Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, NV, October, 2003.

Color Snakes for Dynamic Lighting Conditions on Mobile Manipulation Platforms

Hanspeter Schaub
ORION Int. Technologies
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1003
hschaub@sandia.gov

Christopher E. Smith
Dept. of Electrical and Computer Engineering
MSC01 1100
1 University of New Mexico
Albuquerque, NM 87131-0001
chsmith@ece.unm.edu

Abstract

Statistical active contour models (aka statistical pressure snakes) have attractive properties for use in mobile manipulation platforms as both a method for use in visual servoing and as a natural component of a human-computer interface. Unfortunately, the constantly changing illumination expected in outdoor environments presents problems for statistical pressure snakes and for their image gradient-based predecessors. This paper introduces a new color-based variant of statistical pressure snakes that gives superior performance under dynamic lighting conditions and improves upon the previously published results of attempts to incorporate color imagery into active deformable models.

1. Introduction

Active contour models, also known as snakes, have experienced relative popularity in the computer vision community since their introduction. Various formulations and techniques have been investigated in an attempt to improve upon performance under a variety of conditions. In general, image gradient-based snakes [10] are used in the majority of this work with less than satisfying results.

Our work is motivated by ongoing research on using vision and visual servoing techniques to automate basic tasks on mobile manipulation platforms. Such platforms are used by public safety organizations worldwide for tasks such as Urban Search And Rescue (USAR), security and surveillance, explosive disposal, and inspection. The typical environments for these tasks present a host of difficulties for computer vision and image processing techniques. The most critical among these are dynamic lighting conditions, shadowing, and spectral variability (for example, neon vs. incandescent vs. sunlight).

Furthermore, safety organizations have uniformly found that pure teleoperation requires intensive training and quickly produces operator fatigue and errors. Any vision or visual servoing technique must be easy to initialize, intuitive to adjust, and provide reliable operation under the aforementioned conditions.

To address these problems and requirements, we have adapted previous work in statistical pressure snakes [1][8][11] to color imagery. Color has been used previously in snakes under several different methods. The most common is the use of edge detection in each color plane (the Red, Green, and Blue of RGB) that is then combined to give a color "gradient". A traditional Kass, et al. [10] snake is then placed into the resulting gradient image. Rather than a true color snake, this is simply an alternate formulation for the image gradient. These snakes inherit all of the problems associated with initialization and performance of gradient-based snakes without providing color selectivity.

We have taken previous work in statistical pressure snakes and adopted a color-based pressure term that uses Hue-Saturation-Value (HSV) color space. Our color pressure term matches non-uniformly in the three components of the color vector, unlike previous color gradient approaches. We use a narrow match in hue and broader statistical matches in saturation and value. The result is a snake that produces striking performance under dynamic lighting conditions and during shadowing. Furthermore, we can also use only hue and achieve performance over a greater dynamic lighting range; however, this also degrades performance in color discrimination for similar hues. The latter choice in color selection is particularly useful for targets with preselected color fiducial markers.

2. Previous Color Snakes

One way to address illumination and shadowing problems is to use a color matching model that allows pixels from a brightly illuminated region to statistically match pixels from a shadowed or poorly illuminated region. We turned to color models as a potential solution to the types of changing illumination observed in common mobile manipulation environments. These included illumination intensity changes, shadowing, and illumination spectra changes. Many prior efforts to develop color snakes basically fall into one large group of similar techniques, with a small number of notable exceptions.

Most prior efforts use a standard color cube model with RGB bitplanes to produce an image gradient (typically edge detection) in each bitplane. The results were then combined to form a single "color gradient" image that was then used to drive a traditional Kass, et al. snake. Clearly, illumination changes will affect the strength of the gradient and sharp shadows will produce strong gradients that are not related to the target occluding contour. A variation of this method uses three snakes, one for each bitplane, and then attempts to merge the results into a unified snake. These particular forms of color snakes do not address our design requirements and were therefore rejected.

A different method that used RGB looked at color angular maps [4]. Edge detection was then used on the angular maps to produce a gradient image. Again, problems with initialization required the use of techniques to improve the robustness of the snakes. This effort also failed to address changing illumination and shadowing.

