Lauri Heinonen

WHITE BALANCE QUANTIZATION WITH CLUSTERING ALGORITHMS

Faculty of Information Technology and Communication Sciences
Bachelor’s Thesis
May 2024

ABSTRACT

Lauri Heinonen: White balance quantization with clustering algorithms
Bachelor's thesis
Tampere University
Bachelor's Programme in Computing and Electrical Engineering
May 2024

Computer vision and machine learning are evolving applications in the field of information technology. Video and image recognition models utilize information about the shapes and colours of the environment given by the source image. However, image capturing presents a challenge in collecting the hue of colours which are often distorted by the illumination of the scenery. The human visual system can correct the distortions of shades to be realistic, regardless of the light source's hue. However, cameras must use a method known as white balancing to colour correct the image. This phenomenon is called colour constancy, and it is a challenge encountered in the field of image and video processing.

This thesis examines the possibility of finding a discrete set of white balance values for a series of photographs using various clustering algorithms and explores the differences between these methodologies. Clustering algorithms form centroids within a set of white balance points, and these centroids can be used as new white balance values for photographs. To achieve this, the error rate must fall below the pre-determined threshold, so the human eye cannot visually differentiate if the image has an unnatural hue.

The tests used two sets of 568 white balance values defined by different methods from the same image set. Error values in the study are determined by angular error, which is a commonly used metric in the relative field.

Three different clustering algorithms were experimented with in the thesis: K-means, Spectral, and Agglomerative. These models are from the open-source Scikit-Learn clustering library and operate in the Python programming language. A previous analysis of colour constancy estimation done by Google was used to evaluate the clustering algorithms for this task.

The results of the study indicate that clustering algorithms can create a new, smaller set of white balance values to replace a predefined set. Clustering for one of the sets provided results suggesting that 40 generated centroids could sufficiently replace 568 white balance values. The best results in the tests were obtained with the K-means algorithm, while the weakest algorithm was spectral clustering. However, the clustering models provided significantly different results for both sets of white balance values, and not all results could be considered realistic. Although the results obtained so far seem promising, further tests would be necessary to confirm the findings.

Keywords: computer vision, machine learning, colour constancy, clustering, white balance, white balance evaluation

TIIVISTELMÄ

Lauri Heinonen: Valkotasapainoarvojen määrittäminen klusterointialgoritmeilla
Kandidaatintyö
Tampereen yliopisto
Tieto- ja sähkötekniikan kandidaattiohjelma
Toukokuu 2024

Konenäkö sekä koneoppiminen ovat relevatteja sovellusmuotoja tietotekniikan alalla. Jatkuvasti kehittyvät video- ja kuvatunnistemallit hyödyntävät kameran antamaa tietoa ympäristön muodoista ja väreistä. Kameroissa esiintyy kuitenkin ongelma värien todellisten sävyjen tallentamisessa, johtuen kameran sensorin toimintamallista. Ihmisen näköjärjestelmä pystyy korjaamaan sävyjen vääristymät realistisiksi riippumatta valonlähteen sävystä, kun taas kameroiden täytyy hyödyntää valkotasapainoarvoja vääristymien korjaamiseen. Tätä kutsutaan värikonstanssiksi (eng. colour constancy) ja se onkin keskeinen ongelma kuvan ja videonkäsittelyssä.

Tässä tutkielmassa tarkastellaan mahdollisuutta löytää diskreetti määrä valkotasapainoarvoja erilaisilla klusterointialgoritmeilla sekä tarkastellaan eroavaisuuksia näiden menetelmien välillä. Työssä klusterointialgoritmit muodostavat ryppäitä valkotasapainoarvopisteistä. Näiden ryppäiden joukkoon muodostuvia pisteitä voidaan käyttää uusina valkotasapainoarvoina valokuville, virherajan laskiessa riittävän matalaksi.

Tutkimuksessa käytettiin kahta 568 valokuvan valkotasapainoarvojoukkoa, jotka olivat määritelty erilaisilla metodeilla samoista valokuvista. Työssä virhearvot määritellään alalla yleisesti käytössä olleen kulmavirheen avulla.

Tutkielmassa tarkasteltiin kolmea eri klusterointimenetelmää: K-means, Spectral sekä Agglomerative. Kaikki mallit ovat avoimen lähdekoodin Scikit-Learn-kirjastosta ja ne toimivat Python ohjelmointikielellä. Tulosten evaluaatioon hyödynnettiin esimerkiksi Googlen suorittamaa tutkimusta värivakion määrittämisestä.

Työn tulokset osoittivat, että klusterointialgoritmit pystyvät luomaan uuden sarjan valkotasapainoarvoja ja niillä korvaamaan ennalta määritetyn arvojoukon. Klusteroinnit toiselle datajoukolle antoivat tuloksia, jotka viittaavat jo 40 luodun ryppään antaman arvon riittävän korvaamaan 568 valkotasapainoarvoa. Testien parhaat tulokset saavutettiin K-means-algoritmilla. Käytettyjen tietoaineistojen antamat tulokset poikkesivat toisistaan huomattavasti, eikä kaikkia saatuja tuloksia voitu pitää realistisina. Aihe vaatii syvempää analyysia sekä kokeita tulosten vahvistamiseksi.

