Shades of Gray and Colour Constancy

Graham D. Finlayson, Elisabetta Trezzi
School of Computing Sciences
University of East Anglia
Norwich, NR4 7TJ
United Kingdom

Abstract

Colour constancy is a central problem for any visual system performing a task which requires stable perception of the colour world. To solve the colour constancy problem we estimate the colour of the prevailing light and then, at the second stage, remove it.

Two of the most commonly used simple techniques for estimating the colour of the light are the Grey-World and Max-RGB algorithms. In this paper we begin by observing that this two colour constancy computations will respectively return the right answer if the average scene colour is grey or the maximum is white (and conversely, the degree of failure is proportional to the extent that these assumptions hold). We go on to ask the following question: *"Would we perform better colour constancy by assuming the scene average is some shade of grey?"*. We give a mathematical answer to this question. Firstly, we show that Max-RGB and Grey-World are two instantiations of Minkowski norm. Secondly, that for a large calibrated dataset L^6 norm colour constancy works best overall (we have improved the performance achieved by a simple normalization based approach). Surprisingly we found performance to be similar to more elaborated algorithm.

Introduction

The colour of an object depends on the spectral power distribution of the incident light as well as on the reflectance property of the object's surface. The human visual system is able to deduce a stable descriptor of the object's colour independent of the illumination. In digital photography we would also like to have colour constancy: we would like to remove the cast due to illuminant from images prior to display or printing them.

Mean and *maxima* play an important role in perceptual processing and computer vision. For example, the adaptation of the eye to the brightness of the prevailing light field is relative to the mean brightness [1] and maxima also mediate colour perception [2]. So it is, perhaps, not surprising that Grey-World and Max-RGB, widely used algorithms for colour constancy, are based on these concepts. In the former the estimate of the colour of the light is the average RGB in the image and in the latter it is the maximum R, G and B.

The starting point of this paper is that when we consider a number of samples in a distribution, the mean and the maximum are returned respectively by the two extremes in the Minkoswki family norm, i.e. the L^1 and L^∞ norms, normalized by the number of samples. Since for every $p \geq 1$, the L^p norm (the pth root of the sum of samples raised to the power p) will return an answer somewhere in between this two extrema, we wondered whether any of them might give reasonable estimates for the illuminant colour. In effect we suggest a new approach for estimating the illuminant, based on a different assumption on the scene average.

Our new algorithm based on the norm p was tested on the *Simon Fraser Image Database*. The results show that usually every norm p works better than Grey-World or Max-RGB algorithm and in particular the best performance is obtained for $p = 6$.

Perhaps, is this result not too surprising? First we note that it is easy to find cases where Grey-World works better than Max-RGB and vice versa. So a norm based somewhere in between should, on average, give better performance. Second, the larger that p becomes, the more important bright pixels become in an image and it can be argued that whites and bright chromatic colours tell us more about the prevailing light. Yet, we shy away from the maximum because in a statistical sense it is a fragile measure. It is for example unreliable when pixel are clipped.

In section 2 we present the Grey-World and the Max-RGB algorithms. In section 3 the Minkoswki family norm is introduced. Experiments and results are reported in section 4.

Background

A Lambertian surface with reflectance $S(\lambda)$ (where λ is wavelength) illuminated by a spectral power distribution $E(\lambda)$ reflects the colour signal $C(\lambda)$

$$C(\lambda) = E(\lambda)\, S(\lambda). \tag{1}$$

A digital colour camera typically samples the incoming signal with three different classes of sensors (red, green and blue), so that the response by the sensor (R, G, B)

is equal to the integral over the visible spectrum ω of its sensitivity function ($R(\lambda)$, $G(\lambda)$, $B(\lambda)$) multiplied by the incoming signal:

$$R = \int_{\omega} E(\lambda)\, S(\lambda)\, R(\lambda)\, d\lambda,$$
$$G = \int_{\omega} E(\lambda)\, S(\lambda)\, G(\lambda)\, d\lambda,$$
$$B = \int_{\omega} E(\lambda)\, S(\lambda)\, B(\lambda)\, d\lambda. \quad (2)$$

In colour constancy we attempt to calculate $E(\lambda)$ given only image RGB's. Or, as an easier alternative, to estimate the 3-dimensional vector $[R_e G_e B_e]^T$ corresponding to the RGB of the prevailing light.

