

**IDENTIFYING HUMPBACK WHALE FLUKES BY  
SEQUENCE MATCHING OF TRAILING EDGE  
CURVATURE**

By

Zachary Jablons

A Thesis Submitted to the Graduate  
Faculty of Rensselaer Polytechnic Institute  
in Partial Fulfillment of the  
Requirements for the Degree of  
MASTER OF SCIENCE  
Major Subject: COMPUTER SCIENCE

Examining Committee:

---

Dr. Charles Stewart, Thesis Adviser

---

Dr. Barbara Cutler, Member

---

Dr. Bülent Yener, Member

Rensselaer Polytechnic Institute  
Troy, New York

April 2016  
(For Graduation May 2016)

© Copyright 2016  
by  
Zachary Jablons  
All Rights Reserved

# CONTENTS

LIST OF TABLES . . . . .	iv
LIST OF FIGURES . . . . .	v
ACKNOWLEDGMENTS . . . . .	vi
ABSTRACT . . . . .	vi
1. Introduction . . . . .	1
1.1 Humpback Whales . . . . .	1
1.2 Current Identification Methods . . . . .	1
1.2.1 Based on Trailing Edge . . . . .	1
1.2.2 Based on general Fluke appearance . . . . .	2
1.2.2.1 Hotspotter . . . . .	2
1.3 The Dataset . . . . .	2
1.4 Method Overview . . . . .	2
2. Background . . . . .	4
2.1 Convolutional Networks . . . . .	4
2.1.1 Facial Keypoint Prediction . . . . .	4
2.1.2 Fully Convolutional Networks . . . . .	5
2.2 Seam Carving . . . . .	5
2.3 Curvature Measures . . . . .	6
2.3.1 Differential Curvature . . . . .	6
2.3.2 Integral Curvature . . . . .	6
2.4 Dynamic Time Warping . . . . .	6
3. Methods . . . . .	8
3.1 Trailing Edge Extraction . . . . .	8
3.1.1 Gradient Information . . . . .	8
3.1.2 Keypoints . . . . .	8
3.1.3 Trailing Edge Scoring . . . . .	8
3.2 Trailing Edge Matching . . . . .	8
3.2.1 Curvature Measurement . . . . .	8

3.2.2	Sequence Matching . . . . .	8
3.3	Alternative Approaches . . . . .	8
3.3.1	Embedding via Convolutional Networks . . . . .	8
3.3.1.1	Raw Images . . . . .	8
3.3.1.2	Trailing Edges . . . . .	8
3.3.2	Aligning Trailing Edges . . . . .	8
3.3.2.1	Keypoint Alignment . . . . .	8
3.3.2.2	Dynamic Time Warping Alignment . . . . .	8
3.3.3	Histogram Matching . . . . .	8
4.	Results . . . . .	9
4.1	Main method . . . . .	9
4.1.1	Characterization of Success cases . . . . .	9
4.1.2	Characterization of Failure cases . . . . .	9
4.1.3	Variations . . . . .	9
4.2	In Combination with Hotspotter . . . . .	9
4.2.1	Failure cases . . . . .	9
5.	Discussion . . . . .	10
5.1	Issues with current method . . . . .	10
5.2	Future work . . . . .	10
5.3	Conclusion . . . . .	10
	REFERENCES . . . . .	11
	APPENDIX	

## LIST OF TABLES

## LIST OF FIGURES

## ACKNOWLEDGMENTS

This project could not have been completed so quickly and thoroughly without the invaluable help from the rest of the IBEIS team. I would like to thank Jon Crall for helping to write the experimentation framework which has saved a lot of time and effort. I would also like to thank Jason Parham and Hendrik Weidemann for the great discussions and advice, as well as the handy `LATEX`template. Additionally, without the development and annotation efforts of Andrew Batbouta, a lot of models could not have been so successfully trained. Importantly, I would like to thank the Wildbook team (Jason Holmberg, Jon van Oast) for providing the dataset and corresponding annotations with which a lot of models were trained and experiments run. Of course, this work could never have been undertaken without the help, advice, and direction of my advisor, Charles Stewart.

## ABSTRACT

Humpback whales (*Megaptera novaeangliae*) are an important part of our ocean's ecosystems [citation needed], and have historically been at risk for extinction [citation needed]. While they are currently rated as 'Least Concern' [citation needed], tracking their migration patterns is important for helping the (currently small) population grow [citation needed]. In order to discern these migration patterns, conservationists need to be able to track individual humpback whales [citation needed]. One of the easiest (and cheapest) ways to do this is to watch for their tails as they breach the water surface [citation needed], giving a clear view of what is known as a Humpback 'fluke'. These are often patterned and scarred in unique ways, allowing conservationists to identify individuals [citation needed]. However, until recently, most automated identification methods still rely on significant manual effort to describe and identify the fluke, severely limiting the amount of humpbacks that can be tracked [citation needed?].