Sapiro [13] used edge detection in a CIE (Comission Internationale de l'Eclairage) standard color space (1976 L*a*b*, or simply CIELAB) to drive diffusion for snake image energy. This is a color opposition model where color is defined on two axes with (a and b) where red is a, green is $-a$, yellow is b, and blue is $-b$. This provides similar information to HSV, where the opposing colors yellow and blue have angles π radians apart and the opposing colors green and red are $2\pi/3$ radians apart.

An alternate method by Gevers, et al. [6] used a color gradient method containing RGB, normalized RGB, intensity, and hue to determine gradient in an attempt to provide color constancy under varying illumination. The method was supported by limited examples with saturated colors and very little true illumination change. Again, the use of RGB-based color models called into question the ability to determine color similarity; however the work did use hue as a partial measure of color gradient.

The work of Zhu and Yuille [14] was based upon region competition and an attempt to base segmentation on albedo. Again, limited results using saturated colors were presented where the true changes in the illuminant were negligible. The appeal of this work is the idea of removing the illuminant and growing regions in multiband data using albedo. When possible, removing the effects of illumination should improve performance; however, it would also appear that this method would be applicable under only a narrow dynamic range of illumination changes and would also become intractable in cases of multiple light sources and interrflections. Other work has investigated the color constancy problem [2][5][9], but these often have strong prior assumptions (e.g. gray-world, smoothly varying illumination, etc.) or are computationally expensive While possibly offering better color matching, use of these techniques would overly constrain our environment.

Zhu and Yuille also point out that balloon snakes [3] and region growing [7] are really variations on the same solution. Since our prior experience with statistical pressure snakes has demonstrated excellent performance in terms of segmentation and tracking bandwidth, we selected these as our underlying model.

3. Color Pressure Snakes

3.1 The Basic Pressure Snake

The traditional deformable model was first proposed by Kass et al. [10] and is a parametric curve of the form

$$S(u) = I(x(u), y(u))', u = [0, 1]. \quad (1)$$

This curve is placed onto a potential field (typically image gradient) derived from the image and allowed to change shape and position, minimizing the energy along the length of the curve. The energy function is defined as:

$$E = \int_0^1 E_{Snake}(S(u))du \quad (2)$$

that includes internal, image, and constraint energies.

$$E = \int_0^1 E_{int}(S(u)) + E_{img}(S(u)) + E_{con}(S(u))du \quad (3)$$

where the internal energy is given by

$$E_{int} = \frac{\alpha}{2}\left|\frac{\partial}{\partial u}S(u)\right|^2 du + \frac{\beta}{2}\left|\frac{\partial^2}{\partial u^2}S(u)\right|^2 du. \quad (4)$$

In the closed contour case ($S(0) = S(1)$), we obtain

$$E = \frac{\alpha}{2}\oint\left|\frac{\partial}{\partial u}S(u)\right|^2 du + \frac{\beta}{2}\oint\left|\frac{\partial^2}{\partial u^2}S(u)\right|^2 du + \oint P(I)du \quad (5)$$

where α and β are weights. P is the potential induced by the image values.

We replaced E_{int} with a single term that maintains a constant third derivative (i.e. a zero fourth derivative) to more accurately reflect the original motivations for active deformable models [11]. The implementation of this energy term includes an even spacing constraint required for discrete derivatives, making the tension term in (3) redundant. We also replaced P with a statistical pressure term

$$E_{Pressure} = \rho\left(\frac{\partial S}{\partial u}\right)^{\perp}(\varepsilon - 1) \quad (6)$$

where

$$\varepsilon = \frac{|I(S) - \mu|}{k\sigma} \quad (7)$$

and is based upon the mean μ and standard deviation σ of pixel values from a seed region or target model appearance [8][11][12]. The k factor determines the spread of acceptance on pixels values (e.g. 1.0σ, 2.3σ, 3.25σ, etc.) The pressure is applied perpendicular to the derivative of the snake curve S and is weighted by ρ. The seed region or target appearance identifies positive vs. negative pressure regions. In other words, image regions that are statistically similar to the seed region yield positive pressure while image regions that are some number of standard deviations away from the seed yield negative pressure. When a portion of the contour is in a positive region, it expands. When a portion of the contour is in a negative region, it contracts. It follows that the minimum energy of the contour lies on the pressure boundary between the positive and negative regions.

