AUTOMATIC SKELETON EXTRACTION AND SPLITTING OF TARGET OBJECTS

Sang Min Yoon†, Holger Graf‡
†GRIS, Informatik, TU Darmstadt
‡Fraunhofer IGD
Fraunhoferstrasse 5, Darmstadt, Germany

ABSTRACT
The understanding of object’s kinematic structure is one of main challenges in the area of computer vision. Especially, skeleton of deformable objects, which is familiar with human visual perception, visualizes its characteristic using few data. This paper describes an efficient approach for automatic skeleton extraction and its splitting in the space of diffusion tensor fields, which are generated from normalized gradient vector flow fields of a given image. Our method is based on two steps: Skeleton extraction using second order diffusion tensor fields, Splitting skeleton using dissimilarity measure between neighbor elements. The evaluation proves the efficiency of our technique which might be applied to object retrieval, pose estimation and action recognition, object registration and visualization.

Index Terms— Skeleton, Tensor fields, NGVF, Kinematic Structure

1. INTRODUCTION
The structural analysis of deformable objects is one of important issues in computer vision and image processing because it can be used in many applications, including shape matching [1][2], computer animation [3], and object registration and visualization [4]. Especially, skeleton is a compact one-dimensional representation of complex and deformable objects and also describes an object’s geometry and topology by using few data. A precise definition of the skeleton or Medial Axis Transform [MAT] in the continuum was given by Blum [5], who postulated the well-known prairie fire analogy. Since the first study of skeleton technique through estimating the MAT, the skeletonization of shapes has attracted attention from many researchers in various application domains. Previous skeleton computing methodologies can be roughly classified into three categories by approaches: topological thinning [6], distance transform based extraction [7], and geometric methods [8]. However, existing skeleton extraction and splitting algorithms are still weak because of their high computational complexity, noise sensitivity, centeredness inside the underlying complex shape, partial occlusion or artifacts within a singular region of a given shape.

Most previous skeleton extraction methodologies are based on vector fields which are generated from a given image by different physical properties. Few work are investigated to extract the features of diffusion tensor fields from a topological point of view. Some researchers spent on visualizing two dimensional slices of MR diffusion tensor data by color mapping the direction of principal eigenvectors. Within this paper, we develop the skeleton extraction and splitting methodology using a novel topological analysis for deformable target objects’ investigating the space of associated gradient vector flow(GVF) fields. As we analyze the tensor fields of a normalized gradient vector flow(NGVF) within a given image, the proposed methodology has the following advantages comparing previous vector field based skeleton extraction technique: (1) There is no need to determine a priori information from a given image,(2) Our proposed methodology shows an improved skeleton extraction within a singular region of the shape. (3) The algorithm is robust against noise and partial occlusion. (4) We can analyze the motion of deformable object by splitting the skeleton. Figure 1 displays our methodology which we will explain in detail further on.

2. SKELETON EXTRACTION IN DIFFUSION TENSOR FIELDS

2.1. Normalized Gradient Vector Flow of image
Originally, the GVF field was proposed to solve the problem of initialization and poor convergence to boundary concave objects yielding a traditional snake form [9]. The GVF is a vector diffusion approach on Partial Differential Equa-
tions(PDEs). It converges towards the object boundary when very near to the boundary, but varies smoothly over homogeneous image regions extending to the image border. The main advantage of GVF fields is that it is able to capture a snake from a long range and could force it into concave regions. Mathematically defined, the GVF is the vector field $\mathbf{v}$ that minimizes the following energy functional,

$$
\varepsilon = \int \int \mu (u_x^2 + u_y^2 + v_x^2 + v_y^2) + ||\nabla f||^2 ||\mathbf{v} - \nabla f||^2 dx dy
$$

where $\mathbf{v} = [u(x, y), v(x, y)]$, and the initial value of $\mathbf{v}(x,y)$ is determined by $\nabla f(x,y)$. $\nabla f(x,y)$ is the gradient image derived from a given image. $\mu$ is a regularization parameter to be set on the basis of noise present in image. Minimizing this energy will force $\mathbf{v}(x,y)$ nearly equal to the gradient of the edge map where $||\nabla f(x,y)||$ is large. Nevertheless, the general GVF method cannot efficiently extract the medial axis as a weak vector has very little impact on its neighbors that have much stronger magnitudes.

A NGVF can tremendously affect a strong vector, both in magnitude and in orientation by normalizing the vectors over the image domain during each diffusion iteration [10]. Figure 2 shows the NGVF field from a given image. The traditional GVF has difficulty preventing the vectors on the boundary from being significantly influenced by the nearby boundaries and thus causes a problem such that the “snake” may move out of the boundary gap. NGVF avoids this problem. From figure 2-(c), we can see the detail of the NGVF in the vector around the boundary gap point.

### 2.2. Skeleton extraction in second order diffusion tensor field

In this section, we will explain an automatic skeleton extraction and refinement by a topological analysis in NGVF fields. The diffusion tensor field, which is defined as a topological representation of a 2D symmetric, second-order tensor field is shown as:

$$
T(\bar{x}) = \begin{pmatrix}
T_{11}(x,y) & T_{12}(x,y) \\
T_{21}(x,y) & T_{22}(x,y)
\end{pmatrix}
$$

$T(\bar{x})$ is fully equivalent to two orthogonal eigenvectors

$$
\tilde{T}(\bar{x}) = \lambda_i(\bar{x})\tilde{e}_i(\bar{x})
$$

where $i=1,2$. $\lambda_i(\bar{x})$ are the eigenvalues of $T(\bar{x})$ and $\tilde{e}_i(\bar{x})$ define the unit eigenvectors [11]. The topology of a tensor field $T(\bar{x})$ is the topology of its eigenvector fields $\tilde{e}_i(\bar{x})$.

