Umanistica Digitale – ISSN: 2532-8816 – n. 10, 2021
N. Redfern – *Colour palettes in US film trailers: a comparative analysis of movie barcodes*
DOI: http://doi.org/10.6092/issn.2532-8816/12468

Colour palettes in US film trailers: a comparative analysis of movie barcodes

Nick Redfern

Independent scholar
nickredfernres@outlook.com

Abstract

This article analyses the smoothed movie barcodes of 173 trailers nominated for a Golden Trailer award between 2016 and 2019 across nine genres: action, animation/family, comedy, documentary, drama, fantasy/adventure, horror, romance, and thriller. The results show that colours in the nine genres are similar, dominated by dark, unsaturated colours in the orange (hue = 30° ± 30°) and azure (hue = 210° ± 30°) regions of the HSL colour wheel. Colour palettes for each genre have similar colours but show some differences in the diversity and evenness of the distributions of these colours.

Questo articolo analizza i codici a barre di 173 trailer nominati per un Golden Trailer award tra il 2016 e il 2019 in nove generi: azione, animazione/famiglia, commedia, documentario, dramma, fantasy/avventura, horror, romantico e thriller. I risultati mostrano che i colori nei nove generi sono simili, dominati da colori scuri e insaturi nelle regioni arancione (tonalità = 30° ± 30°) e azzurro (tonalità = 210° ± 30°) della ruota dei colori HSL. Le tavolozze dei colori per ogni genere hanno colori simili ma mostrano alcune differenze nella diversità e nell'uniformità delle distribuzioni di questi colori.

Introduction

Taylor Arnold and Lauren Tilton ([1]) describe *distant viewing* as a methodological and theoretical framework for the computational analysis of large collections of images such as films, television programmes, photographic archives, and art collections. It is an approach to studying culture grounded in a different epistemic viewpoint to the close analysis of texts, embedding individual objects in a broader context: 'distant viewing is the presentation of connections between multitudes of images (Big Image Data) in meaningful spatial proximity to each other in order to gain an overview of a corpus that would not be possible without digital means' ([21]: 35). As an analytical method, distant viewing involves the extraction of formal elements from visual materials, followed by the aggregation and visualisation of those elements, which may include low-level syntactic features, such as colour, lighting, camera movement, shot scale, shot duration, audio events, and higher-level semantic features, such as face recognition and object detection ([1]; [9]).

In this article I adopt a distant viewing approach to analyse a sample of 173 trailers nominated for a Golden Trailer award between 2016 and 2019 in nine genres in order to determine what similarities and differences exist between their colour palettes at the level of genre, with specific attention paid to hue, saturation, and lightness. Flueckiger and Halter ([11]) argue that computational approaches to analysing film colours are capable of going beyond the hermeneutical interpretation of individual films to visualise and explore the different attributes of colour in the cinema at the level of the corpus. I therefore apply computational methods to the task of distant viewing using movie barcodes as representations of the colour data in the trailers. To date the only study of colour in film trailers is Feng ([10]), who considered the trend in luminosity in trailers from 1951 to 2015 but did not analyse other colour attributes such as saturation or hue and did not look at differences in colour for trailers in different genres. This article is therefore the first to examine how colour varies across a corpus of film trailers.

Methods

Movie barcodes as method

Computational analysis of film colour is resource-intensive and so the workflow comprises two stages to make the process manageable:

- *sampling*: frames are selected from a film based on their timecode (e.g., every *n*-th frame or *n* frames per second) or their representativeness of a sequence of frames based on a segmentation of the film into coherent scenes.
- *data reduction*: the pixel data in each frame is reduced to a single set of values that are representative of the whole image, and which are usually based on either the average or dominant colour of the pixels, or the representativeness of an exemplar colour based on cluster analysis.

A movie barcode is one way of sampling and reducing the colour data in a film and is the most commonly used method of analysing and visualising colour in cinema ([4]; [25]). Isola et al. ([19]) used movie barcodes to segment the structure of films based on colour that correlated well with DVD chapters. Burghardt et al. ([5]) and Burghardt, Kao, and Walkowski ([6]) explored the relationship between colour, dialogue, and character in motion pictures using movie barcodes as a fingerprint of the chromatic structure of a film. Hohman et al. ([18]) explored the relationship between colour, dialogue, and emotion in *Game of Thrones* using a method similar to movie barcodes to visualise the structure of the show's episodes. Bhardwaj ([3]) employed barcodes to explore the macro-structure of colour in anime films.

The sampling and reduction workflow of computational approaches inevitably results in a loss of data that imposes limitations on any analysis. A significant limitation of using movie barcodes is that they use only a small proportion of the available data. For example, a 100-minute feature length film with a frame rate of 24 frames per second has 144000 frames, which means that selecting 1000 frames for inclusion in a barcode uses only 0.7% of the available data. There are additional limitations to using movie barcodes as a research method. Halter et al. ([13]) argue the use of movie barcodes beyond distant watching is limited due to a loss of granularity through the averaging process. Kuhn et al. ([22]) note that while visualisation methods such as movie barcodes are informative, they lack explanatory power in themselves. It is therefore necessary to extract data from the smoothed barcodes that can be further processed in order to draw conclusions about colour in film trailers.

Sample

The sample used in this study comprises 173 trailers nominated for a Golden Trailer Award (GTA) in nine genres between 2016 and 2019, inclusive. The nine genres are action, animation/family, comedy, documentary, drama, fantasy/adventure, horror, romance, and thriller (Golden Trailer Awards 2020). The GTAs also have similar categories for non-US films designated as foreign action, foreign horror, etc., but these are not included in the present study. The sample only includes trailers nominated in the cinematic categories and does not include trailers produced for television or online distribution, though it does include a handful of trailers produced for ComicCon that were nominated in cinematic categories. The use of genres in this paper does not therefore reflect Platonic conceptualisations of a specific genre – no such conceptualisation being possible, in any case (see [24]: 24-27). Nor are these categories based on the genre of the film promoted by a trailer. Rather they are based on the ways in which the producers of the trailers themselves consciously positioned their work within the marketplace by submitting a trailer for an award.

Trailers receiving multiple nominations are not included in the sample, ruling out three trailers nominated in more than one category (*Long Shot* – 'This Guy', *A Star is Born* – 'Not Alone', and *Us* – 'International Trailer'). One trailer (*Adrift* – 'Bon Voyage') was nominated in the romance category in consecutive years and is included in the sample once only.

The trailers were downloaded from YouTube, the Clios website, or the websites of the trailer production companies as mp4 files and edited to remove MPAA rating tag screens. Some trailers are topped and tailed by additional promotional material either announcing the trailer itself, additional social media information not part of the cinematic release, and/or advertising the YouTube channel hosting the trailer and this was also removed prior to analysis. I cropped each trailer to remove letterbox blanking bars in order to ensure the colour analyses in the study were not distorted by having large parts of the frame dominated by a colour unrelated to the trailers themselves.

Generating the barcodes

The smoothed barcode of each trailer was generated using Melvyn Laïly's Movie Barcode Generator (v.1.6; [23]). This software takes an input video file and extracts 1000 individual frames which are then reduced to a single pixel's width and stacked together, with each column representing a frame from the film. The Movie Barcode Generator generates unsmoothed and smoothed barcodes from a video file, with the bars in the latter the average colour of the bars in the original version. The barcodes of the trailer *Ralph Breaks the Internet: Wreck It Ralph 2* – 'Wired Refresh' in Figure 1 are examples of the outputs of this process. The complete set of 173 smoothed barcodes used in this study is available online under a Creative Commons Attribution 4.0 International License ([27]).

