \) is an invariant chain under the flow operator \( \Phi(D_{\alpha}) = D_{\alpha}. \) The discrete ascending manifold can be defined analogously. The discrete Morse–Smale complex of a cell complex \( K \) is formed by the intersection of its descending and ascending manifolds.

4. Method

4.1. Generating the Morse–Smale complex

To compute the Morse–Smale decomposition of a 2D transfer function domain any of the existing techniques can be used (Lewiner et al., 2004; Bremer and Pascucci, 2007; Robins et al., 2011). In this paper we adapt the technique for the computation of a combinatorial gradient field proposed by Robins et al. (2011). In the terms of discrete Morse theory, this algorithm is proven to be correct up to the volumetric functions. To extract and simplify the extremal structure of the function and subsequently segment the manifold into Morse–Smale cells we use the method described in Reininghaus et al. (2010). Based on this structure we label the separate regions and load the pixel values into a 2D texture, where each label is represented by a unique color.

Morse theory assumes the underlying manifold is continuous, but for many classifications this is not the case. Therefore, in some cases we apply a filter to the transfer function image before computing the segmentation. We primarily use smoothing filters, such as a discrete Gaussian or anisotropic diffusion. The advantage of using a pre-filtered image instead of the original one is that the discontinuities in the gradient field are smoothed out, which helps to avoid common segmentation artifacts. The filter kernel size and other parameters depend on the transfer function topology and have to be estimated manually. However, one has to be careful with smoothing filtering, sometimes it can result in undesired segmentation when applied strongly, resulting in poor segmentation of the original transfer function and an ineffective querying mechanism for peak finding.

4.2. Modifications to peak finding

Applying MS-decomposition to the peak finding algorithm at render time is then simple. At each sample \( t \) along the ray, we compute the interpolated two-variate field function \( f = (f(\hat{R}(t)), g(\hat{R}(t))) \). We then query the MS decomposition texture at \( f(t) \).

On a ray segment with endpoints \( t_0, t_1 \), we compute \( f_0 = f(t_0) \) and \( f_1 = f(t_1) \), and note that:

- if \( f_0 = f_1 \), then the transfer function is monotonic on that segment, and we can omit peak finding;
- if \( f_0 \neq f_1 \), then the transfer function is (potentially) non-monotonic, thus we perform peak finding.

The test is not perfect; it does not consider whether the field functions themselves are monotonic. Indeed, peak finding is often applied in sampling at sub-Nyquist rates, illustrating isosurfaces where they may exist. Here, the goal is not to ensure robustness, but to more aggressively reproduce high-frequency features at lower sampling rates. MS decomposition is a way of determining where such features do not exist, without having to subsample in transfer function space. However, in most cases, the MS-decomposition provides an effective query for regions in transfer function that have monotone behavior. By detecting these regions along the ray, we can avoid expensive sampling in transfer space peak finding.

Peak finding with MS-decomposition can equally be employed in a modality similar to preintegration, where the integral in the transfer domain is computed dynamically as opposed to statically form a lookup table. In practice, we find that peak finding is qualitatively preferable in cases of high-frequency classification, regardless of whether MS-decomposition is used.

5. Results

The overall goal of MS-decomposition is to perform peak finding only where necessary, and (ideally) to achieve frame rates closer to standard volume rendering at the same world-space sampling rate while peak finding. Peak finding is especially beneficial for classifications with high frequencies (i.e. sharp peaks). At the same time, 2D classification is most commonly applied to univariate volumes with gradient magnitude modality. For our experiments, we consider CT and microscopy data with 2D gradient-magnitude classification and relatively sharp features. We note, however, that it is difficult to apply topological analysis to these kinds of functions. A two-dimensional discrete function, like the one shown in Fig. 2, cannot be approximated by a continuous Morse function. Large areas of constant function value contradict one of the conditions for a function to be Morse. Moreover, sharp peaks mean, that the function is not \( C^1 \) differentiable, in some cases it is not even \( C^0 \). Unfortunately, these properties are typical for transfer functions. This puts some limits on the application of the concepts, which originate from the differential topology. Nonetheless, as a means of partitioning transfer function space into regions of uniform gradient, the MS complex proves useful.

