Abstract

Colour is an important aspect of art. Not only does it give richness to images, but it always provides a means to highlight certain objects. This idea of spot colour has been used extensively in both fine art and commercial illustrations. Many non-photorealistic rendering (NPR) algorithms produce grayscale or monochromatic images with low saturations. In this paper we introduce the idea of spot colour to NPR and propose a simple and automatic algorithm to add spot colour to these rendering styles. The hue is thresholded into colour layers and the most appropriate layer is automatically determined based on factors such as layer region shape and salience. We also consider using an edge-based criterion to colourise the background, which is an effective means of making the foreground stand out. We demonstrate the effectiveness of our approach by adding spot colour to a diverse set of NPR styles.


Keywords: non-photorealistic rendering, image abstraction, colour, segmentation

1 Introduction

Colour is obviously an important component of fine art. Not only can it be used to faithfully reproduce a scene, but it has been manipulated by artists in various ways to produce interesting effects. For instance, the Fauves exaggerated saturated colours and also used complementary colours to express emotions, while the Pointillists combined small points of pure colour to produce a vibrant effect. Of interest in the paper is the technique of restricting the palette for the majority of the painting, typically to dull colours, so that a small patch of distinct and bright colour can be applied to highlight an object in the scene. An example from fine art is given in “The Dance Lesson” by Edgar Degas in which the scene is made up from a range of shades of brown and gray, except for one of the girls who is painted wearing a bright red/orange shirt, which instantly draws the viewer’s attention.

Commercial illustration has extensively applied the idea of working with a reduced palette, and many such styles have been developed. Examples of one such style are shown in figure 2, in which the colour photographs have been taken or manipulated with this idea in mind. Figure 2a is naturalistic in that it has been composed such
that the image is dominated by green, making the coloured flowers stand out. Figure 2b has been modified such that the majority of the image is desaturated, with the original colours restricted to the masked region. Figures 2c&d show examples in which the complete image was desaturated, and then flat colour reapplied to one or more selected regions.

1.1 Related Work

Computer graphics and non-photorealistic rendering also process images to modify their colour palette for effect. For instance, Cohen-Or et al. [2006] performed “colour harmonisation” to produce more attractive palettes. Heckbert developed the median cut algorithm to perform adaptive colour quantisation [Heckbert 1982].

As colour gives perceptually useful information, existing research considers colourisation of grayscale images. This is generally an ill-posed problem and some additional information is often needed. This involves using an example image and transferring the colour information of the example image to similar regions based on local texture similarity [Welsh et al. 2002]. To further control the correspondence of regions, swatches can be manually specified to improve the transfer results. An improved example-based colourisation is considered by Iwata et al. [2005] that performs image segmentation and makes the decision of correspondence on a region basis, leading to more coherent colourisation. Another direction is to incorporate user scribbles to guide the image colourisation. Levin et al. [2004] model colourisation as optimising a quadratic energy with user scribbles as the boundary conditions, which requires solving a linear system and is quite efficient. A simpler and more efficient approach is proposed by Yatziv and Sapiro [2006] which produces colourisation by luminance weighted chrominance blending from user scribbles. Liu et al. consider example-based transfer in the illumination neutral domain [Liu et al. 2008] which works better for images with varying lighting conditions. Chia et al. use widely available Internet images for colourisation with semantics [Chia et al. 2011]. This work requires user labelling to help find the semantically relevant Internet images. The opposite problem of converting colour image to grayscale is generally straightforward. However, some methods [Gooch et al. 2005; Grundland and Dodgson 2007] convert colour image to grayscale with chromatically different pixels better differentiated.

Another relevant problem studied in previous literature is colour transfer where the colour style of one image is transferred to another image. Reinhard et al. [2001] use a linear mapping in the $l_a\beta$ colour space, based on pixel statistics. Abadpour and Kasaei [2007] obtain a correlated linear model between 3-channel colour vectors using PCA-based statistics. These methods make the colour mapping on an individual pixel basis. To incorporate spatial smoothness and consistency with the input image, Tai et al. [2005] use an expectation-maximisation (EM) scheme to take both spatial and colour smoothness into account. Xiao et al. [2009] introduce gradient domain consistency with the input image into the energy function to improve the perceptual fidelity.

