Manifold Learning Approach to Curve Identification with Applications to Footprint Segmentation

Conference Paper · December 2014

DOI: 10.1109/CIMSIVP.2014.7013288

6 authors, including:

Namita Lokare
North Carolina State University
9 PUBLICATIONS **8** CITATIONS

SEE PROFILE

Zoe Jewell
SAS Institute and Duke University
46 PUBLICATIONS **260** CITATIONS

SEE PROFILE

Edgar Lobaton
North Carolina State University
69 PUBLICATIONS **977** CITATIONS

SEE PROFILE

Manifold Learning Approach to Curve Identification with Applications to Footprint Segmentation

Namita Lokare*, Qian Ge*, Wesley Snyder*, Zoe Jewell†, Sky Allibhai† and Edgar Lobaton*
*Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC 27606
Email: ndlokare@ncsu.edu, qge2@ncsu.edu, wes@ncsu.edu, edgar.lobaton@ncsu.edu
†Wildtrack.org, Monichque, Portugal 8550-909
Email: wildtrack.footprinting@gmail.com

***Abstract*—Recognition of animals via images of their footprints is a non-invasive technique recently adopted by researchers interested in monitoring endangered species. One of the challenges that they face is the extraction of features from these images, which are required for this approach. These features are points along the boundary curve of the footprints. In this paper, we propose an innovative technique for extracting these curves from depth images. We formulate the problem of identification of the boundary of the footprint as a pattern recognition problem of a stochastic process over a manifold. This methodology has other applications on segmentation of biological tissue for medical applications and tracking of extreme weather patterns. The problem of pattern identification in the manifold is posed as a shortest path problem, where the path with the smallest cost is identified as the one with the highest likelihood to belong to the stochastic process. Our methodology is tested in a new dataset of normalized depth images of tiger footprints with ground truth selected by experts in the field.**

I. Introduction

Identification and tracking of endangered species is of great concern to environmental biologists due to the constant exploitation of global biodiversity [1]. In addition to techniques for animal identification from natural images [2], researchers have recently started using images of footprints [3]–[7] to monitor endangered species. An inexpensive non-invasive Footprint Identification Technique (FIT) [4], [7] makes use of features selected from images of animal footprints in order to identify individual animals (see Fig. 1). One of the main obstacles to efficient and cost-effective data collection is the manual extraction of these image features located along the boundary of the footprint which introduces human variability and requires training, limiting the number of volunteers involved with these efforts.

We think of this problem as the identification of a curve of interest in an image, which we call *Interest Curve*. For our motivating application, the interest curve is a boundary of a footprint. Similar tasks are also present in other application areas such as the medical (e.g., segmentation of the interface between various biological tissues) and environmental fields (e.g., tracking weather patterns such as hurricanes using remote sensing data). All of these problems involve the identification of a pattern in some imaging data along a curve of interest.

In this paper, we introduce a framework for solving this identification problem by modeling the interest curves as a stochastic process on a manifold specified by the space of variations of the imaging pattern around the curves. Fig. 2 illustrates the pipeline for our method. We locally approximate the shape of the manifold by using an eigen-basis decomposition of previously observed patterns and pose the pattern identification problem as a shortest path problem in a graphical representation. The constructed combinatorial model captures imaging information as well as geometric information (arc-length and curvature in particular) of a set of feasible curves in the space. An interest curve is selected by finding the path with the highest likelihood to be generated by the defined stochastic process. This path is then refined using variational method. We test our method in a unique database made available to the scientific community; it is composed of depth (and color) images of tiger footprints with ground-truth segmentation provided by experts in the field [12].

Fig. 1. Animal Identification via footprint recognition. Color and depth images of footprints are captured (left) and features along their boundary are selected (right). We only make use of the depth images for our analysis.

The rest of the paper is organized as follows: section II discusses some relevant work in the area of segmentation, section III presents the procedure used to capture and normalize the footprint images, section IV provides a formal mathematical description of our problem, section V introduces our stochastic model of interest curves on a manifold, section VI provides a discrete formulation of the general approach,

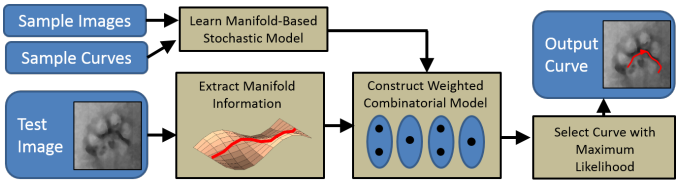

Fig. 2. Pipeline of the method. Training images and interest curves are used to learn a manifold model of the data. The identification of an interest curve in an image is posed as a shortest path problem in a graphical model.

*This work was partially supported by the National Science Foundation under award CNS-1239323.

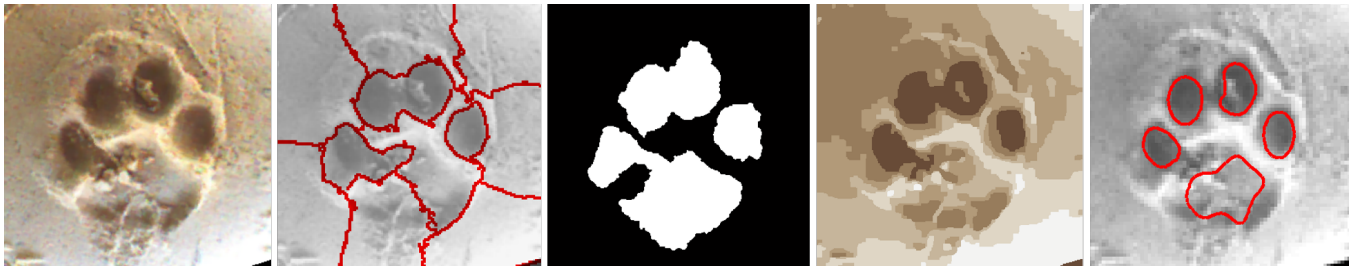

Fig. 3. Segmentation of footprint by traditional algorithms (from left to right): Input color image, Normalized Cuts [8], Graph Cuts [9], KM [10], and DRLSE [11] results. Similar results are observed for depth images.

section VII illustrates how the identification of interest curves is posed as a shortest path problem, section VIII describes the refinement process to obtain a continuous curve, section IX presents our experimental results, and section X provides some final remarks.

