CPSC525
Correction of Chromatic Aberration from a Single Image Using Keypoints

Benjamin T. Cecchetto

Department of Computer Science
The University of British Columbia

Abstract

In this paper, we describe a method to correct for chromatic aberration in a single photograph. What is proposed is a method for replicating what a user would do in a photo editing program to account for this defect. We obtain matching keypoints in each colour channel and align them automatically, as a user would manually.

1 Introduction

Chromatic aberration (also known as colour fringing) is a phenomena where the different wavelengths of light refract differently through different parts of a lens system. Thus the colour channels may not align properly as they reach the sensor/film. This is most notable in cheaper lenses, however it is becoming more noticeable with higher resolution photos. Why we would want an image free of chromatic aberrations is simple, it is clearly not the optimal image as not all of the planes are in focus. For a simple lens system this amounts to the three colour channels (red, green and blue) being misaligned by a uniform scaling and translation. For more complicated lenses, this may not be the case as distortions may be introduced. Also, there are many types of other chromatic aberrations, however we only propose a method to deal with the simple, more common scenario and possible ways to deal with more complex ones.

2 Previous Work

Previous work in correcting chromatic aberration has been initially used in domains where scientific accuracy is a must. This usually takes place in a lab setting where there are heavily constrained scenarios. One such method is in active optics [Willson and Shafer 1991]. In active optics, instead of just taking one exposure for all the colour channels simultaneously (like most digital cameras), they take three. One for each channel with slightly different focal points, accounting for the aberration in the lens so the three channels will be aligned in the final result. The main drawback in this system is it's limitation to need to alter the hardware, as well as take 3 different exposures at 3 different points in time. This is limiting especially for dynamic scenes, as the object may be moving in that timespan.

Another method that has been proposed is to model the edge displacements with cubic splines to compute a non-linear warp of the image [Boult and Wolberg 1992]. This method cannot handle blurring/defocus in the image plane along with saturation effects. It breaks down in regions that are either underexposed or overexposed.

A calibrated technique for lateral chromatic aberration has been proposed where one directly calculates the aberrations and uses warping to compensate [Mallon and Whelan 2007]. However, this only works for calibrated cameras/images which may not be feasible for a more simple user. It does employ colour plane realignment, which we shall utilize as well.

Somewhat related is using chromatic aberrations as a cue to retrieve depth information [Garcia et al. 2000]. This demonstrates that it is a detectable and useful source of information without prior calibration.

Most recently, a full analysis of how aberrations are not a simple shift of colour planes has been studied and can be computed by using a non-linear warping to obtain the desired result [Kang 2007]. This algorithm handles many different cases of aberrations, which it categorizes in the paper. Its only drawbacks are minor artifacts caused by saturations, as well as being a non-linear solution so it may be quite slow.

Since we want to deal with the simple case, we may want to remove perspective distortions. This has been done in the past by a few different algorithms [Swaninathan et al. 2003]. If we remove the radial distortion cause by the lens, we should end up with only needing to solve a translation/scaling of the different colour channels to align them.

Most recently, a metric L has been proposed in a paper on extracting a depth matte from a colour filtered aperture [Bando et al. 2008]. This metric measures how collinear and correlated points are in 3D (essentially colour space). This metric essentially is how misaligned colours are in RGB-space, and they use it to find a disparity from their specially designed colour filtered aperture.

Figure 1: Left, a photo with chromatic aberration. Right is a possible correction for it. The original image is from Wikipedia.

3 Overview

Instead of concentrating on the optical derivation, we shall instead consider a different one. The problem is the task that an artist can manually align the colour channels of an image with chromatic aberration to minimize the effect. They align these different channels by moving the corners of the different channels and essentially translating and scaling the image. We shall attempt to replicate this artist driven result.

The main idea is to align two of the colour channels to the third one. Usually, the green channel has the least amount of aberration as it is in the middle of the colour spectrum and won't refract too much or too little. In some of the results, we match to the red channel for simplicity. The result of not choosing green is just a larger or smaller version of the image instead of the correct scale, but it will still be aligned properly.

We solve the alignment by finding keypoints in the fixed channel, and finding where those keypoints are in the other channel while trying to minimize a suitable metric. This is similar to finding disparities in stereo, however in 3 channels instead of two images. Once we have these matchings, we can prune the results based on how well the matchings are. Then we need to find a transformation from the non-fixed planes to the fixed one, with restriction the warp be a scale and translation.