There are a number of non-photorealistic rendering methods that modify the source image to produce a reduced palette. In the extreme case, a number use only two tones (generally black and white) [Mould and Grant 2008; Xu and Kaplan 2008; Inglis et al. 2012], while others allow a larger palette, including colour [Kyprianidis et al. 2009; Song et al. 2008; Rosin and Lai 2013]. More details are given in a recent review [Lai and Rosin 2013].

1.2 Motivation and Problem Statement

Various non-photorealistic rendering methods have been proposed. Among these there are quite a few styles that are naturally monochromatic or that use very low colour saturation. This is largely due to the artistic styles these algorithms mimic, such as charcoal, pen and ink, etc. Grayscale rendering styles are also suitable for low cost (black and white) printing.

Inspired by both fine art and commercial illustration, we consider the problem of generating non-photorealistic rendering with spot colour. A selected region of an image is highlighted with colour while the remaining image is left monochromatic or with low saturation, and the original rendering style is preserved. The idea can be effectively combined with various NPR styles; a few examples are given in figure 1. Aesthetically pleasing results with key objects highlighted are achieved for all the rendering styles. To the best of our knowledge, this has not been studied in previous research, and opens up a new topic in NPR.

Although traditional methods such as the median cut can be used to produce an image with relatively sparse colour (figure 3b), and three tone drawing by Rosin and Lai [2013] can be modified to have the mid-tone mapped to a specific colour (figure 3c), these are not spot colour as they do not capture some reasonable compact object. In contrast, figure 3d&e show that this is achieved by our automatic algorithm.

More specifically, we define spot colour to be the application of colour to a monochromatic image such that the colour is applied sparsely so that its effect is to highlight significant portions of the image. Thus, possible criteria for selecting regions of spot colour include:

1While our usage of spot colour is common, in offset printing this term also refers to the use of a number of solid blocks of colour, where each colour is applied with a separate ink and block. Another related concept in printing is halftoning, in which spots of colour (often CMYK, but other choices are possible [Power et al. 1996]) are combined to achieve continuous tones and colours.

Figure 3: Comparison of traditional NPR with spot colour. (a) original image, (b) the result of median cut with three colours [Heckbert 1982], (c) the result of three tone drawing [Rosin and Lai 2013] with the mid tone mapped to a colour, (d) the result of dithering with spot colour, (e) the result of ink pen with spot colour.
2 Method

Our algorithm takes a colour image as well as a monochromatic NPR rendering as input and produces an NPR image with spot colour.

The colour image is first converted from RGB colour space to hue, saturation and value to better match the perception of colours. To reduce the effects of noise the colour channels are smoothed with a Gaussian kernel ($\sigma = 4$).

Thresholding is applied to the hue – the aim is to initially represent all the image with a small number of colours. Making an analogy with the number of primary colours, we have chosen to threshold the hue layer into three classes. In fact, rather than use all the image pixels, we follow the approach of Rosin and Lai [2013], and consider the image to consist of three types of regions: dark (black), intermediate (gray) and light (white). Spot colour will only be applied to regions which have been classified as belonging to the intermediate type. Thus, the threshold values are computed from hues corresponding to the intermediate class pixels. A further restriction is applied, and intermediate class pixels with low saturation\(^2\) are excluded.

The three colour layers corresponding to hue are analysed to determine which is the most appropriate for spot colour. This is achieved by incorporating criteria described in section 1. More details are given in the following subsections.

The circular mean hue and the mean saturation of the selected hue layer are calculated. We apply the mean hue and saturation to mid-tone pixels belonging to the selected hue layer. Both the mean hue and the mean saturation are copied and their intensities are preserved.

The pipeline given in figure 4 illustrates how mid tone blocks and salience maps are used to select thresholded hue layers to provide spot and background colour overlays.

