Red is the new black* - Or is it?

David Austin and Nick Barnes
Autonomous Systems and Sensing Technologies Programme
National ICT Australia, Locked Bag 8001
Canberra, ACT 2601.
david@syseng.anu.edu.au, nick.barnes@nicta.com.au

Abstract

A common assumption in the robotics community is that objects have a constant colour and, hence, can be recognised by their colour. Humans have a powerful vision system that normalises away colour shifts and causes objects to appear to have constant colour. Unfortunately, there is no such system for digital imaging devices. Experiments are performed with a digital camera that demonstrate that typical lighting conditions cause large, non-linear colour shifts. We conclude with a discussion of the problem and a suggested approach to a constant colour system for robots.

1 Introduction

When an object appears in front of a colour CCD camera, the resultant image largely depends on three things: the reflectance properties of the object, the imaging properties of the CCD, and the incident light on the object. In the typical case, the underlying object has the same reflectance properties over a series of images, and for a good quality camera in typical conditions, the CCD imaging properties remain reasonably constant. However, the same cannot generally be said of lighting.

Colour constancy in humans is the constancy of the perceived colours of surfaces under changes in the intensity and spectral composition of the illumination [Foster *et al.*, 1997]. With human colour constancy and the unchanging nature of the objects being viewed, many researchers assume that an object in the world has a colour, and this colour can be used for recognition. This is a tempting idea for robotics. Colour-based classification can be performed extremely quickly, making it highly suitable for in the loop navigation, while other properties of objects, such as shape, can only be recovered with more computational expense. Current vision-based robot competitions may give the false impression that colour is a simple cue. However, Robocup, and much of the current colour-based robotics work engineers the problem away by keeping lighting constant. In industrial vision this can be performed by constant temperature light boxes, in Robocup by bright controlled lighting and careful calibration. Another common approach is to perform robot trials with some natural lighting, but over short time spans so that light does not change too greatly.

Clearly, such assumptions are ultimately unworkable if robot vision systems are required to perform in environments with sunlight over extended periods. To achieve this we require robot colour constancy. Obviously this can be achieved by continual manual calibration, or by including a template in every image (such as a colour chart), but this is not generally acceptable. This paper addresses the problem of what is required for autonomous robot colour constancy, without modifications to the environment. We do not attempt to solve the problem at this stage, but to define the problem in a precise way, and show that it cannot be taken for granted, or to be assumed a simple problem.

2 Human colour constancy

It is easy to assume that digital imaging devices deliver constant colours in a similar way to humans[Foster *et al.*, 1997]. It is therefore useful to present a couple of examples of human colour constancy at work, to demonstrate the weakness of the assumption of constant colours.

Firstly, consider the scene shown in Figure 1. This is a typical office corridor. We all know from personal experience that the colours of the walls appear constant as we walk along a corridor. However, it is clear from Figure 1 that the colours vary significantly. The powerful human colour constancy is normalising the colours without our being aware of it.

As a second example, consider the two short videos included with this paper. Video 1 shows a time lapse video of an office as the sun sets. Video 2 shows the same scene, but has been altered. If the reader looks at

[1]With appologies to Rei Kawakubo.

Figure 1: Simple scene with constant colours. Note the constant colour of the carpet, walls, ceiling and even the nearby door!

the videos and takes to guess what has been altered, we predict that the guess will be wrong.

In Video 2, it appears that the colour of the door gets lighter. In fact, the colour of the door remains exactly constant (the door pixels from the first image have been inserted in the following images). This is a demonstration of human colour constancy at work: our vision system sees that the whole scene is changing colour and normalises for this change. Removing the colour change in the door leads to the appearance that the colour of the door is changing.

3 Prior work

The most basic approach to colour-based identification is linear thresholding of colour channels in some colour space (e.g., [Bruce *et al.*, 2000]). This is problematic as rectangular colour subspaces are not sufficient to describe the appearance of typical single colour objects. Alternative approaches such as decision trees allow arbitrary concave colour classes. Alternatively, a 3D lookup table across colour space can be employed to allow arbitrarily shaped clusters at the cost of storage space.