This work can be viewed as the application of statistical segmentation or region growing to the snake image force. The advantage is that image pressure is defined across the extent of the image, whereas image gradient is a highly localized feature. Pressure snakes therefore converge faster and have very loose constraint upon initial placement.

3.2 Statistical HSV Color Pressure

The use of greyscale pixel values in the snake pressure term produces significant sensitivity to illumination changes. Intuitively, as illumination increases, the object luminance increases accordingly, resulting in a shift of greyscale values toward white (camera saturation). As illumination decreases, the greyscale values shift toward black (minimum camera lux).

In order to track objects in greyscale across illumination changes, the value k can be expanded, automatic gain control can be used, or post processing techniques such as histogram stretching or equalization can be utilized. Each of these methods has significant problems when used to achieve constant object pixel values. Each of these methods also fail to handle partial shadowing where an object of constant surface color may appear to have many different shades.

Color information provides a better statistical measure across these changes. The pressure force calculation only requires a scalar measure indicating how well the current pixel values match those of the target. To make the snake routine work with a full color image, the target is specified in terms of all three color channels. Here the target error signal is defined as

$$\varepsilon = \sqrt{\left(\frac{p_1 - \tau_1}{k_1\sigma_1}\right)^2 + \left(\frac{p_2 - \tau_2}{k_2\sigma_2}\right)^2 + \left(\frac{p_3 - \tau_3}{k_3\sigma_3}\right)^2} \tag{8}$$

where p_i are local average pixel color channel values, τ_i are the target color channel values, and σ_i are the target color channel standard deviations. The change to the pressure force subroutine code was minimal to incorporate a full three color channel image target specification.

While the RGB color space is routinely used when manipulating and printing images, it isn't ideal for the snake algorithm code. The goal is to have the target be identified under a range of lighting conditions. This means that if the target is specific type of red, for example, then we would like to register all shades of this red as the target color. It is possible to compute such a set of target colors in RGB space, but it is a very complicated function. Instead, it is convenient to map the image color channels from RGB space to Hue-Saturation-Value (HSV) color space. A similar color space is the Hue-Saturation-Intensity (HSI) color space. In the HSV color space, the hue determines the general color of the pixel (think of the color selection in a rainbow), the saturation determines how rich or washed out the color is, and the value determines the brightness or darkness of the color.

The HSV color space is illustrated in Figure 1. Assuming the hue begins with blue at 0, the hue values cycle through all possible colors until they return to blue at 360 degrees. For example, a pure green color would have a hue value of 120, while a pure red would have a hue of 240. If the saturation value S is 0, then the value V control the gray scale of the pixel. For gray scale images, the hue value is ill-defined.

Reconsider Equation (8). To achieve a hue-only snake, the weights k_2 and k_3 are set to maximum value (functional infinity). As these factors approach infinity, their respective terms approach zero and the target error becomes

$$\varepsilon = \left|\frac{p_1 - \tau_1}{k_1\sigma_1}\right|. \tag{9}$$

Equivalently, if one desires the original greyscale pressure snakes, k_1 and k_2 should be set to maximum value (functional infinity). This is useful, for instance, when the saturation values of the target region are extremely low (approaching black), extremely high (approaching white), or when the saturation corresponds to RGB colors where $R = G = B$ (i.e. greyscale).

Let us define our target color in the HSV color space. If we specify the target to only have a specific hue value (with associated standard deviation), then the target error function will return 0 if the current image pixels are of this hue type, regardless of how light or dark, saturated or washed out the current pixel color is. All shades and variations of a color will be accepted. If the target is surrounded by a color with a clearly different hue (for example a blue box surrounded by a red frame), then the snake will be

