QUANTITATIVE METHODOLOGY FOR THE ANALYSIS OF COLOUR AND ITS MEANING: AN APPLICATION TO CAREER GIRLS

LAURA CORTÉS-SELVA

Universidad Católica San Antonio de Murcia
Spain

lcortes@ucam.edu

M.MERCEDES CARMONA-MARTÍNEZ

Universidad Católica San Antonio de Murcia
Spain

mcarmona@ucam.edu

Keywords: Colour, Cinema Style, Visual Style, Quantitative Methodology

Several researchers are dealing with colour, as a basic expressive element of cinema style. Nevertheless, the difficulty in analyzing a large number of images and describing their nature and characteristics in non subjective terms are still among the problems faced in this area.

This research presents a quantitative methodology that enables a systematic, objective and verifiable study of colour, by automatically obtaining accurate colour data from large amounts of images. The proposed scheme also allow us to find possible meanings when relating quantitative information obtained to the narrative content.

In order to test its efficiency, the methodological tool is applied to Career Girls (Mike Leigh, 1997). As a result, hue is found to follow an expressive program evolving within the narrative content, as it is closely linked to the different periods present in the film.

xCoAx 2014

2014.xCoAx.ORG PORTO, PORTUGAL COMPUTATION, COMUNICATION, AESTHETICS AND X

1. INTRODUCTION

Nowadays, image analysis is required in any communication process in the field of journalism, advertising or media, and images are composed of a great number of expressive elements among which, colour stands out due to its high communicative possibilities.

Specifically, studying movie images colour from the perspective of the creators'/authors' decisions and how they are printed on a film, enables access to information which doesn't depend upon viewers, but stays invariable throughout time as the legacy of an author. Those decisions may indicate the presence of some descriptive elements that become patterns, which enable to detect a possible style related to one or several works. Similarly, those decisions may also show possible meanings connected to the use of colour in one or several films.

Numerous scholars have dealt with colour study (Street, 2012; Misek, 2010; Dalle Vache and Price, 2006; Bordwell, 2006, 1998; Bordwell *et al.*, 1997; Haines, 1993) basing their research of movie images on subjective meanings. But Edward Branigan (2006: 170) state that colour has no identity by itself but acquires meaning through its consistency with other elements of the film *mise en scène* such as space, time or characters, and that it is possible to demonstrate those links both numerically and visually.

One of the most innovative and effective methodologies for image analysis was proposed by Barry Salt (1992, 2006) in his stilometry research. Later on, other scholars such as David Bordwell (*http://www.cinemetrics.lv/bordwell.php*), Warren Buckland and Thomas Elsaesser (2002), Charles O'Brian (*http://www.cinemetrics.lv/obrien.php*) and Yuri Tsivan (*http://www.cinemetrics.lv/tsivian.php*) apply and develop this empirical methodology, as can be seen in the collective project entitled Cinemetrics.

Salt uses statistics to obtain an average of several quantifiable cinema variables that depend directly upon the cinema director, such as shot scale, shot duration, camera movements, reverse angle cuts, point of view shots and insert shots. Although in his early research Salt analyzes each film as whole, years later he chooses a sample of the first thirty or forty minutes of each film, which he considers representative of the complete film. Then, more recently, he chooses shots susceptible to analysis in a sequential way throughout the film. (*http://www.cinemetrics.lv/salt.php*).

This statistical-type methodology faces a specific problem when applied to the analysis of colour in films: the need of processing and obtaining accurate colour data from large amounts of images.

Introduction and constant improvement of digital technology enables to obtain, in real time, substantial, objective quantitative data which allow a deeper understanding of visual data (Manovich and Douglas, 2009). In this regard, David M. Berry (2011) notes that digital technology is fundamentally changing research procedures, especially those traditional methods of analysis, and transforming our ability to use and understand information in a different way from traditional knowledge structures.

Therefore, the important work carried out by the California University Department of Visual Arts research group led by Lev Manovich must be considered. They have developed certain methods and techniques – named Cultural Analytics – that use digital tools to create visualizations that allow for the exploration and research of large quantities of visual material (*http://lab.softwarestudies.com/2008/09/cultural-analytics.html*).

This research shows that it is possible to objectively define and measure colour in cinema, and find possible meanings when linking it to the narrative content. To do so, we present a quantitative methodological tool that enables to measure colour variables of a vast array of images, and to automatically obtain objective, verifiable and generalizable results. Moreover, with this tool it is also possible to show visual and numerically what the colour is like, its evolution throughout narrative content development, and how it is related to other variables such as narrative time, specifically the different periods that coexist inside the film.