Using the smoothed barcode of each trailer means that every row of the image has the same information and so we can pick any individual row of pixels in order to analyse the colour information contained within the barcode. This improves the computational efficiency of the analysis while accurately reflecting the colour information available in the smoothed barcode. I extracted the 100th row of pixels of each smoothed barcode and concatenated the data for each trailer in a genre to create a single dataset for each genre used in the subsequent analyses.

Some of the limitations of using movie barcodes as a method are mitigated in this study. The sample size problem is reduced because I am working with trailers with numbers of frames that range from 1176 (*Our Souls at Night* – 'Sunday Love Teaser', 49.0 seconds at 24 fps) to 4702 frames (*Fantastic Beasts: Crimes of Grindlewald* – 'Rare ComicCon,' 195.9 seconds at 23.976 fps). The barcodes used in this study therefore represent between 85.0% and 21.3% of the frames in a trailer, with a median of 29.0% frames sampled from each film. The loss of data resulting from the sample-and-reduce workflow is not an issue because I am interested only in variations in the dominant colours between genres.

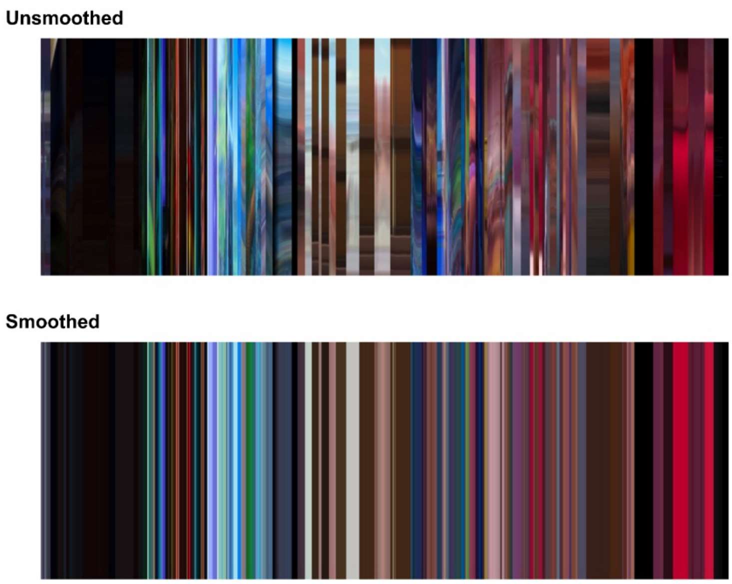

Figure 1: Barcodes for Ralph Breaks the Internet: Wreck It Ralph 2 – 'Wired Refresh.' The bars in the smoothed barcode are the average colour of the bars in the unsmoothed barcode.

HSL colour data

To compare the different colours each genre I converted the RGB image of each barcode to the HSL colour space ([29]) using the RGB2HSL() function from the imager package ([2]) for the statistical programming language R (v. 4.0.3; [26]), with the saturation and lightness channels multiplied by 100 to express them as percentages.

The HSL colour space (Figure 2) is defined by three colour-making attributes:

- Hue: the natural colour expressed as a degree on the colour wheel, where red = 0°, green = 120°, and blue = 240°.
- Saturation: the purity of the colour expressed as a percentage and measured as its distance from grey (S = 0%). Pure hues occurs when S = 100%.
- Lightness: the brightness of a colour relative to the brightness of a similarly illuminated white expressed as a percentage. Full colour occurs when L = 50%, with colours becoming darker as $L \rightarrow 0\%$ and lighter as $L \rightarrow 100\%$

In HSL colour space, hue and saturation values are arbitrary for black and white, which depend solely on their lightness value, where L = 0% is black and L = 100% is white. Hue values are arbitrary for grey, which describes all colours with a saturation of 0% and where 0% < L < 100%. The HSL colour space is not perceptually uniform but it is easy to interpret the values of H, S, and L, unlike the $L^*a^*b^*$ colour space which is preferred for calculations because of its perceptual uniformity (see below) but which has attributes that are harder to interpret.[1]

In the HSL colour space, black and white are defined only by their lightness and have arbitrary saturation values. To limit the data set to meaningful saturation attributes I filtered black and white pixels using an expanded definition of each to account for the fact that it is not possible to visually distinguish between colours that are close to black or white. For example, I cannot perceive any meaningful distinction between black, with L = 0% and arbitrary H and S values, and very dark grey, where S = 0%, L = 1%, and H is arbitrary. The definitions used here are:

- Black: a lightness value less than or equal to 5% (L ≤ 5%)
- White: a lightness value greater than or equal to 96% (L ≥ 96%)

Filtering the data in this way leaves us with only those saturation values that are meaningful.[2]

1 The perceptual uniformity of a colour space ensures that the perceived difference between two colours is proportional to the Euclidean distance between them in a colour space, so that difference in the values of the attributes correspond to differences we actually see between colours.

2 The perception of colour is subjective and depends on a range of features, including the available light in an environment, the way objects reflect light, the way light reacts with a perceiver's eyes, and the way a perceiver's brain processing information about colour. I may be able to perceive shades, tints, and tones (the mixture of a colour with black, white, and grey, respectively) others may not recognise and so filtering at 5% will seem too low for them. Likewise, others may find my 5% cut off value too generous as they are able to distinguish between shades, tints, and tones below this level.

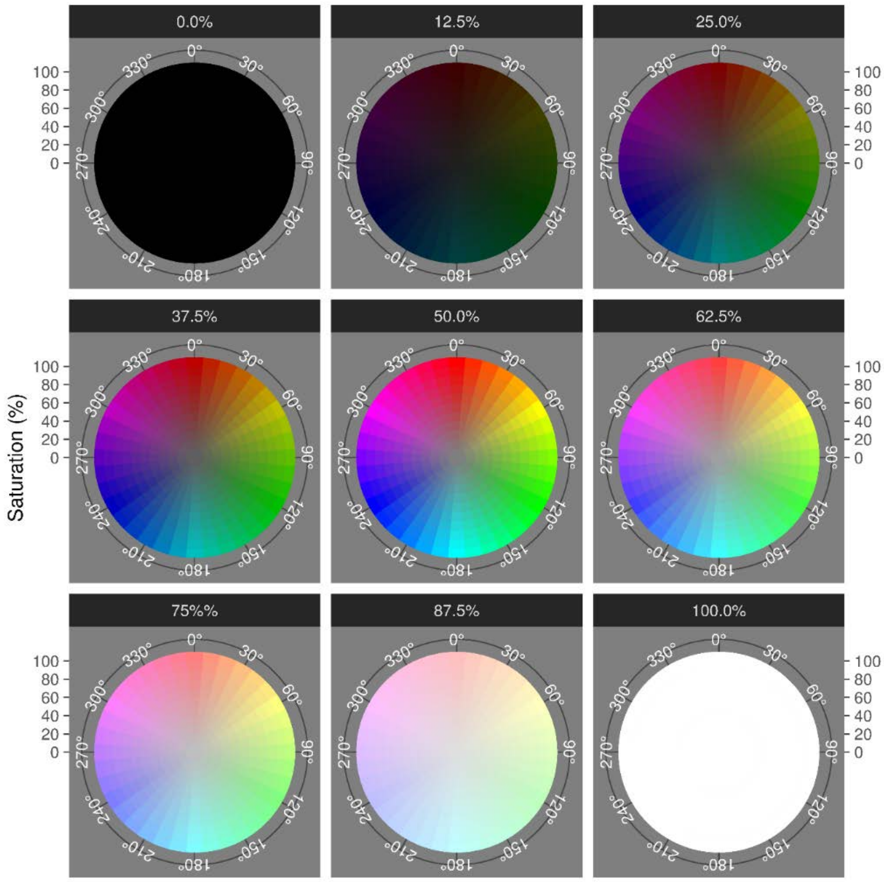

Figure 2: The HSL (hue, saturation, lightness) colour space, showing hue (the angle around the colour wheel) and saturation (the distance from the centre of the colour wheel, where grey is at S = 0%) at nine different lightness values.