One approach to colour constancy is to take a colour distribution at some initial time, and map colour regions back to their initial distribution. Without *a priori* knowledge, this can only be performed in colour space. Austermeier *et al.* [Austermeier *et al.*, 1996] exemplify this approach by using self-organising feature maps to take a cloud of colour points in three space under the initial illumination, then look for a similar cloud in a second image that minimises some distance metric in mapping back to the original cluster. Such a technique assumes small moves in colour space of colour clusters, otherwise it will be difficult to map back to the starting cluster. Colour shifts over a series of images where the total distance moved in illumination space is large could be handled by a series of transformations, each of which is small. The problem here is that colours will drift without bound. Given that in some images, some pixels are sometimes incorrectly identified as being of a particular colour, this error will be propagated and eventually is likely to lead to the cluster drifting from the truth.

The standard alternative approach is to use some form of calibration chart (e.g., [Legenstein *et al.*, 2000]), which is not acceptable. The approach that we have taken previously is to take advantage of known geometry of objects in the scene, and take actions of the robot to increase the probability of viewing these objects. This can be used to ground the drift in tracking colour for particular objects, leading to robot colour constancy [Cameron and Barnes, 2003]. However, such research does not tell us what minimal models are adequate for colour transitions. It also only tracks the colour of a set of objects, but without some understanding of the change to colour space over time, this gives us no additional information about regions of colour space that are not included within the known set of objects.

4 Models of colour imaging

The visible spectrum is a large continuous space and so representing the whole spectrum is prohibitive. Human vision is generally accepted as having three main colour reception channels, and most standard cameras have three also. There is no clear reason why we should assume that three discrete channels are sufficient to encode all the information present. However, given that most CCD devices encode only three, we may take it as a definition that we are representing a discrete three dimensional space of colour. There are a plethora of encoding schemes that are used for the three channels. Some of these schemes, including YUV and HSV encode light intensity in a single dimension, and colour across the other two. We now examine a number of possible models of the colour and intensity change that is possible in viewing a static scene with an unchanging camera, given changes in the apparent lighting.

We may say that the camera is a function that, for each pixel, measures the incoming light. We further assume that the incoming light is generated by reflection from some object along the ray of the pixel. The reflected light is a function of the object's reflectivity properties and the incident light. The result is that the camera measures a three-dimensional colour value:

$$
\begin{aligned}
L_h &= f_h(I(\lambda), R(\lambda)) && (1)\\
L_s &= f_s(I(\lambda), R(\lambda)) && (2)\\
L_v &= f_v(I(\lambda), R(\lambda)), && (3)
\end{aligned}
$$

where L_h, L_s and L_v are the hue, saturation and brightness values respectively, and I and R are the incident light and reflectivity properties of the object, both continuous functions of wavelength. Note that for a camera, without *a priori* knowledge, we know nothing about I, R, or the functions f_x, so clearly the problem is underdetermined. Let us assume that we know that a single pixel corresponds to a particular object, and that we know R for that object. To achieve colour constancy, we must have a mapping that reverses the effects of changing I.

4.1 Constant colour

The simplest model is that for all incident light, the object will appear the same colour:

$$
\begin{aligned}
L_h &= f_h(R(\lambda)) && (4)\\
L_s &= f_s(R(\lambda)) && (5)\\
L_v &= \alpha_v \int_\lambda I(\lambda) d\lambda. && (6)
\end{aligned}
$$

This model assumes that colour is constant regardless of lighting, and that all that is necessary to recover the colour of an object is calibration of the functions f_h and f_s. This corresponds to a one dimensional look-up table for each colour channel. Clearly this model is violated by zero light conditions when nothing is visible. However, it is reasonable to assume that the incident light is within the dynamic range of the camera.

4.2 Linear in intensity

A more general model would be that the apparent colour varies linearly with the intensity and the reflective properties of the object:

$$
\begin{aligned}
L_h &= f_h(R(\lambda)) + \alpha_h \int_\lambda I(\lambda) d\lambda && (7)\\
L_s &= f_s(R(\lambda)) + \alpha_s \int_\lambda I(\lambda) d\lambda && (8)\\
L_v &= \alpha_v \int_\lambda I(\lambda) d\lambda, && (9)
\end{aligned}
$$

where $I(\lambda)$ is the intensity of the incident light, and α_r is a constant.

In this model, as the intensity of the incident light varies, the apparent colour will change linearly (because, in general, α_r will be different to α_g and α_b). In this case, one simply needs to calibrate the α_x for the camera, then for any image, recover the incident light intensity and colour constancy can be achieved.

