Medical Diagnostics over Video

Extracting the pulse and oxygen levels in blood from video of a subject

D Deriso, A Fallou, N Banerjee

Abstract

In both developing and developed countries, reducing the cost of medical care is a primary goal of science and government. In this project we seek to find and extract information from a video of a human that tells us the pulse rate and the oxygen level saturation of the blood. We therefore aim to create a virtual pulse oximeter: the ultimate non-invasive, equipment-free medical diagnostics tool, which could be deployed to anyone with video recording capabilities. Features were chosen to be related to the three color channel intensity values, with the idea that changing color of the video would relate to blood flow around the body. Extensive pre-processing was required on both the video data and the pulse oximeter data to enable training. Early results showed that feature selection was vital in reducing the mean-squared error of the output, but plenty of further work can be done.

Introduction

Cardiovascular health is the *sin qua non* of human life. Early detection of cardiovascular disease is of paramount importance in public health. This project aims to develop a method to visualize the perfusion of blood through the skin via pulse oximetry. Pulse oximetry is a technique that exploits the fact that oxygenated and deoxygenated hemoglobin changes the color of red blood cells. The technique maps these changes in rgb color of the visible skin to the invisible presence of oxygenated vs deoxygenated blood in the local vasculature underneath the skin.

Previous studies have shown that video obtained from an ordinary webcam can be used to visualize perfusion by selectively amplifying temporal frequencies in video (see http://people.csail.mit.edu/mrub/vidmag/). A study by the MIT CSAIL [1] showed that this technique can also be used to infer heart rate from the person being taped. The present project aims to extend this work to detect the relative changes in oxygenated vs deoxygenated blood and reconstruct the pulse oximeter waveform from an ordinary webcam video.

Main Objectives

1. Build software that can simultaneously record pulse oximeter wave values and time values while recording from a video.
2. Extract pixel color information and pulse oximeter data.
3. Estimate the number of training examples needed to implement linear regression using learning theory.
4. Implement regression in a variety of ways and train the weight matrix for the features from the video.
5. Test the newly learned weight matrices on further videos.
6. Create an error vs technique graph to compare techniques.
7. Conclude with the best technique.