In order to focus attention on the hues of the pixels, the generic data sets were further filtered so that all grey pixels were removed. Again, I used an expanded definition of grey to reflect the difficulty in visually distinguishing between grey and colours close to grey:

- Grey: a saturation value of less than or equal to 5% ($S \leq 5\%$) and a lightness that is not already classed as black or white ($5\% < L < 96\%$).

The remaining colours are therefore defined by their hue, $S > 5\%$, and $5\% < L < 96\%$. This removes all pixels with arbitrary H values that would otherwise lead us to misinterpret the distribution of hues in a genre. It also removes datapoints with hues that are difficult to interpret as they are perceptually indistinguishable from black, white, or grey.

Generic palettes

To get an overall view of the characteristic colours of a genre I construct a palette for each genre of trailers using the unfiltered generic data sets. Defining a palette for a genre can be treated as an unsupervised machine learning problem of dividing n objects (i.e., pixels) into k distinct groups (i.e., colours) using cluster analysis ([20]). To construct a palette for each genre, I pooled the pixel data for the 100th row of each of the barcodes in a genre into a single data set and converted the RGB colours to the $L^*a^*b^*$ colour space using grDevices:: convertColor() ([26]), a perceptually uniform colour space in which a colour is defined in terms of its luminosity (L^*) and it positions on the red-green (a^*) and blue-yellow (b^*) axes, with the reference white at the CIE Standard Illuminant D65 ([30]). I then applied a z-score transformation[3] to scale the data before applying CLARA, a clustering method suitable for partitioning around medoids (PAM) with large data sets, using the fpc package for R ([16]) to find the optimal number of clusters for the data across a range of possible values of k from 10 to 20, with 50 samples of 250 pixels for each value of k. PAM selects values from the data set as medoids to function as exemplars of a cluster and assigns the values in a data set to the nearest cluster based on the minimization of a distance measure between a value and its exemplar ([28]). As the distance between two colours in $L^*a^*b^*$ colour space is the Euclidean distance between them, I used this distance measure for the clustering algorithm. An advantage of the k-medoids approach over k-means clustering, the standard method used for cluster analysis of colours in images, is that the cluster centres are drawn from the colours in the image which cannot be guaranteed using the k-means approach. After clustering, I reversed the z-score transformation to return the medoids to the original scale of the $L^*a^*b^*$ colour space. The resulting colour palettes for each genre are plotted as a tree map, where the area of a branch is proportional to the number of pixels assigned to a medoid, as an efficient way of visualising colour palettes, allowing for easy comparison without discarding information about the prevalence of a particular colour.

I calculate a range of summary statistics describing the palettes. The *richness* of a palette is the number of colours (k) it contains, though this is a crude measure of the difference between palettes. The *diversity* and *evenness* of a genre's palette can be quantified based on the relative frequency of pixels in a cluster: $p_i = n_i/k$, where n_i is the number of pixels assigned to a cluster and k is the total number of clusters in a palette. Shannon's entropy quantifies the degree to which data is distributed over its possible values and is sensitive to rarer pixels:

$$H = -\sum_{i=1}^{k} p_i \log_2 p_i \,.$$

The greater the value of H, the greater the diversity of a palette. As different genres have different numbers of colours in their palettes, Pielou's J' statistic is a measure of the evenness of a palette, where J' is the ratio of the observed entropy to the maximum possible entropy for k colours: $J' =$

3 A z-score transformation indicates how many standard deviations (σ) a value is above or below the mean (μ) and is equal to $z = (x - \mu)/\sigma$. This scales the observations in a vector so that μ = 0 and σ = 1. To reverse the transformation, we multiply an observation by the standard deviation and add the mean.

$H/\log_2 k$. Simpson's reciprocal diversity index also measures how pixels are distributed across colours but is more sensitive to the most abundant cluster in a palette, with values in the range [1, k] that increase as diversity increases:

$$S' = \frac{1}{\sum_{i=1}^{k} {p_i}^2}.$$

Dividing S' by k provides an additional measure of evenness: $S_E = S'/k$. Both J' and S_E take values in the range [0, 1], where 0 is maximally uneven (i.e., all pixels would be assigned to just a single colour out of k possible colours) and 1 is the maximum possible evenness, where all individuals are equally distributed across all categories and the number of pixels assigned to a colour would be equal to $1/k$. See [15] for a discussion of these indices.

Hue, saturation, and lightness

Lightness

From the distribution of the total set of lightness values for each barcode in Figure 3, we see that the barcodes for the trailers in each genre tend to be dark with lightness heavily concentrated at low values. There are few pixels in each genre with a lightness greater than 50%, ranging from just 4.7% in the case of the horror genre to 17.7% for documentary, and there are very few white pixels (L = 100%) in the sample for any genre. Cutting et al. ([8]) reported that films Hollywood films have become progressively darker over time as the mean luminance of films has decreased from 1935 to 2010. It is therefore unsurprising that the trailers for contemporary Hollywood films such be so dark given the declining trend in mean luminance in Hollywood cinema, though Feng ([10]: 67) did not find a similar decline in colour luminosity of Hollywood trailers released since 1950.

The horror genre is the darkest genre, with a median lightness value of 7.3% and black (L = 0%) accounting for 14.8% of the pixels in this genre. The brightest barcodes are to be found in the comedy genre, with a median lightness value of 23.7% and black pixels accounting for just 3.8% of the total pixels for this genre. This genre has the most even distribution of lightness values across the nine genres in the sample. This might seem obvious, given these genres are associated with different types of emotions, but there is, however, no consistent pattern across the other genres in the sample and other factors have to be considered when considering the lightness of a trailer. The animation/family genre has some very bright barcodes, as can be seen in the example in Figure 1, but it also has several dark ones. These are largely Disney's non-animated productions: *Christopher Robin* – 'Into the Wood' (see Figure 7.A), *Dumbo* – 'Courage', *The Nutcracker and The Four Realms* – 'Clara', *Beauty and the Beast* – 'Tale', and *The Jungle Book* – 'Jungle', which has one of the darkest trailers in the whole sample with a median lightness of 3%. Of the drama trailers in the sample, the barcode for *Megan Leavey* – 'Mission' stands out as having a median brightness (38%) much greater than other films in this genre. This is generally true of the barcodes of trailers for films with desert settings (including *Mad Max:*

Fury Road – 'Survive', *Sicario: Day of the Soldado* – 'Drive', *Sicario* – 'Trailer'), which tend to have among the highest median brightness values in their genre. Setting also dominates the lightness of the brightest of the romance trailers, *Mamma Mia 2* – 'Memories', which is set on a sunny Greek Island. Another factor determining brightness is the use of vivid colours as part of the marketing function of the trailer, and this can be seen in the bright pink of the marketing screens for the *Love Simon* – 'Courage' trailer (see Figure 7.B), the bright yellow in the trailer for *Three Identical Strangers* – 'Reunited', the bright blues of the *Eddie the Eagle* – 'Sport' trailer, and the neon blues of *John Wick: Chapter 2* – 'Vengeance.'