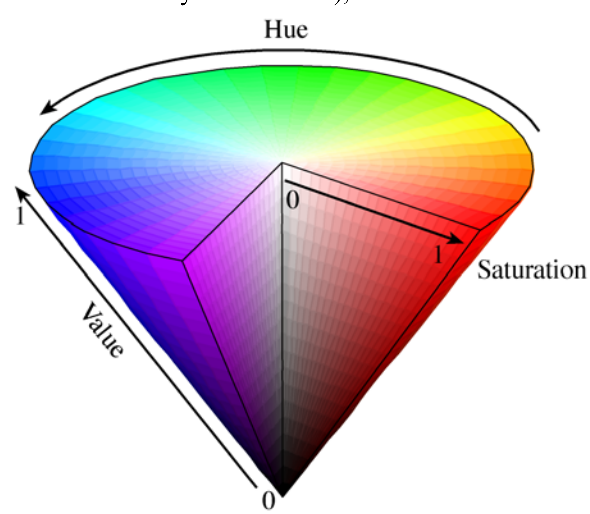

Figure 1 HSV Color Space

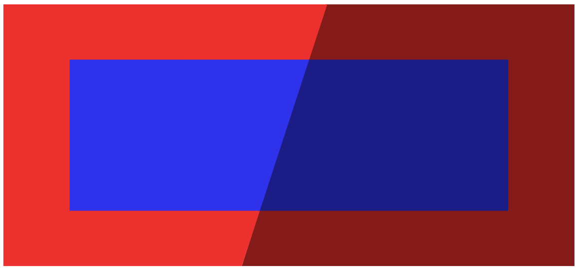

Figure 2 Shadowed boxes

able to easily track the blue box, even though the target might have drastically different lighting conditions across its surface. This idea is illustrated in Figure 2. Similarly, if portions of the box are more brightly lit, then these blue zones with lower saturation value would also still be registered as the correct color. Such robustness to lighting changes between images, as well as across images, is not possible with classical gray-scale images. By moving to the HSV space, it is possible to make the snake routine a lot more robust to lighting changes as would be experienced in a typical unstructured outdoor environment (shadows being cast across object, clouds move in, etc.).

If the snake is to outline a general target, however, then only looking at the hue value to determine the error signal ε to the target color could be to forgiving. For example, consider that a blue suit-case is to be acquired. If only the hue value of a pixel is considered to generate ε, then a very white pixel, with only a slight hint of blue, would still be considered to be the target color. However, in practice this very light or dark color is often the color of another object in the scene. While the hue proximity is typically clipped with a gain of $k_1 = 2$, it is found that for general target tracking the saturation values S and intensity values V should also be clipped with a gain of about $k_{2,3} = 6$. Thus, we are still weighting mostly on the hue parameter, but also penalizing for drastic departures from the target saturation and intensity values.

Using this HSV color space target selection scheme will have issues when the target color is near white or near black. For example, assume the target is nearly white, but has a slight shade of red to it. In this case it is possible that a different portion of the target might have a slightly blue shade (from a different light source, reflection off another target, etc.). If the target is very dark (near black), our method continues to work reasonably well with the same choice of gains. Principally, this is due to dark objects having rather large standard deviations, desensitizing the calculation of ε to hue since σ_1 is large. The result is that our algorithm automatically reverts to using mostly value instead of hue for dark objects.

4. Experimental Results

We have implemented our color snakes using the OpenCV libraries on both laptops and workstations running Microsoft Windows 2000 and XP. The snakes use DirectX registered video sources and a TCL/Tk interface to allow intuitive user input and interaction (a mouse click on the target initializes the snake, and all parameters can be adjusted while the snake is running via input sliders). We have tested our models on both indoor and outdoor imagery to validate the performance under the types of conditions mobile manipulation platforms operate.

The first set of experiments involve indoor scenes using both fluorescent and natural lighting. These were performed on an 3 Ghz Pentium 4 workstation with a Pyro 1394 Firewire WebCam. Performance in a structured office environment (see Figure 3) was excellent with respect to segmentation and speed (frame-rate), with the system consistently tracking objects during illumination changes (only ambient light from a window, one set of fluorescent lights on, both sets of fluorescent lights on). The snake is shown, as well as the major and minor axes of the snake. These experiments used $k_1 = 2$, $k_2 = 6.0$, and $k_3 = 8.0$.