II. Related Work

Application areas such as medical imaging [13], [14], object detection [15] and recognition [16], [17] make image segmentation a well-studied problem in image processing and computer vision. Numerous segmentation algorithms have been proposed over the last few decades and they differ depending on different applications or different modalities. Some particularly popular techniques are active contours, level set methods, and graph-cut based approaches.

A level set method was proposed by Osher and Sethian [18] for fronts propagating in 1988. By representing the surface as a level set, the topological changes of boundaries can be easily handled [18]. Inspired by this, Caselles et. al. [19] presented a geodesic active contours model for object detection of boundaries, and Malladi et. al. [20] introduced a shape modeling scheme for complex shape recovering. Both algorithms require solving an energy minimization problem. However, they may converge to a local minimum. In [21], Chan and Vese proposed a new active contours model for binary segmentation, which has fewer number of local minima by using a region-based stopping function instead of an edge-based stopping function as [19], [20]. In [22], the authors proposed a modification to their method having a multiphase level set framework to segment images with more than two regions. In 1995, Cootes et. al. [23] introduced Active Shape Models to incorporate model specificity along with an iterative refinement algorithm to fit the data consistently with the training set. Their method was different from the active contour model in that it could only deform to fit the data in ways consistent with the training set. Sundararamoorthi et. al. [24] introduced an automatic region based active contour algorithm, using an adaptive "lookout" region which depends on the statistics of the data, estimated during the detection process.

Graph based segmentation considers segmentation as a graph partitioning problem [25]. In 1997, Shi and Malik [8], [25] first proposed normalized cuts criteria for segmenting a graph. Recently, a modification of the normalized cuts method with priors was proposed in [26]. In 2001, another graph based segmentation, interactive graph cuts segmentation, which incorporates user hard constraints and soft constraints to find a globally optimal segmentation through a min-cut/max-flow algorithm in [27] was proposed by Boykov and Jolly [9]. Later on, several segmentation algorithms based on graph cuts have been developed. Rother et. al. [28] introduced an extension to graph cuts, GrabCut, which requires fewer interactions. In the work of Salah et. al. [10], images are transformed implicitly by a kernel function, which makes unsupervised graph cut segment applicable. There are also several other approaches that fall in this category [16], [17], [29], [30].

Footprint segmentation can also be viewed as a pattern recognition problem for which template matching could be applied. These techniques apply cross-correlation algorithms to find the location of a reference image or an object of interest in a scene. The traditional methods are however computationally expensive. Choi and Kim [31] propose a fast, rotation invariant algorithm through a hierarchical approach. Another fast pattern matching scheme introduced by Hel-Or et. al. [32] allows matching under non-linear tone mapping.

In recent years, because of the increasing use of depth cameras, a large number of segmentation algorithms combining depth and color information have been proposed. In [33], color information is used to enhance depth image and then a graph cut based segmentation method is applied for image retargeting. In [34]–[36], RGB-D image segmentation is used for indoor scene understanding. In [37], a planar segmentation algorithm using RGB-D images is proposed.

In Fig.3, we show segmentation results of four previous algorithms. The test image shown in the first column is one of the color images in our dataset. Columns 2-5 show the segmentation results of Normalized Cuts [8], Graph Cuts [9], KM [10] and DRLSE [11]. As shown in Fig. 3, the first three methods failed to segment the footprint. DRLSE segmented the toes correctly but not the pad. Also, this method is highly dependent on its initialization, which can lead to incorrect results unless a good initial contour is chosen. Similar patterns were observed when segmenting the depth images. These techniques are not suitable for our goal of automatic segmentation of a footprint without any prior information built into the model.

In [38], M. Turk and A. Pentland introduced an unsupervised approach for face recognition using Eigenfaces. They projected the face images onto a feature space to capture the variation among the training images. In [39] Yang et al. propose a color eigen-structure segmentation algorithm which is designed to extract the desired objects with color values close to the training samples. Our approach obtains the segmentation of a footprint by learning local variations

of images using an eigen-based approach and refining it to capture the boundaries. We propose an algorithm which does not require initialization and is fully automated.

III. Image Capture and Pre-Processing

The images of tiger footprints were collected at the Carolina Tiger Rescue [40] from animals of known identity, sex, and age to form a reference database. At this center, a path near a fence was prepared using sand. The sand was prepared and leveled before tigers were guided along this path using a bait. Images were collected after the tigers were safely relocated.

A Microsoft Kinect depth sensor was used in the near mode setting to capture RGB-D images. The images were captured with color and depth image resolutions of 1280×960 and 640×480, respectively. The imaging sensor was mounted on a tripod stand pointing to the floor and covered with a tarp to avoid interference due to external light. A small battery operated light source was also mounted on the tripod in order to avoid total darkness. The software provided a small bounding box in the middle of the displayed images allowing for manual centering and vertical alignment of the footprints. This was done to facilitate the footprint identification procedure while streamlining the data capturing.

Due to the variation of the perspective angle of the sensor with respect to the ground plane of the footprint, we normalize the images via a pre-processing step. First, a proper calibration of the depth sensor is obtained using the toolbox by [41]. A plane is fitted to the recovered 3D point cloud. We virtually rotate the camera sensor so it is placed at a distance of $0.5\ m$ away from the plane and looking straight down onto it. New color and depth images are rendered by using this normalized perspective in order to guarantee that the footprint is centered and the images are taken from approximately the same perspective. This normalization also reduces the amount of training that would be needed for volunteers when acquiring this data.

IV. Problem Formulation

Our objective is to use a set of training images with a preselected set of sample curves, which we call *Interest Curves*, in order to identify similar patterns in a new test image. For our application, we aim to identify the segmentation contour for a footprint using depth image information.