Avainsanat: konenäkö, koneoppiminen, värivakio, klusterointi, valkotasapainoarvo, kuvankäsittely

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

CONTENTS

ABBREVIATIONS AND NOTATIONS

HVS	*Human Visual System*
REC	*RECommended ColorCheker dataset 2018*
RGB	*Red, Green, Blue colour model*
SFU	*Shi and Funt ColorChecker dataset 2010*
d	*Angular Error*
$\boldsymbol{e}_e$	*Estimated Illuminant / Generated centroid*
$\boldsymbol{e}_s$	*Ground truth /* Datapoint from dataset

1. INTRODUCTION

Colour constancy is one of the key phenomena of photography and visual inspection of images. It describes the human ability to perceive the colours of objects accurately, regardless of changes in the colour properties of the light source. Colour constancy is usually connected to the Human Visual System (HVS), but the exact details remain unclear [1]. In practice, the physical colour of scenes may differ depending on the varying external light sources. HVS can easily detect the correct colour of the object, but pictures and videos must be adjusted with a correction of white balance [2].

This study examines the possibility of finding a discrete number of estimated illuminants with different clustering methods, where the number of light sources is defined by the number of cluster centroids. In addition, investigating the presence of significant differences between clustering modules for this kind of evaluation is done. The datasets used in the study come from Lilong Shi and Brian Funt, "Re-processed Version of the Gehler Colour Constancy Dataset of 568 Images" (SFU) [3] and G. Hemrit et al., “Rehabilitating the ColorChecker Dataset for Illuminant Estimation.” (REC) [4]. These datasets are two variations of Gehler’s [5] ColorChecker dataset, containing 568 high-dynamic-range linear images, including a variety of indoor and outdoor scenes. All images are captured with a high-quality digital SLR camera in RAW format and are free of any colour correction. Every image contains the colour checkerboard, which has achromatic squares, that are used to measure the illumination [5]. In this study, pre-calculated ground truths are used instead of the actual images.

Clustering methodologies used are K-Means, Spectral and Agglomerative Clustering, and all these modules come from the Scikit-Learn Clustering Library [6]. In our analysis, we utilise SFU’s and REC’s measured illuminations to calculate the angular error between dataset points and generated centroids. Testing aims to pinpoint a threshold where the angular error drops to around 3 degrees while maintaining a relatively low number of centroids. In the field of colour constancy research, it is widely acknowledged that when the error in white balance correction approaches 3 degrees, the resulting image appears natural to the human eye, rendering any differences in colour correction practically undetectable.

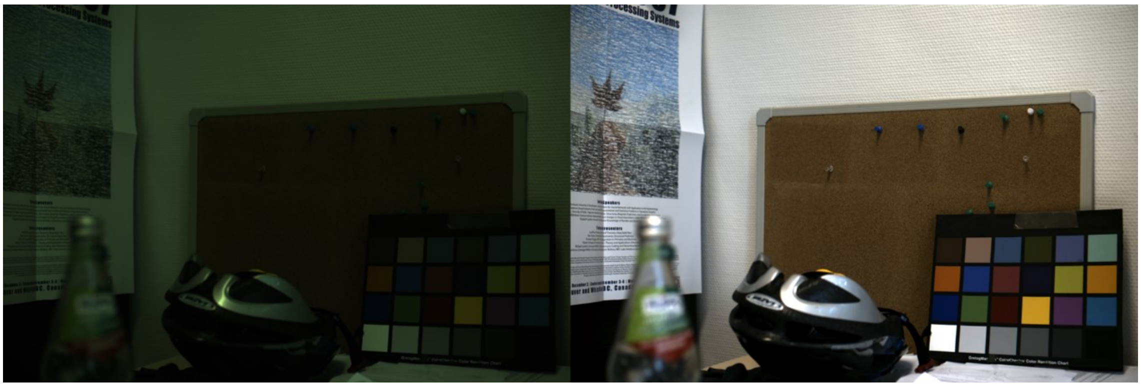

Figure 1. *RAW image from Gehler's dataset containing Macbeth ColorChecker that is Auto White Balanced with photo editing software.*

As demonstrated in Figure 1, the image has been auto white balanced using photo editing software. By utilizing the clustering algorithms, new white balance values could be used to achieve close to similar results.

This thesis is structured the following way: Chapter 2 presents the background of colour constancy and investigates the benefits and usage of it in computer vision-related tasks, while also looking at how clustering methods can be used for white balance value evaluation. The clustering-related modules used in the experiments are discussed in the third chapter. The fourth chapter reviews the experiments done and the results related to the evaluation. The closing chapter provides an overview of the results and future suggestions on related topics.