A second set of indoor experiments was conducted with a less powerful laptop (700 Mhz P-III) using a USB 1.0 Labtech Webcam. This camera's color and clarity are severely degraded from the 1394 camera used in the previous experiments. We have included four images from an experiment involving incandescent lighting on a dimmer switch. To characterize performance of the method only, we shutoff the camera's Automatic Gain Control (AGC). We have observed that AGC can compensate for a large range of illumination changes and the final system will no doubt use cameras with AGC and auto-white balance.

In the first image, the lighting is set to a medium level and the snake is initialized (a simple mouse click on the region of interest). Figure 4 shows the lighting at initialization. The target is a multi-colored patchwork quilt, with two adjacent blocks that can be confused under low-light

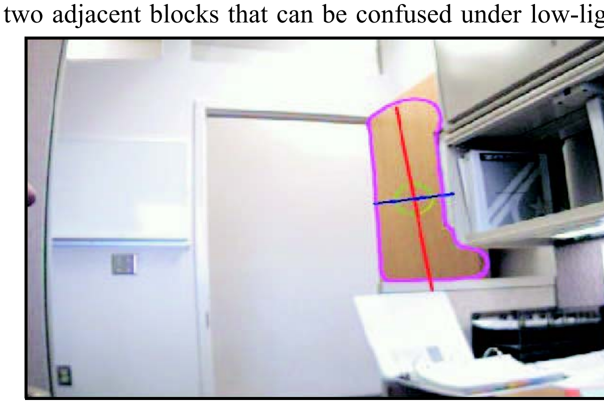

Figure 3 Sample indoor experiment

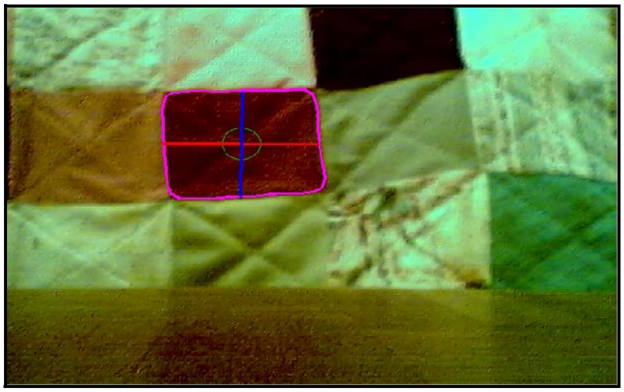

Figure 4 Medium lighting conditions

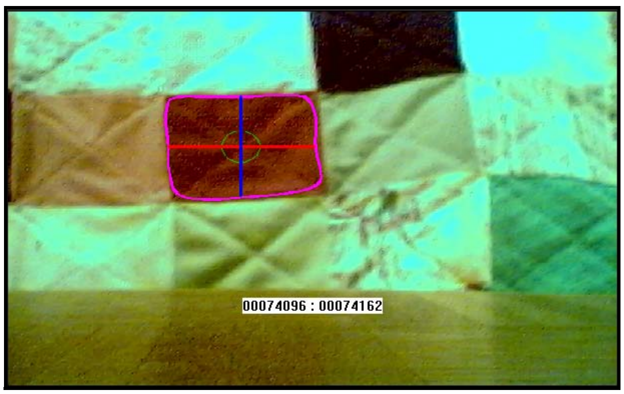

Figure 5 Brighter lighting conditions

conditions. As lighting increases (see Figure 5) the snake continues to segment the selected block correctly. The same is true for reduced illumination conditions (see Figure 6). Finally, in Figure 7, the increasingly dark lighting conditions have led to an inability to differentiate between the adjacent blocks. Indeed, the hue of the non-target block is virtually identical to the target block's hue under the initial lighting conditions. The target block is still recognized as belonging to the initial identified color of interest, even though the color is approaching black.

It is worth noting that in this set of experiments, the target was selected to give less than ideal regions. Earlier color snake research used predominantly imagery with highly saturated, flat matte targets. The blocks of the target quilt provide local variations due to shadowing and contain some specular components due to the material used (a polyester blend). Throughout the experiment, the snake parameters were left at $k_1 = 1.5$, $k_2 = 6.0$, and $k_3 = 8.0$.

A third set of experiments was conducted under real conditions. For these experiments, footage was captured from a Sony camcorder using a laptop and a Dazzle DVC-50 USB capture device. This device only supports quarter sized images (320x240) and results in the snake and axes graphics appearing thicker.