Fig. 2 shows the transfer functions we used to render the Christmas tree and zebrafish embryo datasets and their MS segmentations. The transfer functions can be generated using any of the available transfer function generation tools. The
Fig. 2. Transfer functions and their MS decomposition results: Christmas tree (left) and zebrafish embryo (right).

Fig. 3. Method results. Top: Christmas tree dataset. From left to right with standard postclassification (left, 7.0 fps), peak finding (center, 5.5 fps) and peak finding with MS-querying (right, 6.4 fps). Center: closer view of the Christmas tree. Bottom: Zebrafish embryo with performing at 12.5, 8.9 and 10.0 fps, respectively with the above modalities. Renderings captured at 1024 × 768 resolution on an NVIDIA 285 GTX GPU with 5122 transfer functions. Renderings using chordal peak finding with a world-space step size of $\Delta t = 2$ and a transfer-space sampling rate of $\Delta s = 2$.

form of the widgets strongly depends on the data and the corresponding application. The rendering results can be observed in Fig. 3, which illustrates the comparison between the direct volume rendering (DVR), peak finding and peak finding using MS segmentation. All images are generated with equal number of initial samples. Although the DVR method is faster, it misses many of important details in the dataset. The other two methods give approximately the same qualitative result with some performance improvement in the third case. Overall, MS-decomposition allows for peak finding at frame rates nearly as high as postclassified volume rendering, given the same world-space sampling rate $\Delta t$. In turn, this allows for higher transfer space sampling rates $\Delta s$ to be used at lower total performance penalty.

6. Conclusions

Separably-sampled volume rendering with peak finding provides compelling results in the presence of noise or in case of high frequency transfer functions. However as a payoff for a good rendering quality it requires additional computation time for sampling in the transfer function space. In this paper we have discussed an approach to improve on the performance of this method by excluding unnecessary separable classification in regions where it is not needed. Such regions can be represented by cells of a Morse–Smale complex, where the gradient of a function is monotonic. The Morse–Smale segmentation...
of the transfer function in the preprocessing stage helps to substitute the separable classification procedure with the two table lookups, giving the opportunity to the performance gain.

The main drawback of our method is that it is limited to 2D transfer functions. While 3D MS decomposition is possible, it would be an expensive preprocess and would consume a large memory footprint. In addition, MS-decomposition is poorly suited for many transfer functions, and parameters for the segmentation and possible pre-filtering have to be specified for each function individually. There are common classes of features in the presence of which we can decide about the range of the parameters, but the automatic detection of these features is not a trivial task. It might, however, give a possible direction for future research. Moreover, one of the advantages of the original multidimensional peak finding method is that it requires no precomputation. Though MS-decomposition provides some speed advantages for common 2D classifications (enabling faster peak finding performance for popular gradient-magnitude classification, for instance), it comes at some precomputation cost.

Multidimensional transfer functions for volume rendering remain an active area of research. Intuitive classifications depend greatly on the data and the application. Usually a transfer function is drawn by the user and specified as a table of discrete values. The application of continuous MS theory to inherently non-smooth images poses problems. Ironically, high-frequency transfer function for which peak finding is well suited poses the most difficulties for MS decomposition. In the presence of degenerate features, the decomposition can result in oversegmentation with multiple artifacts, the influence of which in the final rendering result highly depends on the case. In our experiments, oversegmentation did not reduce the quality of the rendering, but could decrease performance. One possible direction for future research would be to develop different segmentation strategies for different types of transfer functions. In addition, we are interested in applying MS-decomposition and related methods to automatic generation of multidimensional transfer functions.

References