2. BACKGROUND

2.1 Colour Constancy

There are multiple ways to describe the colour constancy phenomenon. It can be said to be a challenge of determining the original colour of light that illuminates a scene, typically to eliminate the influence of varying conditions such as light source colour. Given the diverse lighting in images and human vision, it is crucial to differentiate accurately between properly coloured scenery, despite the large differences in the illumination. [7], [8]

HVS can adaptively see consistent colours within different illuminations and under different lights. Colour constancy is typically measured as the capacity to extract colour information from materials and objects. With a comprehensive understanding of scene type and relevant object details, the HVS is capable of perceiving the correct colours of objects with accuracy. [9]

The way colour constancy works in HVS is that regardless of the light source, such as natural sunlight or artificial indoor lighting, the hue of a red object remains red, e.g., strawberries. However, it is apparent that the image projected onto the retina is influenced not only by the red hue but also by the spectral properties of the surrounding light source. As a result, changes in ambient lighting conditions can usually cause shifts in the object's apparent colour. Nevertheless, the human brain employs the mechanisms of the visual system to counteract these variations. Through this cognitive process, the brain normalises the perceived colour of the object across different lighting environments, ensuring consistency in visual observation. [10]

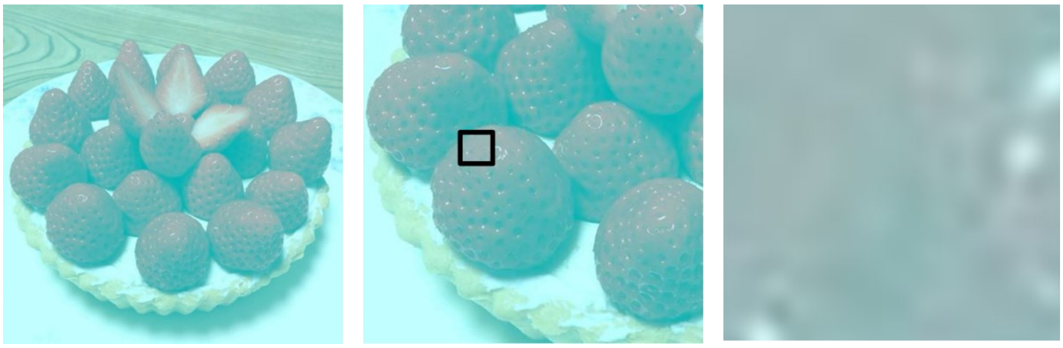

Figure 2. *a visual demonstration of how the HVS works. Images of strawberries, that seem to be red but actually are grey. [10]*

As seen in Figure 2, HVS is "correcting" the strawberries as red, even though the pixels are in a wide scalar of grey. That is because the cyan overlay over the picture is a complementary colour of red, and "red light" does not get through. The human brain makes assumptions about the lighting of the scene, and the visual system takes global colour information into account, allowing humans to maintain constancy in the object's colour [11].

As a problem constantly looking for new improvements, the problem of colour constancy is kept as underconstrained and many algorithms are based on assumptions that require, for example a set of light sources or their spectral characteristics [8].

2.2 Colour Constancy in Computer Vision

Colour features are a common aspect in applications such as image recognition and tracking-related tasks [12]. White balance evaluation in computer vision usually happens in two steps. The illumination parameters are estimated, and the effect of the illumination is removed from the image.[13], [14] Illuminant estimations differ from each other, and the algorithms used can give drastically different results, just as ground truth values for datasets REC and SFU [3], [4] are evaluated differently. Especially in video tracking usage, colour constancy algorithms are used to remove illumination effects and differences for more robust object detection because objects move across scenery and in changing lighting conditions [14].

Improvements for tracking operations can be made with deep learning approaches and achromatic adjustments of the images [13], which can improve the efficiency of the operations. Differential features used in object recognition and tracking are primarily relying on the luminance-based characteristics of the model, making colour corrections important in related tasks [8]. Proper colour correction of collected data is crucial, especially when image data is collectively used in model training or evaluation. In Gu's study [15], the accuracy of neural network objection detection methods were improved with colour priors, which shows the importance of colour correction in related tasks.

2.3 White-balance Evaluation with Clustering

Given the inherent nature of how cluster centroids interact with the data points, it is understood that if we generate one centroid for every ground truth, the error reduces to 0 degrees. However, it is conceivable to execute the clustering algorithms iteratively, starting with fewer centroids and increasing the amount with each run. If the collective mean of angular errors decreases low enough, it could be beneficial to a certain point, to use the generated centroids as new ground truths for the images if continuously calculated white balance values are not used. These clustering methods could be integrated into other machine learning programs, enabling processing of much larger datasets.

3. METHODOLOGY

3.1 Clustering

Data clustering has been addressed as the main methodology of this thesis. Cluster analysis is rather used in large variations in different fields, for example, medicine, physics, biology, and machine learning. A way to describe clustering is that it is sorting data or materials into different types of groups [16]. This grouping can rely on different meaningful components considering the data, such as colour, material, density, or anything useful for that specific scenario. Depending on the task, a variety of different types of clustering methodologies are available for data mining and reaching desired results. Clustering is a common tool in data mining, and it can be described as follows: "Given a set of data points, partition them into a set of groups which are as similar as possible [16]." Clustering techniques can roughly be divided into two groups: hierarchical and partitioning techniques [17].

Cluster-based colour constancy has been studied before in [18], diving deeper into the possibility of cluster-based techniques used in illumination estimation. These results were promising for how clustering can be useful besides illuminant estimation, mainly focusing on grey surfaces' distribution.