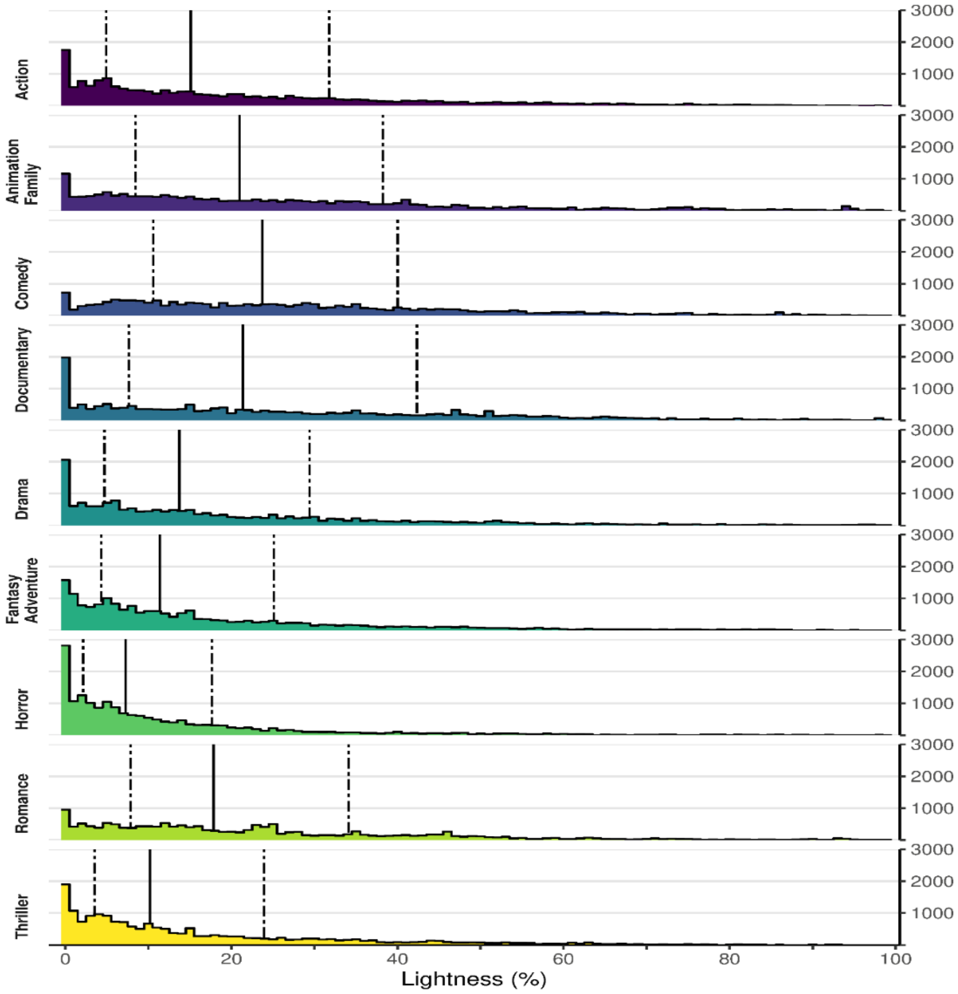

Figure 3: The distribution of lightness values in the HSL colour space for smoothed barcodes of 173 trailers across nine genres, with the median (solid line) and lower and upper quartiles (dashed lines) indicated.

Saturation

Figure 4 presents the distribution of the filtered saturation values in each genre. Like the lightness values, saturation values tend to cluster at lower end of the range indicating that pixels in the barcodes tend to be unsaturated though the number of grey pixels is actually quite small for each genre. Saturation values are more evenly distributed than lightness values but tend to be in the lower half of the range (S < 50%). The median saturation of pixels in the genres ranges from 14.8% for the barcodes of documentary trailers to 28.3% for those of animation/family trailers. The proportion of saturation values in the upper half of the range is approximately 11.7% for comedy, documentary, drama, horror, and thriller trailers, and between 14% and 21% for action, fantasy/adventure, and romance trailers, though for all genres pixels within this range tend to be fully saturated. For animation/family trailers the saturation is greater than or equal to 50% for 25.4% of pixels, though most of these are from just three trailers in the Lego movie franchise: *The Lego Batman Movie* – 'ComicCon,' *The Lego Batman Movie* – 'Domestic Trailer,' and *The Lego Movie 2* – 'Universe.' The barcode of the trailer for *Christopher Robin* – 'Into the Wood' has a median saturation of 9% due to the numerous shots of grey skies that dominates the colour scheme in this trailer.

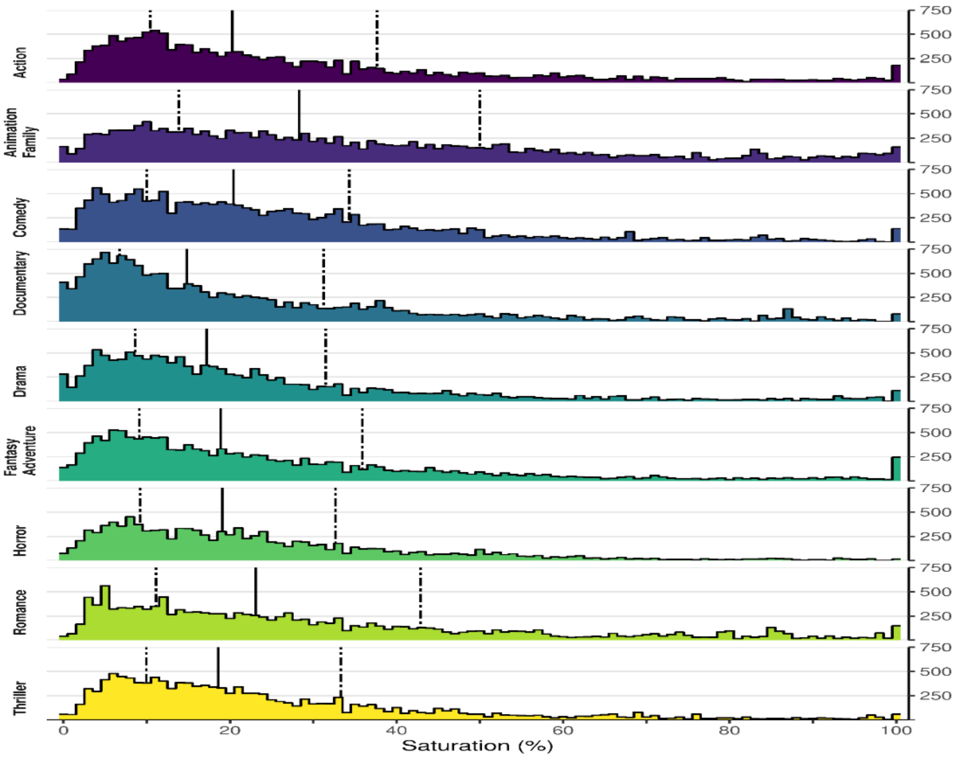

Figure 4: The distribution of saturation values in the HSL colour space for smoothed barcodes of 173 trailers across nine genres, with the median (solid line) and lower and upper quartiles (dashed lines) indicated. Pixels classed as black (L ≤ 5%) or white (L ≥ 9%).

