Improved Blue Sky Detection Using Polynomial Model Fit

Abstract

Blue sky is a material that commonly appears in photographs. Reliable detection of blue sky can aid semantic image retrieval, image understanding, and image enhancement. A novel algorithm is proposed based on a two-dimensional polynomial model of the image of blue sky. An initial sky detection is used to establish some high confidence blue sky regions. Candidate sky regions are found and a two-dimensional polynomial model is used to compute the belief that a given region is also sky. On a database of 83 images, the algorithm correctly classified 31 additional regions as blue sky while adding 8 false positives, most of them reflections of sky. Further, the algorithm correctly increased the belief values on 6 regions found by the initial sky detector.

1. Introduction

Sky is among the most important subject matters frequently seen in photographic images. In a digital color image, a pixel or region represents sky if it corresponds to a sky region in the original scene. In essence, a pixel or region represents sky if it is an image of the earth's atmosphere. Detection of sky can often facilitate a variety of image understanding, enhancement, and manipulation tasks. Sky is a strong indicator of an outdoor image for scene categorization [1][2] (e.g., outdoor scenes vs. indoor scenes, picnic scenes vs. meeting scenes, city vs. landscape, etc.). With information about the sky, it is possible to formulate queries such as "outdoor images that contain significant sky" or "sunset images" etc., as shown by Smith and Li [3]. Thus, sky detection can also lead to more effective content-based image retrieval. Detection of all blue sky regions can be important when formulating queries that depend on the spatial distribution or location of sky regions.

Sky is also helpful for image understanding. For recognizing the orientation of an image, knowledge of sky and its orientation may indicate the image orientation for outdoor images (contrary to the common belief, a sky region is not always at the top of an image). Further, in detecting main subjects in the image, sky regions can usually be excluded because they are likely to be part of the background.

For some image enhancement applications, it is critical to detect all blue sky regions in the image. For example, if a preferred color enhancement is applied to sky regions, then using an incomplete sky classification may result in an image having an unnatural and inconsistent look.

In a work by Saber et al. [4], color classification was used to detect sky. The sky pixels are assumed to follow a 2D Gaussian probability density function (PDF). Therefore, a metric similar to the Mahalonobis distance is used, along with an adaptively determined threshold for a given image, to determine sky pixels.

Figure 1: Example digital image with disjointed blue sky. The algorithm exploits continuity between the separate regions of sky to find all sky regions.

Luo and Etz [5] show that as a result of the physics of light scattering by small particles in the air, clear sky often appears in the shade of deep, saturated blue at the top of the image and gradually desaturates to almost white towards a distant horizon line in the image. Their detector is unlikely to be fooled by other similarly colored objects, such as bodies of water, walls, toys, and clothes. The classification of sky is given to regions exhibiting an acceptable gradient signal. While the method described provides excellent performance, especially in eliminating other objects with similar colors to blue sky, the algorithm may fail to detect small sky regions (e.g. a small region visible between tree branches) because small regions may not exhibit the proper gradient signal.

In this paper, an algorithm based on two-dimensional polynomial fitting is presented to improve on the work of [5] by finding regions of sky that might otherwise be too small to exhibit the expected color gradient signature.

2. Algorithm overview

An example color image is shown in Figure 1. In this work, images were captured with digital cameras and film cameras and are represented in the well-known sRGB color space at a resolution of about 360 by 540 pixels.

An algorithm has been developed to detect disjointed blue sky regions. A block diagram of the algorithm is given in Figure 2. First, an initial belief map is made by using an initial blue sky detection algorithm [5] to determine for each region a belief (i.e., probability) that the region represents blue sky. If at least one non-zero belief region is found, a seed region is determined from the regions having non-zero belief values. In addition, candidate sky regions are found. A polynomial model of sky is used by a classification stage for assigning a belief value to each candidate sky region. A candidate sky region is assigned a non-zero belief value if it is

found to be consistent with the seed region.

The algorithm has an intuitive explanation. Suppose an image is entirely blue sky (i.e. with no occluding objects). An image of sky is slowly varying and one would expect that the image would be well modeled with a low order two-dimensional polynomial. Furthermore, when objects are present in the image and occlusions prevent blue sky from being a single contiguous region, the sky regions can still be modeled in the same fashion. The algorithm begins with a region that is believed to be sky with high confidence (the seed region) and searches candidate sky regions for those with which a consistent sky model can be formed.

3. Initial Blue Sky Detection

The initial blue sky detection algorithm in [5] detects large clear blue sky regions in two stages. In the first stage, a multilayer neural network performs pixel classification based on color and texture features. The sRGB values of each pixel are used directly as the color features (the trained neural network performs the proper color transformation). The output of the pixel classification is a map of continuous "probability" values (not binary yet). Next, an image-dependent adaptive threshold is selected to obtain sky colored regions after connected component analysis. In the second stage, a sky signature validation algorithm is used to eliminate false positive regions. Finally, the algorithm determines a probability to reflect how well the region fits the model.