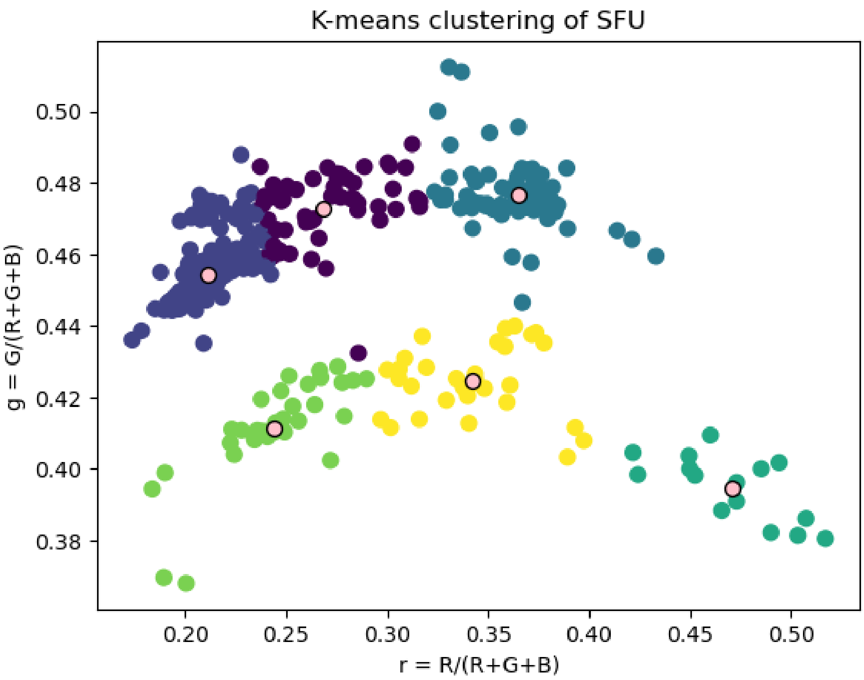

Figure 3. *Partitioning based K-means clustering visualised with SFU dataset where number of clusters is set to be 6. Newly generated centroids are shown as pink dots among the clusters.*

This study contains an illuminant evaluation with various clustering methodologies. Three different clustering algorithms are used to evaluate white balance values for a set of images. Both datasets (REC and SFU) used in the study contain pre-calculated ground truth values in Red-Green-Blue (RGB) colour model. These datasets provide a good foundation for finding differences in the Scikit-Learn [6] Clustering algorithms. Each clustering algorithm takes the dataset and generates an iteratively increasing number of clusters from the dataset. Every iteration runs through all the data points and estimates the closest generated clustering centroid for them, for example shown in Figure 3, where number of generated clusters is set to be 6. The process involves utilising the angular error between each data point and its nearest centroid. After calculating these errors for the entire list of ground truths within an iteration, their mean and median errors are calculated and stored. This methodology relies on clustering centroids and ground truth values from datasets, where each ground truth value is individually used to compute the angular error with the nearest centroid. This approach enables an incremental increase in the number of cluster centroids, resulting in an iteratively lower angular error. Employing this basic method across all clustering algorithms provides simple data for analysis and comparison.

3.2 Error

In colour constancy-related research, there are two metrics frequently used: Euclidean distance and angular error [8]. Angular error has been, to date, one of the most commonly used performance measures for colour constancy [2], [12], [16]. The angular error is a degree-based measurement, and it measures the angular distance between the estimated illuminant and ground truth:

$$d_{angle}(\boldsymbol{e}_e, \boldsymbol{e}_u) = \cos^{-1}\left(\frac{\boldsymbol{e}_e \cdot \boldsymbol{e}_u}{||\boldsymbol{e}_e|| \cdot ||\boldsymbol{e}_u||}\right), \tag{3.1}$$

where $\boldsymbol{e}_u$ is the datapoint (ground truth) and $\boldsymbol{e}_e$ is the generated clustering centroid. The dot product is calculated between $\boldsymbol{e}_u$ and $\boldsymbol{e}_s$. This dot product is then divided by the Euclidean norm of the vectors. Lastly, the inverse cosine function is calculated [8].

In this study, the angular error was calculated between every datapoint and the closest generated centroid to it. Overall, for every number of clusters implemented, there were 568 angular errors calculated. After every iteration, angular errors were used to compute the mean, median, and worst-25% of the means of the angular errors. These metrics were used to compare the results to others in the field [2], [4], [7], [13].

3.3 K-means Clustering features

In the K-means algorithm, data points are grouped into n groups with to minimise the within-cluster sum-of-squares, also referred to as inertia. One advantage of using K-means is its ability to efficiently scale to larger datasets. Inertia is seen as a measurement of how internally coherent clusters are, but it does not come without disadvantages. For example inertia makes assumptions that clusters are convex or isotropic, which is not the case in every module [6]. With its benefits and disadvantages, K-means is a simple and efficient method to use in this type of evaluation.

3.4 Spectral Clustering features

The Spectral Clustering module performs clustering in mainly two steps: It creates a low-dimensional representation of the data based on how similar the data points are to each other. This is achieved by analysing the connections between data points. After that, a standard clustering algorithm like K-means, is applied to this lower-dimensional space to group the data points into clusters. Spectral clustering is beneficial for handling non-spherical clusters but is not as good for larger numbers of clusters.[6]

3.5 Agglomerative Clustering features

Agglomerative clustering is a module of hierarchical clustering that starts with each data point in its own cluster. It then iteratively merges the most similar clusters, building a hierarchy of clusters until all data points are in a single cluster. Beneficial for handling non-spherical clusters but has high computational costs [6]. This study utilised a very basic agglomerative clustering model, and due to its nature, it might not be the most suitable module for this type of experiment.