In choosing a linear warp of the original colour channels, we manage to preserve image details without introducing any new artifacts that were not there. Also, since we use only keypoints, issues arising from saturation will be easy to deal with as we can be careful around those regions and not choose a keypoint nearby.

3.1 Computing the Alignment Metric

Before we start describing how to find the keypoints, and disparities, it is best to define how aligned the colour channels are. It is not suitable to do cross-correlation, as that only tells us how aligned two colour channels are. We want to see how aligned three are. As described above, there exists such a metric L [Bando et al. 2008]. For a given neighbourhood around a pixel (x, y), with eigenvalues of the covariance matrix λ_i and covariance matrix diagonal elements $\sigma_r^2, \sigma_g^2, \sigma_b^2$,

Figure 2: Left, is the L value computed at every pixel of a photo with no chromatic aberration. Right, is the L value computed for a photo with visible chromatic aberration. Note that in the second image, there is a non-zero value for L over the majority of the image.

$$L(x, y) = \frac{\lambda_0 \lambda_1 \lambda_2}{\sigma_r^2 \sigma_g^2 \sigma_b^2}$$

This value is essentially how collinear the colour points in this neighbourhood are in RGB-space. The lower this value is, the more collinear the points are and the higher the less collinear. As mentioned in the appendix of the paper, this can be considered to be related to cross-correlation, and thus is exactly what we want to use. How to choose this neighbourhood size is a different story. The smaller it is, the less statistics we have about that neighbourhood and thus may have a worse matching. The larger it is, the better chance we have of a matching, however the longer it takes. This value is bounded between 0 and 1, as mentioned in the paper. If we show this for every pixel as in figure 2, we can see that images without chromatic aberration show very little misalignment over the whole image.

Another justification of using this metric is that according to the colour lines model [Omer and Werman 2004], colour points in RGB-space of the whole image will lie along different colour lines. If we look at smaller neighbourhoods, then we can assume the points will also lie along either a line, or intersection of lines. This measures the collinearity of those neighbourhoods. If we search for an ideal alignment, then we want to maximize the collinearity, thus minimizing L.

3.2 Finding Keypoints

The first task is to find regions where the keypoints would be useful. We want to find regions where the alignment measure

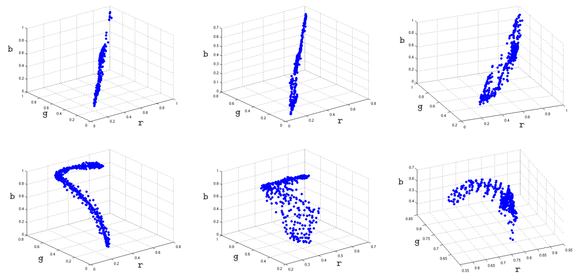

Figure 3: Colour points in neighbourhood clusters around a pixel. The top row are from a control image with no aberrations chosen randomly and the bottom row is from a photo with aberrations, chosen at regions of obvious colour fringing. Note that even though the control image may have an edge, we still have a distinct line whereas the aberration photos generally have a more spread out distribution everywhere.

L is very high, but at the same time we want to be certain there is a good possibility for a good alignment. Choosing regions in the image with high L values is costly, as we have to compute L for every pixel and examine that image. Also, it is not guaranteed to give us a good pixel, such as smooth regions which may have ambiguous results.

A good choice would be to find high gradient regions and use points from those. Specifically where we know there is an edge nearby. This gives us a better match since we know the other two channels should have a high gradient in that region too. We randomly sample from the norm of the gradient image with gradient sufficiently high within a threshold. In addition, if we want to pay the cost of the L image, we can compute it only in regions where the norm of the gradient is sufficiently high. If we multiply L by the norm of the gradient image, and threshold it, we can sample points that are highly unaligned with enough information around them to be aligned.