The initial belief map obtained with this method is shown in Figure 3. Three sky regions are correctly identified by the algorithm. The largest was assigned a belief value of 0.99, and the two smaller regions are assigned belief values of 0.30. Note that several sky regions in Figure 1 are completely missed in the initial belief map because these regions were too small to exhibit the expected gradient signal.

4. Identification of the Seed Region

From all of the non-zero belief regions, a single seed region is selected to avoid having conflicts (which may lead to false positives). The seed region is used to form an initial low order polynomial model of the sky. It is important that the seed region be both large in spatial extent, and have high confidence that it actually represents sky.

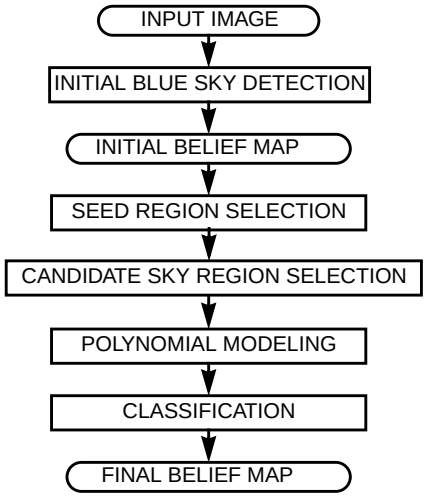

Figure 2: Flow diagram of the algorithm used for improved blue sky detection with polynomial modeling.

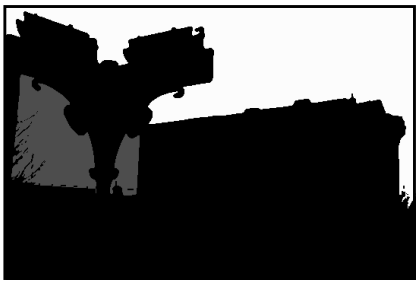

Figure 3: Initial blue sky belief map. Lighter regions indicate higher belief for that region.

For each non-zero belief region of the initial belief map, a score is computed. The score is proportional to the region size and the belief value. The region having the highest score is selected as the seed region. In the case of the belief map shown in Figure 2, the large region along the top and right side of the image is selected as the seed region.

5. Candidate Sky Region Selection

All sky colored regions from the initial blue sky detector (including regions initially rejected but excluding the seed region) are examined to find candidate sky regions. The idea is to find those regions that might be sky even though they were not classified as such in the initial blue sky classification.

A region is called candidate sky region if all of the following conditions are met. The region must be relatively free of texture variation. The region must meet a size requirement. Currently, the region must contain at least 10 pixels.

Figure 4 illustrates the 7 candidate sky regions that were detected for the image of Figure 1.

6. Polynomial Model Fitting

Two stages of model fitting are performed in the algorithm. First, a second-order two-dimensional polynomial is fitted to each color channel of the seed region. The polynomial represents a model of the sky over that high confidence sky region.

The polynomial model can be represented in matrix

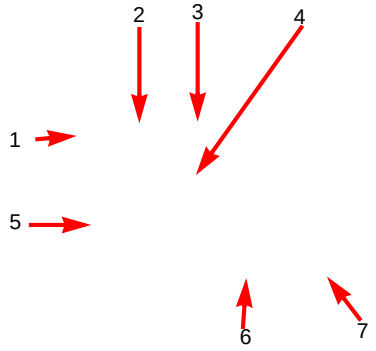

Figure 4: Seven candidate sky regions.

form:

$$\begin{bmatrix} \hat{r}(x,y) \\ \hat{g}(x,y) \\ \hat{b}(x,y) \end{bmatrix} = \begin{bmatrix} r_c^T \\ g_c^T \\ b_c^T \end{bmatrix} \begin{bmatrix} x^2 & xy & y^2 & x & y & 1 \end{bmatrix}^T \quad (1)$$

where:

$\hat{r}(x,y)$, $\hat{g}(x,y)$, and $\hat{b}(x,y)$ are the red, green and blue pixel value estimates of the digital image at position (x,y).

r_c, g_c, and b_c represent the column vectors of polynomial coefficients for the red, green, and blue color channels.

The least-squares polynomial fitting is achieved by forming the Vandermonde matrix V from the *N* pixels of the seed region.