4. EXPERIMENTS

4.1 Testing Environment

The testing environment was built with Python, and it contained iterative algorithms for all the clustering methods used. The same tests were used for both datasets, and the results from them were evaluated. Values from the datasets were normalised between 0 and 1. Every test was done with the same computer and setup, including Python version 3.11.5 and Scikit-Learn version 1.4.2.

4.2 Data

The datasets used in testing were both iterations of Gehler et al.'s originally released ColorChecker dataset. Closer to the present day, the set of images has had a few iterations on how the general ground truth values have been calculated. Every image from the image set contains a Macbeth ColorChecker board that shows the true, predetermined colours that were used to calculate ground truths [3], [4].

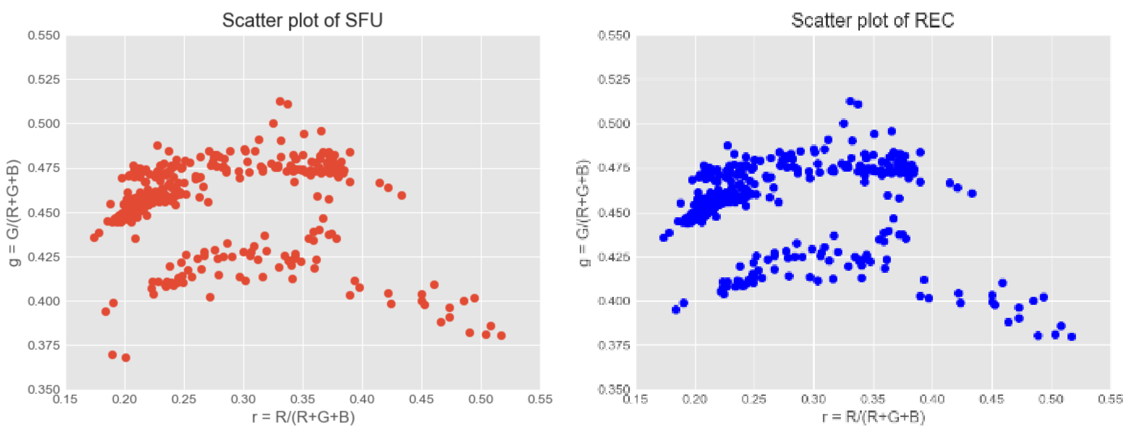

Figure 4. *2D visualisation of the REC and SFU datasets.*

Differences between both SFU and REC were both small, in terms of red and green values. When both sets were computed in 2D, the outliers found in SFU were rather minimal, as shown in Figure 4, But overall, the RGB values are closely fixed and can be used for comparison. Visually compared it is assumed that both datasets give out similar results.

4.3 Experiments with REC

After experimenting with both datasets, the results with REC were unrealistic and could not be used in the evaluation. When K-means clustering tests were done, the angular error dropped faster than expected, resulting in an impossibly good angular error.

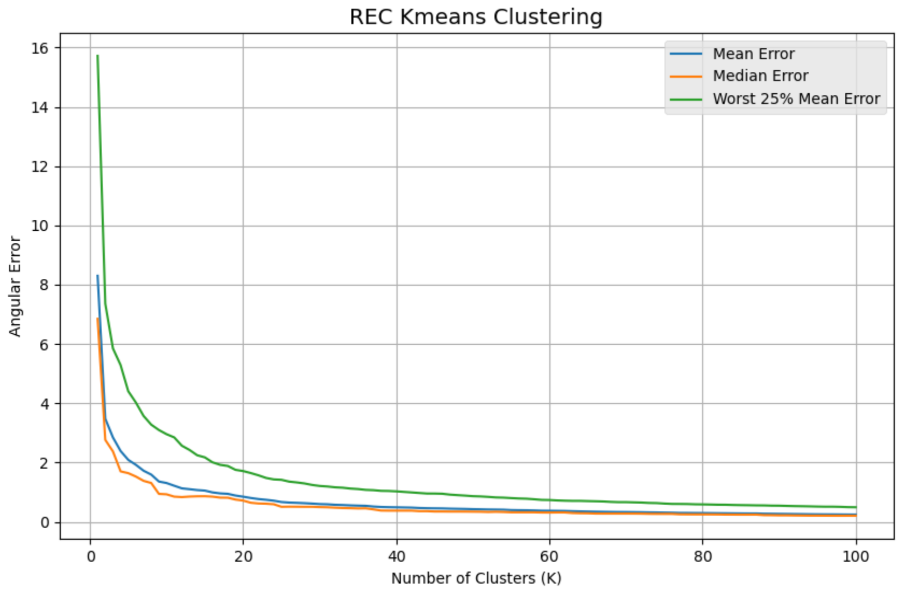

Figure 5. *Result graph of the REC where K-means clustering was used.*

As shown in Figure 5, the mean angular error drops below 3 degrees within few iterations and reaching below 1 degree of error at 40 clusters. When these results are compared to already-done research and other methodologies, results from the REC dataset in this study cannot be used. There was an unknown error in the testing methodology, that was not found while the tests were implemented.

4.4 K-means Clustering

The first implementation of the clustering algorithms used was K-means. With K-means relative simplicity and robustness, it was assumed to be a good method for this kind of evaluation task.