As a formalism, let us define the training grayscale images $\{I_n : \mathbb{R}^2 \to \mathbb{R}\}_{n=1}^{N}$, each with a corresponding interest curve $\gamma_n : [0,1] \to \mathbb{R}^2$. Given a test image I, our task is to identify the location of the corresponding interest curve γ. For a discrete version of the problem, we will consider each curve $\gamma_n = \{\gamma_{n,k}\}_{k=1}^{K}$ as a sampled version of the continuous curves (i.e., $\gamma_{n,k} := \gamma_n(t_k)$ taken at some point t_k), where K is some fixed constant.

V. Stochastic Mathematical Model

The proposed framework considers any interest curve γ as the realization of a stochastic process parameterized by $t \in [0,1]$. This process influences not only the shape of the curve but the grayscale values in a neighborhood around them.

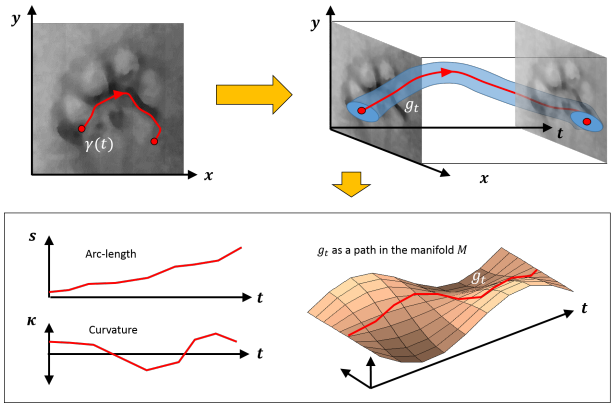

Fig. 4. Interest Curve Representation. Curve $\gamma(t)$ and image I in the image plane (top-left), and the corresponding graph of g_t (top-right). Extracted values of arc-length $s(t)$, curvature $\kappa(t)$ and g_t represented as a path in the manifold M.

First, we assume that any γ has two continuous derivatives and that the images are square integrable (i.e., they belong to L^2). Given these assumptions, we extract from γ the tuple

$$\left(s(t), \kappa(t), g_t(\cdot,\cdot) \right), \tag{1}$$

where $s(t) \in \mathbb{R}^+$ represents the arc-length, $\kappa(t) \in \mathbb{R}$ represents the curvature, and $g_t(\cdot,\cdot) \in L^2(B_r)$ is a square integrable function defined over a disk of radius r (see Fig. 4). The function g_t has the same values as I around a neighborhood centered at $\gamma(t)$. Note that $(s(\cdot), \kappa(\cdot))$ uniquely characterize γ up to translation and rotation, and we think of them as generated by independent stochastic processes.

Furthermore, we assume that all realization of the graph of $g_t(\cdot,\cdot)$ as a function of $t \in [0,1]$ (i.e., the set of points (t, g_t)) lie on a manifold M in $[0,1] \times L^2(B_r)$. By approximating g_t linearly around the mean $\bar{g}_t$ of the stochastic process, we can estimate M using an affine space M_a spanned by a set of orthonormal bases $\{\psi_{t,i}(\cdot,\cdot)\}_{i=1}^{D}$, where D is the dimension of M. That is,

$$g_t(x,y) \approx \bar{g}_t(x,y) + \sum_{i=1}^{D} \psi_{t,i}(x,y) b_i(t) \tag{2}$$

where $b_i(t)$ are stochastic processes corresponding to the coefficients of the basis expansion. Let $b : [0,1] \to \mathbb{R}^D$ be the function with components $b_i(t)$.

In our approach, we generate candidate curves from an image, and select as an interest curve the one that maximizes some likelihood. Hence, given a test curve η and an image I with corresponding functions $(s^\eta, \kappa^\eta, g_t^\eta)$, we will need to determine the probability of it being generated by the previously described stochastic process. This probability can be expressed as:

$$P(\eta|J) = P(g_t^\eta \in M_a\,,\, b^\eta|J)\; P(s^\eta)\; P(\kappa^\eta). \tag{3}$$

VI. Discrete Mathematical Model

In order to make our problem tractable, we generalize the previous framework to a discrete sampled test curve $\eta = \{\eta_k\}_{k=1}^{K}$ in an image I.

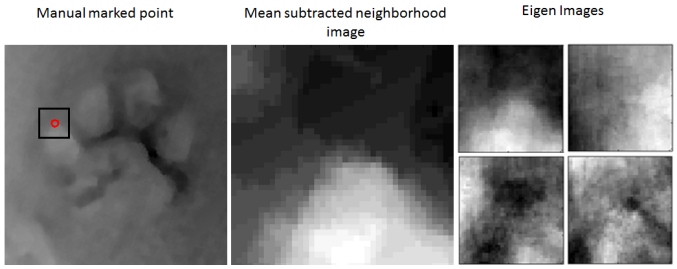

Fig. 5. Image showing the neighborhood for a manually marked point (left), mean subtracted neighborhood image (middle) and the eigen-images corresponding to the largest four eigen-values for that point (right). As expected the first eigen-image captures variation on the depth of the imprint of the toe.

First, we define the discrete version of the tuple $(s^\eta, \kappa^\eta, g^\eta, b^\eta)$. The arc-length $s^\eta = \{s^\eta_{(k,k+1)}\}_{k=1}^{N-1}$ is specified by

$$s^\eta_{(k,k+1)} = ||\eta_k - \eta_{k+1}||. \tag{4}$$

The curvature $\kappa^\eta = \{\kappa^\eta_k\}_{k=2}^{N-1}$ is computed by defining κ^η_k as the Menger curvature [42] between the points $\{\eta_{k-1}, \eta_k, \eta_{k+1}\}$.

The image pattern $g^\eta = \{g^\eta_k\}_{k=1}^N$ is specified by letting g^η_k equal to the image I over a small neighborhood of radius r around η_k. Given that the basis functions $\{\psi_{k,i}\}_{i=1}^D$ are known for all $k = 1, \cdots, K$, then we can specify $b^\eta = \{b^\eta_k\}_{k=1}^N$ by letting

$$b^\eta_{k,i} = \langle g^\eta_k - \bar{g}_k, \psi_{k,i} \rangle \,, \tag{5}$$

where $\bar{g}_k$ is the mean image at the k-th sample point and $\langle f, h \rangle = \int\int_{Br} f(x,y) \cdot h(x,y)\, dx\, dy$.