4.3 The awful truth about CCD sensor responses

Colour varies non-linearly with both incident intensity and incident wavelength, in an interdependent manner. One can express the RGB values as[Thomson and Westland, 2001]:

$$
\begin{aligned}
L_r &= \int_\lambda R(\lambda) I(\lambda) S_r(\lambda) d\lambda && (10)\\
L_g &= \int_\lambda R(\lambda) I(\lambda) S_g(\lambda) d\lambda && (11)\\
L_b &= \int_\lambda R(\lambda) I(\lambda) S_b(\lambda) d\lambda, && (12)
\end{aligned}
$$

where I and R are again the incident light and reflectivity properties of the object and S_r is the spectral response function of the red channel of the camera. Note that I, R and S_r are all continuous functions of λ.

Unfortunately, we can see that varying the parameters of incident intensity and incident wavelength produce both non-linear, and interdependent effects on the resultant image[Thomson and Westland, 2001]. There are also other minor further complexities introduced by gamma correction. Thus, for full recovery of surface reflectance properties, if we assumed that we knew the colour of the surface, and the colour of the light, we must calibrate a three-dimensional space [Thomson and Westland, 2001]. In the extreme this would mean a three-dimensional look-up table. Thomson and Westland [Thomson and Westland, 2001] present an algorithm for recovering this calibration, but it is complex, and, of course, requires calibration charts.

4.4 Further problems with illumination

Perhaps we could say, however, that this expense was a tolerable start up exercise, and a worthwhile expense for a mobile robot. Unfortunately, we are only halfway to a solution to colour constancy for a mobile robot. We do not know the incident illumination for a scene in general. Let us assume that the robot is viewing a uniformly illuminated scene. In this scene, if we know enough surface reflectance properties perhaps we could recover the lighting. However, given the response of the sensor to incident intensity and wavelength is non-linear and interdependent, there is no real guarantee that we can uniquely recover the illumination. Further, the assumption of uniform illumination is far too restrictive for general robotics.

(a) (b)

(c) (d)

(e)

Figure 2: Images obtained with fluorescent lights (and natural lighting)

One question that should be asked at this point is how much do these theoretical models apply to common robot operation. Perhaps the type of lighting variation that mobile robots will typically experience in an environment is actually small enough that a simpler model can apply. This is the key question that this paper aims to address. We take images from a typical indoor robot lab. The illumination arises from fluorescent tubes and sunlight filtered through windows. In this case, we assess the adequacy of the above models of CCD sensor response to changing illumination given constant object reflectance properties.

5 Experiments

To capture the images, a Pentax Optio S 3.2 mega-pixel camera was used, giving images of 2048×1536 pixels. Images were taken of an exterior colour paint chart with auto exposure and auto white balance functions turned off. The extents of each region of colour were manually marked out (in groups of 25 to save time). Each colour blob consisted of approximately 95×95 pixels. A sub-window shrunk by five pixels on each side was used to

(a) (b)

(c) (d)

Figure 3: Images obtained with without lights (natural lighting only)

avoid any inaccuracies in the manual marking and any colour bleed effects. The average colour was simply then computed by averaging over the sub-window (approximately 85×85 pixels). The average colour of each of the colour samples was then recorded during one (cloudless) winter afternoon[1]. Pairs of photos were taken with the fluorescent lights switched on and then with the lights switched off. The photos were taken between 3:30pm and 6:13pm on 28 August 2003. Sunset was at 5:41pm. Figures 2 and 3 show the sequences of images obtained with and without the fluorescent lights, respectively.

6 Results

The circular colour charts shown here use hue-saturation-brightness values with the following (polar) mapping:

$$h = \theta/2\pi \tag{13}$$

$$s = r \tag{14}$$

$$v = 1 \tag{15}$$

where θ is the angle from the horizontal and r is the radius from the centre of the circle (unit circle). The rectangular colour charts have hue along the horizontal axis and saturation along the vertical (again the brightness is set to 1).

Figure 4 shows the individual shifts in colours that occur between pairs of images in Figure 2. Each end of each line is at the average colour of a blob in one of the

[1]Obviously, the experiments were not performed in Melbourne!

photos of Figure 2. i.e. the lines connect blobs of the same colour. The lines represent the shift in *apparent* colour that has occurred because of the change in lighting. It is clear that the colour varies considerably, even with the lights maintaining a reasonable level of illumination. It is also clear from the rectangular colour plots that there is no linear function to predict the variation in hue and saturation as a result of the lighting changes. In the circular plots, the direction of motion appears linear, though the actual displacement varies quite dramatically.