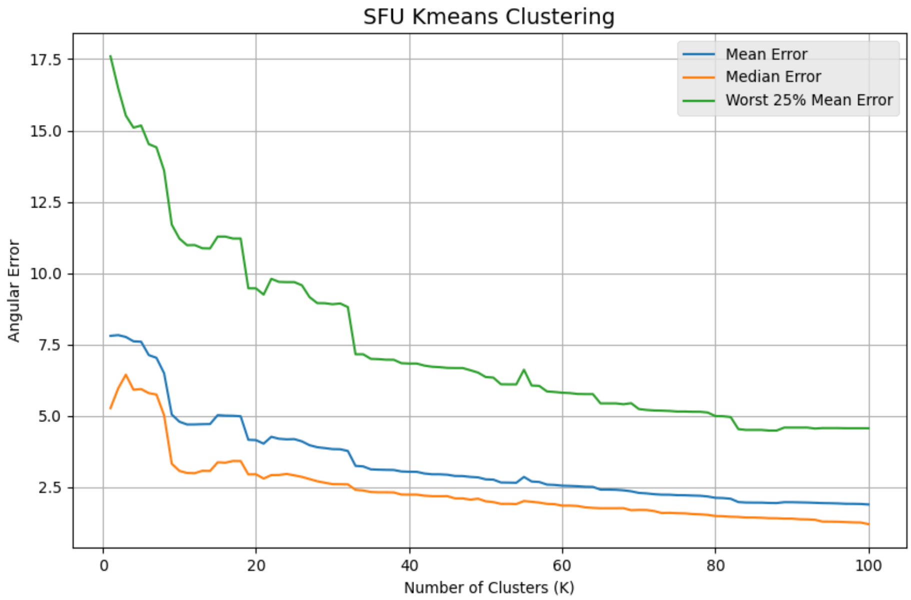

Figure 6. *Result graph of the SFU where K-means clustering was used.*

Visually demonstrated results from the SFU dataset shown in Figure 6, are curves of angular error calculated between data points and generated centroids. Error is calculated in Eq. (3.1), and results display the mean, median, and worst-25% mean errors from the iterations starting with K = 1 and ending with K = 100. When investigating the mean graph, the results seem relatively reasonable. Knowing that 3 degrees of error is a point where images look natural with a generated centroid, reaching it at 40 clusters seems like a good result.

Table 1. *K-means clustering angular errors. K is the amount of generated clustering centroids. REC results included to show the radical difference.*

K-means	REC			SFU		
K	Mean	Median	Worst-25 %	Mean	Median	Worst-25 %
10	1,31°	0,93°	2,96°	4,80°	3,08°	11,21°
20	0,85°	0,72°	1,71°	4,15°	2,96°	9,47°
40	0,49°	0,38°	1,03°	3,05°	2,25°	6,84°
60	0,37°	0,31°	0,74°	2,56°	1,86°	5,81°

Table 1 shows the results from 10 to 60 generated centroids for both datasets. SFU results seem to be more realistic and possible, whereas REC results are unrealistic. A manual evaluation of how newly generated centroids work as white balance values for the images was not conducted and is something that should be done in future experiments.

4.5 Spectral Clustering

As addressed before, spectral clustering performs low-dimensional embedding of the affinity matrix before utilising K-means [6]. Spectral clustering was expected to work better on the evaluation, but results show that this hypothesis was incorrect.

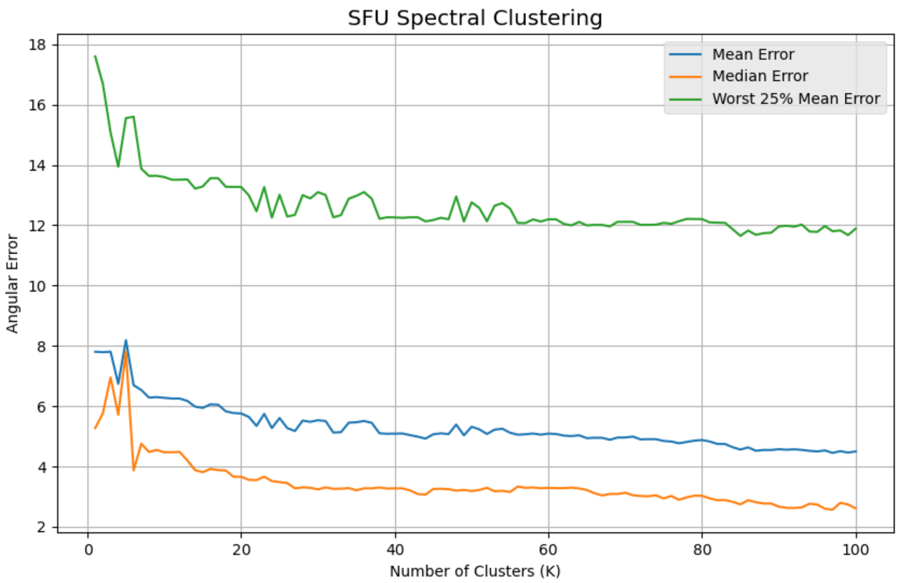

Figure 7. *Result graph of the SFU where spectral clustering was used.*

As shown in Figure 7, mean angular error never reached 3 degrees and instead got stable between 6 and 4 degrees, even though spectral clustering utilises K-means on the actual generation of centroids [6]. All error metrics followed similar pattern.