Buchsbaum [3] made the assumption that the average reflectance of all surfaces in a scene is achromatic. If this Grey-World hypothesis holds, then the average colour of the light leaving the surfaces in a scene will be the colour of the incident light. To see that, let an image I be represented by three N-dimensional vectors $\underline{R}[R_1 \ldots R_N]^T$, $\underline{G}[G_1 \ldots G_N]^T$, $\underline{B}[B_1 \ldots B_N]^T$, where

$$\begin{aligned} R_i &= \int_{\omega} E(\lambda)\, S_i(\lambda)\, R(\lambda)\, d\lambda, \\ G_i &= \int_{\omega} E(\lambda)\, S_i(\lambda)\, G(\lambda)\, d\lambda, \\ B_i &= \int_{\omega} E(\lambda)\, S_i(\lambda)\, B(\lambda)\, d\lambda. \end{aligned} \quad (3)$$

If $g(\lambda) = k$ is a grey surface, i.e. all wavelengths reflect the *kth* part of the light, so

$$\begin{aligned} R &= \int_{\omega} E(\lambda) g(\lambda) R(\lambda) d\lambda = k \int_{\omega} E(\lambda) R(\lambda) d\lambda, \\ G &= \int_{\omega} E(\lambda) g(\lambda) G(\lambda) d\lambda = k \int_{\omega} E(\lambda) G(\lambda) d\lambda, \\ B &= \int_{\omega} E(\lambda) g(\lambda) B(\lambda) d\lambda = k \int_{\omega} E(\lambda) B(\lambda) d\lambda, \end{aligned} \quad (4)$$

the colour of that surface is exactly the colour of the light. Now assume that the scene average is grey, i.e.

$$\mu(S(\lambda)) = \sum_{i=1}^{N} \frac{S_i(\lambda)}{N} = g(\lambda). \quad (5)$$

Now the average for the red channel is:

$$\begin{aligned} \mu(\underline{R}) &= \sum_{i=1}^{N} \frac{1}{N} \int_{\omega} E(\lambda)\, S_i(\lambda)\, R(\lambda)\, d\lambda = \\ &= \int_{\omega} E(\lambda)\, g(\lambda)\, R(\lambda)\, d\lambda = kR_e. \end{aligned} \quad (6)$$

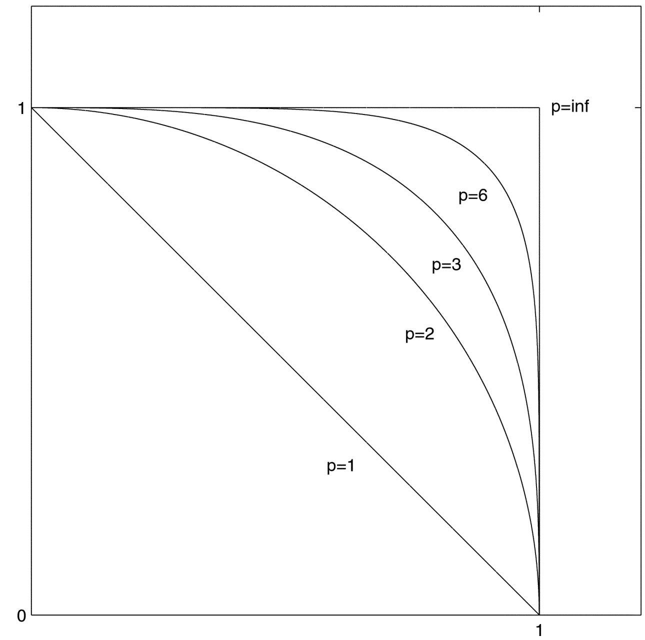

Figure 1. The figure shows a unit ball in R_+^2, that is the set of all positive points whose distance from the origin is less than or equal to one.