The manifold M_a is defined by taking an eigenimage [38] based approach. All neighborhoods for the k-th point in the provided training set of interest curves are used to compute an average image $\bar{g}_k$ and the D eigen-functions $\{\psi_{k,i}\}_{i=1}^D$ that explain most of the variance in the image set are computed. Fig. 5 shows four eigen-images of the neighborhood of a sample point. We define the projection of g^η_k to M_a as

$$\hat{g}^\eta_k = \bar{g}_k + \sum_{i=1}^{D} \psi_{k,i} b_{k,i}. \tag{6}$$

In order to quantify how close g^η is from M_a, we specify the error metric

$$E^2_m(g^\eta) = \sum_{k=1}^{N} \langle g^\eta_k - \hat{g}^\eta_k, g^\eta_k - \hat{g}^\eta_k \rangle \,. \tag{7}$$

We also define measures E_s and E_κ of deviation for arc-length and curvature from their mean values. These quantities are given by

$$E^2_m(s^\eta) = \sum_{k=1}^{N-1} ||s^\eta_{(k,k+1)} - \bar{s}_{(k,k+1)}||^2 \tag{8}$$

and

$$E^2_\kappa(\kappa^\eta) = \sum_{k=2}^{N-1} ||\kappa^\eta_k - \bar{\kappa}_k||^2, \tag{9}$$

where $\bar{s}_{(k,k+1)}$ and $\bar{\kappa}_k$ are the mean arc-length and curvature computed from the training data.

We define each term in Eqn 3 in terms of the error measures as

$$\begin{aligned} P(g^\eta \in M_a \,,\, b^\eta | I) &= C_m \cdot e^{-w_m \cdot E^2_m(g^\eta)} \\ P(s^\eta) &= C_s \cdot e^{-w_s \cdot E^2_s(s^\eta)} \\ P(\kappa^\eta) &= C_\kappa \cdot e^{-w_\kappa \cdot E^2_\kappa(\kappa^\eta)} \end{aligned} \,, \tag{10}$$

where w_m is a normalization weight for the manifold term, C_m is a normalization constant, and similar definition are made for the arc-length and curvature parameters.

Using these definitions, we have

$$\begin{aligned} &\arg\max_\eta P(\eta | I) \\ &= \arg\max_\eta P(g^\eta \in M_a \,,\, b^\eta | I)\; P(s^\eta)\; P(\kappa^\eta) \\ &= \arg\max_\eta \log \left(P(g^\eta \in M_a \,,\, b^\eta | I\, P(s^\eta)\; P(\kappa^\eta) \right) \\ &= \arg\min_\eta \big(\, w_m E_m(g^\eta) + \\ &\quad w_s E_s(s^\eta) + w_\kappa E_\kappa(\kappa^\eta) \, \big) \end{aligned} \tag{11}$$

Thus, finding the curve with the maximum likelihood is equivalent to finding the curve η^* with smallest weighted sum of error measures.

VII. Finding the Optimal Discrete Interest Curve

Let us define the cost function for a discrete test curve η to be

$$C(\eta) = w_m E_m(g^\eta) + w_s E_s(s^\eta) + w_\kappa E_\kappa(\kappa^\eta). \tag{12}$$

Our objective is to identify the curve η^* in an image I that minimizes $C(\eta)$. This is done by formulating the problem as a shortest path problem in a graph. The approach is outlined in Algorithm 1.

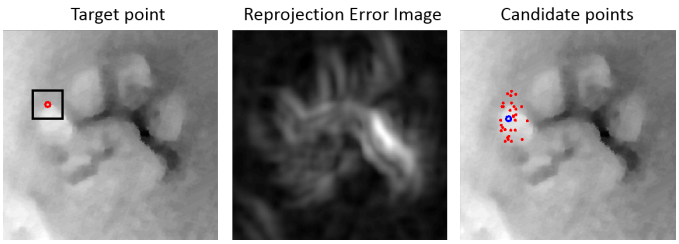

Fig. 6. Manifold error. A target feature point (left), its corresponding reprojection error over the entire image (middle), and the set of candidate points in red around the mean location from the training set in blue (right).

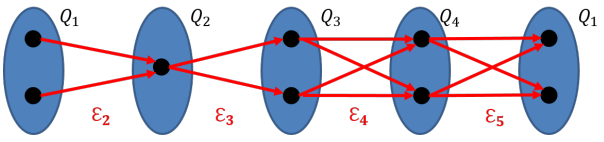

Fig. 7. Graphical model $\mathcal{G}$ for a closed interest curve with $K = 4$ points. The graph is a K partite graph where the sets of candidate segments $\mathcal{Q}_k$ represent the nodes and the set of candidate triplets $\mathcal{E}_k$ specify the edges.

Algorithm 1 Finding Optimal Interest Curve

1: **for** $k = 1$ to K **do**
2: Compute $E^2_{m,k}(x)$ for all points x in the image I
3: Select points $\mathcal{P}_k = \{p_{k,i}\}_{i=1}^{N_k}$ around the mean location of the k-th point in the IC
4: **end for**
5: **for** $k = 1$ to $K-1$ **do**
6: Set $\mathcal{Q}_k = \{\}$
7: Compute $E^2_{s,k}(i,j)$ for all pairs (i,j) where $i = 1 \cdots N_k$ and $j = 1 \cdots N_{k+1}$
8: Store all (i,j) into $\mathcal{Q}_k$ that have values of $E^2_{s,k}$ within the specified bounds
9: **end for**
10: **for** $k = 2$ to $K-1$ **do**
11: Set $\mathcal{E}_k = \{\}$
12: **for** $i = 1$ to N_k **do**
13: Compute $E^2_{\kappa,k}(j_1, i, j_2)$ for all the pairs $(j_1, i) \in \mathcal{Q}_{k-1}$ and $(i, j_2) \in \mathcal{Q}_k$
14: Store all $\{(j_1, i), (i, j_2)\}$ into $\mathcal{E}_k$ that have values of $E^2_{\kappa,k}$ within the specified bounds
15: **end for**
16: **end for**
17: Construct a graph $\mathcal{G}$ with vertices and edges specified by $\{\mathcal{Q}_k\}_{k=1}^{K-1}$ and $\{\mathcal{E}_k\}_{k=2}^{K-1}$
18: Assign as weights to the vertices the corresponding sum of the values of $E^2_{m,k}$ and $E^2_{s,k}$
19: Assign as weights to the edges the corresponding values of $E^2_{\kappa,k}$
20: Let η^* be the path with the lowest cost between all paths starting at Q_1 and ending at Q_{K-1}