In some contexts the audience can certainly perceive that there is a certain hue in a scene, but the description of that hue will be subjective and not accurate enough for a scientific approach. Which kind of hue? To which extent can we make a difference between blue and blue-greenish (both of them subjective terms)? And how can you evidence in an objective way that there is a certain hue in a scene? This quantitative approach allows a systematic, verifiable and objective methodology that yields unbiased conclusions, not dependent of the watcher's subjective perception.

In this paper, such methodology is presented through its application to the analysis of *Career Girls* (Mike Leigh, 1997). Hue is analyzed, along with its evolution throughout the film and the significance it acquires when linked to the variable of time, or more precisely, period.

2. MATERIALS AND METHODS

The film *Carrer Girls* is available on DVD. Files in this format are .m2v, which although involves a certain colour compression compared to celluloid, can easily be handled with simple software, and allows for a simple comparison of results with other films in the same format.

Although nowadays the processing power of computers allow to automatically apply this methodology to each of the frames of a film, there is still the necessity of manually choose and cut each of those frames. In the case of *Career Girls*, that means 4860 frames, so it would require 20 times more than the 243 randomly chosen shots.

In statistical analysis, it is quite usual to select a subset of individuals from within a population to estimate characteristics of and describe the whole population (this is referred to as "inference"). This is a time-saving practice that ensures scientific general results and is widely used in all scientific areas.

Simple random sampling is one of the easiest unbiased sampling methods; moreover it offers a sample that is highly representative of the population – with an *a priori* limited sampling error –, and simplifies analysis of statistical results. Using other sampling methods, such as stratified or cluster sampling, can even improve sample representativeness, but they require further information[1] (auxiliary variables or covariates).

2.1. SAMPLE SIZE AND SAMPLING METHODS

As in any inferential statistical analysis, optimum sample size must be determined, taking into account that that the minimum unit in our research is each shot, and the film has a total number of 966 shots. Then, considering a sample error of 5.44%, the optimal sample is composed of 243 shots.[2]

In order the 2403 shots in the sample to have a homogeneous distribution, the film has been divided into four proportional parts or quartiles, according to its duration in minutes: the first division (from 0 to 20 minutes) includes the first 250 shots of the sample, the second division (from minute 21 to 41) includes from shots 251 to 500.

1 Cf.: Smith (2001); Chambers and Skinner (2003).

2 As we have a finite and known universe, we follow this formula (Pérez, 2009):

$$n = \frac{Z^2 \times N \times p(1-p)}{(N-1) \times e^2 + Z^2 \times P(1-p)}$$

Where "n" is the optimum sample size; "N" is the film total number of shots; "Z" is the Normal distribution value that guarantees a 95% confidence level (so Z=1.96), which is the probability that results obtained are true; "p" is the population percentage that has certain characteristics under study (as it is an unknown value, it is assumed that p= 0.5); and "e" is the considered sample error (e=0.0544), i.e. the maximum accepted difference between results obtained by sample analysis and those obtained if the whole film would have been analyzed.

In the third division (from minute 42 to 61) are included shots from 501 to 750, and finally, fourth division includes from 751 to 966 shots (from minute 62 to 81).

Then, a representative frame of each of those 243 shots has been randomly chosen and saved as a .tiff picture.

Afterwards, analysis of formal variables related to colour in each selected frame is carried out employing ImagePlot[3] and its macro ImageMeasure, that transforms images in a RGB colour model into an HSB one, and extract from them statistical data related to hue, among other variables (*http://lab.softwarestudies.com/p/imageplot.html*).

2.2. MEASUREMENT

For performing the quantitative analysis, two variables have been considered, one related to colour and one variable related to time: hue and period.

Image hue measurements are obtained from ImageMeasure on a 0 to 255 scale, but they have been classified along a scale from 0° to 360°, which coincides with the chromatic circle. These results were afterwards also transformed in a qualitative scale including six categories (red, orange-yellow, green, cyan, blue and magenta), according to Table 1.

Results related to hue were also included in a broader scale, which distinguishes between cold and warm hues, according to Table 2.

The temporal variable period is related to narrative time, and has been determined by observation of each one of the chosen shots. In this case, *Career Girls* is composed by flashbacks sequences (representing previous decades in relation to the shooting time, and labeled as "past") and other related to present time (labeled as "present", and in which narrative time coincides with the time of the shooting).