We draw attention to the value of k in equation (4). It tells us we can recover the light colour up to some multiplicative constant. This is normal in colour constancy research. It is hard to discriminate between bright lights impinging on dark surfaces and the converse. Let us now consider the Max-RGB approach. In an early version of Land's [4] retinex algorithm it is tacitly assumed that every image contains a white surface. Furthermore this patch is assumed to be the maximally reflective surface in the scene, so that its location can easily be found by searching for the maximum response in every single channel. Let $w(\lambda) = 1$ a white surface, i.e. all wavelengths reflect 100% of the light, then

$$\begin{aligned} \max_i |R_i| &= \int_{\omega} E(\lambda)\, w(\lambda)\, R(\lambda)\, d\lambda = R_e, \\ \max_i |G_i| &= \int_{\omega} E(\lambda)\, w(\lambda)\, G(\lambda)\, d\lambda = G_e, \\ \max_i |B_i| &= \int_{\omega} E(\lambda)\, w(\lambda)\, B(\lambda)\, d\lambda = B_e, \end{aligned} \quad (7)$$

that is the RGB of that patch equals the colour of the incident light.

Of course the average reflectance in a image is not grey and white is not always present. So, Grey-World and Max-RGB colour constancy algorithm can fail. However, given the simple statistical formulation, it is perhaps surprising how well they can work. Moreover, of importance to this paper, the results are not always correlated: Max-RGB can give good result when Grey-World fails and vice versa.

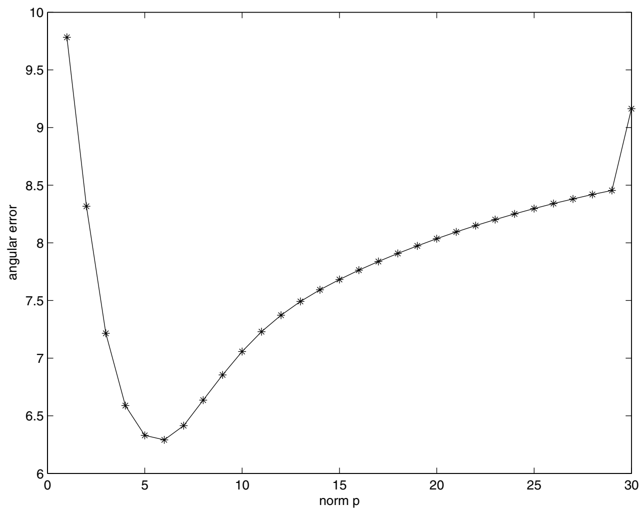

Figure 2: *The figure shows the angular error of the group A images for 30 values of p; in particular for p from 1 to 29 and* $p = \infty$.

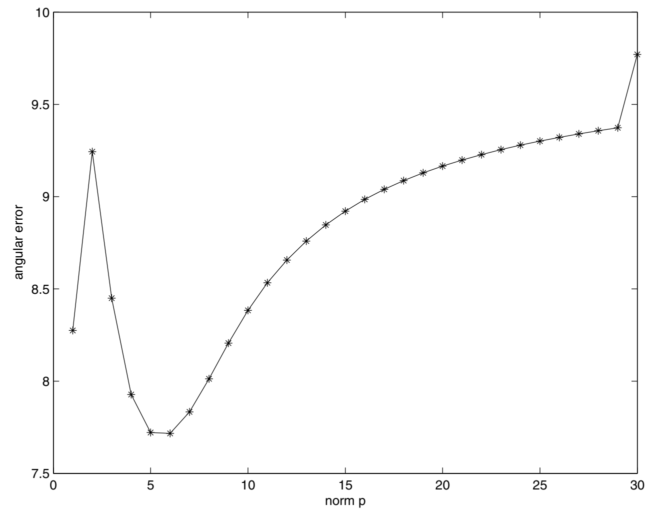

Figure 3: *The figure shows the angular error of the group B images for 30 values of p; in particular for p from 1 to 29 and* $p = \infty$.

The Minkowski family norm

Let $\underline{X} = [X_1 ... X_N]^T$ a vector in R^N. For every $p \geq 1$ the quantity

$$\|\underline{X}\|_p = \left\{ \sum_{i=1}^{N} |X_i|^p \right\}^{1/p} \tag{8}$$

defines a norm on R^N because of the Minkowski inequality [5]:

$$\left\{ \sum_{i=1}^{N} |X_i + Y_i|^p \right\}^{1/p} \leq$$
$$\left\{ \sum_{i=1}^{N} |X_i|^p \right\}^{1/p} + \left\{ \sum_{i=1}^{N} |Y_i|^p \right\}^{1/p}.$$