Figure 5 shows the individual shifts in colours that occur between pairs of images in Figure 3. Here we have a much more significant lighting shift because the total level of illumination is dropping sharply. This is particularly true for Figure 5(c), where the lighting shifts are large and nonlinear in both the circular and the rectangular representation.

The combined results for all of the with-lights images of Figure 2 are shown in Figure 6. Again, each red point indicates the average colour of one colour sample of the paint chart under varying lighting conditions and the lines connect the same paint sample together. Again, the rectangular chart (Fig. 7(b)) shows that there is no linear mapping to explain the colour variations. The circular mapping (Fig. 7(a)) again illustrates a clear general trend, though the magnitudes vary and in the aggregated results we see more variations.

Figure 7 shows the aggregated results for the without-lights images of Figure 3. Here it is clear that the colour variations are highly nonlinear.

The results of these experiments are perhaps surprising, given the human perception of colour constancy. However, given the non-linearity of the digital imaging process, we should not be surprised at the results. Some clear conclusions can be drawn:

- Colour variation in scenes with artificial lighting is significantly affected by any natural lighting.
- Natural lighting variations are dramatic. A sunset leads to large, non-linear colour shifts (as perceived through the digital imaging process).
- On cloudy days and in outdoor scenes with shadows, we can expect rapid and huge colour changes.

We can see that the constant colour model of Section 4.1 is acceptable for small lighting changes, but that switching on or off lights causes significant colour shifts. Also, the linear colour shift model of Section 4.2 is invalid, though one could propose a linear transformation in the polar hue-saturation space which would be a reasonable fit. Clearly, more attention needs to be paid to modelling and compensating for colour shifts.

7 Discussion

This paper has explored the issue of colour constancy and the lack of colour constancy for digital imaging devices (i.e. for robots). Humans perceive colours as being constant[Foster *et al.*, 1997] and so there is often an assumption that the same holds for digital imaging devices. However, our experiments have shown that this assumption is unsound. In particular, any natural sunlight entering the scene causes large, non-linear colour shifts (as perceived through the digital imaging process), as a result of sun elevation and clouds. These colour shifts cannot be compensated for without recourse to some external information (as discussed in Section 3).

The proposed solution is to track the colours of stationary colour blobs from a stereo camera pair mounted on a robot. In this way, the robot can maintain a mapping from the current perceived colour space to some reference (initial) colour space. The cameras mounted on a robot gives additional information: we are able to estimate the world position of each colour blob and by tracking stationary colour blobs we can be reasonably sure that changes are due to lighting variations. The goal is to provide colour constancy for robots with performance similar to that of humans, allowing quick and robust colour segmentation of objects. Our earlier work[Cameron and Barnes, 2003], is a step in this direction, but needs to be generalised.

It would be interesting to perform further experiments with different cameras. It is expected that the Pentax camera is representative of RGB cameras, but we have not verified this. Also, the paint chart does not cover the whole colour space[2]. Use of a "Macbeth Color Checker" [Scientific,] or a custom chart with better colour spread would give more complete coverage of the colour space.

References

[Austermeier *et al.*, 1996] H Austermeier, G Hartmann, and R Hilker. Colour-calibration of a robot vision system using self-organising feature maps. In *ICANN'96 Proc. Int Conf on Artificial Neural Networks*, pages 257–62, 1996.

[Bruce *et al.*, 2000] J Bruce, T Balch, and M Veloso. Fast and inexpensive color image segmentation for interactive robots. In *Proc IEEE/RSJ Int Conf on Intelligent Robots and Systems (IROS'00)*, volume 3, pages 2061–2066, 2000.

[Cameron and Barnes, 2003] D Cameron and N Barnes. Knowledge-based autonomous dynamic colour calibration. In *Proc. Robocup Symposium*, 2003. in press.