3.3 Finding Disparities

If we want to align two channels to the remaining one (say align green and blue to the red channel), we need to shift over all combinations of possible windows, with different scales. The idea is we want the neighbourhood of the green and blue channels to be correlated with the red, and those channels we know may be elsewhere with a different scale. Thus, we want to minimize the misalignment $L(x, y)$ subject to a shifted and scaled window in both green and blue channels so we can write

$$L(x, y; d_x^G, d_y^G, \sigma^G, d_x^B, d_y^B, \sigma^B)$$

With disparities d_x, d_y and relative scales σ. Where we iterate over all acceptable disparities (as in stereo) and all acceptable scales. We know however, that the disparities and scale difference shouldn't be too large (unless one has a truely horrible lens) so we can limit the search to local neighbourhoods in that sixtuple. In fact, if the scale difference is decently small (which it usually is in the case of aberrations) we can simply look for disparities to find the scale aspect of the transform. Thus we can write L as

$$L(x, y; d_x^G, d_y^G, d_x^B, d_y^B)$$

To be able to handle different scales, one could simply do an image pyramid based approach as in other computer vision papers. This is unnecessary, as most aberrations are not that distant in the scale domain.

The reason why this works is if we have a perfect edge (all one colour on one side, then all the other colour on the other side of the edge) in an image then the colours in the neighbourhood will cluster into two distinct regions. Since we are dealing with a natural image and edges are not perfect, these clusters will connect in a line as there will be a gradient from one colour to the other. If the image is misaligned, then this region will become more spread out. Now if we consider a multicoloured region, we get a more complicated shape. However, if we find disparities that minimizes this cluster's collinearity, we should get a better aligned image since we're minimizing the spread of the whole shape.

3.4 Pruning Keypoints

Although we've found keypoints and disparities for those keypoints in the other two channels, they might not give us good information. For example, if the best matching L value for the point neighbourhood was high, we shouldn't want to use that point as it is not a very well aligned point. We only want to use points that have gone from high L value (unaligned) to low L value (aligned). Since we only chose points that are unaligned, we just need to prune away the points that remain unaligned. Thus we only choose points with a low enough new L value.

Alternatively, since we know we want to do a scale and translation for each channel (3 degrees of freedom for each) we only need 2 keypoints (each being 2 dimensional). Thus we could choose the 2 keypoints with the lowest L value. Other methods could include weighting the keypoints based on the L parameter and haven't been fully explored. In practice, thresholding by the right amount is sufficient for good results.

3.5 Computing Image Transforms

Now that we have a set of good points, we can solve for the transformation pretty easily. Let us consider the red and green channels for now. We have (p_x^R, p_y^R) chosen in the fixed red channel and a point translated by disparity in the green channel (p_x^G, p_y^G). We have the equation:

$$\begin{pmatrix} \sigma^G & 0 & t_x^G \\ 0 & \sigma^G & t_y^G \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_x^R \\ p_y^R \\ 1 \end{pmatrix} = \begin{pmatrix} p_x^G \\ p_y^G \\ 1 \end{pmatrix} \quad (1)$$

Thus we can rearrange it for this point pair as:

$$\begin{pmatrix} p_x^R & 0 & 1 \\ p_y^R & 1 & 0 \end{pmatrix} \begin{pmatrix} \sigma^G \\ t_x^G \\ t_y^G \end{pmatrix} = \begin{pmatrix} p_x^G \\ p_y^G \end{pmatrix} \quad (2)$$

If we have a second point we can solve this system for σ^G, t_x^G, t_y^G, the scale, and translation respectively. In practice, we'd want more than just two points because the points might have only a good local solution and not global. Thus we'd want points from different regions of the image.

4 Results

So far this seems to work quite well as seen in figure 1 and more closely in figure 4. In figure 5, there is an image with lens distortion where our algorithm is expected to fail. In figure 6 we have a synthetic image being corrected. As we can see the correction is quite similar to the original image. A lot of the high frequency details are a little blurred because the misaligned image has a lower resolution than the original image and thus we cannot get those details back.

In figure 7 we have a larger scale change between the different channels. Our algorithm solves this pretty well too. We can see the keypoints in the third image of this figure and some matchings aren't always correct such as the one in the bottom left pointing in the wrong direction.

In figure 8, we have taken some photos with a lens assembly with chromatic aberrations. These photos are blurry because it is very hard to focus with this assembly. On the left we have the original photos and the right we have the corrected versions. Notice that the box is actually worse than the original whereas the tripod has a slight improvement. These photos are hard to deal with because the aberration is not as apparent since all of the colour channels are blurry.

Figure 4: Zoomed in area from figure 1 with original on the left and corrected on the right.