Hue

Figure 5 plots hue against lightness and Figure 6 plots hue against saturation in the HSL colour space for each genre in the sample. It is clear from these plots that two colours dominate all genres, with points clustered around the complimentary colours orange (hue = 30° ± 30°, ranging from red to brown) and azure (hue = 210° ± 30°, ranging from cyan to blue).[4] The orange-teal look of Hollywood films has long been noted ([14]; [17]), though the results presented here suggest a bluer hue with less green is a key component of colour in these trailers rather than teal (H = 180°, S = 100%, L = 25%; see Figure 2), which marks the lower limit of the azure region in Figure 5 and Figure 6. That these complementary colours are dominant across all genres in the sample indicates a common approach to colour in Hollywood cinema irrespective of the genre of film promoted. Two types of trailers contribute to the distribution of hues in a genre: those in which both orange and azure occur, as they do in the case of *Escape Room* – 'Boxes' (Figure 7.C), or those dominated by one orange or azure such as *Mad Max: Fury Road* – 'Survive' (Figure 7.D).

Pixels in the green (hue = 120° ± 30°) and magenta (hue = 300° ± 30°) regions are much rarer. The barcodes for trailers in the action, drama, fantasy/adventure, horror, romance, and thriller categories have some green pixels but not magenta hues. The barcodes for drama film trailers, in particular, have notably fewer pixels between 240° and 345° than those in other genres. In contrast, the barcodes for trailers in the animation/family, comedy, and documentary genres tend to have more pixels in the green and magenta regions, though it is still the case that green hues tend to be less saturated than magenta hues. Furthermore, when pixels do occur in these two regions, the saturation tends to be below 25%, unlike the orange and azure regions where a much broader range of saturation values is used. Green hues tend to be unsaturated, and vivid greens are much rarer and tend to come from the barcode of a single trailer. For example, the light saturated lime-greens with hues between 153° and 158° in the horror genre come from the barcode for *Insidious: The Last Key* - 'Family', where this colour comes from point-of-view shots seen as though through night-vision goggles. The magenta regions for both animation/family and comedy trailers use a broader range of saturation values than other genres, and in the latter case show a shift towards the redder pole of the colour space and away from azure for highly saturated colours without a corresponding increase in complementary saturated greens. One feature notable in Figure 6 is the absence of red hues with saturation in the range 30% to 75% in barcodes from the drama trailers, indicating that this genre has a much more restricted range of hues than other genres in the sample.

4 Colour names are taken from https://www.colorhexa.com, assuming that the saturation is 100% and the lightness is 50%.

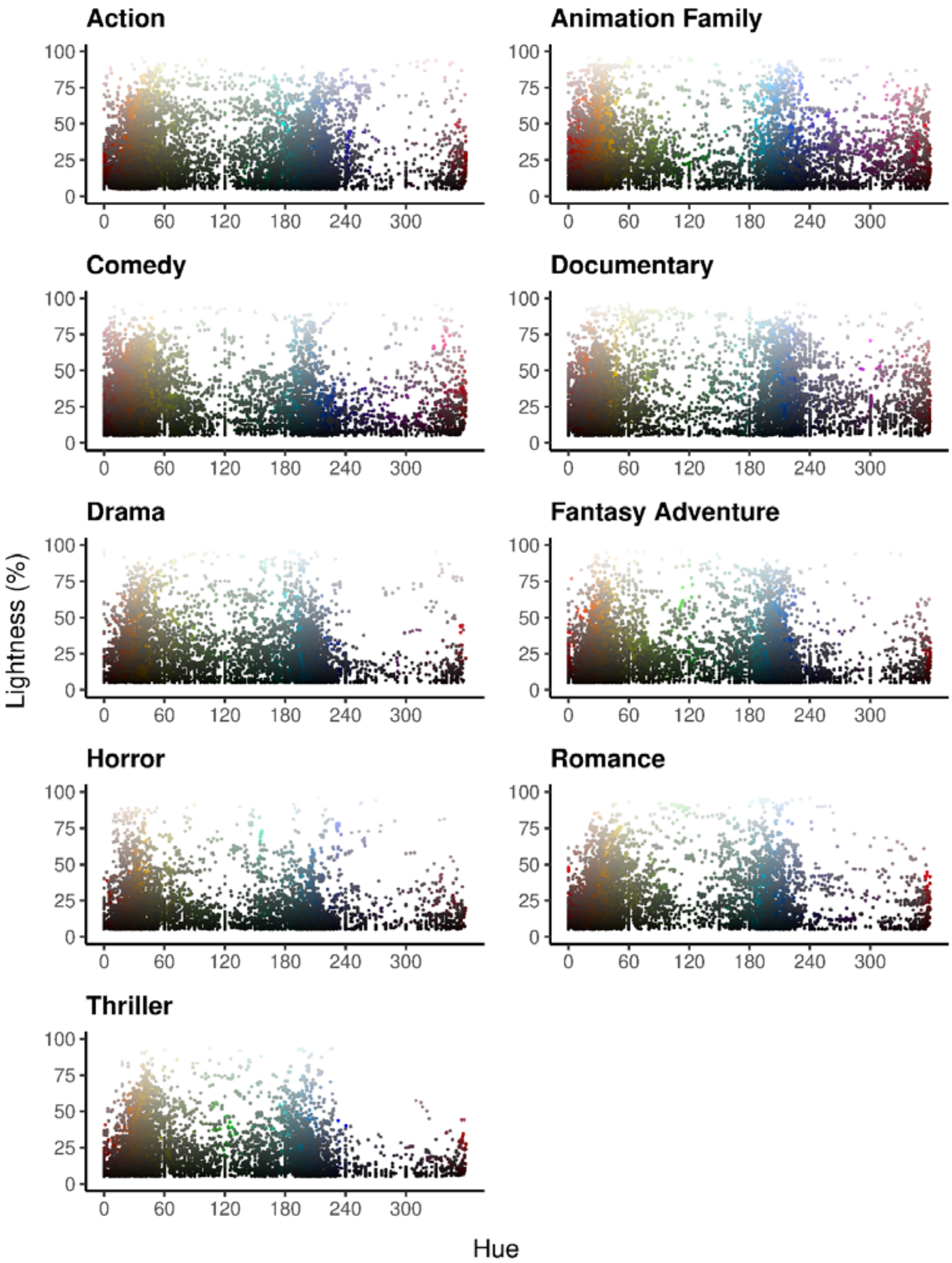

Figure 5: Hue v lightness in HSL colour space for smoothed barcodes of 173 trailers across nine genres. Pixels classed as black (L ≤ 5%), white (L ≥ 96%), and grey (S ≤ 5%, 5% < L < 96%) have been filtered out.