Thus, depending on the chosen p, there is a different definition of the distance between two points, or the length of the vector $\underline{X}$. There are some useful relationship between these norms. First of all it is easy to see from figure 1 that $\|\underline{X}\|_q \leq \|\underline{X}\|_p$ for every $p \leq q$, then the sequence $\alpha_n = \|\underline{X}\|_n$ is monotonically **decreasing**. The limit is also a norm, called norm infinity and is equal to the maximum of the coordinates:

$$\|\underline{X}\|_\infty := \lim_{n \to \infty} \alpha_n = \max_{i \leq N} |X_i|. \tag{9}$$

It is also true that for every $p \leq q$,

$$\frac{\|\underline{X}\|_p}{N^{1/p}} \leq \frac{\|\underline{X}\|_q}{N^{1/q}}. \tag{10}$$

Given N sample values $\{X_1, ..., X_N\}$, for every $p \geq 1$ we can calculate the quantity

$$\mu_p\left(\underline{X}\right) = \frac{\|\underline{X}\|_p}{N^{1/p}}. \tag{11}$$

Therefore, for equation (10), μ_p is a monotonically **increasing** sequence. We see that $\mu_1\left(\underline{X}\right) = mean\left(\underline{X}\right)$ and $\mu_\infty\left(\underline{X}\right) = \max_i |X_i|$. If the samples are the data of a colour image, then the vector $\left[\mu_1\left(\underline{R}\right) \mu_1\left(\underline{G}\right) \mu_1\left(\underline{B}\right)\right]^T$ corresponds exactly to the estimated illuminant whether the Grey-World algorithm is used. On the other hand, with Max-RGB computation the evaluated light is the vector $\left[\mu_\infty\left(\underline{R}\right) \mu_\infty\left(\underline{G}\right) \mu_\infty\left(\underline{B}\right)\right]^T$.

The p shade of grey algorithm evaluates the RGB of the unknown illuminant whit the p norm of the data, that is

$$\begin{aligned} R &= \mu_p\left(\underline{R}\right), \\ G &= \mu_p\left(\underline{G}\right), \\ B &= \mu_p\left(\underline{B}\right). \end{aligned} \tag{12}$$

Yet, under which assumptions does this computation return the correct result? We give a mathematical answer, explaining how the scene average is a shade of grey. In particular we will show that the following relationships:

$$\begin{aligned} \mu_p\left(\underline{R}\right) &= k \int_\omega E(\lambda) R(\lambda) d\lambda, \\ \mu_p\left(\underline{G}\right) &= k \int_\omega E(\lambda) G(\lambda) d\lambda, \\ \mu_p\left(\underline{B}\right) &= k \int_\omega E(\lambda) B(\lambda) d\lambda \end{aligned} \tag{13}$$

are valid when the scene average of a modified version of the original image is grey.

First of all we apply a (non linear) invertible transformation at every single pixel in each channel. Moreover we consider the inverse transformation applied to the illuminant. Then, with respect these conditions, we assume that the scene average is grey. The following diagram represent the process in a schematic way.

$$\begin{array}{ccc} I & \xrightarrow{f} & f(I) \\ & & \downarrow gw \\ \left[g(R)\,g(G)\,g(B)\right] & \xleftarrow{f^{-1}=g} & [R\,G\,B]\,. \end{array} \quad (14)$$

If we choose as function f the raising to the power p (and g is the pth root), thus the equations (13) are satisfied. In fact, let consider the image I_p such that $R_{p,i} = (R_i)^p$, $G_{p,i} = (G_i)^p$ and $B_{p,i} = (B_i)^p$. Let $E_p(\lambda)$ the illuminant so that each component of the correspondent RGB vector (using the same digital camera sensitivities) is the unknown illuminant raised to the power p; for example for the red component we have:

$$R_p = \int_\omega E_p(\lambda) R(\lambda) d\lambda = \left\{ \int_\omega E(\lambda) R(\lambda) d\lambda \right\}^p . \quad (15)$$