Pre-processing steps

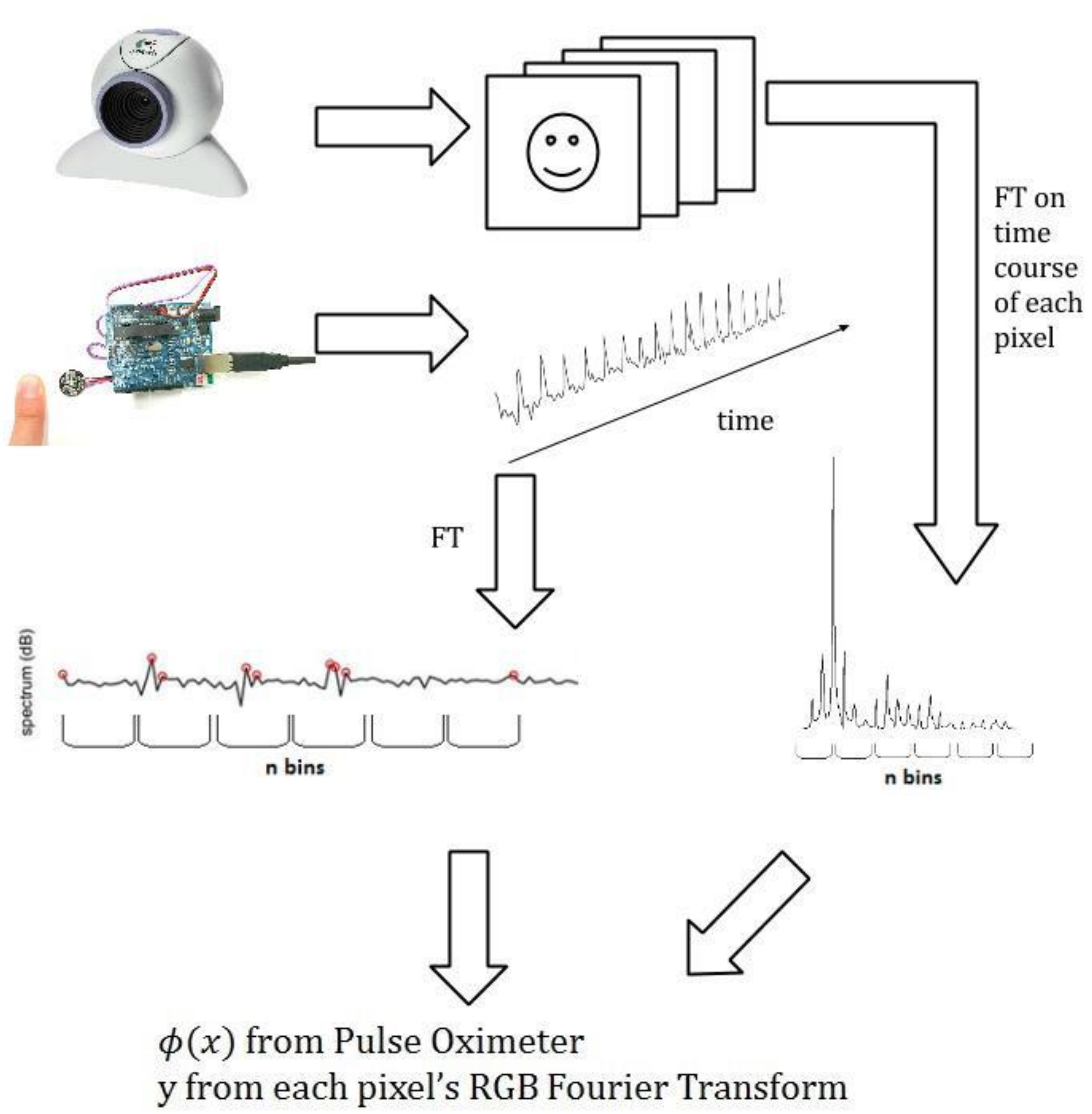


Figure 1: Chart showing the steps involved in pre-processing

Results

Features were extracted from each pixel over time, and included the pixel's [i,k] location, the FFT of the pixel's time course binned to 4 buckets per frequencies 0-6Hz [fq1, fq2, …, fq24], and phase [ph1, ph2, …, ph24]. Linear least squares regression models were trained on the following combination of features:

Freq	Phase	Pixel location	MSE	Cross Correlation
•	•	•	0.00884	0.0936
•	•	o	0.114	0.0943
•	o	•	0.00903	0.0936
•	o	o	0.116	0.0943
o	•	•	0.0115	0.0936
o	•	o	0.748	0.0967
o	o	•	0.0117	0.0936

These results suggest that each of the features, frequency, phase, and pixel location; play a role in the prediction. Furthermore, the least error was obtained when all three features were used (MSE =0.00884), followed by just frequency and pixel (MSE=0.00903), and pixel location (MSE=0.0117). As expected, for a single video, with numerous pixels serving as training examples, the residuals were low.