Before we can find the optimal η^* it is necessary to learn the appropriate model from the training data. Parameters that must be estimated include mean values for arc-length and curvature, and the eigen-images used to represent M_a locally.

Algorithm 1 proceeds by first finding a set of candidate points (lines 1:4) in the image, finding candidate segments between these points (lines 5:9), determining all feasible triplets that give valid curvatures (lines 10:16), and then constructing a graphical representation from which a shortest path solution can be found corresponding to η^* (lines 17:20).

The candidate points $\mathcal{P}_k$ are obtained by randomly selecting points using a gaussian distribution around the mean location for a point x found from the training set. The error $E^2_m(x)$ is associated with how far a neighborhood around x is from M_a. Fig 6 shows the error and candidate points for a point x.

The set of candidate segments $\mathcal{Q}_k$, where each entry corresponds to segments $(p_{k,i}, p_{k+1,j})$, are obtained by computing a measure $E^2_{s,k}(i,j)$ of their offset from their mean value. This quantity is essentially one of the terms from Eqn 8. The candidate triples $\mathcal{Q}_k$ are computed in a similar way where $E^2_{\kappa,k}(j_1, i, j_2)$ corresponds to the offset of the curvature of the triplet $(p_{k-1,j_1}, p_{k,i}, p_{k+1,j_2})$ from its means (i.e., one of the terms in Eqn 9). We remove some triples by computing the angle between the edge segments formed by them, if the angle does not satisfy the bound.

The graph $\mathcal{G}$ that is constructed is the dual graph of the graph connecting all points in the collection $\{\mathcal{P}_k\}$. This dual construction is needed to accommodate the curvature cost and constraints. Assuming that $|\mathcal{P}_k| = N$, the graph $\mathcal{G}$ is a $K-1$ partite graph with at most N^2 vertices in each group and at most N^3 edges between adjacent groups. As observed in Fig 7, we can also generalize this formulation to closed curves by adding $\mathcal{P}_1$ at the end of $\mathcal{G}$ which turns it into a K partite graph.

The final stage of our algorithm is to find the path with the smallest cost in $\mathcal{G}$. Note that we assign weights to the vertices and edges corresponding to the different terms in the cost function $C(\eta)$. We then use a shortest-path algorithm to find η^*. In order to accommodate for different sizes of toes between smaller and larger animals we repeat the process over a few scaled images and use the result which returns the minimum cost.

VIII. Refining the Interest Curve

So far, we have described how to provide an initialization to the toes and the pad via the identification of a discrete interest curve. Next, we describe how these contours are deformed based on prior knowledge and gradient information.

We represent a family of contours in polar coordinates as $\gamma(\theta) = (\theta, r(\theta))$, where $r(\theta)$ is expressed in terms of a linear combination of basis functions of the form:

$$r(\theta) = r_0(\theta) + c_0 + \sum_{n=1}^{N} c_{2n-1}\cos(n\theta) + c_{2n}\sin(n\theta). \quad (13)$$

The curve $(\theta, r_0(\theta))$, which we refer to as the "base curve", is the polar representation of the interpolated η^* using cubic splines with respect to its center of mass. The vector of coefficients $c := [c_0, c_1, c_2, \cdots, c_{2N}]^\top$ contains the values that will be optimized in our approach. We emphasize the dependency of γ on these coefficients by writing $\gamma(\theta; c)$.

The set of coefficients that determine the maximum a posteriori (MAP) estimate for the contour is defined by

$$\hat{c}_{MAP}(I) = argmax_c \log\left(P(I|\gamma(\cdot;c))P(\gamma(\cdot;c))\right), \quad (14)$$

where I represents the depth image observed.

We make the assumption that

$$\log\left(P(I_k|\gamma(\cdot;c))\right) \propto \int_0^{2\pi} w_{E,k}(\gamma(\theta;c))\, d\theta \quad (15)$$

and

$$\log\left(P(\gamma(\cdot;c))\right) \propto -\lambda\left(c_0^2 + \sum_{n=1}^{N} \frac{c_{2n-1}^2 + c_{2n}^2}{n^2}\right), \quad (16)$$

where $w_{E,k} : \Omega_k \to \mathbb{R}^+$ is a weight function associated with edge information in the k-th footprint image. The parameter λ was set to $1e-4$ for all our experiments.