For a two-dimensional second order polynomial, the Vandermonde matrix has *N* rows and 6 columns where each row corresponds to the position coordinates of one pixel:

$$V = \begin{bmatrix} x_0^2 & x_0 y_0 & y_0^2 & x_0 & y_0 & 1 \\ x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ f & f & f & f & f & f \\ x_{N-1}^2 & x_{N-1} y_{N-1} & y_{N-1}^2 & x_{N-1} & y_{N-1} & 1 \end{bmatrix} \quad (2)$$

Additionally, for each color channel, a vector A is defined with the actual pixel values from the digital image at the corresponding location:

$$A = \begin{bmatrix} c(x_0,y_0) \\ c(x_1,y_1) \\ f \\ c(x_{N-1},y_{N-1}) \end{bmatrix} \quad (3)$$

where $c(x,y)$ represents the value of a particular channel of the digital image at position (x,y). The least squares solution for the coefficients is:

$$c_c = (V^T V)^{-1} V^T A \quad (4)$$

The model error for each color channel is a measure of the goodness of fit between the polynomial model and the actual image data. The model error is determined by computing the square root of the mean squared difference between the vector A (actual pixel values from the digital image) and the vector Vc_c, which is the model estimate of pixel color for a particular channel. When the polynomial model perfectly describes the sky's image variation, the model error is due only to image noise. The model error in the red, green, and blue channels is represented as E_{1r}, E_{1g}, and E_{1b}, respectively. For the seed region in the example image, the values of E_{1r}, E_{1g}, and E_{1b} are 2.2, 1.4 and 0.9 respectively.

In the second model fitting stage, a second polynomial is fitted to the pixel values of both the seed region and a particular candidate sky region. This step is repeated for each of the candidate sky regions. The Vandermonde matrix is an *N+M* by 6 matrix formed from the *N* pixels of the seed region and the *M* pixels of the candidate sky region.

This dual-region polynomial model represents blue sky assuming that both the seed region and the candidate sky region are sky. If in fact the assumption is correct, one would expect that a polynomial model exists that would represent both regions well. If the assumption is incorrect and the candidate sky region does not in fact represent sky, then one would expect that an attempt to represent both regions with the low-order polynomial would result in a large amount of error. The model error is used to either confirm that the candidate sky region is in fact sky, or provide evidence that the candidate sky region is not sky.

The model error of this polynomial is measured only over the pixels belonging to the candidate sky region. We chose to do so because including pixels from the seed region often leads to an unrealistically low model error simply because the seed region is often much larger than the candidate sky region. This model error relates to the goodness of fit of the polynomial model to the candidate sky region. Again, the model error is computed for all color channels and designated as E_{2r}, E_{2g}, and E_{2b}, respectively.

Table 1: Algorithm results for image in Figure 1

Region	E_{2r}	E_{2g}	E_{2b}	Result	Correct?
1	3.0	2.3	1.6	promoted	yes
2	1.8	1.0	0.8	included	yes
3	2.2	1.2	1.0	included	yes
4	1.9	0.8	0.7	promoted	yes
5	1.6	0.8	0.8	included	yes
6	16.2	18.6	16.0	not included	yes
7	112	121	95	not included	yes

7. Classification

The candidate sky region is classified as sky if the following conditions are met:

1. The model errors for the candidate sky region must not be greater than T_1 (default $T_1 = 4$) times the model errors for the seed region. This condition ensures that the polynomial model is capable of fitting both the candidate sky region and the seed region with a similar model error. The threshold ensures that the texture in the seed region is not significantly different than the structure in the candidate sky region.

2. The model errors for the candidate sky region must not exceed a maximum error E_M (default $E_M = 9.5$). This ensures that the polynomial model fits the candidate sky region with a specific quality in an absolute sense. If both of the conditions are met, then the candidate sky region is clas-

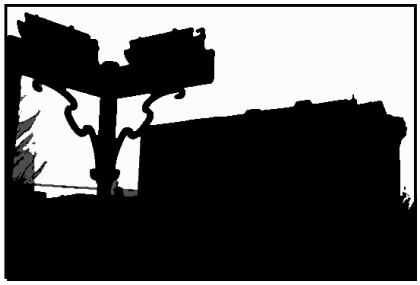

Figure 5: Final belief map computed by the algorithm.

sified as sky and assigned the belief value of the seed region.

Table 1 shows the model errors for the 7 candidate sky regions using the same region numbers as shown in Figure 4. The first five regions meet all of the thresholding requirements. The associated candidate sky regions are either added to the belief map or promoted (if the region has a lower but non-zero belief in the initial belief map) to have the belief value of the seed region. Regions 6 and 7 of the building wall are correctly left out.

8. Experimental Results

The algorithm has been applied to 83 images that each contained at least 1 region classified as blue sky by the initial sky detector. In total, the initial sky detector classified 104 regions as blue sky. Of these, 88 were actually blue sky and 16 were false positives. The polynomial fitting algorithm resulted in the classification of 39 more regions as blue sky. In actuality, 31 of these regions were actually blue sky and 8 were false positives. In addition, 6 regions were correctly promoted in their belief by the algorithm.