2.1 Circular Statistics

Since hue is a circular quantity, circular statistics need to be applied rather than linear statistics. This applies to the smoothing and thresholding of the hue layer described above. The circular mean $\mu$ of a set of $n$ orientation samples $\theta_\nu$ is defined as [Mardia and Jupp 1999]:

$$\mu = \begin{cases} 
\mu' & \text{if } S > 0 \text{ and } C > 0 \\
\mu' + \pi & \text{if } S < 0 \text{ and } C > 0 \\
\mu' + 2\pi & \text{if } S > 0 \text{ and } C < 0 
\end{cases}$$

where

$$\mu' = \tan^{-1} \frac{S}{C}, \quad C = \frac{1}{n} \sum_{x=1}^{n} \cos \theta_x, \quad S = \frac{1}{n} \sum_{x=1}^{n} \sin \theta_x.$$  

Also, the circular variance is calculated as $\sigma = 1 - \sqrt{S^2 + C^2}$. Note that $0 \leq \sigma \leq 1$.

2.2 Circular Thresholding

Although there is a large literature containing hundreds of algorithms for image thresholding [Sezgin and Sankur 2004], there has been very little work on thresholding circular quantities such as hue. Moreover, many of the published algorithms are both iterative and non-optimal [Tseng et al. 1995; Wu et al. 2006]. Instead we use the standard (non-circular) Otsu algorithm [Otsu 1979], which aims to minimise the within-class variance. This algorithm can be directly applied to hue data by simply changing from the linear statistics definition of variance to that from circular statistics.

For the two class thresholding problem it is adequate to exhaustively test all possible thresholds and choose the one minimising the within-class variance. However, this is time consuming when more threshold values are required. Instead, to more efficiently determine the values for three thresholds we use dynamic programming [Otsu 1980].

2.3 Tonal Blocks

Tonal blocks have been shown to be effective in capturing the overall tonal balance of an image [Rosin and Lai 2013]. We similarly threshold the input image into dark, light and mid tones. Those pixels in the range of mid tones are more suitable for spot colours as they are not too bright nor too dark.

\(^2\)We define low saturation as $\leq 64$ where values are in the range 0–255.
2.4 Hue Selection

Once the image has been thresholded into separate colour layers the goal is to select the most appropriate colour layer. As described in section 1, several criteria are considered. We start by describing the criteria based on the compactness of the layer shape, which also has some impact on the size of the shape. After that, we consider more semantic related criteria, using salience. We finally consider colourising the background as an alternative to spot colour since it also helps the foreground objects to stand out.

2.4.1 Shape Based Criteria

Two criteria were initially considered for finding compact regions: spread and relative variance. Spread is a geometric moment invariant defined as

$$\frac{M_{20} + M_{02}}{M_{00}}$$

where $M_{pq}$ are the $pq$'th central moments, and the spread should be minimised to provide compact regions. A problem exists for regions consisting of multiple components, since as the components become more widely separated the value of spread can increase without limit. Thus, even a colour layer consisting of two very compact regions will be rated poorly if the regions are not close. As a potential solution, only the largest region was retained and analysed for each colour layer. However, this proved to be sensitive to minor segmentation errors that altered connectivity, and generally produced poor results.

Relative variance is the ratio of the sample variance to the sample mean, and is used to test for a random spatial distribution modelled as a Poisson distribution [Upton and Fingleton 1985]. It was later used to select the least random, i.e. most clustered, spatial distribution for threshold selection by maximising relative variance [Rosin 2002]. It is calculated by counting the number of observations (i.e. the pixels belonging to a specified colour layer) in $n$ windows: \(\{c_i\}_{i=1}^n\). Relative variance is then computed as \(V_r = \frac{v^2}{\bar{c}}\), where $\bar{c}$ and $v^2$ are the mean and variance of $c_i$. We have divided the image into an $8 \times 8$ grid of windows. Although the test is dependent on the window size and point density it works adequately as long as $\bar{c}$ is sufficiently large. When regions are translated within the image, the relative variance is largely unaffected, although the alignment of regions with grid does have some effect in the obtained value. The colour layer that maximises the relative variance has better compactness, hence it is chosen for spot colour.