Consequently to solve the colour constancy problem for the image I, determining the RGB of the illuminant $E(\lambda)$ it is enough working out the RGB of the illuminant $E_p(\lambda)$. Hence let $\underline{\sigma}_p(\lambda)$ the function vector of the surface reflectances associated with the image I_p and the illuminant $E_p(\lambda)$[1]:

$$\begin{aligned} R_{p,i} &= \textstyle\int_\omega E_p(\lambda)\,\sigma_i(\lambda)\,R(\lambda)\,d\lambda = R_i^p, \\ G_{p,i} &= \textstyle\int_\omega E_p(\lambda)\,\sigma_i(\lambda)\,G(\lambda)\,d\lambda = G_i^p, \\ B_{p,i} &= \textstyle\int_\omega E_p(\lambda)\,\sigma_i(\lambda)\,B(\lambda)\,d\lambda = B_i^p. \end{aligned} \quad (16)$$

If the scene average of I_p is grey, i.e. $\mu\left(\underline{\sigma}_p(\lambda)\right) = k_p$, then we can apply the Grey-World algorithm to find the vector $[R_p G_p B_p]^T$:

$$\begin{aligned} \mu_1\left(\underline{R}_p\right) &= k_p R_p, \\ \mu_1\left(\underline{G}_p\right) &= k_p G_p, \\ \mu_1\left(\underline{B}_p\right) &= k_p B_p. \end{aligned} \quad (17)$$

Because $\mu_1\left(\underline{R}_p\right) = \{\mu_p(\underline{R})\}^p$ (and also for the green and the blue coordinates), it follows that equations (13) are satisfied. Since the Grey-World technique usually does not give the value of the constant k_p, the p shade of grey algorithm do not recover the intensity of the illuminant.

Note we do not set forth any technical argument to support the idea that $\mu\left(\underline{\sigma}_p(\lambda)\right) = k_p$ holds more for some p than others. But we do test this in the next section.

Experiments and results

To investigate the performance of the p norm algorithms we tested them on a large dataset made up of many colourful objects measured under different lights whit a calibrated camera [6]. In particular we consider two distinctive collections: one consisting of 321 images of a variety of 32 scenes and another of 220 images of a variety of 22 scenes, both groups taken under 11 coloured illuminant. Details about the images, how they were collected and the illuminants measured can be found in [6] and in [7] chapter 11. All images are available on-line [8]. We call group A the 321 images [9] and group B the 220 images [10]. In both sets each image is stored with the correct light RGB measurement. Thus we can assess a colour constancy algorithm by determining how close the estimated light colour is to the measured one. Here we compare measured and estimated light illuminants using an angular measure. Given the image with the measured light $\underline{q}_L = [R_l\, G_l\, B_l]^T$, we run the colour constancy algorithm to recover the estimated illuminant $\underline{q}_E = [R_e\, G_e\, B_e]^T$. The error is then

$$ang\_err = angle\left(\underline{q}_l, \underline{q}_e\right) = \frac{\cos^{-1}\left(\underline{q}_l \cdot \underline{q}_e\right)}{\left|\underline{q}_l\right| * \left|\underline{q}_e\right|}. \quad (18)$$

We also calculate the distance error in the chromaticities space $r = R/(R+G+B)$ and $g = G/(R+G+B)$:

$$dist\_err = \left\{(r_l - r_e)^2 + (g_l - g_e)^2\right\}^{1/2}. \quad (19)$$

Both measures are, by choice, independent of intensity.

The results obtained are shown in the figures.

In figure 2 the mean of the angular error of the group A images is plotted for 30 values of p; in particular for p from 1 to 29 and $p = \infty$.

In the same way figure 3 shows the mean over all images in the group B. For both databases, the L^6 $norm$ performs best. Note also that this is not a small effect: performance is surprisingly better. The table 1 shows all results, for both groups of images and for both error measures.

It is interesting to note that the performance of L^6 norm can be compared with the results of algorithms that make use of statistical information about the chromaticities related to the illuminants. For example in [11] are reported the mean angular errors for different algorithms tested on the 321 images dataset. The value that we obtained for the $norm$ 6 algorithm is comparable to many advanced colour constancy algorithm. With respect to the results in [11] only one or two algorithms perform better. Yet this improvement over prior approaches is achieved at significant computational cost.

Conclusions

In this paper we have introduced a new technique for illuminant estimation. Exploiting the Grey-World algorithm we have improved it with a different approach, based on a simple normalisation. We have demonstrated the efficacy of the p shade of grey algorithm by testing it on a set of real images.