Table 2. *Spectral clustering results. K is the number of generated centroids.*

Spectral	SFU		
K	Mean	Median	Worst-25 %
10	6,28°	4,47°	13,59°
20	5,76°	3,66°	13,27°
40	5,09°	3,28°	12,26°
60	5,09°	3,29°	12,19°

Table 2 shows that these metrics are not great, when compared to K-means clustering or state-of-the-art methods in the Barron’s [7] study. Spectral clustering was expected to perform well in this type of task, but it turned out to be bad.

4.6 Agglomerative Clustering

The third algorithm tested in this thesis was agglomerative clustering, which is a hierarchical clustering technique. Agglomerative clustering required a linkage parameter, which determined how the clusters were formed. For this test, it was determined to be a “ward”, which is a good default choice and tends to produce clusters that are similar in size [6]. Ward linkage turned out to outperform other possible linkages in this test.

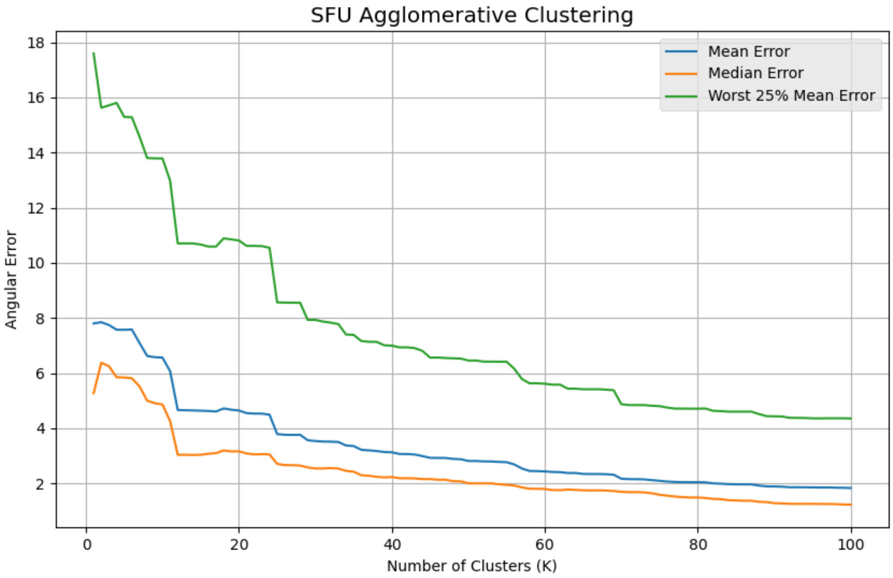

Figure 8. *Results graph of Agglomerative clustering where the SFU dataset was used.*

When comparing the agglomerative clustering result graph of the SFU dataset in Figure 8, to previous results, K-means and Agglomerative clustering share similar elements on how angular error acted in the generation of new centroids. Even though it was expected that Agglomerative would not perform that well in this type of task, it turns out it was the best of the algorithms used. The mean error descended to be around 3 degrees at 40 centroids, which is a considerable realistic outcome.

Table 3. *Agglomerative clustering results.*

Agglomerative	**SFU**		
K	**Mean**	**Median**	**Worst-25 %**
10	6,57°	4,86°	13,79°
20	4,65°	3,17°	10,81°
40	3,13°	2,23°	7,00°
60	2,44°	1,80°	5,62°

Table 3 shows that the results of Agglomerative clustering performed similarly to K-means, where the angular error was steadily decreasing and reached 3 degrees after 40 clusters.

4.7 Results

As a result of running clustering tests for both datasets, results from REC were way too good to be realistic and were discarded, because the causing error in the testing setup was not found. SFU gave more realistic angular errors, which was not expected as it is an older and less refined variant of Gehler's ColorChecker dataset [5]. Spectral clustering's performance was the worst of the tested algorithms and never reached better than 5 degrees of error. The results from both K-means and agglomerative clustering show that using 40 generated centroids can achieve an error of 3 degrees and these centroids could potentially be used as alternative values for white balancing.

When comparing the results of implemented clustering methods to other proven methods, we can conclude that K-means and agglomerative of the SFU dataset were better and the results were more realistic than other outcomes.

Table 4. *Results of tested methods where the generated centroids amount is 40. Barrons [7] "CCC" was added for comparison. The evaluation methodology was different, but the metrics are the same.*

All methods	**SFU**		
K=40	**Mean**	**Median**	**Worst-25 %**
K-means	3,05°	2,25°	3,05°
Spectral	5,09°	3,28°	5,09°
Agglomerative	3,13°	2,23°	3,13°
CCC [7]	1,95°	1,22°	1,95°

In Barron's study [7], the same SFU dataset [3] was used to evaluate different colour constancy algorithms. Barrons's method was based on 2D spatial localisation, and their algorithm is referred to as "CCC". As shown in Table 4, Barrons results are excellent when compared to the clustering methods implemented. Barron used a 2D spatial localization task for the evaluation [7].

These results can serve as a foundation for future improvements on the topic. Integrating clustering with other machine learning modules could be beneficial if illuminant estimations are needed and they do not have to be exact. However, this might be a waste of computational resources since cameras can seamlessly white balance images with a low margin of error.

