WASHINGTON UNIVERSITY IN ST. LOUIS

School of Engineering and Applied Science

Department of Computer Science and Engineering

Thesis Examination Committee:
Dr. Ron Cytron Dr. Robert Pless
Dr. William Smart

USING REGRESSION TECHNIQUES TO PREDICT WEATHER SIGNALS FROM IMAGE SEQUENCES

by

Richard Speyer

A thesis presented to the School of Engineering
of Washington University in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

May 2009
Saint Louis, Missouri

ABSTRACT OF THE THESIS

Using Regression Techniques to Predict Weather Signals from Image Sequences

by

Richard Speyer

Master of Science in Computer Science

Washington University in St. Louis, 2009

Research Advisor: Dr. Robert Pless

Webcams are cheap sensors that capture a potentially large amount of information about a scene. This thesis considers the use of regression and correlation techniques such as Canonical Correlation Analysis (CCA) to convert these webcams into environmental sensors and predict the values of weather signals. Local environmental properties often directly affect the images we collect from the webcams; whether it is cloudy or sunny is visible by the presence of shadows; wind speed and direction is visible in smoke, flags, or close up views of trees; particulate density is reflected in haziness and the color spectrum during sunset. Using the AMOS database, which has been archiving nearly 1,000 webcams every 30 minutes for the last 3 years, we explore relationships between the amount of training data and the accuracy with which we are able to infer the values of certain weather signals including wind speed & direction and vapor pressure from inherent properties in the image. This allows the webcams *already* installed across the earth to act as generic sensors to improve our understanding of local weather patterns and variations.

Acknowledgments

First and foremost, I would like to thank Dr. Robert Pless, who has been my research and faculty adviser for the last four years. His constant attention and guidance have been crucial in the construction of this thesis and I am extremely grateful for all of his help. My experience working under him in the Media & Machines Lab is one of the most valuable of my college career and has taught me skills that I will surely use after I graduate.

I would also like to thank Nathan Jacobs and many other students who helped review this thesis and the related research as well as Dr. William Smart and Dr. Ron Cytron for sitting on my Thesis Committee along with Dr. Pless.

Finally, I would like to thank the Department of Computer Science & Engineering for promoting and providing many opportunities for undergraduates to get involved in research projects. I have benefited greatly from working in a research lab during my time here and that would not have been possible without the department's commitment to undergraduate research.

Richard Speyer

Washington University in Saint Louis
May 2009

Contents

Abstract . . . . . **ii**

Acknowledgments . . . . . **iii**

List of Tables . . . . . **v**

List of Figures . . . . . **vi**

1 Introduction **1**
1.1 Background Information 1
1.2 Related Work 1

2 Canonical Correlation Analysis (CCA) **2**

3 Results & Analysis **3**
3.1 Wind Speed & Direction 3
3.2 Vapor Pressure 3
3.3 Error Analysis 3

4 Conclusion **4**

5 Future Work **5**

Appendix A Camera Information **6**

Appendix B Weather Data **7**

References **8**

Vita **9**

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background Information

The Archive of Many Outdoor Scenes (AMOS) database [1] has been collecting images from 835 webcams every 30 minutes since March 2006 and now contains over 40 million images. This database is the largest known collection of natural scenes collected from static cameras and as such offers a wealth of data to test our methods against. While there are cameras located across the world, we focus on those located within the continental United States so that we can collect accurate ground truth weather data.

1.2 Related Work

Chapter 2

Canonical Correlation Analysis (CCA)

Chapter 3

Results & Analysis

3.1 Wind Speed & Direction

3.2 Vapor Pressure

3.3 Error Analysis

Chapter 4

Conclusion

Chapter 5

Future Work

Appendix A

Camera Information

Appendix B

Weather Data

References

[1] Nathan Jacobs, Nathaniel Roman, and Robert Pless. Consistent temporal variations in many outdoor scenes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, Minneapolis, MN, June 2007.

[2] Dongge Li, Nevenka Dimitrova, Mingkun Li, and Ishwar K. Sethi. Multimedia content processing through cross-modal association. In *MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia*, pages 604–611, New York, NY, USA, 2003. ACM.

[3] S.G. Narasimhan, C. Wang, and S.K. Nayar. All the Images of an Outdoor Scene. In *Europian Conference on Computer Vision (ECCV)*, volume III, pages 148–162, May 2002.

Vita

Richard Speyer

Date of Birth	October 11, 1986
Place of Birth	Johannesburg, South Africa
Degrees	B.S. Computer Science, May 2009 M.S. Computer Science, May 2009
Professional Societies	Association for Computing Machinery Sigma Xi Scientific Honorary
Publications	N. Jacobs, S. Satkin, N. Roman, R. Speyer, and R. Pless. Geolocating static cameras. In *Proc. IEEE International Conference on Computer Vision*, Oct. 2007.

May 2009

Learning to Find Signals in Image Sequences, Speyer, M.S. 2009