In order to obtain $w_{E,k}$ we observe that we expect edges in the depth map to overlap with the locations of the contour. Furthermore, the gradient at these locations must be pointing toward the center of the contour since points outside the region of interest (either a toe or a pad) must have a depth value

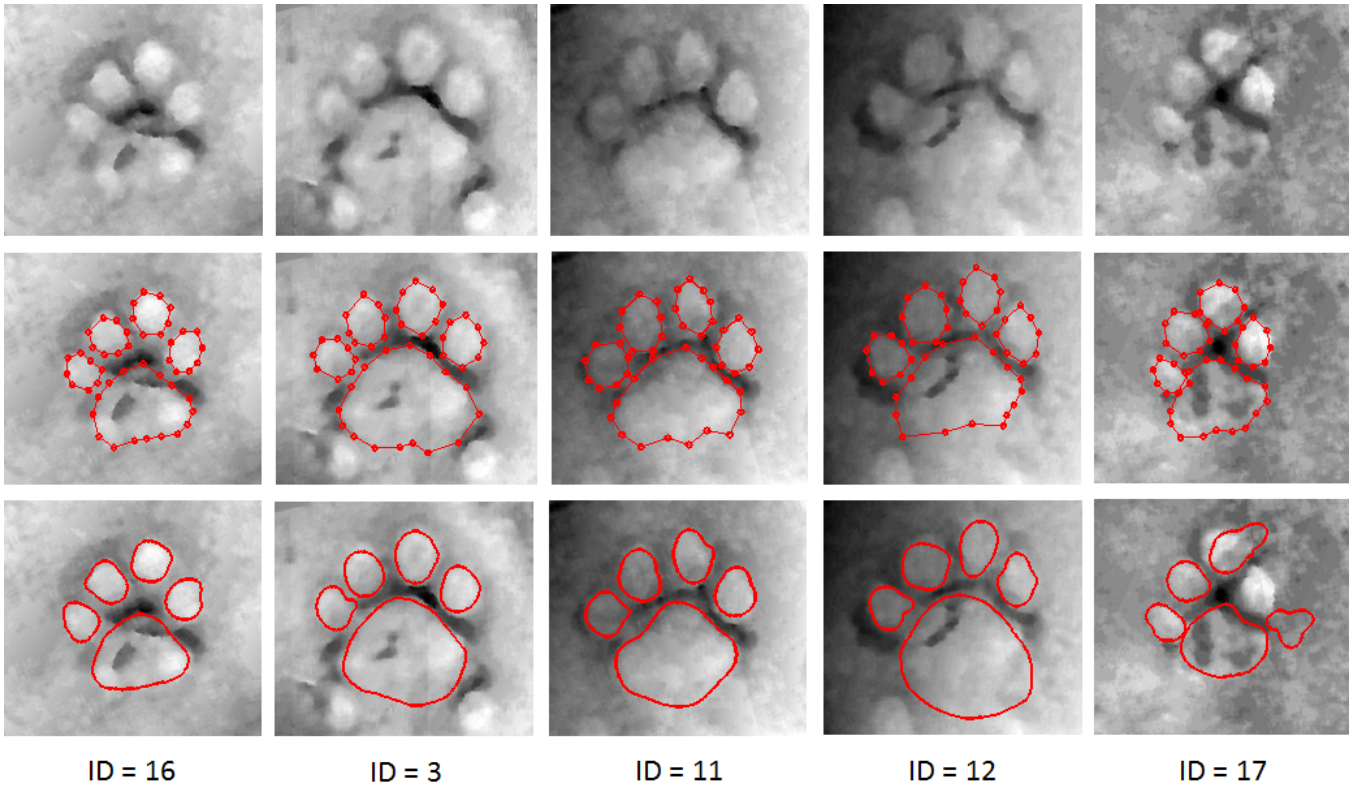

Fig. 8. Samples and results from the footprint dataset (from top to bottom): depth images, manual segmentation and automatic segmentation result. The ID refers to the image identification number. ID 17 is an outlier.

smaller than values inside the region of interest. Hence, we define this edge weight function to be:

$$w_{E,k}(x) = 1 - e^{-\alpha(\nabla_r I_k(x))}, \tag{17}$$

where $\nabla_r I_k(x)$ is the gradient of the depth image I_k at location x in the radial direction. The parameter α was experimentally selected to be 0.05.

The optimal set of coefficients $\hat{c}_{MAP}$ are obtained by minimizing the cost functional via numerical optimization.

IX. Experimental Results

The validation of our method was performed on a dataset of 20 left hind tiger footprint depth images normalized using the protocol described in Section III. The resulting images have size 300×300 pixels. The captured dataset contains features manually selected by several volunteers, and a growing number of footprint segmentation contours selected by experts in the wildlife monitoring field. A total of 8 points were selected around each toe, 16 points around the pad, and the number of eigen functions was chosen to be 4. We perform a leave one out analysis due to small size of the dataset. Fig. 8 shows some of the images from the dataset including manual and automatic segmentation results. If we exclude the outliers, image with ID 12 has the worst ADE for the entire footprint, and ID 16 and 3 have the best.

We compare results for the pad using two approaches: (1) by first aligning the test image using the centroid of the left and right most toes of a reference image, and (2) without doing any alignment. We essentially follow a sequential segmentation approach. This additional alignment accommodates for any scale issues we would have due to different sizes of footprints and to get a better initial estimate. Fig. 9 shows examples of how the additional alignment gave a better initialization and refinement. We can see that in the nonaligned case, the shortest path result obtained was not estimated correctly whereas for the same case when aligned with the training reference, the estimate was much better and hence the refinement was also good.

In order to evaluate the accuracy of the approach, we compute the average distance error (ADE) between the ground truth points $\{\beta_n\}_{n=1}^{N}$ and our results γ:

$$ADE = \frac{1}{N}\sum_{n=1}^{N} dist(\beta_n, \gamma)), \tag{18}$$

where $dist(\beta_n, \gamma)$ is the minimum distance between β_n and the curve. Table I summarizes our results for the aligned case. We note that the overall average distance error for the toes, pad and entire footprint is still within 8 pixels (considering the outliers).