5. CONCLUSION

Colour correction is a crucial part of machine learning and computer vision-related tasks. It ensures that models can utilise data accurately and effectively. In this study, different clustering methods were used to find a discrete number of white balance values for the dataset to be used as an alternative method of white balancing. Additionally, different clustering methods were compared to find the best algorithm for a related task.

The evaluation of clustering algorithms on the datasets revealed notable differences in performance. From used clustering algorithms K-means and agglomerative clustering performed similarly and gave reasonable results. Spectral clustering was anticipated to perform better but gave worse results when compared to other algorithms. When all the results were compared with Barron's "CCC" algorithm [7], it became evident that the values obtained from the REC dataset, particularly with K-means and agglomerative clustering, exceeded expectations for this type of evaluation. The unrealistically high precision with the REC dataset implies issues within the testing environment or data preparation. The notable results from the SFU dataset indicate the potential for using clustering algorithms for this type of illuminant estimation task.

The algorithm and evaluation setup used were basic and relied on few premade clustering algorithms. For future applications on similar tasks, more clustering methods could be harnessed for more comprehensive results. In addition, incorporating visual evaluations of the generated white balance values could enhance the reliability of these methods.

REFERENCES

[1] 'What is color constancy? | Color Constancy'. Accessed: Apr. 09, 2024. [Online]. Available: http://colorconstancy.com/color-constancy/what/./index.html

[2] A. Gijsenij, T. Gevers, and J. Van De Weijer, 'Computational Color Constancy: Survey and Experiments', *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2475–2489, Sep. 2011, doi: 10.1109/TIP.2011.2118224.

[3] 'Shi's Re-processing of Gehler's Dataset'. Accessed: Apr. 09, 2024. [Online]. Available: https://www.cs.sfu.ca/~colour/data/shi_gehler/

[4] G. Hemrit *et al.*, 'Rehabilitating the ColorChecker Dataset for Illuminant Estimation', *Color Imaging Conf.*, vol. 26, no. 1, pp. 350–353, Nov. 2018, doi: 10.2352/ISSN.2169-2629.2018.26.350.

[5] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, 'Bayesian color constancy revisited', in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2008, pp. 1–8. doi: 10.1109/CVPR.2008.4587765.

[6] 'Scikit Learn - Clustering', scikit-learn. Accessed: Apr. 09, 2024. [Online]. Available: https://scikit-learn/stable/modules/clustering.html

[7] J. T. Barron, 'Convolutional Color Constancy', in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile: IEEE, Dec. 2015, pp. 379–387. doi: 10.1109/ICCV.2015.51.

[8] T. Gevers, *Color in computer vision fundamentals and applications*. in Wiley-IS&T Series in Imaging Science and Technology. Hoboken, NJ: Wiley, 2012.

[9] H. Heidari-Gorji and K. R. Gegenfurtner, 'Object-based color constancy in a deep neural network', *JOSA A*, vol. 40, no. 3, pp. A48–A56, Mar. 2023, doi: 10.1364/JOSAA.479451.

[10] R. Salowe, 'R01 Grant Investigates Color Constancy - Penn Medicine'. Accessed: Apr. 09, 2024. [Online]. Available: https://www.pennmedicine.org/layouts/PennMedicine/AcademicDepartments/Academic%20Departments%20Layout.aspx

[11] O. Rinner and K. R. Gegenfurtner, 'Cone Contributions to Colour Constancy', *Perception*, vol. 31, no. 6, pp. 733–746, Jun. 2002, doi: 10.1068/p3352.

[12] F. Laakom *et al.*, 'Bag of Color Features for Color Constancy', *IEEE Trans. Image Process.*, vol. 29, pp. 7722–7734, 2020, doi: 10.1109/TIP.2020.3004921.

[13] G. D. Finlayson, 'Colour and illumination in computer vision', *Interface Focus*, vol. 8, no. 4, Jun. 2018, doi: 10.1098/rsfs.2018.0008.

[14] V. Agarwal, B. Abidi, A. Koschan, and M. Abidi, 'An Overview of Color Constancy Algorithms', *J. Pattern Recognit. Res.*, vol. 1, pp. 42–54, Jun. 2006, doi: 10.13176/11.9.

[15] P. Gu, X. Lan, and S. Li, ‘Object Detection Combining CNN and Adaptive Color Prior Features’, *Sensors*, vol. 21, no. 8, p. 2796, Apr. 2021, doi: 10.3390/s21082796.

[16] C. C. Aggarwal and C. K. Reddy, Eds., *Data clustering: algorithms and applications*. in Chapman & Hall/CRC data mining and knowledge discovery series. Boca Raton: Chapman and Hall/CRC, 2014.

[17] A. Saxena *et al.*, ‘A review of clustering techniques and developments’, *Neurocomputing*, vol. 267, pp. 664–681, Dec. 2017, doi: 10.1016/j.neucom.2017.06.053.

[18] W. Xiong, J. Kuang, X. Wang, J. Kuang, and X. Wang, ‘Cluster Based Color Constancy’, *Color Imaging Conf.*, vol. 16, pp. 210–214, Jan. 2008, doi: 10.2352/CIC.2008.16.1.art00040.