Accepted to 10th Congress of the International Colour Association, AIC (Association Internationale de la Couleur) Colour 2005. Granada, Spain. In press.

Rapid and Robust Algorithms for Detecting Colour Targets

J. Coughlan, R. Manduchi*, M. Mutsuzaki* and H. Shen
Smith-Kettlewell Eye Research Institute, San Francisco, CA
****Dept. of Computer Engineering, Univ. California Santa Cruz, Santa Cruz, CA***

Introduction

Colour is a distinctive feature that can vastly simplify visual search for objects. However, some form of colour constancy is needed to cope with unknown lighting conditions. Unfortunately, current computer vision colour constancy algorithms[1] generally make restrictive assumptions about illumination conditions and/or require significant computational resources. We propose a computer vision algorithm for the rapid and robust detection and localization of specially designed colour targets. The algorithm draws on properties of colour gradients which are invariant to a wide range of illumination conditions, without needing to implement full colour constancy. It is intended to be used for an environmental labelling system for helping visually impaired people navigate their environment. Since the algorithm needs to run in real-time on an inexpensive handheld PDA/cell phone with colour camera (see Figure 1a), which has limited computational power, algorithmic efficiency is of paramount importance.

Our algorithm for detecting colour targets exploits the fact that, although the apparent target colours depend strongly on lighting conditions, colour gradients within them vary much less. Moreover, the spatial and spectral properties of the targets make them easy to distinguish from typical background clutter. Colour targets are located in an image through a series of simple colour gradient tests. Targets may be placed adjacent to signs or barcodes, which can be rapidly read given the location of the target (and conveyed to the user by synthetic speech). The targets are small and can be detected from a distance; the maximum range is proportional to the size of the target and is a few meters from the camera using a 4 cm diameter target. Detection takes 0.12-0.5 sec per image in our prototype system, demonstrating its feasibility for real-time use.

Method

While a variety of OCR (optical character recognition) techniques are effective for reading scanned-in text, they fail when the text appears in natural, cluttered scenes because of their inability to localize the text in the image before reading it. Recently developed algorithms address text localization[2], but they currently require more CPU power than is available in an inexpensive, portable unit. Barcodes suffer from a similar limitation in that they must be localized -- typically by hand -- before they can be read.

We have devised a solution to this problem by designing special colour targets to be placed near barcodes or text of interest. These targets are designed with a distinctive combination of colours arranged in a particular configuration (see Figure 1b). A series of simple and very rapid tests performed on an input image (Figure 1c) will quickly detect and localize the colour targets (Figure 1d). The tests exploit invariants based on colour gradients that we have derived empirically under a variety of indoor and outdoor lighting conditions for our colour pattern. While the precise colour gradients among the three colour patches vary depending on illumination and noise, some aspects of the gradient are highly predictable and create a nearly unique signature of the target.

Specifically, four sequential tests suffice to rule out all but a small fraction of the image pixels that do not lie on a colour target. Given any candidate target center with pixel coordinates (x,y), three points p_1, p_2 and p_3 are computed with fixed offsets relative to (x,y), corresponding to nearby, equidistant locations in the red, green and blue patches, respectively. The four tests are the following inequalities: $R_1 > R_2 + T$, $G_1 + T < G_2$, $G_2 > G_3 + T$ and $R_1 > R_3 + T$. Here the subscripts refer to the RGB intensity channels (ranging from 0-255) at points p_1, p_2 and p_3, and T is a threshold (empirically chosen to be 20). The tests are applied in a loop that scans the entire image to test each possible candidate center location. Testing of a location ends as soon as any of the inequalities fail. The overwhelming majority of pixel locations fail the first or second tests, which means most of the image requires very little processing. Locations that pass all four tests are classified as on a target.

Figure 1: (a) PDA with camera worn by visually impaired user. (b) Colour target consists of three colour patches. (c) Target shown in cluttered environment. (d) Detection result from (c) indicated by red square.

We implemented the algorithm on a Dell Axim X3 (624 MHz CPU) with Veo Photo Traveler 130S 1.3 megapixel camera. The algorithm detects multiple targets in about 0.12 sec – 0.5 sec (depending on camera resolution). From our preliminary tests we estimate a detection rate of 95%. The current false positive rate is non-negligible (about one false positive target per every 4 images), but we plan to add additional colour target tests to reject the false positives (and simultaneously allow us to increase the detection rate). The additional tests will only be applied to candidate locations which pass the first four tests, and hence will slow the algorithm by only a negligible amount. The detection is invariant over a range of scales (from about 0.5 m to a few meters), and accommodates significant rotations (up to about 30 degrees in the camera plane) and slant.

Results and Conclusions

Our experiments demonstrate the feasibility of our proposed approach, which allows us to implement an inexpensive (under $300), real-time environmental labelling system that requires no infrastructure beyond simple colour targets (and bar codes) that can be printed on a laser printer.

Basing the search algorithm on the identification of specific colour gradient features, rather than less distinctive grayscale features such as edges, is crucial to the efficiency of our algorithm. Past research in computer vision and image processing has used colour for such purposes as skin detection[3] and colour gradients for certain tasks such as contour fitting[4] and edge detection[5]. However, to our knowledge our approach is the first application of colour gradients for rapid detection and localization of specific targets.

References

1. D. Forsyth and J. Ponce. Computer Vision: A Modern Approach. Prentice Hall. 2002.
2. X. Chen and A.L. Yuille. “Detecting and Reading Text in Natural Scenes,” IEEE Conf. Computer Vision and Pattern Recognition. Washington, DC. June 2004.
3. M. Jones and J. Rehg. “Statistical Color Models with Application to Skin Detection.” International Journal of Computer Vision. Vol. 46 , Issue 1, pp 81 - 96 . January 2002.
4. T. Gevers, S. Ghebreab and A. Smeulders. “Colour Invariant Snakes,” British Machine Vision Conference. 1998.
5. J. Geusebroek , A. Dev , R. Boomgaard , A. Smeulders , F. Cornelissen , H. Geerts, Color Invariant Edge Detection, Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision, p.459-464, September 26-27, 1999.