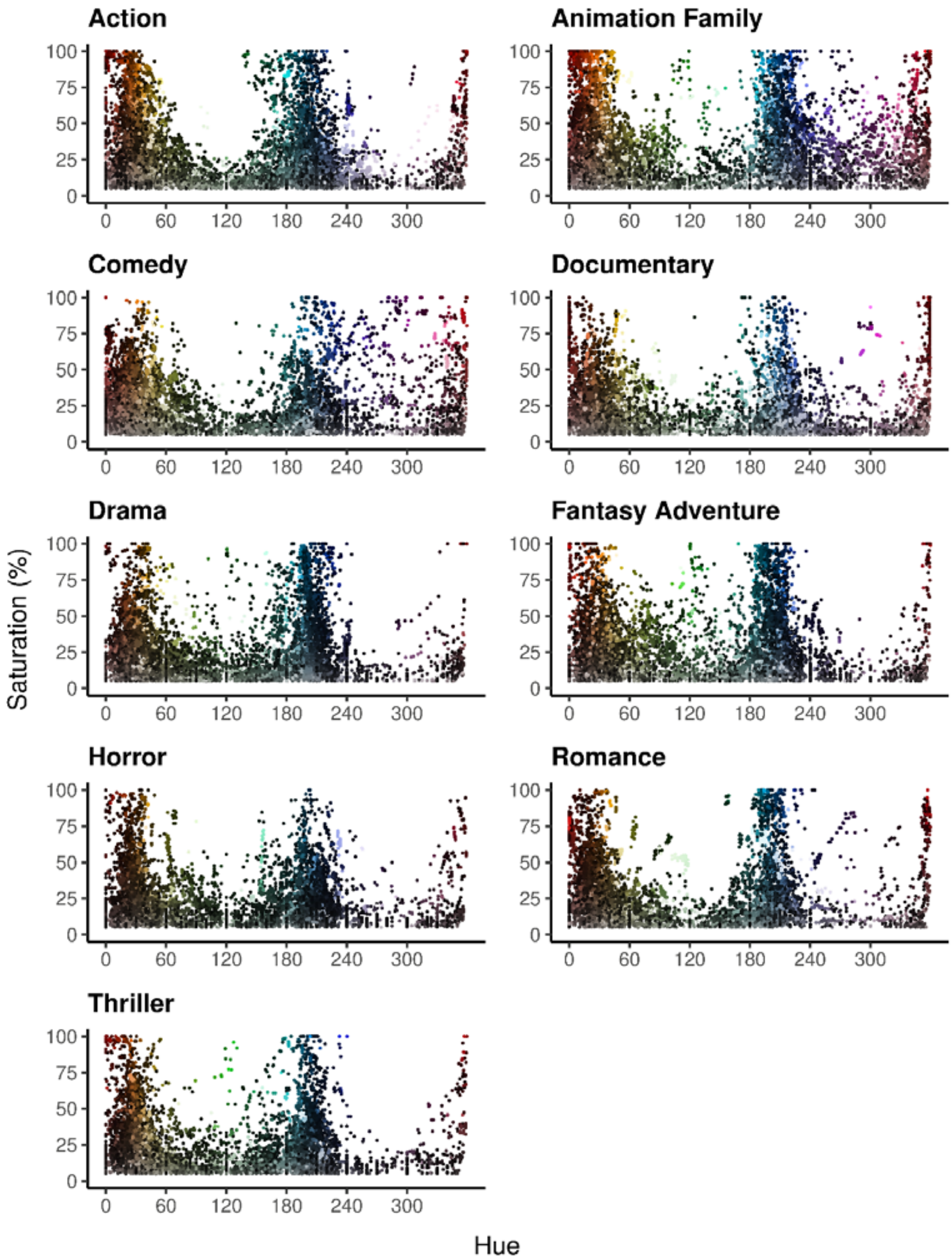

Figure 6: Hue v saturation in HSL colour space for smoothed barcodes of 173 trailers across nine genres. Pixels classed as black (L ≤ 5%), white (L ≥ 96%), and grey (S ≤ 5%, 5% < L < 96%) have been filtered out.

Figure 7: Smoothed barcodes for four trailers. (A) The dominance of unsaturated colours in Christopher Robin – 'Into the Wood.' (B) Bright marketing sections in Love Simon – 'Courage.' (C) Orange and azure hues in Escape Room – 'Boxes.' (D) Monochromatic orange hues in Mad Max Fury Road – 'Survive'.

Generic palettes

From the plots of hue against saturation and lightness in Figure 5 and Figure 6 it can be difficult to distinguish the degree to which pixels in a barcode have a particular colour due to the tendency of points to be overlaid on one another. The palettes generated by applying cluster analysis to the unfiltered data from the barcodes in each genre shows the extent to which pixels belong to a particular colour.

The palettes of the barcodes in each genre are dominated by the same colours – blacks, greys, and dark, desaturated browns and blues. However, despite the palettes being largely comprised of the same colours, there is variation between the palettes arising from the extent to which they are dominated by single cluster and the presence or absence of particular hues. The dominant cluster in each palette is black, due in part to the tendency of some trailers to fade to black at the end, but the range of the size of this cluster is quite large. In the horror barcodes, this large cluster of black pixels accounts for 44.2% of all pixels, while in the barcodes for the comedy trailers it accounts for only 16.4% of pixels. The relatively large number of black pixels in the animation/family palette (30%) largely derives from the trailers of the live-action Disney remakes included in the sample, which, as noted above, tend to be much less colourful than other trailers in this genre. Most of the variation in the palettes arises from the differing proportions of browns, blues, and greys, though in each case, they tend to be the same light and dark browns and greys, and the same cold, metallic blues in each genre. The palettes for drama and fantasy/adventure trailers have more green hues than the other genres, while action and romance trailers have more reds. Light, saturated colours account for a much smaller percentage of pixels across the sample, and in the case of the horror genre are completely absent. The animation/family, comedy, documentary, and romance genres stand out as having more brightly coloured pixels than other genres, but these clusters represent for only a small proportion of the total number of pixels, accounting for between 6% and 8% of the total.

Looking at the richness of the palettes in Figure 8 and the statistical summary in Table 1 we see two groups emerge. Four genres – comedy, documentary, drama, and fantasy/adventure - have palettes with between 17 and 20 colours, while the remaining five genres have between 12 and 14 colours. Surprisingly, animation/family trailers belong to the group with smaller palettes due to the fact that this genre has several relatively large clusters of saturated colours and relatively fewer clusters comprising brown and grey colours indicating less variation between pixels. Comedy trailers, by contrast, have the richest palette due to greater variation among brown colours, though this genre too has some clusters of highly saturated, light colours. With 19 colours, drama also has a rich palette despite having a more restricted range of hues than other genres. Variation in this genre arises from more fine-grained distinctions between a limited set of colours, unlike the comedy genre where there is greater variation in hues. The horror and thriller genres have essentially the same palette of 12 colours, albeit with those colours in differing proportions: the thriller trailers tend to have more brown pixels while the horror trailers tend to have more blue pixels, accounting for the difference in the distribution of warm/cold pixels in these genres. A key difference between the two genres lies in their greens, with horror trailers having much darker hues than the thriller trailers which have a small number of lighter and more saturated greens, but this accounts for a relatively small number of pixels (<3%) in each case.