TABLE I. Average Distance Error Results

Image ID	Toe 1	Toe 2	Toe 3	Toe 4	Pad	Footprint
16 (Best)	2.26	2.56	1.93	1.79	3.63	2.641
12 (Worst)	4.09	4.51	3.31	3.27	15.66	7.75
3 (Second Best)	2.45	2.56	2.02	3.07	5.81	3.62
17 (Outlier)	4.43	3.96	6.66	35.26	7.87	11.01
Overall (Average)	3.62	3.95	4.67	6.89	7.20	5.59

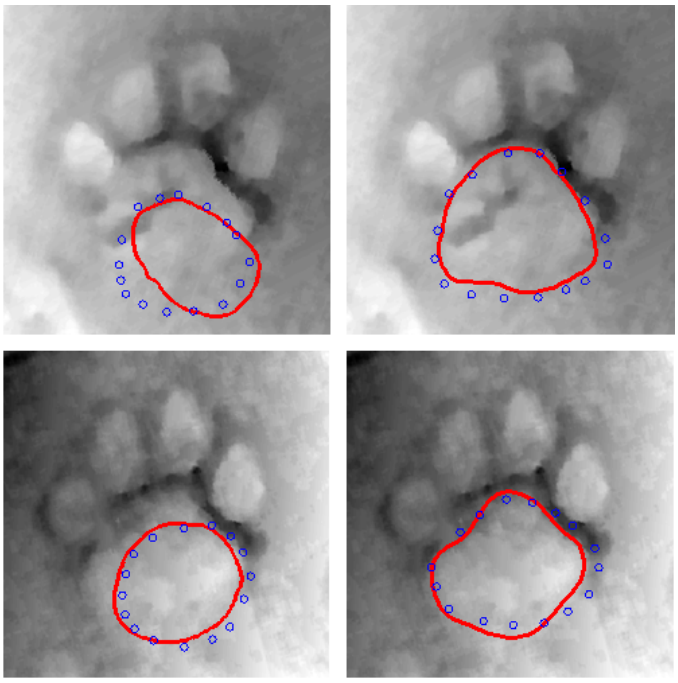

Fig. 9. Results showing the alignment effects on the initialization and final result. Blue circles are the results of the shortest path and red curve is the result of the refinement. Column 1: Results for test images not aligned to the reference image. Column 2: Results for test image aligned with the reference image.

Fig. 10 shows the average distance error for the different parts of the footprint. We observe that the ADE for the toes is smaller than the ADE for the pad. This is expected as the boundary edges of the lower portion of the pad are not clearly distinguishable. Experts had to rely on color image information in order to obtain their segmentation in these cases. Overall, the ADE for the entire footprint is less than 6 pixels. We also show the ADE for the pad without alignment. It clearly demonstrates that the error is more than the aligned case and hence justifies the need for alignment.

X. Conclusion and Future Work

This paper introduces a methodology for the identification of curves of interest in an image with applications to the segmentation of animal footprints from depth images. The approach models each curve as a stochastic process in a manifold and searches for the curve with the maximum likelihood to be generated by such a process. The search is done via a shortest-path formulation of the problem in a graphical model, and this shortest path estimate is then refined. The approach is trained and tested using a dataset of images with feature points manually selected by experts.

Improving efficiency: The method can be made more efficient by using more target feature points and getting a much better initialization for the refinement process. For the training, we used only one pair of the toes for aligning the other images; we could improve the training by considering all combination of pairs of toes and choose the combination which gave the least variance in the model. This could lead to better performance. The optimization process could also be sped up by applying a particle-filter search. Instead of randomly

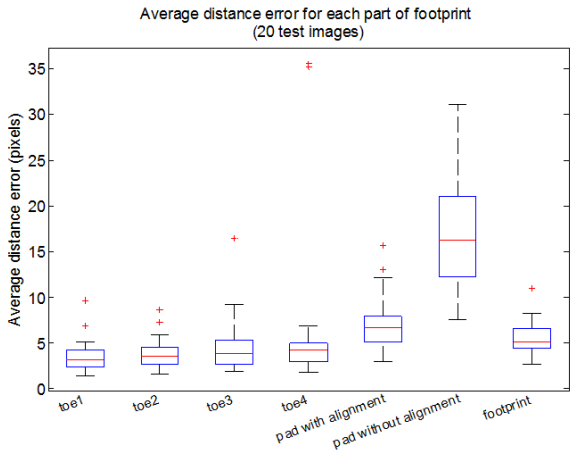

Fig. 10. Box plots of average distance error for different regions of the footprints. Errors are shown for the different toes, pad with alignment, pad without alignment and whole footprint.

searching for potential candidates around the mean locations for all the target points, we could fix a given point to be our initial point. For this location we perform a random search around the mean location. And for all potential candidates we search the next target point. In this way we only find feasible paths and do not perform an exhaustive search. We can also find target feature points in the image using our eigen-based approach instead of random selection.

Generalization: Several applications in medical imaging require the segmentation of biological tissue. Our approach for footprints can be generalized as an object segmentation method in the following way. We consider a sequential segmentation approach, finding objects based on saliency of features.

In the future, we plan to expand the footprint dataset in order to perform a more in-depth evaluation of our method and incorporate this process with non-invasive footprint identification techniques [43]. The final goal would be to classify animals using the footprint images.

References

[1] P. Sukhdev, H. Wittmer, C. Schrter-Schlaack, C. Nesshver, J. Bishop, P. t. Brink, H. Gundimeda, P. Kumar, and B. Simmons, "Mainstreaming the economics of nature: A synthesis of the approach, conclusions and recommendations of teeb," *TEEB*, 2010.

[2] H. Lex, L. Phil, P. Narendra, K. N. Samba, G. Arjun M., and K. K. Ullas, "A tiger cannot change its stripes: using a three-dimensional model to match images of living tigers and tiger skins," *Biol. Lett.*, vol. 5, no. 3, pp. 383–386, 2009.

[3] Z. C. Jewell, "Effect of monitoring technique on quality of conservation science," *Conservation Biology*, vol. 27, pp. pp. 501–508, 2013.

[4] Z. Jewell and S. Alibhai, "Identifying endangered species from footprints," *The International Society for Optics and Photonics (SPIE) Newsroom*, 2013.

[5] P. Riordan, "Unsupervised recognition of individual tigers and snow leopards from their footprints," *Animal Conservation*, vol. 1, no. 4, pp. 253–262, 1998. [Online]. Available: http://dx.doi.org/10.1111/j.1469-1795.1998.tb00036.x

[6] S. Sharma, Y. Jhala, and V. B. Sawarkar, "Identification of individual tigers (panthera tigris) from their pugmarks," *Journal of Zoology*, vol. 267, no. 1, pp. 9–18, 2005.