According to [12], we can build a topological analysis of diffusion tensor fields from the concept of degenerated points, which play the role of critical points in vector fields. Streamlines in vector fields never cross each other except at critical points and hyperstreamlines in tensor fields meet each other only at degenerated points. Thus, the degenerated points are the basic singularities underlying the topology of tensor fields. Mathematically, those points are defined as the two eigenvalues of $T(\bar{x})$ which are equal to each other. Degenerated points in tensor fields are the basic constituents of critical points in vector fields. There are various types of critical points - such as nodes, foci, centers, and saddle points - that correspond to different local patterns of the neighboring streamlines. Delmarcelle [13] has proven that the local classification of line fields or degenerate points can be determined by constraints.

From the degenerated point, $x_0$, the partial derivatives are evaluated according to

$$
a = \frac{\partial(T_{11} - T_{22})}{\partial x}, \quad b = \frac{\partial(T_{12})}{\partial x}, \quad c = \frac{\partial(T_{12})}{\partial y}, \quad d = \frac{\partial(T_{22})}{\partial y}
$$

An important quantity for the characterization of degenerated points is

$$
\delta = ad - bc
$$

So a simple point topologically should be classified into two types: trisector if $\delta < 0$, and wedge if $\delta > 0$ [14]. Within the target object, these points are assumed as trisector [15]. Thinning the skeleton, connected by continuous degenerated points, can be very efficiently done by using the fact that a point within the object which has not at least one background point as an immediate neighbor cannot be removed, since this would create a hole. Therefore, the only potentially removable points are at the border of the object. Once a border point is removed, only its neighbors may become removable.
3. AUTOMATIC SKELETON SPLITTING USING DIFFUSION TENSOR SIMILARITY MEASURE

After obtaining a skeleton of deformable objects, the skeleton is split into several branches by analyzing its tensorial characteristics. Within an extracted skeleton, we can separate the elements by using following definition. A branch point is the pixel inside the skeleton that connects each branch. A end point is the pixel inside the skeleton with only one neighbor. A joint point is the pixel inside a branch that separate the neighbor. End points can be interpreted as the polar points in the space of diffusion tensor fields and branch points are also can be understand as the combination of various eigenvalues in neighbor pixels. Within a branch, we split the branch with similarity measure between neighbor skeletal elements. For each pixel $I_i$ which is recognized as the skeleton, we measure the dissimilarity between neighbor skeleton pixel elements and measure the dissimilarity using tensorial dissimilarity function. Given two tensors $T_i$ and $T_j$, there are some dissimilarity measures that might be used to compare them. The tensor can be represented by an ellipsoid, where the main axis lengths are proportional to the square roots of the tensor eigenvalues $\lambda_1$ and $\lambda_2(\lambda_1 \geq \lambda_2)$ and their direction correspond to the respective eigenvectors. With this properties, we can measure the dissimilarity between neighbor elements. The simplest one is the tensor dot product [15]:

$$d_1(T_i, T_j) = \sum_i \sum_j \lambda_i^1 \lambda_j^2 (e_i^1 \cdot e_j^2)^2$$  \hspace{1cm} (6)

It uses not only the principal eigenvector direction, but the full tensor information. Another dissimilarity measure that uses the full tensor information is the Frobenius norm [16]:

$$d_2(T_i, T_j) = \sqrt{\text{Trace}((T_i - T_j)^2)}$$  \hspace{1cm} (7)

The dissimilarity measure between two elements is the multiplication of $d_1$ and $d_2$. Joint points are determined by comparing the similarity measure between neighbor points and joint points are decided when the direction of NGVF changes and scale of main and sub eigenvalue is over the threshold. In figure 4, we display the extracted skeleton using ellipsoid representation method and the end points are painted red, branch points green, and joint points blue which are decided by tensorial dissimilarity measure.

4. EXPERIMENTS

We lead some experiments in order to extract the skeleton and split the kinematics of deformable objects using our proposed methodology. Before we generate the NGVF, several input images are converted to binary format due to performance and comparison issues with previous approaches. Afterwards, we calculate the eigenvectors and eigenvalues which are extracted from tensor fields for identifying the degenerated points.

Having demonstrated the performance of our proposed skeleton extraction under various visual transformations, we extract the skeleton of input images of public database¹ shown in Figure 5. Figure 6 is the extracted skeleton of human body parts. Splitted areas are displayed by different colors. We compared our proposed skeleton extraction system with previous technique with previous methods such as morphological approach, and skeleton pruning using contour partition [17] in figure 7. Our proposed skeleton extraction method can efficiently represent the characteristic of target object, but very robust in noise effect.

5. CONCLUSION AND FUTURE WORK

Skeletonization is a very efficient method for visualizing deformable objects using one dimensional data. In this paper, we have shown a novel method for extracting and splitting the skeletons of target objects by a robust, accurate, and computationally efficient technique using diffusion tensor fields and associated topologies. The essential idea is to analyze a tensor topology in order to extract and connect degenerated points from a NGVF fields. We have illustrated our approach on a variety of 2D deformable objects by comparing it to previous techniques. We also showed our method into a framework for refining the skeleton.

We will focus in particular on a 3D skeleton extraction from volume data in view of motion analysis of deformable objects. An additional similarity measure would help us to analyze the objects’ motion and retrieve related motion patterns from within a database. Thus, our future work will address fast graphical representation methods of extracted skeletons combined with shape recognition techniques.

¹http://www.lems.brown.edu/vision/software
6. REFERENCES