From the indices presented in Table 1, we see that, despite the fact the different palettes in Figure 8 exhibit varying degrees of diversity, they tend to have similarly even distributions of pixels across their colours once the number of colours in a palette is accounted for, with most values of J' between 0.82 and 0.88 and of S_E between 0.44 and 0.54. Comedy has the most diverse palette, being far less dominated by its most prevalent colour and containing several small clusters of rarely occurring colours. In contrast, the least diverse palette is the horror genre, which, as noted above, is dominated by a single colour and has fewer small clusters than other palettes. The palette for the thriller trailers is more diverse and more evenly distributed than those in the horror genre despite the similarities in the colours of the palettes of these genres. Although the Shannon entropy of the thriller palette is similar to that of the horror genre because it is also dominated by a single category that accounts for 35.5% of pixels, the size of the other clusters in its palette is relatively even. This is reflected in the greater value of the S_E index. Despite the richness of its palette, the indices for the fantasy/adventure genre indicate that its palette is dominated by a small number of colours resulting in an S_E value similar to that of the horror genre. The drama palette, by contrast, shows much greater diversity and evenness given the richness of its palette. The difference between the indices of these two genres demonstrates that the richness of a colour palette can be misleading and that it is necessary to take diversity and evenness into account when comparing genres.

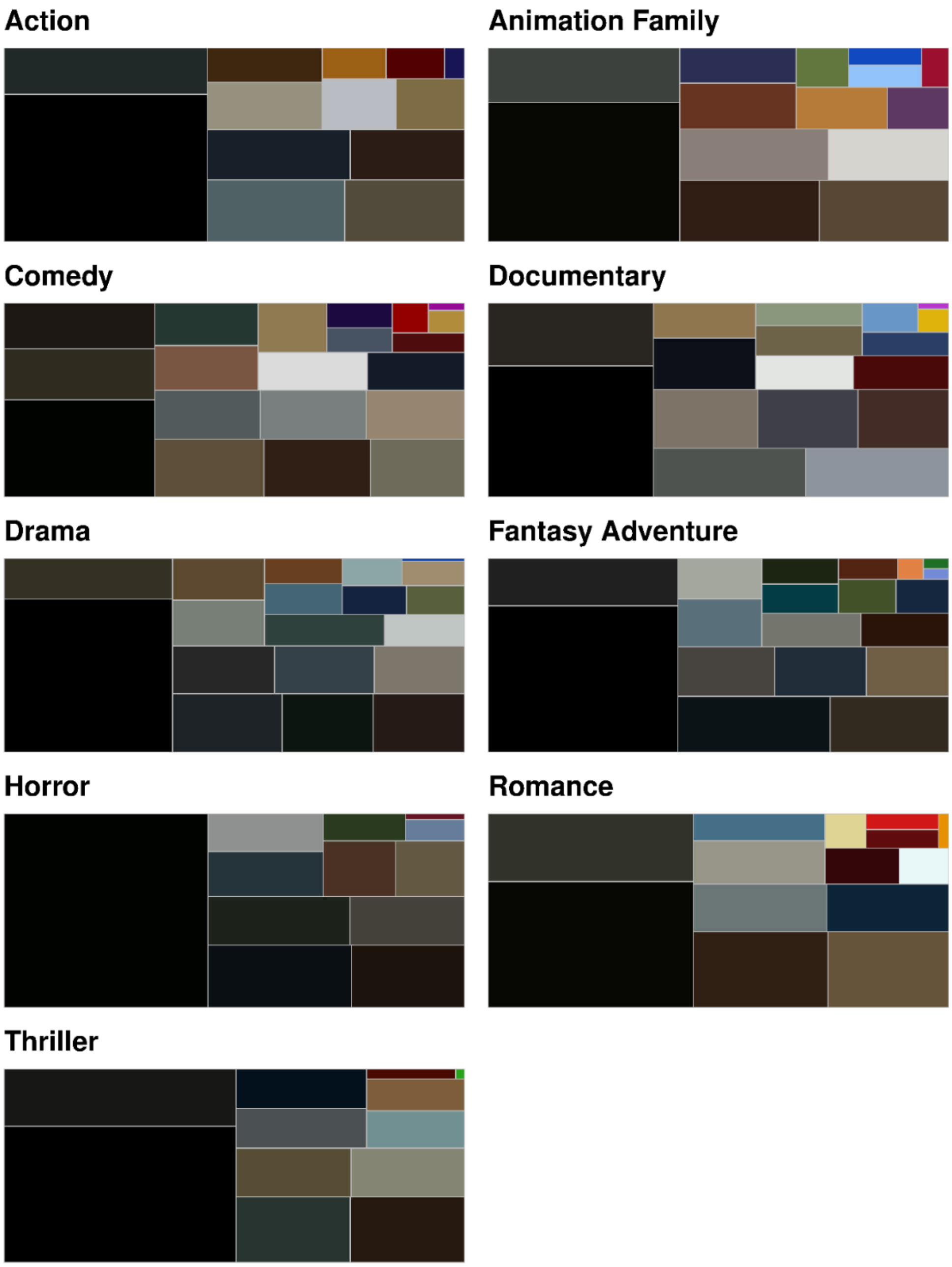

Figure 8: Treemaps of colour palettes for 173 barcodes of US film trailers across nine genres.

Genre	k	H	J'	S'	S_E
Action	13	3.15	0.85	6.26	0.48
Animation Family	14	3.27	0.86	7.05	0.50
Comedy	20	3.98	0.92	13.36	0.67
Documentary	17	3.60	0.88	9.18	0.54
Drama	19	3.68	0.87	8.40	0.44
Fantasy Adventure	19	3.50	0.82	7.29	0.38
Horror	12	2.78	0.78	4.31	0.36
Romance	14	3.16	0.83	6.75	0.48
Thriller	12	2.95	0.82	5.56	0.46

Table 1: The number of colours in a genre's palette (k), Shannon entropy (H), Pielou's evenness index (J′), and Simpson's diversity (S′) and evenness (SE) indices. (NB: figures were rounded to two decimal places after calculation of the indices).

Conclusions and future work

In this article I analysed the colour of smoothed barcodes for 173 trailers for Hollywood films across nine genres. The results show that the trailers in this sample are similar in their colour palettes. In each genre, the same two orange and azure regions of the HSL colour wheel dominate, and although there are variations in the green and magenta regions for some genres, these differences are relatively small. Furthermore, the colours in each genre are predominantly dark and unsaturated. The colour palettes of the different genres are similar, with the key difference between them being the evenness of the palette. Colours in trailers for the horror and thriller trailers show many similarities to the extent that they could be considered a single genre, though there is a difference in the evenness of their palettes. A similar relationship is evident between the action and fantasy/adventure genres.

As a trailer inherits its colour scheme from its reference film, these results provide some limited information about the use of colour in different film genres. However, the analysis presented here does not generally support previous research on colour and genre in film. Chen, Wu, and Lin ([7]) analysed colour in five genres, noting that romance and comedy films were brighter overall and richer in red and yellow hues than horror, science fiction, action films. While I found that comedy trailers generally do have greater lightness than other genres and that horror trailers do tend to be darker, there were no large differences between the other genres and no large differences in the distribution of hues. Existing studies of genre and colour in film have focussed on emotion as a key mediating variable, but this may not be the determining factor when it comes to the colours of a trailer. The screens presenting marketing information to the viewer are often made up of blocks of bright that are not typical of the general colour scheme of the trailer, while setting also determines the palette of trailer to a significant degree. These features are rarely accounted for when classifying films based on their colour and may lead to key differences being overlooked: for example, the darkness of live-action Disney films would suggest they are closer to the drama genre that to other films in the animation family genre.