This thesis lays out a method that automates the identification of Humpback flukes directly from still images thereof, using the 'trailing edge' of the fluke. Using this method, we achieve a fairly high top-1 ranking accuracy on a large dataset (consisting of about 400 identified individuals). We also show that this method significantly helps the accuracy of a pure appearance based method, Hotspotter [citation needed], giving 89% top-1 accuracy.

To our knowledge, this is the first method that can achieve this level of accuracy on Humpback fluke identification without any manual effort at test-time.

TODO: Put in citations



# CHAPTER 1

## Introduction

### 1.1 Humpback Whales

TODO: Put in something about the importance of identifying humpback whales, with citations and all

### 1.2 Current Identification Methods

Photo-identification of Humpback whale flukes has been attempted since the early 90s [21] using computational aid. While early efforts mostly relied on a manual description of the fluke that would then be matched, later efforts have involved matching flukes based on automated analysis of both the patterns on the ventral side of the fluke. It is worth noting however that as shown in [3], the trailing edge changes less with age than surface patterns on the fluke, which means that it can (potentially) be a more reliable identifier over time. That said, trailing edges are hard to photograph well, requiring high resolution imagery and a consistent angle between fluke and photographer.

Each of these annotation methods can be separated into three categories:

- Manually annotated – a human must manually annotate or catalogue features in the fluke
- Semi-automated – a human must guide an algorithm (e.g. by setting control points or highlighting interesting regions) that then automatically identifies the individual
- Fully-automated – the algorithm can identify individuals from raw images

#### 1.2.1 Based on Trailing Edge

In [21], information about the trailing edge and fluke patterns are manually catalogued and used to match individual whales. This falls under a manual an-

notation approach. In the I3S contour system [10], the user must input start and end points on the contour, which are then extracted and checked manually, giving a semi-automated system. At the time of writing no published results on this system applied to Humpback whales could be found. Automatically identifying Humpback whales by their entire trailing edge contour is done experimentally in [13], using a technique that is originally designed for Great White Sharks. However the results published on that technique are for a much smaller dataset than the one worked on in this thesis. While trailing edge matching has seen limited use in Humpback whale identification, it is a much more common technique in Sperm whale (*P. macrocephalus*) identification [12], [2] [33], with varying levels of manual effort.

### **1.2.2 Based on general Fluke appearance**

The primary method for identifying Humpback whale flukes is to use the ventral fluke pattern, as seen in [21], [4], [3], [9], and (in an semi-automated fashion) [15].

#### **1.2.2.1 Hotspotter**

Hotspotter [6] is an automated photo-identification algorithm based on SIFT features that has been used in identifying Grevy’s Zebras, Plains Zebras, Giraffes, and Elephants [24]. This work is the first to our knowledge of Hotspotter being applied to Humpback whale flukes, and the results are presented in chapter 4.

## **1.3 The Dataset**

TODO: Get information from Jason (and Ted?) about how the dataset(s) I’m using were collected and sourced.

## **1.4 Method Overview**

The method for trailing edge identification put forth in this thesis is fully automated, requiring no human annotation when used (although manual annotation is necessary for training the machine learning models used). On its own, it achieves decent results on a (relatively) large dataset, comparable with the fully automated

method used in [13]. Ultimately we suggest that this method be used in combination with an automated pattern matching method (e.g. Hotspotter) to provide very high accuracy matches. We also explore alternative methods based on more recent advances in deep learning for identification, however it appears that the dataset is too small to properly train these methods.

## CHAPTER 2

### Background

In this chapter, we provide a series of sections detailing background information on the algorithms from which this method was developed.

#### 2.1 Convolutional Networks

In recent years, convolutional neural networks have provided state of the art results in several challenging computer vision tasks, including general image classification [16], [32], image segmentation [20], [5] and individual identification (specifically for human faces) [7], [28].

The essential idea of a convolutional network is that it we can use the gradient of an error signal to learn hierarchies of convolution kernels separated by nonlinear activation functions, providing a meaningful prior to neural networks when applied to data with spatial invariances (e.g. image data). Convolutional networks have been around for a long time, but the current incarnation of these networks can be traced back to [18]. More recently however, convolutional networks have grown deeper and are often modeled after the architectural decisions made in [30], [29], and [16]. These decisions are specifically the use of Rectified Linear Units (ReLUs) as activation functions, Dropout [11] after fully connected layers for regularization, and small square kernels with "same" padding alternated with 2x downsampling layers (max pool).

The convolutional networks used in this thesis use the above architectural decisions, and also use batch normalization [14] at every layer.