Eulerian Video Magnification: where we fit

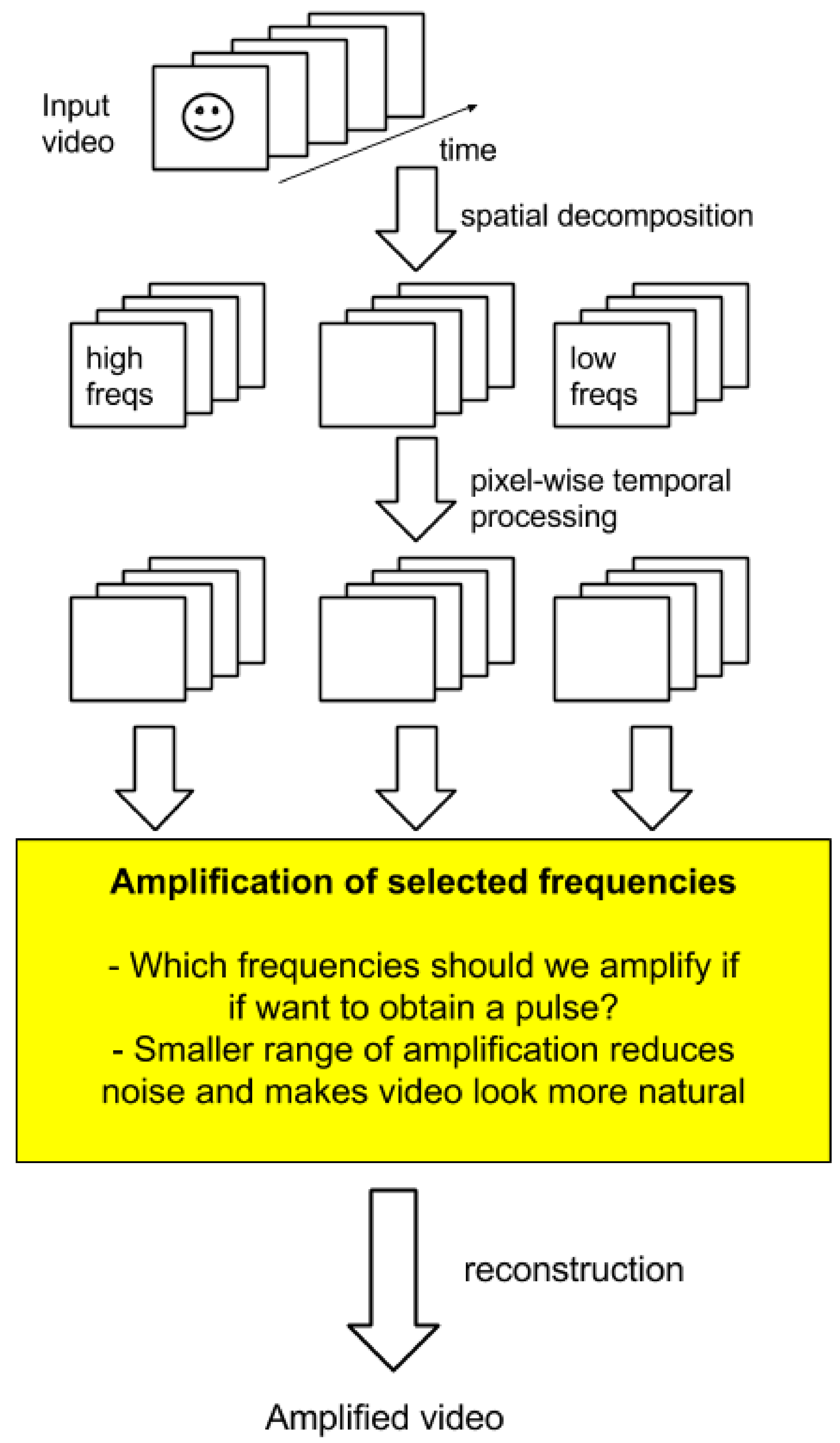
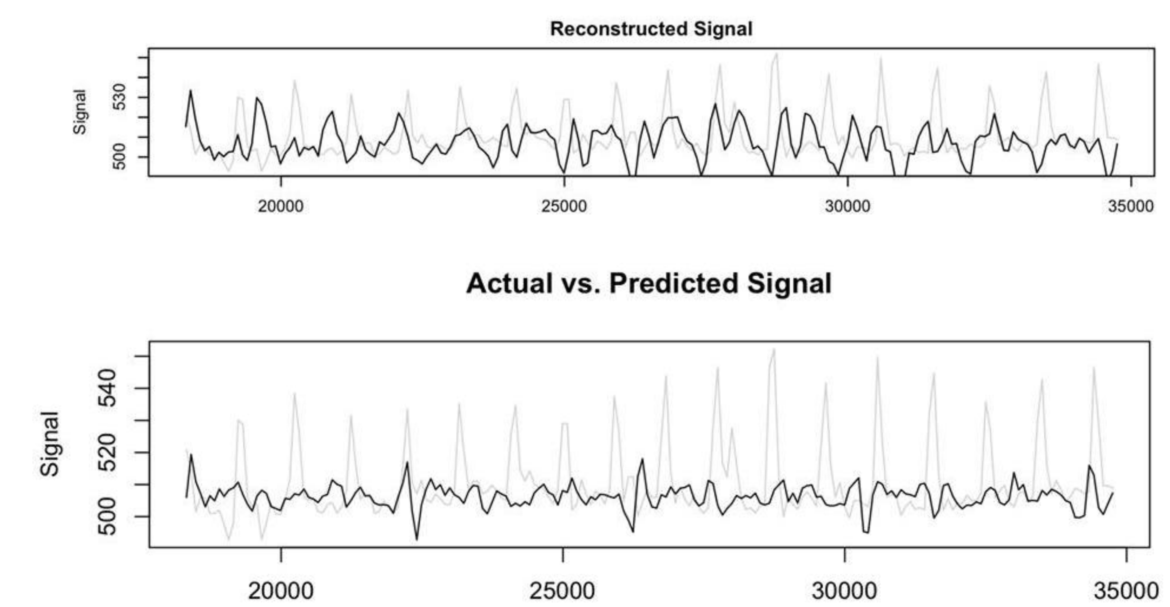


Figure 2: Flow chart showing how our pulse frequency extraction can be used to enhance Eulerian Video Magnification

Conclusions

- Using a binned Fourier Transform, we could efficiently reduce the video data to a finite dimensional feature space. This may not be the most ideal method of dimensionality reduction.
- We built an excellent program to enable training of the data: the program could record a series of images and pulse oximeter readings for those images.
- Increasing the variety of features included in the video reduced our mean squared error, however the cross correlations were little affected.



Further Work

- An ability to better reconstruct our signal is our top priority. The compression technique we used to encode our pulse oximeter signal was lossy.
- Feature selection is a weak point in our data analysis; we cannot get a very accurate understanding of the performance of our model until we have better features.
- More training examples are needed to better train our matrix and perform more meaningful cross correlations.

References

[1] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Duran and William T. Freeman. Eulerian vida magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (Proc. SIGGRAPH 2012)*, 31(4), 2012.

Acknowledgements

We would like to thanks Stephan Junek for his MatVis tool which allowed us to view the individual color channel intensities of videos; invaluable for the debugging process.