Whereas spread is scale invariant, relative variance is not since the numerator in $V_r = \frac{v^2}{\bar{c}}$ is squared. We therefore also tested an alternative scale invariant version, $S_r = \frac{v^2}{\bar{c}}$, but found this to perform poorly. It tended to often select very small and insignificant regions. In contrast, relative variance performed better for the images that we tested, which typically contain 0.5–1M pixels, since the measure is biased to favour larger regions.

Although it is possible that the compactness measures could select outsized regions (either too small or large), this is limited in practice. First, since the potential colour layers are defined on just the mid tone regions this tends to prevent regions becoming too large. Second, even very small regions (as small as 1% of the image size) are still perceptually acceptable for spot colours.

2.4.2 Salience Based Criteria

When spot colours are used by artists they tend to be used to emphasise key foreground objects. Criteria based on shape are not sufficient to capture this and can quite often choose regions which are not so significant. To improve this, we combine image salience with the shape criteria to guide the hue selection. Many salience detection algorithms have been proposed, based on colour rarity, location, and spatial distribution etc. We use the approach described by Cheng et al. [2011] which is efficient and generally produces good results. Alternative salience measures could also be used. This approach first segments the input image into regions, and calculates the colour histogram for each region. The salience $s$ for pixels in the region is calculated as the sum of pairwise histogram distances between this region and all the other regions, weighted based on their spatial distances so that neighbouring regions have more effect. The salience values are normalised to be within $[0, 1]$.

Although it is possible that the compactness measures could select outsized regions (either too small or large), this is limited in practice. First, since the potential colour layers are defined on just the mid tone regions this tends to prevent regions becoming too large. Second, even very small regions (as small as 1% of the image size) are still perceptually acceptable for spot colours.

We incorporate the salience to form a salience weighted spread criterion as follows. For pixels $(x_i, y_i)$ in the colour layer, we define the moments to be weighted by the salience $s_i$, making salient pix-
els more significant in spread calculation:

$$M_{pq} = \sum_i s_i x_i^p y_i^q,$$

and the salience weighted spread is defined as

$$\frac{M_{20} + M_{02}}{M_{00}^2 \sum_i s_i} = \frac{M_{20} + M_{02}}{M_{00}^2}.$$

Similar to the spread criterion, the colour layer with the lowest salience weighted spread is selected. This criterion favours regions that have an overall higher salience, and whose relatively salient pixels are compact. Some examples are shown in figure 5. The spread criterion is not able to capture regions which semantically need to be highlighted (figure 5b). Salience maps obtained using Cheng et al.’s method [Cheng et al. 2011] effectively capture the salient objects (figure 5c), and by using salience weighted spread, spot colours are appropriately applied to significant foreground objects (figure 5d).

### 2.4.3 Background Selection

An alternative to selecting small foreground objects for colouring is to select the background. If the background is large, then this has a related effect to spot colour, except that now the foreground is highlighted by being grayscale. We model the background as containing less structure than the foreground, and measure this by computing the mean edge magnitude (using the Sobel edge detector). Furthermore, to penalise small regions and encourage large background regions the mean is down-weighted by the region size. Thus the measure is

$$\sum_{p_i \in R} E_{p_i} |R|^2$$

where $p_i$ is a pixel in region $R$ and $E_{p_i}$ is the edge magnitude at $p_i$. In our experiments we have set a threshold of 0.00007 on the background measure to identify those images with sufficiently large and uniform backgrounds to merit colouring in this manner.

Examples are shown in figure 6 to demonstrate the effect of various spot colour criteria as well as background colourisation. Colourising the background in these examples (figure 6b) makes the foreground more visible, as they are uncoloured. Salience weighted spread (figure 6c) successfully adds spot colour to the foreground objects, whereas the relative variance criterion (figure 6d) often chooses the background (the first example) but may also choose some foreground object (which may not be the most significant one) – see the example on the second row.

