2012

Using Image Processing Techniques to Estimate the Air Quality

Ernesto Zamora
University of Nevada, Las Vegas

Using Image Processing Techniques to Estimate the Air Quality

Ernesto Zamora, McNair Scholar, Computer Science Major

Dr. Evangelos Yfantis, Faculty Mentor, Computer Science Department

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

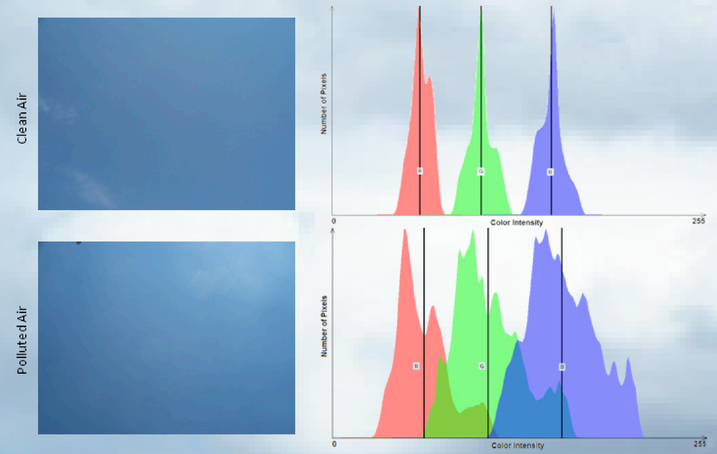
The Center
CENTER FOR ACADEMIC ENRICHMENT & OUTREACH

ABSTRACT

The color of the sky varies depending on the composition of the air. The particles lingering in the atmosphere scatter the components of the light based on their wavelength in relation to particle size. Thus, the color of the sky depends on what wavelengths are scattered while the light travels through our atmosphere.

By analyzing the color of the sky using computer image processing, it is possible to determine the quality of the air in an area.

It is possible to classify pictures of the sky into clean or polluted air and have a computer estimate the category any other picture belongs to.



INTRODUCTION

The purpose of this research is to use computer image processing techniques and statistics to study the irradiated tonality of the sky and its relation to air pollution. The sky color variations from shades of blue to shades of gray indicate higher levels of concentration of particles bigger than clean air normally contains. Using image processing, a computer can recognize the differences in color and determine the quality of the air in the area where a picture of the sky was taken.

Larger molecules scatter longer wavelength colors. When the concentrations of larger molecules such as pollutants increase, the scattering of these longer wavelength colors becomes more significant. The scattered wavelengths mix, making the blue sky lighter even making it look completely white or a shade of gray when pollution levels or water vapor molecules that cling together or attach to other molecules are high.

METHODOLOGY

Color histograms of the tristimulus values (color channels red, green and blue) were used to obtain data and analyze images throughout this study. Through observation of the resulting color histograms it was possible to determine how clean or polluted the air was in the area where the benchmark pictures were taken: note how on clean air histogram there is little to no overlap of the color components, while on less-clean to polluted air, the histogram colors are closer together and there is mayor overlap between the colors. We observed that the more polluted the air is, the closer the peaks get and the higher the overlap is.

A computer program was developed to generate the color histograms for the pictures as well as reports based on these histograms.

The sample benchmark pictures taken were analyzed and separated into clean and polluted categories based in their tristimulus expected values, forming two sets in the Euclidian space. We calculate the Mahalanobis distance from the expected color value of a picture of the sky to each of the benchmark sets' center of mass.

RESULTS

The theory discussed was applied and implemented in a computer program with good results. The software was able to distinguish pictures where the human eye could not appreciate a marked difference in color.

This is an ongoing study. A larger data collection will increase the accuracy of the software. Further analysis is being done and theory is being researched with plans to enhance the software to be able to determine the kind of pollutants existing in the air from the color changes in the sky.