##### 2.1.1 Facial Keypoint Prediction

Facial keypoints are used in a lot of identification pipelines, as well as in motion capture and expression recognition. There has been recent work in using convolutional networks for facial keypoint prediction [31], [23].

In this work, fluke keypoints are predicted using a very similar technique but

with different underlying convolutional architectures. The essential idea is that the convolutional network predicts points (rather than classifications) in the form of  $(x, y)$  coordinates, and is treated as a regression network (i.e. the RMSE loss is used). This will be explained in more detail in the next chapter.

### 2.1.2 Fully Convolutional Networks

Classically, convolutional networks for classification predict a single class for an image, ignoring the possibility of multiple objects of varying classes being present throughout the image. These classification networks usually have fully connected (or dense) layers towards the end, which forces the size of the network input to be fixed. However, this allows the network to make learned decisions over the entire input without any sort of spatial pooling. When dealing with arbitrarily sized images, it is typical to use networks that are 'fully convolutional', in which case the entire network consists of convolution kernels. This technique can also be used for segmentation, as we can simply replace the dense part of the network with convolutional parts that can make class predictions on every pixel of their input. By upsampling and combining different stages of prediction, the authors in [20] produce high quality image segmentations.

In this work, fully convolutional networks are used for predicting the 'trailing edginess' of an image, which allows us to refine the trailing edge contour extraction. This is explained in greater detail in the next chapter.

## 2.2 Seam Carving

Seam carving is a technique that uses image saliency information to resize images without warping or distorting the objects shown in the image [1]. This technique uses dynamic programming to find minimal saliency paths through an image, where saliency is often defined as the gradient. The motivation for this is that these minimal saliency paths are not important to the image, so they can be removed to reduce its size.

While this method is not directly used in this work, the underlying algorithm for trailing edge extraction is based on a single iteration of the seam carving algo-

rithm, using gradient information.

## 2.3 Curvature Measures

Contour curvature allows one to characterize the overall shape of an contour by looking at its edge. A lot of work has been done on using curvature information for detection [22], classification [8], and identification [17]. This curvature information can be broadly broken down into either integral or differential curvature, and is usually computed at multiple scales.

### 2.3.1 Differential Curvature

Differential curvature can generally be seen as measuring the measures the angle of the tangent normal of the gradient at each point in an image [8]. For our purposes, we can then take only those points that lie on the contour and use their curvature. While doing this directly can be fast to compute, it tends to be noise sensitive and we found that integral curvature (below) works better for our purposes.

### 2.3.2 Integral Curvature

Integral curvature works (conceptually) by sliding a circle of some radius  $r$  along the contour [25], and measuring how much of the circle is 'inside' the contour. This measurement is usually taken at multiple scales, and has the appealing property of being invariant to rotation and translation. In this work, we approximate the circular curvature with a square, which appears to perform just as well but can be computed faster. This is further explained in the next chapter.

## 2.4 Dynamic Time Warping

In deciding a sequence comparator, one criterion that is often important is ensuring that small shifts in the sequence do not balloon into large distances. Dynamic Time Warping (DTW) is a sequence comparison method that, roughly, finds the optimal matching between all sets of points in the two given sequences that minimizes the overall distance between the matched points, while keeping the locality of the

points intact. This allows for shifts and some warps in the two sequences to be compensated for, and results in a nonlinear mapping of one sequence onto another. The algorithmic complexity of dynamic time warping can be limiting in large datasets, as it is quadratic in both space and time – making a one-to-one comparison a bit daunting.

There are several variants on DTW that give faster speeds [27] [19], however we only use the Sakoe-Chiba bound [26], which both constrains the neighborhood in which points can be matched and gives a complexity of  $O(nw)$ , where  $w$  is a user set parameter.

Sequences of curvature measures have been used with DTW for signature verification [? ], however this combination has not been used for matching trailing edges to our knowledge.