The relevance of our results is that we have achieved the enhancement, keeping the simplicity of the Grey-World

[1] Note, we do not actually specify how to find $E_p(\lambda)$ or $\sigma_i(\lambda)$. But, since $\left[R_i^p G_i^p B_i^p\right]^T$ are 3-dimensional vectors, we can find the required form using standard techniques.

norm	ang_err		$dist\_err$	
p	group A	group B	group A	group B
1	9.78	8.27	0.0788	0.0624
2	8.32	9.24	0.0640	0.0668
3	7.22	8.45	0.0535	0.0593
4	5.59	7.93	0.0476	0.0549
5	6.33	7.72	0.0448	0.0530
6	6.29	7.71	0.0440	0.0527
7	6.42	7.83	0.0445	0.0533
8	6.64	8.01	0.0456	0.0543
9	6.86	8.21	0.0469	0.0555
10	7.06	8.38	0.0481	0.0566
11	7.23	8.53	0.0492	0.0576
12	7.37	8.66	0.0502	0.0584
13	7.49	8.76	0.0510	0.0590
14	7.59	8.85	0.0517	0.0596
15	7.68	8.92	0.0523	0.0601
16	7.76	8.99	0.0523	0.0605
17	7.84	9.04	0.0534	0.0609
18	7.91	9.09	0.0539	0.0612
19	7.97	9.13	0.0544	0.0615
20	8.04	9.17	0.0549	0.0618
21	8.10	9.20	0.0553	0.0620
22	8.15	9.23	0.0557	0.0622
23	8.20	9.26	0.0561	0.0624
24	8.25	9.28	0.0565	0.0626
25	8.30	9.30	0.0568	0.0627
26	8.34	9.32	0.0571	0.0629
27	8.38	9.34	0.0574	0.0630
28	8.42	9.36	0.0577	0.0631
29	8.46	9.37	0.0580	0.0632
∞	9.16	9.77	0.0630	0.0659

Table 1: *Results for the p shade of grey algorithm on two databases considered: the firsts two columns are the mean of angular errors and the lasts two report the distance error in the chromaticities space.*

algorithm and without using any sophisticated method. In the future we will use the mathematical tool of the Minkowsky family norms in combination with more elaborate techniques. We also believe that the L^p norms will prove insightful in considering other aspects of colour processing.

References

1. Eduardo-José Chichilinski and Brian A. Wandell, Photoreceptor sensitivity changes explain colour appearance shifts induced by large uniform backgrounds in dichoptic matching, Vision Research, pp. 239–254 (1995).
2. John J. McCann, Suzanne P. McKee and Thomas H. Taylor, Quantitative studies in retinex theory, Vision Research, 16, 239 (1976).
3. G. Buchsbaum, A spatial processor model for object colour perception, Journal of the Franklin Institute, 310, 1 (1980).
4. Edwin H. Land, The retinex theory of color vision, Scientific American, 237, 108 (1977).
5. A.N. Kolmogorov and S.V. Fomin, Introductory real analysis, Dover Publications (1975).
6. Kobus Barnard, Lindsay Martin, Adam Coath and Briant Funt, A comparison of computational color constancy algorithms; part two: Experiments with image data, IEEE Transactions on Image Processing, 11(9).
7. Kobus Barnard, Practical Colour Constancy, Ph.D. thesis, Simon Fraser University, School of Computing Science (2000).
8. Http://www.cs.sfu.ca/˜colour/data/.
9. Http://www.cs.sfu.ca/˜colour/pub/pub/data/colour_constancy_test_images/mondrian/mondrian_8_bit.tar.gz http://www.cs.sfu.ca/˜colour/pub/pub/data/colour_constancy_test_images/specular/specular_8_bit.tar.gz.
10. Http://www.cs.sfu.ca/˜colour/pub/pub/data/objects_under_different_lights/objects_8_bit.tar.gz.
11. G.D. Finlayson, S.D. Hordley and I. Tastl, Gamut constrained illuminant estimation, Proceedings of the 9th International Conference on Computer Vision, 2 (Oct 2003).

Biography

Elisabetta Trezzi graduated from Università degli Studi in Milan with a Laurea degree (M.Sc.) in Mathematics. Then she obtained a MSc in Information Technology at Image and Audio Processing Area at CEFRIEL, Education and Research Center in Information Technology from the Milan Polytechnic. She is currently a Ph.D. student working on Colour Constancy with the Colour Research Group at the School of Computing Sciences, University of East Anglia in Norwich.