ROBUST COLOR CORRECTION FOR STEREO

S.F. Hasan*[∞], J. Stauder*, A. Trémeau[∞]

[*] Technicolor Research & Innovation, France{Hasan.Sheikh-Faridul,jurgen.stauder}@technicolor.com
[∞] Université Jean Monnet, Saint-Etienne, France, Alain.Tremeau@univ-st-etienne.fr

Abstract

Color difference between views of a stereo pair is a challenging problem. Applications such as compression of stereo image demands the compensation of color differences which is typically done by methods called color mapping. Color mapping is based on geometric feature correspondences. From these feature correspondences, color correspondences are generated which is used for fitting a color mapping model. This paper focuses on detection of outliers in the geometric feature correspondences. We propose a novel iterative outlier removal method which exploits colors in the neighborhood of the feature correspondences. From the analysis of our experimental results and comparing with existing methods we conclude that the proposed outlier removal method outperforms other geometric and color based methods.

Keywords: color mapping, stereo color correction, color correspondence, 3D

1 Introduction

Applications involving multiple views and 3D modeling of the same scene such as stereo imaging may suffer from color differences between corresponding images. Color differences can come from non-calibrated cameras, non-calibrated film scanners, inconsistent color decisions in post-production workflow, physical light effects in the scene and other effects.

Multiple view based applications usually require the compensation of these color differences. For example, in compression of a stereo image pair, color differences between left and right images, will reduce the correlation between the images and thus will decrease the performance of compression. Another example is the generation of 3D object models from still images. In this case, the texture of the object is extracted from the still images and color differences between the still images have to be modeled and compensated. The challenge is similar for image stitching which is analogous to wide baseline stereo. In this paper, we will assume that color differences come from global effects such as camera calibration.

In a stereo workflow, we propose to do color difference compensation as the first step so that further steps such as disparity estimation or compression can benefit as shown in Figure 1. And, we want to use as less geometric cues or geometric assumption as possible.

One approach for the compensation of color differences between images is *color mapping*. It assumes that left and right views show the same scene. In other words, there is a strong semantic relationship between the views. Figure 2 shows the basic steps of color mapping.

Figure 1: The positioning of color difference compensation in a stereo workflow

Color mapping typically starts by finding *geometric feature correspondences* [11, 12, 13] using methods such as Scale Invariant Feature Transform (SIFT) [6] or simple normalized cross correlation [4]. As geometric feature correspondences computation is not free from errors, some of the corresponding feature points are wrong and are called outliers. So, those outliers are usually removed in the next step. Afterwards, colors (typically encoded in R, G, and B coordinates for red, green and blue) at the geometric feature correspondences in the left and right views are retrieved. We call these retrieved pairs of colors *color correspondences*. Finally, these *color correspondences* are used to fit a color mapping model. This model describes the relationship between the *reference* (correct) view and the *test* view (to be corrected). Therefore, if this model is applied to the *test* view, it will compensate the color differences. In stereo imaging, the right view could be corrected with respect to the left view or vice versa.

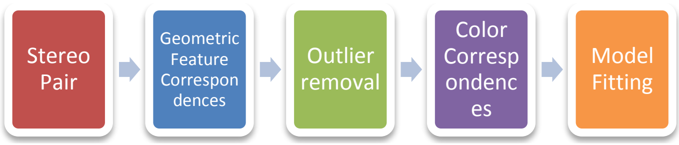


Figure 2: Basic steps of color mapping

In the literature, different color mapping models and respective estimation methods are proposed. Jung et al. [4] propose a global, parametric gamma-offset-gain model which is based on the camera characteristics. Tehrani et al. [11] as well as Yamamoto and Oi [13] use a global, data-driven, look up table based model. Wang et al. [12] propose a local, region-based color mapping model with just a simple, constant color coordinate offset per region. Tai et al. [10] propose probabilistic segmentation-based region mapping using Gaussian mixture models and expectation-maximization. In this paper, we will focus on the removal of outliers among the geometric feature correspondences. Efficient outlier removal is significant because small errors in feature correspondences may cause large errors in *color* *correspondences* and finally may cause large model errors. Note that, other methods for compensation of color differences exist that does not depend on geometric feature correspondences. These methods are usually called *color transfer*. One example is the transfer of probability density functions (PDF) of colors [7]. *Color transfer* methods can be used for low baseline stereo only. It is not optimal for wide baseline stereo as the non-overlapped region might contain very different colors. Therefore, for wide baseline PDF transfer of colors could result in strong artifacts on the non-overlapped region.