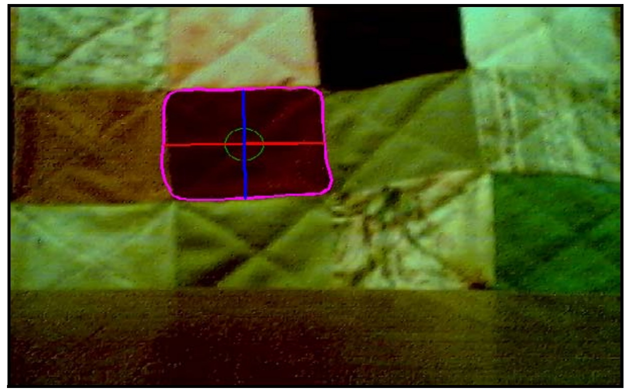

Figure 6 Darker lighting conditions

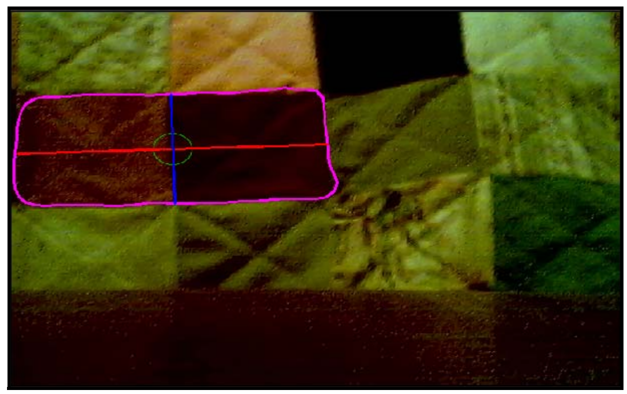

Figure 7 Segmentation bleeds to adjacent block

In this set, an unattended piece of luggage is sitting partially in the shadow of a vehicle. Three images (Figure 8 through Figure 10) are shown as the camera view changes upon approach to the suitcase. The snake parameter settings were the same as the second set of experiments; however the camcorder uses AGC and auto white balance.

The performance and stability of our new color snakes allow their use for supervised and autonomous guidance of mobile manipulation platforms by using established visual servoing approaches. The result will be robots that can be easily integrated into public safety teams without extensive training and exhaustive teleoperation.