2.3. STATISTICAL METHODOLOGY

Once data related to each variable is obtained, SPSS[4] v.21 software is used to perform a statistical analysis on two levels. The first analysis is based on descriptive statistics and examines the basic behavior of the colour variable, as well as its progress in each quartile and how it develops throughout the film. Quantitative variable information refers to mean, median, standard deviation, minimum and maximum; while qualitative variable information refers to percentages values offered by contingency tables.

Table 1 Hue subdivision codes

Scale (grades)	Hue
330° - 30°	Red
31° - 90°	Orange - Yellow
91° - 150°	Green
151° - 210°	Cyan
211° - 270°	Blue
271° - 330°	Magenta

Table 2 Warm and cold hue codes

Scale (grades)	Hue
315° - 134°	Warm Hues
135° - 314°	Cold Hues

3 ImagePlot is a free software tool, developed by Software Studies Initiative, which enables to create high resolution two dimensional visualizations from a large quantity of images. It is implemented as a macro, which works with ImageJ64 (*http://rsb.info.nih.gov/ij/docs/intro.html*), an open source image-processing program.

4 SPSS is a proprietary software tool, developed by IBM Corporation, for statistical analysis in social science.

When the descriptive stage of analysis is complete, a second multivariate analysis based on inference statistics is performed in order to study possible links between hue and period, and therefore, a feasible meaning of colour.

Since none of the quantitative variables in our study satisfies the assumption of normality,[5,6] Kruskal-Wallis contrast[7] is used to test if there is any relationship between the qualitative (period) and the quantitative variable (hue).

3. RESULTS

In this section, results obtained from the application of the methodological tool are shown.[8] In the first place, the descriptive study of hue is presented, in the second place, inference analysis will show if there is a relationship between hue and period.

3.1. DESCRIPTIVE ANALYSIS

Table 3 shows that *Career Girls* has a mean hue of 134,70°. Furthermore, the highest values for hue are in the second and first quartile.

With regards to hue evolution throughout the narrative development in *Career Girls*, it grows from the first to the second quartile and diminishes in the other two consecutive quartiles in a progressive way.

Table 3 Average hue quartile percentages

	1st Quartile	2nd Quartile	3rd Quartile	4th Quartile	Mean
Mean	136.65	180.82	123.76	97.57	134.70
Standard Deviation	57.36	52.30	67.84	47.29	56.20
Minimum	58	61	35	40	49
Maximum	229	223	313	202	242

The scatterplot in Fig. 1 represents the evolution of average hue in each of the shots analyzed in *Career Girls*. The "X" axis is divided into four quartiles according to the film duration in minutes, while the "Y" axis represents the average hue of each shot on a 0° to 360° scale. This figure visually shows the film structure according to the hue of the shots.

It also offers a clear pattern of the mean hue distribution, visible in all the quartiles, especially in the first and in the second one. The first quartile is divided into two

5 The assumption of normality implies that data offer a symmetrical distribution with a single mode.

6 To verify this Kolmogorov-Smirnov tests were applied.

7 Kruskal-Wallis is a non-parametric test to contrasts if quantitative variable behavior is the same in each one of the groups defined by qualitative variables, considering independent and more than two. (Siegel and Castellan, 1988: 206-215; Barreiro *et. al.* 2006: 133-136).

8 Due to space limitations, only a part of what this methodological tool is capable of , is shown. In a more extended research – which is part of a PhD (*https://www.educacion.gob.es/teseo/imprimirFicheroTesis.do?fichero=41358*) – evidence is presented to support the relationship between hue and the following variables: brightness, saturation, shot length, shot scale, camera movement, location, exterior, interior, day and night and characters.

parts: the first one is dominated by cold hues and the second one by warm hues; in the case of the second quartile, most of the hues are cold. The third quartile could be also divided into two parts: the first one is dominated by warm hues and the second one by cold ones. Finally, the forth quartile is mostly dominated by warm hues, an opposite tendency to the second quartile.