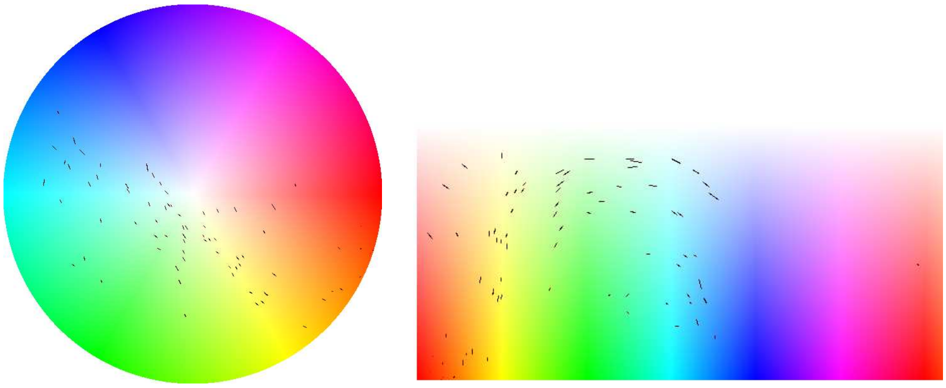
[2]Strangely, no-one wishes to paint their house bright pink, bright blue or bright green!

[Foster *et al.*, 1997] D H Foster, S M C Nascimento, B J Craven, K J Linnel, F W Conelissen, and E Brenner. Four issues concerning color constancy and relational colour constancy. *Vision Research*, 37(10):1341–1345, 1997.

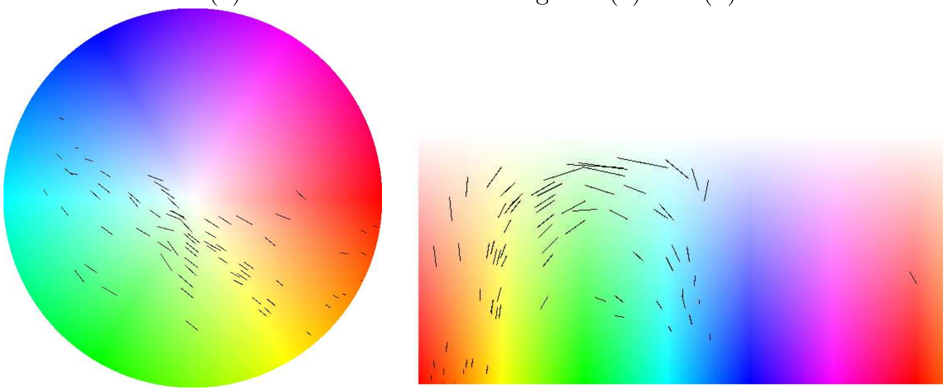
[Legenstein *et al.*, 2000] D Legenstein, M Vincze, and S Chroust. Finding colored objects under different illumination conditions in robotic applications. In *Proc SPIE Vol 4197, Intelligent Robots and Computer Vision XIX:Algorithms, Techniques, and Active Vision*, 2000.

[Scientific,] Edmund Scientific. Macbeth color checker. http://www.edmundoptics.com/.

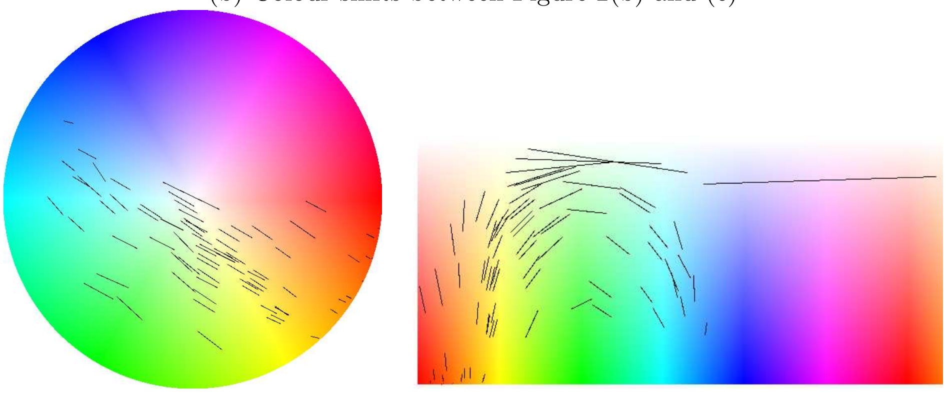
[Thomson and Westland, 2001] M Thomson and S Westland. Colour-imager characterization by parametric fitting of sensor responses. *Color Research and Application*, 26(6):442–449, Dec. 2001.