Figure 7 shows an example with different hue selection criteria. The relative variance criterion (figure 7b) in this example selects the background (although as shown in the previous example, this is not always the case). The spread criterion chooses some of sportsmen and salience weighted spread chooses the other sportsmen, and in this example both criteria give reasonable results. The background selection gives the correct background regions. From these and other examples we find that salience weighted spread gives consistently good spot colour regions and thus we use this as the default choice. The background selection method is effective when colourising the background is preferred.

### 2.5 GrabCut Refinement

The colour layer obtained from hue thresholding may not necessarily align region boundaries with structures. This is generally acceptable but we have found by using GrabCut [Rother et al. 2004] the region boundary can be improved by taking the spatial consistency of pixels into account. We use the colour layer to initialise the

---

**Figure 6:** Colourising background and spot colour results. (a) input images, (b) rendering with background colourised, (c) rendering with spot colour using salience weighted spread criterion, (d) rendering with spot colour using relative variance criterion.
GrabCut refinement, with pixels close to the layer boundary (within 50 pixels in all our experiments) considered as unknown regions, and label the remaining pixels based on whether they belong to the layer or not. Some examples of GrabCut refinement are shown in figure 8 in which GrabCut helps to align the layer boundary to significant image features. However, in general we have found that it often makes errors that detract from the final result. For example, in the lower example in figure 8 the refined result covers part of the belt, which is semantically incorrect. In other cases we have found GrabCut to substantially grow regions, such that they are no longer suitable for spot colour. Therefore, we have not made GrabCut part of the default pipeline.

3 Experiments

We have already shown a variety of non-photorealistic rendering results with spot colour. Some more results are shown in this section. Figure 9 gives a spot colour example using spread, relative variance and salience weighted spread criteria. Although different colour layers are selected, they actually all look reasonable. Figure 10 shows examples of spot colour as well as background colourisation. Both of them help to make the foreground objects stand out. Figure 11 compare the rendering results of five different styles, without (second row) and with (third row) spot colour. Our method works
well for a variety of styles, and effectively makes the foreground plane a focus of attention. Finally, figure 12 shows more results of our method. The second example in the top row shows a monochromatic (but not gray) rendering that has been spot coloured. The top two examples both use GrabCut refinement of the spot regions. In the bottom example the salience weighted spread criterion has selected the background since the cloud has been assigned a very high salience value.

The current unoptimised code takes about 2 seconds to add the spot colour to a 0.6M pixel image for the basic version of the pipeline. Programs were run on a 3.40 GHz Intel Core i7.

4 Conclusion

Inspired by fine art and commercial illustration, we introduce the idea of spot colour to non-photorealistic rendering. A simple and automatic algorithm is proposed which extracts a salient and compact layer with consistent colour, and aesthetically pleasing results are obtained when applied to various rendering styles, either grayscale or of low saturation. Various (subjective) criteria can be incorporated to decide which regions should be highlighted and there are certainly scope to improve our current approach. Moreover, rather than fully automate the process, it may be desirable to allow the possibility of some interaction to enable artists to control and adjust the results. As all the other NPR problems, it is generally difficult to evaluate results quantitatively. For this, user studies may be a practical approach.

Acknowledgements

The source images have been made available on Flickr under the creative commons license by the following members: figure 1: Klearchlos Kapoutsis; figure 2: Dun(can), Samantha R(20129), GonchoA; figure 3: mikebaird; figure 4: Fimb; figure 5: szekemike; figure 6: chefranden; figure 7: familymw; figure 8: Paul Stevenson, hoyasmeeg; figure 9: flikr; figure 10: Sabrina Campagna, dagnygr; figure 11: Armchair Aviator; figure 12: Drew Avery, Alex E. Proimos, SergioTudela, mikebaird, flatworldedge.

References


Figure 11: Non-photorealistic rendering results without and with spot colour. (a) input image, (b) three tone rendering [Rosin and Lai 2013], (c) PhotoShop Conté Crayon, (d) GIMP ink pen, (e) PhotoShop Plaster, (f) dithering with polygons, (g-k) corresponding results of (b-f) with spot colour.


Figure 12: Gallery; left column contains input images. Middle and left columns contain spot colour applied to various rendering styles.