Figure 6 shows examples of the experimental results. The 3rd and 4th images contain 6 of the 8 total false positives. In both cases, the detected regions were actually reflections of sky. These regions are nearly uniform with almost no gradient, otherwise they would be rejected if exhibiting the opposite gradient signal of the seed region.

Figure 7 illustrates the criticality of detecting all sky regions. Sky regions are enhanced to a preferred saturated blue. When the initial sky map is used, a non-uniform and unacceptable image results. When the final belief map is used, the sky appears to be uniformly blue. Note that simply applying a color transform to all blue pixels would cause problems by undesirably modifying the text on the building.

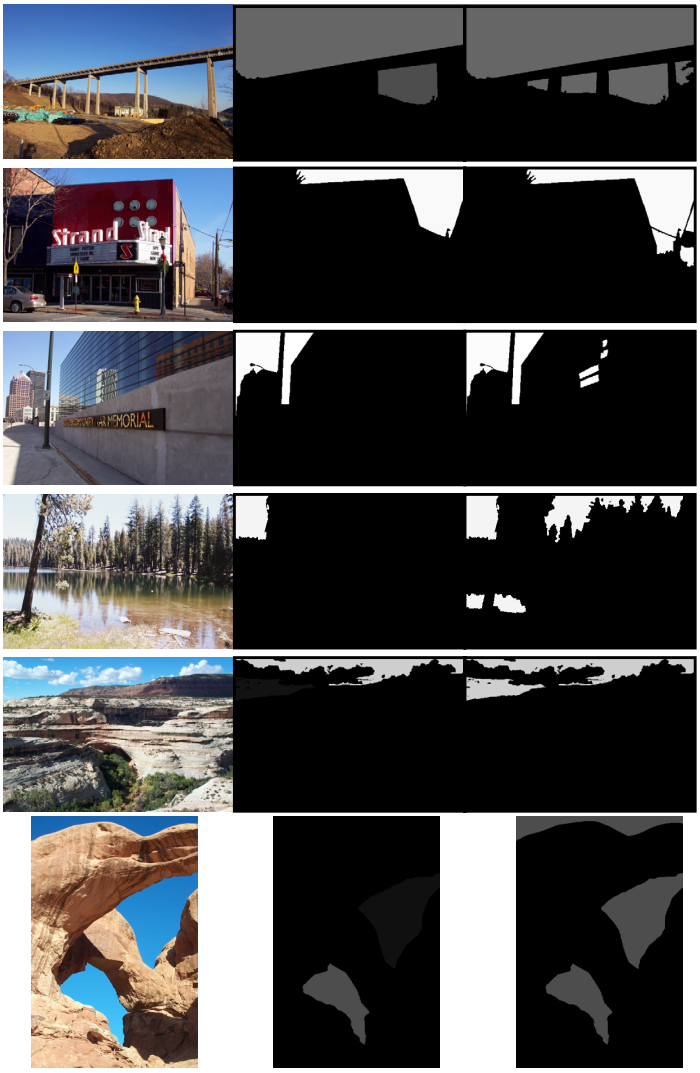

Figure 6: Algorithm results. **LEFT:** Original images. **MIDDLE:** Initial sky belief maps. **RIGHT:** Final belief maps from the algorithm.

Figure 7: Using the belief map to improve sky saturation. **LEFT:** Initial sky belief map produces a poor result. **RIGHT:** Sky appears uniformly blue using the algorithm's final belief map.

References

[1] M. Szummer and R. Picard, "Indoor-Outdoor Image Classification," Proc. IEEE Intl. Workshop on Content-based Access of Image and Video Database, 1998.

[2] A. Vailaya, A. Jain, and H. Zhang, "On Image Classification: City vs. Landscape," Proc. IEEE Intl. Workshop on Content-based Access of Image and Video Database, 1998.

[3] J. Smith and C. Li, "Decoding Image Semantics Using Composite Region Templates," Proc. IEEE Intl. Workshop on Content-based Access of Image and Video Database, 1998

[4] E. Saber, A. Tekelp, R. Eschbach, and K. Knox, "Automatic Image Annotation Using Adaptive Color Classification," CVGIP: Graphical Models and Image Proc., vol. 58, pp. 115-126, 1996.

[5] J. Luo and S. Etz, "A Physical Model-based Approach to Sky Detection in Photographic Images," IEEE Transactions on Image Processing, vol. 11, no. 3, pp. 201-212, 2002

[6] J. Warnick, R. Mehrotra, and R. Senn, "Method and System for Detection and Characterization of Open Spaces in Digital Images," U.S. Pat. No. 5,901,245, May 4, 1999.