Future research will need to expand in two dimensions: *horizontally*, to look at how trailers for US films in these genres compare to those trailers nominated for a Golden Trailer Award in the international categories in order to determine the extent to which colour palettes differ between, for example, action movies produced in Hollywood or elsewhere; and, *vertically*, to track changes in colour over time in order to determine if the declining trend in the luminosity of films is evident in their trailers or not.

References

[1] Arnold, Taylor, and Lauren Tilton. 2019. “Distant Viewing: analyzing large visual corpora,” *Digital Scholarship in the Humanities* fqz013. DOI: https://doi.org/10.1093/digitalsh/fqz013

[2] Barthelme, Simon. 2020. imager: Image Processing Library Based on 'CImg'. R package version 0.42.3. https://CRAN.R-project.org/package=imager

[3] Bhardwaj, Nalin. 2020. “Shadows of stained glass: an analytical look at animated horror.” *Journal of Geek Studies* 7, no. 2: 61-78.

[4] Buonocore, Tommaso. 2019. “Exploring chromatic storytelling in movies with R.” towardsdatascience.com, 8 April 2019. https://towardsdatascience.com/exploring-chromatic-storytelling-with-r-part-1-8e9ddf8d4187

[5] Burghardt, Manuel, Katharina Hafner, Laura Edel, Sabrin-Leila Kenaan, and Christian Wolff. 2017. “An Information System for the Analysis of Color Distributions in Movie Barcodes.” Humboldt-Universität zu Berlin.

[6] Burghardt, Manuel, Michael Kao, and Niels-Oliver Walkowski. 2018. "Scalable MovieBarcodes–An Exploratory Interface for the Analysis of Movies." In IEEE VIS Workshop on Visualization for the Digital Humanities.

[7] Chen, I-Ping, Fang-Yi Wu, and Chih-Hsiang Lin. 2012. "Characteristic color use in different film genres." *Empirical Studies of the Arts* 30, no. 1: 39-57. https://doi.org/10.2190/EM.30.1.e

[8] Cutting, James E, Kaitlin L Brunick, Jordan E DeLong, Catalina Iricinschi, and Ayse Candan. 2011. “Quicker, Faster, Darker: Changes in Hollywood Film over 75 Years.” *I-Perception* 2, no. 6: 569–76. https://doi.org/10.1068/i0441aap

[9] Dorai, Chitra, and Svetha Venkatesh. 2002. "Bridging the semantic gap in content management systems." In *Media Computing: Computational Media Aesthetics*, edited by Chitra Dorai and Svetha Venkatesh, 1-9. Springer, Boston, MA.
[10] Feng, Ping Feng. 2016. "Examination of the Hollywood Movie Trailers Editing Pattern Evolution over Time by Using the Quantitative Approach of Statistical Stylistic Analysis." MA diss., Temple University.
[11] Flueckiger, Barbara, and Gaudenz Halter. 2020. "Methods and Advanced Tools for the Analysis of Film Colors in Digital Humanities." *DHQ: Digital Humanities Quarterly* 14, no. 4: http://www.digitalhumanities.org/dhq/vol/14/4/000500/000500.html.
[12] Golden Trailer Awards (2020) "Home." Accessed February 15, 2021. https://goldentrailer.com/home/
[13] Halter, Gaudenz, Rafael Ballester-Ripoll, Barbara Flueckiger, and Renato Pajarola. 2019. "VIAN: A visual annotation tool for film analysis." *Computer Graphics Forum* 38, no. 3: 119-129. https://doi.org/10.1111/cgf.13676
[14] Hart, Hugh. 2013. "Infographics: charting Hollywood's odd obsession with orange and teal." *Wired*, 23 January 2013. https://www.wired.com/2013/01/orange-teal-infographics/
[15] Heip, Carlo HR, Peter MJ Herman, and Karline Soetaert. 1998. "Indices of diversity and evenness." *Oceanis* 24, no. 4: 61-88.
[16] Hennig, Christian. 2020. "fpc: flexible procedures for clustering (Version 2.2-9)." http://cran.r-project.org/web/packages/fpc/index.html
[17] Hoad, Phil. 2010. "Hollywood's new colour craze." *The Guardian*, 26 August 2010. http://www.guardian.co.uk/film/2010/aug/26/colour-grading-orange-tealhollywood
[18] Hohman, Fred, Sandeep Soni, Ian Stewart, and John Stasko. 2017. "A viz of ice and fire: Exploring entertainment video using color and dialogue." In Proceedings of the 2nd Workshop on Visualization for the Digital Humanities (VIS4DH), Phoenix, Arizona.
[19] Isola, Phillip, Daniel Zoran, Dilip Krishnan, and Edward H. Adelson. 2015. "Learning visual groups from co-occurrences in space and time." arXiv preprint arXiv: 1511.06811.
[20] Kaufman, Leonard, and Peter J. Rousseeuw. 2009. Finding Groups in Data: An Introduction to Cluster Analysis. Chichester: John Wiley & Sons.
[21] Klinke, Harald. 2020. "The Digital Transformation of Art History." In *The Routledge Companion to Digital Humanities and Art History*, edited by Kathryn Brown, 1st ed., 32–42. New York : Routledge. https://doi.org/10.4324/9780429505188-5
[22] Kuhn, Virginia, Ritu Arora, Alan Craig, Kevin Franklin, Michael Simeone, Dave Bock, and Luigi Marini. 2012. "Large Scale Video Analytics: On-demand, iterative inquiry for moving image research." In 2012 IEEE 8th International Conference on E-Science, 1-5. https://doi.org/10.1109/eScience.2012.6404446
[23] Laïly, Melvin. 2020. "Movie Barcode Generator, version 1.6.0.0." https://zerowidthjoiner.net/movie-barcode-generator
[24] Moine, Raphaëlle. 2008. *Cinema Genre*, trans. Alistair Fox and Hilary Radner. Malden, MA: Blackwell.

[25] Otto, Isabella, Alice Plutino, Matteo Lanaro, and Alessandro Rizzi. 2018. "All the colours of a film: study of the chromatic variation of movies." In Proceedings of the International Colour Association (AIC) Conference 2018, Lisbon, Portugal.

[26] R Core Team. 2020. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org.

[27] Redfern, Nick. 2021. "US_Film_Trailers_Barcodes_Sample." Zenodo, 18 February 2021. https://doi.org/10.5281/ZENODO.4548435

[28] Schubert, Erich, and Peter J. Rousseeuw. 2019. "Faster K-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms." In *Similarity Search and Applications*, edited by Giuseppe Amato, Claudio Gennaro, Vincent Oria, and Miloš Radovanović, 11807: 171–87. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-32047-8_16

[29] Tkalcic, M., and J.F. Tasic. 2003. "Colour Spaces: Perceptual, Historical and Applicational Background." In The IEEE Region 8 EUROCON 2003. Computer as a Tool., 1:304–8. Ljubljana, Slovenia: IEEE, 2003. https://doi.org/10.1109/EURCON.2003.1248032

[30] Zheng, Xin, Qinyi Lei, Run Yao, Yifei Gong, and Qian Yin. 2018. "Image segmentation based on adaptive K-means algorithm." *EURASIP Journal on Image and Video Processing* 68. https://doi.org/10.1186/s13640-018-0309-3