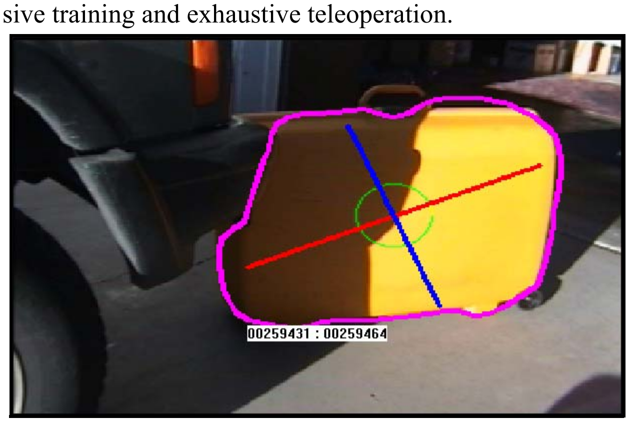

Figure 8 Suitcase with mixed lighting

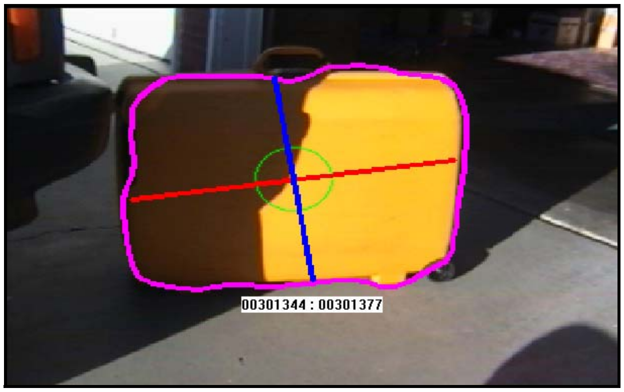

Figure 9 Suitcase with mixed lighting

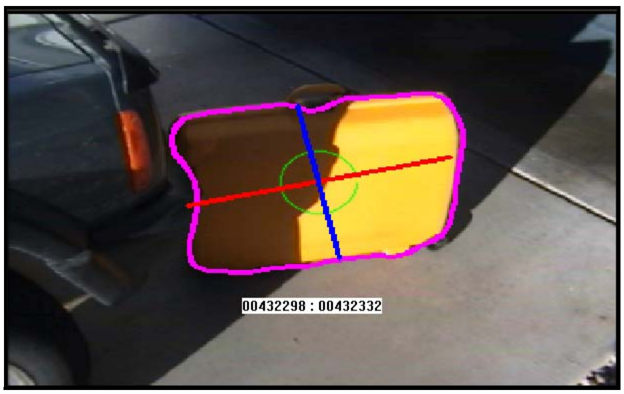

Figure 10 Suitcase with mixed lighting

We have placed color figures from this paper at http://www.ece.unm.edu/~chsmith/IROS2003 since color is crucial to the presentation of this work.

5. Conclusion

The standard statistical pressure snake routine was modified to function on a color image in HSV color space. If the target color is only specified in terms of hue, then large variations in target lighting and shading are permitted. This provides a significant increase in the robustness of tracking a specific target in an unstructured, outdoor environment. One requirement here is that the target is surrounded by a frame of a clearly different hue to avoid the snake spilling over to other image objects with similar hues. By specifying the target color to contain hue, saturation and intensity values, it is possible to a establish reasonably robust method to track general image features of a common color. Here it is important to penalize more strictly hue departures from the target color and only mildly penalize the saturation and intensity departures.

Our current plan is to incorporate more powerful pattern recognition techniques that we have previously applied to greyscale snakes [1]. We also plan to investigate other measures of color constancy, including uniform color spaces and albedo.

6. Acknowledgements

This material is based upon work supported by the Department of Electrical and Computer Engineering at the University of New Mexico, and by Sandia National Laboratories under the Sandia-University Research Program (SURP).

7. References

[1] Abd-Almageed, W. and Smith, C. "A pattern recognition approach to active deformable models," to appear, *International Journal of Image and Graphics*, 2003.

[2] Brainard, D. and Freeman, W., "Bayesian color constancy," *Journal of the Optical Society of America: Part A*, Vol. 14, pp. 1393-1411, 1997.

[3] Cohen, L. and Note, D., "On active contour models and balloons," *CVGIP (Image Understanding)*, Vol. 53, No. 2, pp. 211-218, 1991.

[4] Dumitras, A. and Venetsanopoulos, A., "Angular map-driven snakes with application to object shape description in color images," *IEEE Transactions on Image Processing*, Vol. 10, No. 12, 2001.

[5] Finlayson, G., Schiele, B., and Crowley, J., "Comprehensive colour image normalization", *Proceedings of ECCV'98 Fifth European Conference on Computer Vision*, 1998.

[6] Gevers, T., Ghebreab, S., and Smeulders, A., "Color invariant snakes," *Proceedings of the Ninth British Machine Vision Conference*, 1998.

[7] Horn, B., *Robot Vision*, MIT Press, Cambridge, MA, 1986.

[8] Ivins, J. and Porrill, J., "Active region models for segmenting medical images," *Proceedings of the 1st International Conference on Image Processing*, pp. 227-231, Austin, TX, 1994.

[9] Jobson, D., Rahman, Z., and Woodell, G., "A multi-scale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing*, pp. 965 - 976, 1997.

[10] Kass, M., Witkin, A., and Terzopoulos, D., "Snakes: active contour models," *International Journal of Computer Vision*, Vol. 1, No. 4, pp. 321-331, 1987.

[11] Perrin, D. and Smith, C. "Rethinking classical internal forces for active contour models," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, 2001

[12] Ronfard R. "Region-based strategies for active contour models" *International Journal of Computer Vision*, Vol. 13, No. 2, pp. 229-251, Oct. 1994.

[13] Sapiro, G., "Color snakes," Computer Vision and Image Understanding, Vol. 68, No. 2, 1997.

[14] Zhu, S. and Yuille, A., "Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.18, No.9, 1996