## **CHAPTER 3**

### **Methods**

#### **3.1 Trailing Edge Extraction**

##### **3.1.1 Gradient Information**

##### **3.1.2 Keypoints**

##### **3.1.3 Trailing Edge Scoring**

#### **3.2 Trailing Edge Matching**

##### **3.2.1 Curvature Measurement**

##### **3.2.2 Sequence Matching**

#### **3.3 Alternative Approaches**

##### **3.3.1 Embedding via Convolutional Networks**

###### **3.3.1.1 Raw Images**

###### **3.3.1.2 Trailing Edges**

##### **3.3.2 Aligning Trailing Edges**

###### **3.3.2.1 Keypoint Alignment**

###### **3.3.2.2 Dynamic Time Warping Alignment**

##### **3.3.3 Histogram Matching**



## **CHAPTER 4**

### **Results**

#### **4.1 Main method**

##### **4.1.1 Characterization of Success cases**

##### **4.1.2 Characterization of Failure cases**

##### **4.1.3 Variations**

#### **4.2 In Combination with Hotspotter**

##### **4.2.1 Failure cases**

## **CHAPTER 5**

### **Discussion**

**5.1 Issues with current method**

**5.2 Future work**

**5.3 Conclusion**



## REFERENCES

- [1] S. Avidan and A. Shamir. Seam carving for content-aware image resizing. *ACM Trans. Graph.*, 26(3), July 2007.
- [2] B. Beekmans, H. Whitehead, R. Huele, L. Steiner, and A. G. Steenbeek. Comparison of two computer-assisted photo-identification methods applied to sperm whales (*physeter macrocephalus*). *Aquatic Mammals*, 31(2):243, 2005.
- [3] A. L. Blackmer, S. K. Anderson, and M. T. Weinrich. Temporal variability in features used to photo-identify humpback whales (*megaptera novaeangliae*). *Marine Mammal Science*, 16(2):338–354, 2000.
- [4] C. A. Carlson, C. A. Mayo, and H. Whitehead. Changes in the ventral fluke pattern of the humpback whale (*megaptera novaeangliae*), and its effect on matching; evaluation of its significance to photo-identification research. *Rep. int. Whal. Commn (special issue)*, 12:105–11, 1990.
- [5] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv preprint arXiv:1412.7062*, 2014.
- [6] J. P. Crall, C. V. Stewart, T. Y. Berger-Wolf, D. I. Rubenstein, and S. R. Sundaresan. HotSpotter - patterned species instance recognition. In *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, pages 230–237. IEEE.
- [7] H. Fan, Z. Cao, Y. Jiang, Q. Yin, and C. Doudou. Learning deep face representation. *arXiv preprint arXiv:1403.2802*, 2014.
- [8] P. Fischer and T. Brox. Image descriptors based on curvature histograms. In *Pattern Recognition*, pages 239–249. Springer, 2014.
- [9] N. Friday, T. D. Smith, P. T. Stevick, and J. Allen. Measurement of photographic quality and individual distinctiveness for the photographic identification of humpback whales, *megaptera novaeangliae*. *Marine Mammal Science*, 16(2):355–374, 2000.
- [10] d. J. Hartog and R. Reijns. I3S Contour MANUAL, 2013.
- [11] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*, 2012.
- [12] R. Huele, H. U. De Haes, J. Ciano, and J. Gordon. Finding similar trailing edges in large collections of photographs of sperm whales. *Journal of Cetacean Research and Management*, 2(3):173–176, 2000.

- [13] B. Hughes and T. Burghardt. Automated identification of individual great white sharks from unrestricted fin imagery. In M. W. J. Xianghua Xie and G. K. L. Tam, editors, *Proceedings of the British Machine Vision Conference (BMVC)*, pages 92.1–92.14. BMVA Press, September 2015.
- [14] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.
- [15] E. Kniest, D. Burns, and P. Harrison. Fluke matcher: A computer-aided matching system for humpback whale (*megaptera novaeangliae*) flukes. *Marine Mammal Science*, 26(3):744–756, 2010.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [17] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. Soares. Leafsnap: A computer vision system for automatic plant species identification. In *Computer Vision–ECCV 2012*, pages 502–516. Springer, 2012.
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [19] D. Lemire. Faster retrieval with a two-pass dynamic-time-warping lower bound. *Pattern recognition*, 42(9):2169–2180, 2009.
- [20] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.
- [21] S. A. Mizroch, J. A. Beard, and M. Lynde. Computer assisted photo-identification of humpback whales. *Report of the International Whaling Commission*, 12:63–70, 1990.
- [22] A. Monroy, A. Eigenstetter, and B. Ommer. Beyond straight linesobject detection using curvature. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 3561–3564. IEEE, 2011.
- [23] M. E. Munich and P. Perona. Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 108–115. IEEE, 1999.
- [24] D. Nouri. Using convolutional neural nets to detect facial keypoints tutorial, 2014.

- [25] J. R. Parham. Photographic censusing of zebra and giraffe in the nairobi national park.
- [26] H. Pottmann, J. Wallner, Q.-x. Huang, and Y.-l. Yang. Integral invariants for robust geometry processing. 2007.
- [27] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):43–49, 1978.
- [28] S. Salvador and P. Chan. Fastdtw: Toward accurate dynamic time. *Warping in Linear Time and Space*, 2007.
- [29] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 815–823, 2015.
- [30] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*, 2013.
- [31] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [32] Y. Sun, X. Wang, and X. Tang. Deep convolutional network cascade for facial point detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3476–3483, 2013.
- [33] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015.
- [34] H. Whitehead. Computer assisted individual identification of sperm whale flukes. *Reports of the International Whaling Commission*, 12:71–77, 1990.

## APPENDIX