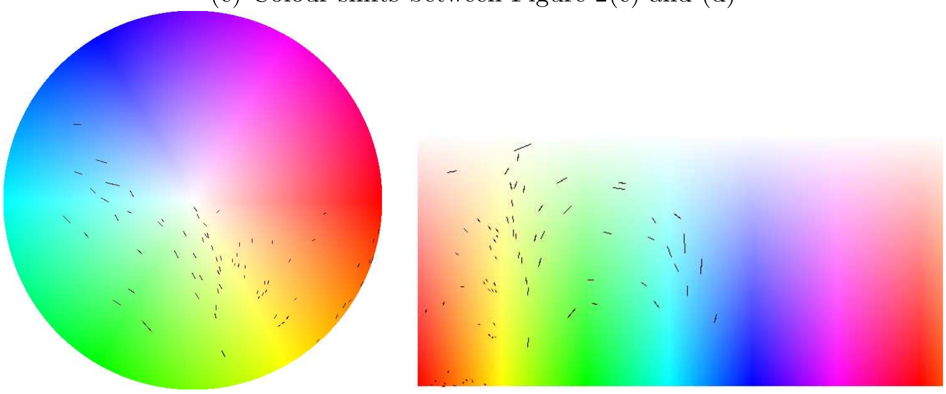
(a) Colour shifts between Figure 2(a) and (b)



(b) Colour shifts between Figure 2(b) and (c)

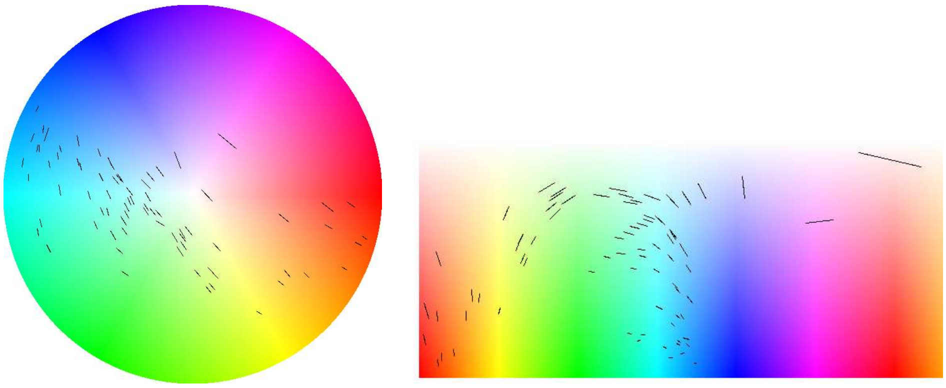


(c) Colour shifts between Figure 2(c) and (d)

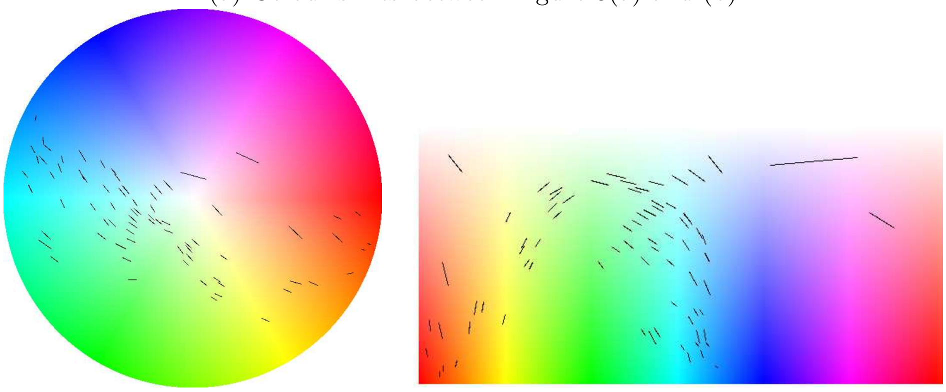


(d) Colour shifts between Figure 2(d) and (e)