5 Discussion

If we reduce the neighbourhood size to compute L, we get more false-positives in the correlations. This is true because more disparities give a low number. The colours will cluster into a spherical region in a smooth neighbourhood, whereas we want lines. It will just choose a random disparity in this case. For the blurry images in figure 8, this is a similar phenomena and there isn't enough information in the image for this method to work well and reliably. Many disparities in this image regardless of direction gave L a value below 0.01 where less than that is considered a 'good' alignment in regions with more detail.

It was mentioned earlier that we may weight the keypoints and their appropriate disparities based on how much they reduced the alignment measure. Each row of equation 2 would be multiplied by a function of its associated L. Different linear and squared error weighting based on L have been attempted with little change in results. One could try to normalize the weights somehow instead of just using L directly.

Also, there might be other statistical methods to explore to prune the keypoints such as computing the translation/scaling and getting rid of the outliers using RANSAC. The computationally expensive part of this algorithm is determining the disparities where we have a 4D loop (without handling extremely large scale changes). After we have the disparities it's much quicker to deal with the keypoint data, especially since we need so few keypoints.

Another option to consider speedups is perhaps a hierarchical approach. One could try to solve the problem with a reduced resolution image and gradually work our way to the full resolution image using the lower resolution information.

What isn't clear but is worth exploring is a statement made earlier saying that unwarping a distorted image will yield an image with a chromatic aberration that can be corrected with a scale/translation. This is worth exploring, as the undistortions relatively fast as well as this algorithm.

Another approach to find keypoints would be to segment the image into many cells and pick an appropriate keypoint from each as in section 3.3. This will allow enough information from different parts of the image to make a global warp more accurate than just choosing random points in the acceptable regions.

One other thing that has been explored is doing the same thing in gradient domain. Some initial results have shown that it didn't work as well as the natural image formations as in our figures. It is worth exploring trying to align the gradients in a different way, perhaps using chamfer alignment on the edges of the channels.

6 Conclusions

We have presented a method that corrects chromatic aberrations in a single image without any use of calibration. Also, since this method is keypoint based, we don't need to spend

Figure 5: An image from http://www.tlc-systems.com/artzen2-0047.htm on the left and on the right a 'corrected' version. Note the properly aligned stand but misaligned horse head.

Figure 6: Top: The original image from Google Images. Left: The same image with a different translation shift in two of the colour channels. Right: A corrected version using our algorithm.