[7] S. K. Alibhai and Z. C. Jewell, "Identifying white rhino by a footprint identification technique, at the individual and species levels," *Endangered Species Res.*, vol. 4. [Online]. Available: #http://www.intres.com/articles/esr2008/4/n004p205.pdf#

[8] J. Shi and J. Malik, "Normalized cuts and image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 888–905, 2000.

[9] Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary amp; region segmentation of objects in n-d images," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1, 2001, pp. 105–112 vol.1.

[10] M. Salah, A. Mitiche, and I. Ayed, "Multiregion image segmentation by parametric kernel graph cuts," *Image Processing, IEEE Transactions on*, vol. 20, no. 2, pp. 545–557, 2011.

[11] C. Li, C. Xu, C. Gui, and M. Fox, "Distance regularized level set evolution and its application to image segmentation," *Image Processing, IEEE Transactions on*, vol. 19, no. 12, pp. 3243–3254, 2010.

[12] "Tiger rgb-d dataset," [URL TO BECOME AVAILABLE IF PAPER IS ACCEPTED].

[13] A. Elnakib, G. Gimelfarb, J. Suri, and A. El-Baz, "Medical image segmentation: A brief survey," in *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, A. S. El-Baz, R. Acharya U, A. F. Laine, and J. S. Suri, Eds. Springer New York, 2011, pp. 1–39. [Online]. Available: http://dx.doi.org/10.1007/978-1-4419-8204-9_1

[14] S. Parisot, H. Duffau, S. Chemouny, and N. Paragios, "Graph-based detection, segmentation amp; characterization of brain tumors," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 988–995.

[15] A. Mishra, A. Shrivastava, and Y. Aloimonos, "Segmenting simple objects using rgb-d," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, 2012, pp. 4406–4413.

[16] A. Hernandez-Vela, N. Zlateva, A. Marinov, M. Reyes, P. Radeva, D. Dimov, and S. Escalera, "Graph cuts optimization for multi-limb human segmentation in depth maps," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 726–732.

[17] A. Quattoni and A. Torralba, "Recognizing indoor scenes," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 413–420.

[18] S. Osher and J. A. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations," *JOURNAL OF COMPUTATIONAL PHYSICS*, vol. 79, no. 1, pp. 12–49, 1988.

[19] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *Computer Vision, 1995. Proceedings., Fifth International Conference on*, 1995, pp. 694–699.

[20] R. Malladi, J. Sethian, and B. Vemuri, "Shape modeling with front propagation: a level set approach," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, no. 2, pp. 158–175, 1995.

[21] T. Chan and L. Vese, "Active contours without edges," *Image Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 266–277, 2001.

[22] L. A. Vese, T. F. Chan, Tony, and F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International Journal of Computer Vision*, vol. 50, pp. 271–293, 2002.

[23] T. Cootes, C. Taylor, D. Cooper, and J. Graham", "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38 – 59, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1077314285710041

[24] G. Sundaramoorthi, S. Soatto, and A. Yezzi, "Curious snakes: A minimum latency solution to the cluttered background problem in active contours," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, June 2010, pp. 2855–2862.

[25] J. Shi and J. Malik, "Normalized cuts and image segmentation," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, 1997, pp. 731–737.

[26] S. Maji, N. Vishnoi, and J. Malik, "Biased normalized cuts," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 2057–2064.

[27] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 9, pp. 1124–1137, 2004.

[28] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," in *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 309–314.

[29] S. Vicente, V. Kolmogorov, and C. Rother, "Graph cut based image segmentation with connectivity priors," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008, pp. 1–8.

[30] J. Carreira and C. Sminchisescu, "Constrained parametric min-cuts for automatic object segmentation," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010, pp. 3241–3248.

[31] M.-S. Choi and W.-Y. Kim, "A novel two stage template matching method for rotation and illumination invariance," *Pattern Recognition*, vol. 35, no. 1, pp. 119 – 129, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320301000255

[32] Y. Hel-Or, H. Hel-Or, and E. David, "Fast template matching in non-linear tone-mapped images," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, 2011, pp. 1355–1362.

[33] M. Dahan, N. Chen, A. Shamir, and D. Cohen-Or, "Combining color and depth for enhanced image segmentation and retargeting," *The Visual Computer*, vol. 28, no. 12, pp. 1181–1193, 2012. [Online]. Available: http://dx.doi.org/10.1007/s00371-011-0667-7

[34] S. Gupta, P. Arbelaez, and J. Malik, "Perceptual organization and recognition of indoor scenes from rgb-d images," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 564–571.

[35] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from rgbd images," in *Proceedings of the 12th European conference on Computer Vision - Volume Part V*, ser. ECCV'12. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 746–760. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-33715-4_54

[36] X. Ren, L. Bo, and D. Fox, "Rgb-(d) scene labeling: Features and algorithms," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 2759–2766.

[37] C. Erdogan, M. Paluri, and F. Dellaert, "Planar segmentation of rgbd images using fast linear fitting and markov chain monte carlo," in *Computer and Robot Vision*, 2012. [Online]. Available: http://frank.dellaert.com/pubs/Erdogan12crv.pdf

[38] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, Jun 1991, pp. 586–591.

[39] J.-F. Yang and S.-S. Hao, "Image segmentation with eigen-subspace projections," in *Video Segmentation and Its Applications*, K. N. Ngan and H. Li, Eds. Springer New York, 2011, pp. 25–57. [Online]. Available: http://dx.doi.org/10.1007/978-1-4419-9482-0_2

[40] (2009) Carolina tiger rescue. http://www.carolinatigerrescue.org/. [Online]. Available: http://www.carolinatigerrescue.org/

[41] (2013) Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/

[42] J. C. Leger, "Menger curvature and rectifiability," *Annals of Mathematics*, vol. 149, no. 3, pp. pp. 831–869, 1999. [Online]. Available: http://www.jstor.org/stable/121074

[43] (2004) Wildtrack. http://www.carolinatigerrescue.org/. [Online]. Available: http://wildtrack.org/