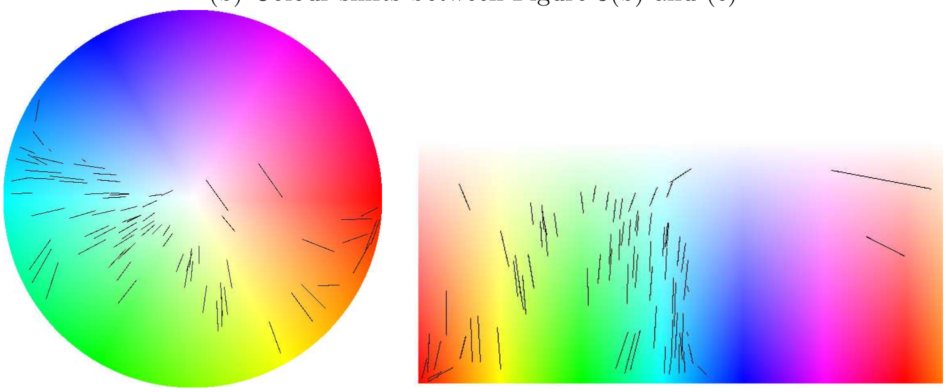
Figure 4: Colour changes with fluorescent lights and diminishing natural lighting



(a) Colour shifts between Figure 3(a) and (b)

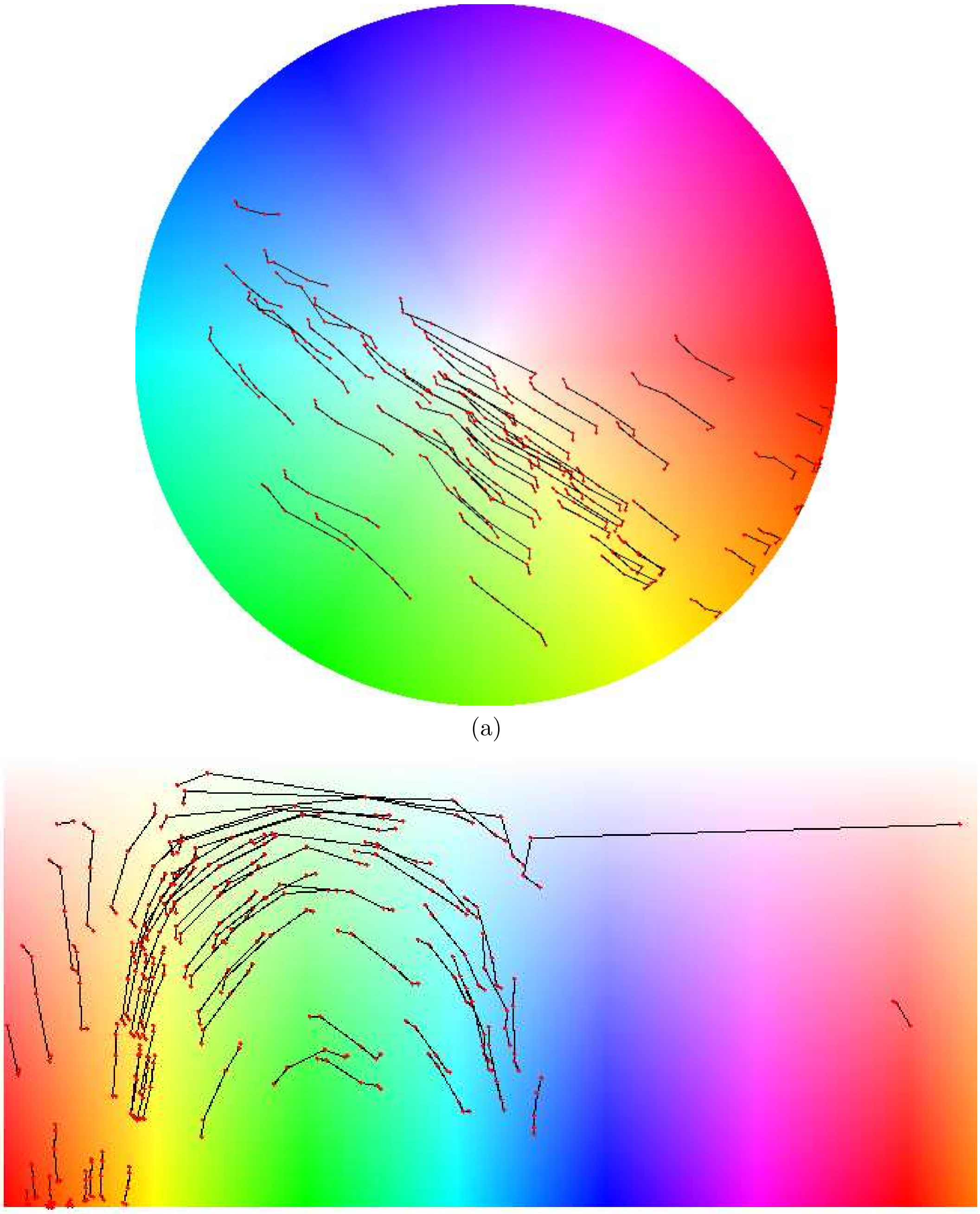


(b) Colour shifts between Figure 3(b) and (c)