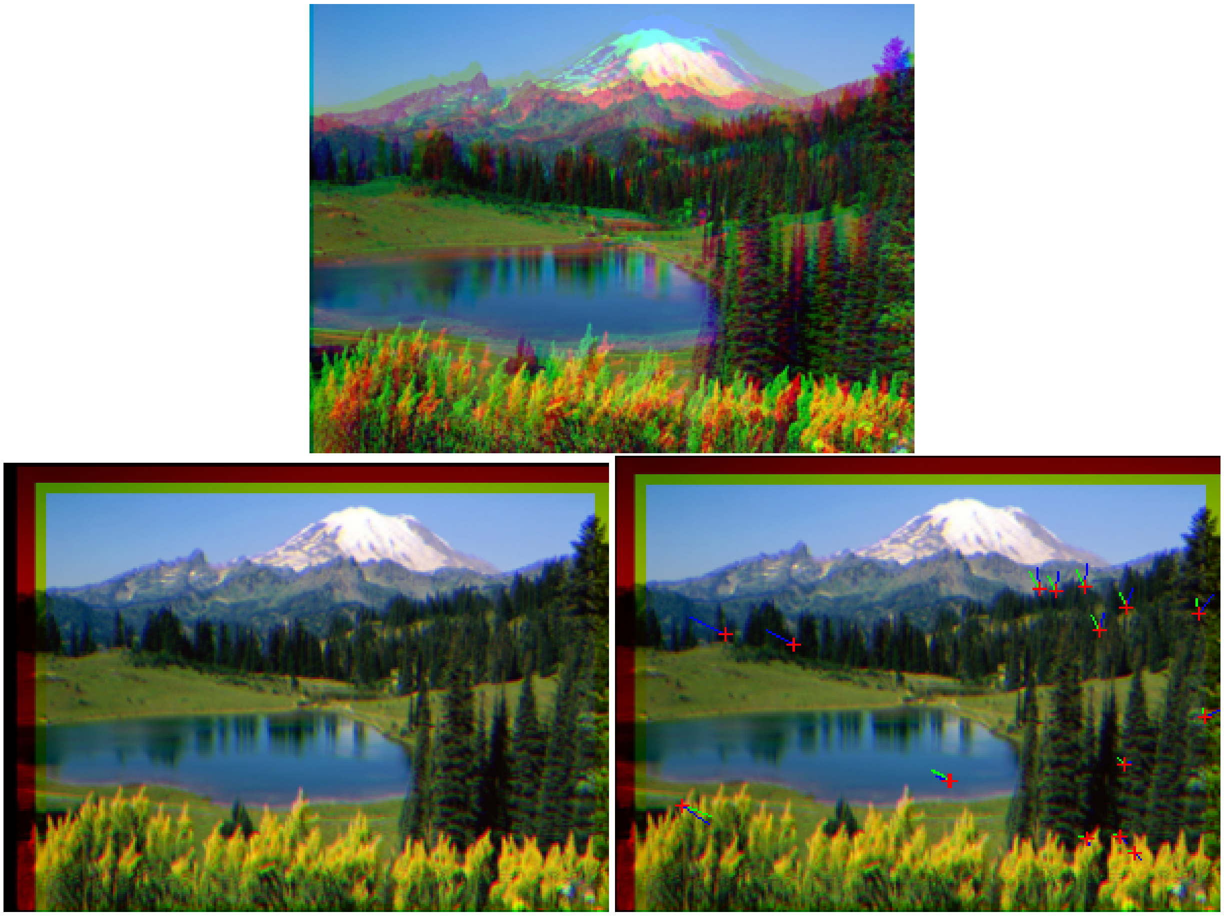

Figure 7: Top: An image with different translation and relatively large scale shift in two of the colour channels. Left: A corrected version using our algorithm. Right: The same corrected version showing disparities and keypoints

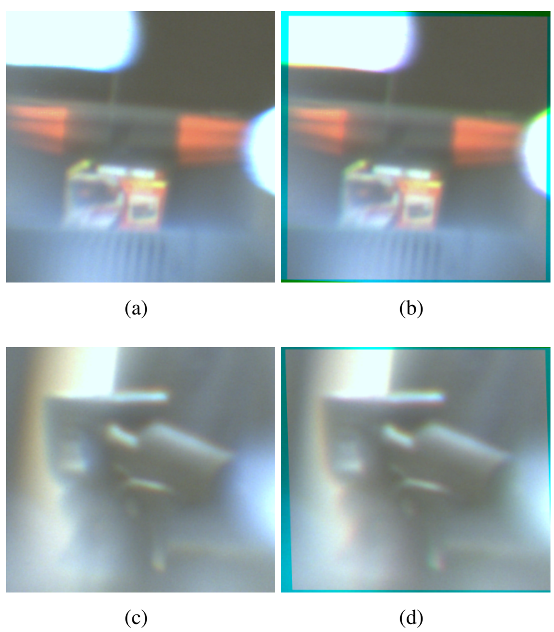

Figure 8: a) An image taken with a lens with large chromatic aberrations. b) A corrected version after using our algorithm. Similarly for c) and d).

much time to correct a single image. This method also works well with saturations as it can just ignore those regions by not choosing keypoints near them.

7 Acknowledgements

Wolfgang Heidrich for the idea for the project. Gordon Wetzstein for his lens.

References

BANDO, Y., CHEN, B., AND NISHITA, T. 2008. Extracting depth and matte using a color-filtered aperture. In *International Conference on Computer Graphics and Interactive Techniques*, ACM New York, NY, USA.

BOULT, T., AND WOLBERG, G. 1992. Correcting chromatic aberrations using image warping. In *1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92.*, 684–687.

GARCIA, J., SANCHEZ, J., ORRIOLS, X., AND BINEFA, X. 2000. Chromatic aberration and depth extraction. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 1.

KANG, S. 2007. Automatic removal of chromatic aberration from a single image. In *IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota, USA*, 1–8.

MALLON, J., AND WHELAN, P. 2007. Calibration and removal of lateral chromatic aberration in images. *Pattern Recognition Letters 28*, 1, 125–135.

OMER, I., AND WERMAN, M. 2004. Color lines: Image specific color representation. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2.

SWANINATHAN, R., GROSSBERG, M., AND NAYAR, S. 2003. A perspective on distortions. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*, vol. 2.

WILLSON, R., AND SHAFER, S. 1991. Active lens control for high precision computer imaging. In *1991 IEEE International Conference on Robotics and Automation, 1991. Proceedings.*, 2063–2070.