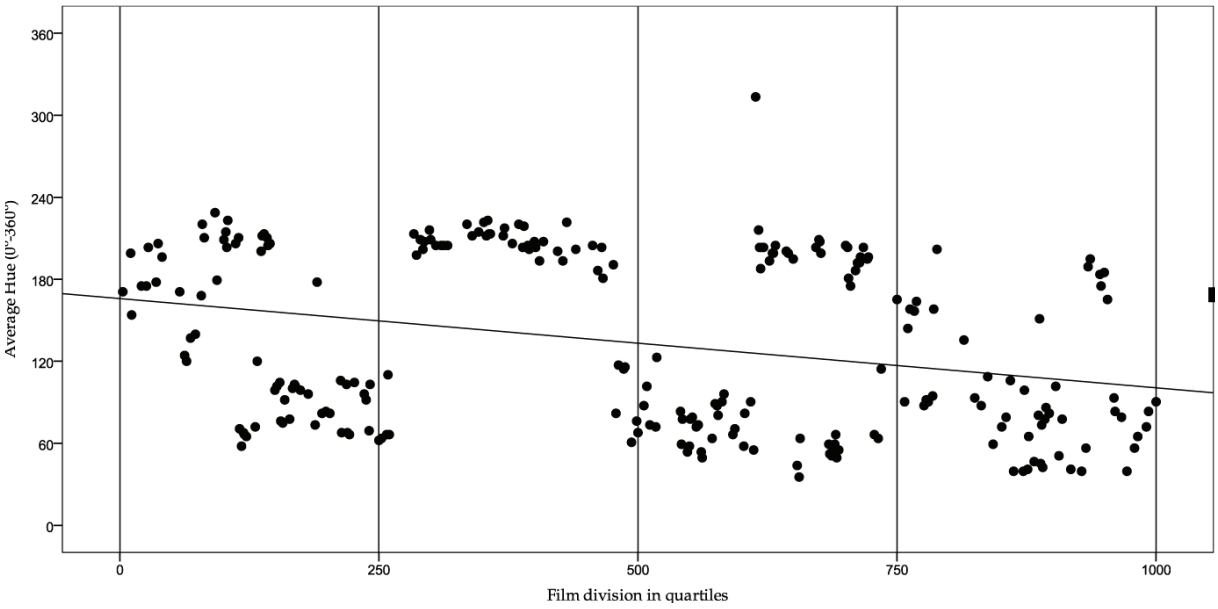

Fig. 1 Hue evolution throughout *Career Girls*

Table 4 shows the quartile-distribution of hue, considering the qualitative six-category scale for this variable.

Table 4 Hue scale quartile percentages

		1st Quartile	2nd Quartile	3rd Quartile	4th Quartile	Shots Total Number
Orange- Yellow	nº.of shots	18	7	38	32	95
	percentage	18.9%	7.4%	40.0%	33.7%	39.1%
Green	nº.of shots	20	4	4	10	38
	percentage	52.6%	10.5%	10.5%	26.3%	15.6%
Cyan	nº.of shots	23	25	27	12	87
	percentage	26.4%	28.7%	31.0%	13.8%	35.8%
Blue	nº.of shots	6	15	1	0	22
	percentage	27.3%	68.2%	4.5%	0,0%	9.1%
Magenta	nº.of shots	0	0	1	0	1
	percentage	0.0%	0.0%	100.0%	0.0%	0.4%

The most outstanding hue in *Career Girls* is orange-yellow with 39.10% of shots in the sample, followed by cyan with 35.80%. Green is next one with 15.60%, blue with 9.10%, magenta with 0.4% and red with 0%.

All the hues are present in all the quartiles with the exception of blue – present in three: first, second and third –, magenta – only present in the third one –, and red –not present in the sample –. Furthermore, the evolution of the different hues throughout the four quartiles is irregular, meaning that a behavioral pattern among them does not exist.

Table 5 shows the quartile-distribution of hue, considering now a binomial scale for this variable (cold-warm).

Except for the second quartile in which warm hues predominate, and in the fourth one in which cold hues prevail, there is a balance between cold and warm hues in the rest of the quartiles.

Table 5 Cold and warm hues quartile percentages scale quartile percentages

		1st Quartile	2nd Quartile	3rd Quartile	4th Quartile	Shots Total Number
Cold hues	No. of shots	36	11	42	40	129
	Percentage	53.7%	21.6%	59.2%	74.1%	53.1%
Warm hues	No. of shots	31	40	29	14	114
	Percentage	46.3%	78.4%	40.8%	25.9%	46.9%

3.2. RELATIONSHIP BETWEEN HUE AND PERIOD

In this section the relationship between the chromatic variable and the time-related variable is numerically shown. Finding statistically significant differences between the mean hue obtained for each of the categories of period ("past" and "present") implies a certain association between both variables.

Table 6 shows that in this film, shots related to a past time have an average hue of 202°, while shots related to the present time have an average hue of 85.51°.

The Kruskal-Wallis test performed shows (sig.= 0.000) that hue of each of the analyzed shots is related to the different temporal dimensions of the film. This means that in *Career Girls* there are statistically significant differences between present hue colour and past hue colour.