(c) Colour shifts between Figure 3(c) and (d)

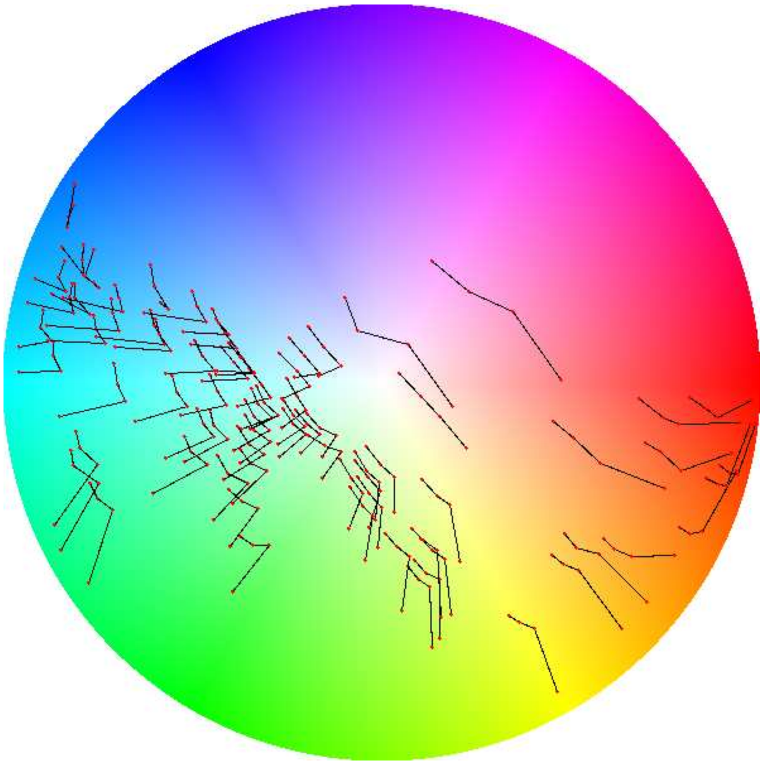
Figure 5: Colour changes with just diminishing natural lighting



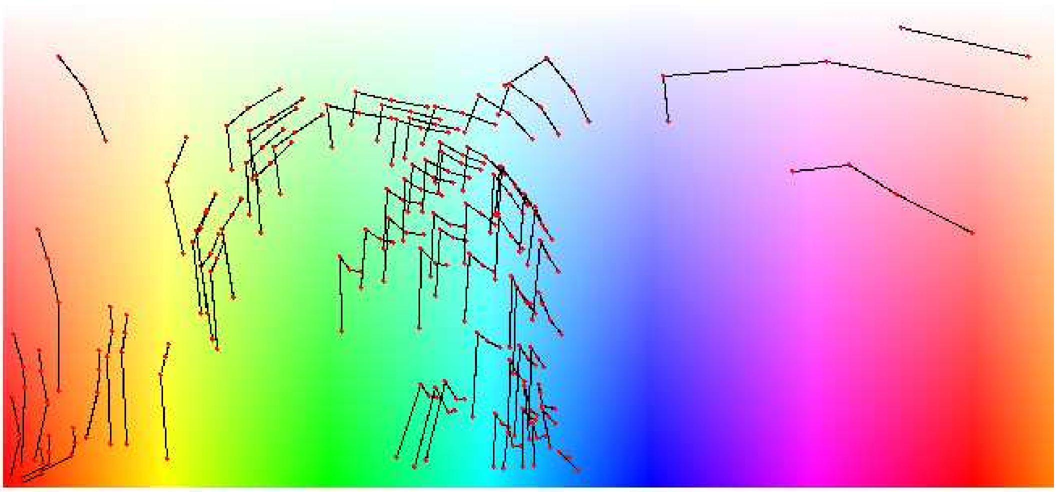
(a)

(b)

Figure 6: Aggregated results (with lights)



(a)



(b)

Figure 7: Aggregated results (without lights)