Table 6 Hue and period descriptive results

	Hue (°)		
	Mean	**Min.**	**Max.**
Past	202.05	154	313
Present	85.51	35	213
Total	133.47	35	313

4. CONCLUSIONS

Conclusions obtained in this research are divided into those related to methodology and those concerning film analysis.

With regards to the proposed methodological tool, it overcomes traditional analysis focused on a limited number of images, since it can cover a higher number of images. Furthermore, it offers an alternative and a complement of subjective analysis to describe colour and also to find its possible meanings.

Such analytical methodology enables the obtaining of data related to colour in a systematic, objective and verifiable way, as well as its visualization from a huge volume of images. As a consequence, it allows us to describe and depict colour, to observe its evolution throughout the film, and to find possible meanings when linking it to the narrative content.

Regarding conclusions about the film analysis, it is shown visually and numerically that, from a descriptive point of view, *Career Girls* has an average hue of 134,70°, which according to table 1 and 2, means that it is a green and warm film. Nevertheless, although *Career Girls* is, on average, a warm film, there is quite a similar proportion between warm and cold hues, with 53% and 47%, respectively.

Concerning hue scale present in the film, the lack of red is noteworthy. Then, from the lowest to the highest percentage, magenta shows only a 0.4%, blue a 9.1%, and green a 1.6% of the sample. Cyan is the next hue (35.8%), followed closely by orange-yellow (39.1%).

Besides, colour evolves with the narrative content of the film: if we take into account the beginning and the end of the film, it becomes warmer. Regarding colour meaning, and as a consequence of statistical inference, it is demonstrated that hue is closely related to the different periods present in the film, as it is possible to associate flashbacks with an average hue of 202° (cyan and cold), while those sequences related to the present have an average hue of 85.51° (orange-yellow and warm).

The described statistical methodology yields unbiased objective results that are basic for the explanation of how period (past and present) is connected to hue, and that´s a meaning for colour: blue means past and orange-yellow, present.

The methodological proposal can be applied to different fields such as art history, design, photography or cinematography, as directions of further research. In the field of cinematography, it allows for the study of the relationship between several chromatic variables (such as hue, brightness and saturation) and other narrative content variables (such as characters, space and time) in any film.

ACKNOWLEDGEMENTS

Authors are specially grateful to Barry Salt, who provided them with his valuable comments and useful remarks.

REFERENCES

Barreiro, A. *et al.* *Tratamiento de datos.* Madrid: Díaz de Santos, 2006.

Berry, D. M. The Computational Turn: Thinking about the Digital Humanities. *Culture Machine* 12: pp.1–22, 2011.

Bordwell, D. *On the History of Film Style.* Cambridge, MA: Harvard University Press, 1998.

———. *The Way Hollywood Tells It: Story and Style in modern movies.* Berkeley and Los Angeles: University of California Press, 2006.

Bordwell, D. *et al.* *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960.* New York: Columbia University Press, 1997.

Branigan, E. The Articulation of Colour in a Filmic System. In: Dalle Vacche A and Price B (eds) *Colour, the Film Reader.* New York: Routledge, pp.170–182, 2006.

Chambers, R.L. and C.J. Skinner (Eds). *Analysis of Survey Data.* London: Wiley, 2003.

Elsaesser, T. and W. Buckland. *Studying Contemporary American Film: A Guide to Movie Analysis.* London: Arnold, 2002.

Haines, R. *Technicolour Movies: The History of Dye Transfer Printing.* Jefferson, N.C.: McFarland & Co, 1993.

Manovich, L. and J. Douglas. *Visualizing temporal patterns in visual media*, 2009. Available at: http://softwarestudies.com/cultural_analytics/visualizing_temporal_patterns.pdf (accessed 8 April 2013).

Misek, R. *Chromatic Cinema.* Oxford: Wiley-Blackwell, 2010.

Pérez, C. *Técnicas de muestreo estadístico.* Madrid: Garceta, 2009.

Salt, B. *Film Style and Technology: History & Analysis.* London: Starword, 1992.

———. *Moving into pictures.* London: Starword, 2006.

Siegel, S. and J. Castellan. *Nonparametric Statistics for the Behavioural Sciences.* New York: McGraw Hill, 1988.

Smith, T. M. F. Biometrika centenary: Sample surveys. *Biometrika*, 88 (1), pp.167–243, 2001.

Street, S. *Colour Films in Britain. The Negotiation of Innovation 1900–55.* London: British Film Institute, 2012.