In this context of robust color correction for stereo, we propose a novel method to remove outliers from the geometric feature correspondences exploiting the color information in their spatial neighborhood. Unlike the existing methods, the proposed method avoids the use of geometric information. The motivation for this is that we want to address all kinds of stereo set ups.

This paper is organized as follows. Section 2 reviews existing outlier removal techniques for color mapping. The proposed method is presented in Section 3. Experimental results are shown in Section 4 and the conclusions are given in Section 5.

2 Existing outlier removal methods

Jung et al. [4] propose a color mapping method that identifies geometric feature correspondences using cross-correlation independently in three channels. Then, they detect the outliers by observing inconsistencies between channels. After that, they propose color compensation for each channel using a so-called Gamma-Offset-Gain (GOG) model which is based on camera characteristics and it is very often called camera characteristic equation as shown in equation 1.

$$C_{ref} = G\{C_{test}/(2^n - 1)\}^{\gamma}(2^n - 1) + b \qquad (1)$$

Here, C_{ref} is the color coordinate of a *reference* view whereas C_{test} is the color coordinate of a *test* view with parameters G being a gain, γ a gamma and b an offset. Finally, 2^n refers to the total number of quantization levels for each color channel.

Another classical way of dealing with outliers is robust estimation. First, all color correspondences are used for estimation of a model such as shown in Figure 3. A confidence corridor is defined assuming that the initial estimation result is close to the ground truth. All corresponding colors outside the corridor are considered to be outliers and are removed from estimation. Here binary weight is used, i.e. either use observations for estimation or declare them outliers. The remaining corresponding colors are the so called *valid samples* and estimation is iterated. The limitation of this method is as follows: if the initial estimation is far from ground truth, this method will miss a wide range of *true outliers* as well as it will mistakenly detect some truly valid samples as outliers, *false positives*. Another limitation is that outlier removal is done color channel-wise which may raise inconsistency between channels. For example, different sets of color correspondences are used to estimate the three partial color mapping models for R, G and B.

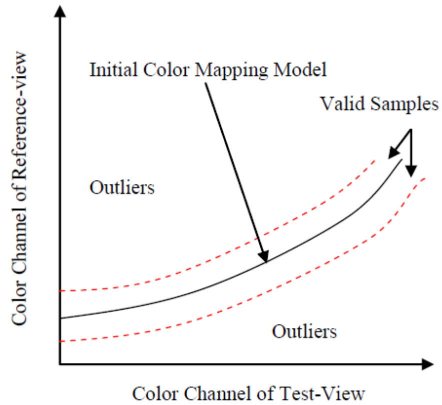


Figure 3: Classical outlier removal; the solid curve represents the initial model using all pairs of corresponding colors. Then dotted region defines confidence corridor to remove outliers

Another outlier removal technique is proposed by Tehrani et al. [11]. The authors assume that geometric feature correspondences should have disparities parallel to the base line of the stereo pair. To exploit that, firstly they propose to connect corresponding geometric feature points by lines. Then, a histogram of the angles between all the lines and the base line is calculated. Assuming the angles are normally distributed, all correspondences within a given range are considered as valid samples whereas the others are considered as outliers. The method of Tehrani et al. [11] is limited mainly in two aspects. First, when the baselines between cameras are not parallel, it will accept very large range of corresponding pairs as valid and the valid range has to be manually chosen. The second limitation is that being parallel to baseline cannot always ensure that the correspondences will be valid. For example, in case of horizontal baseline, mismatches in horizontal direction can never be detected. These limitations imply that the method of Tehrani et al. [11] is not efficient for all camera set-ups such as general camera sensor networks.

For color compensation, Yamamoto and Oi [13] as well as Tehrani et al. [11] propose to compute the color correspondences not only from the reference and test views but also from Gaussian-blurred versions of those views. Yamamoto and Oi [13] claim that information from correct color correspondences accumulates whereas information from the outliers scatters during estimation. However, the method of Yamamoto and Oi [13] suffers from two main limitations. First, the color correspondence information by the Gaussian blurring averages colors and the model will lose precision. Second, they do not propose any outlier detection method.

RANSAC [2] is a popular method for outlier detection and removal of geometric feature correspondences based on geometric constraints. The model assumes a planar scene and uses a planar homography for disparity estimation from geometric feature correspondences. For example, to find planar homography RANSAC starts with a random four correspondences and instantiate the model. Then, geometric feature correspondences within a distance threshold provide the consensus set S. If S is large enough re-estimation using all correspondences in S is done and RANSAC terminates. But if S is not large enough, another subset is chosen and repetitions of the above steps are done. Finally, after N trials the largest consensus set is selected and the model is re-estimated.

3 Proposed method

3.1 Motivation

The idea is to exploit the color of the spatial neighborhood of the geometric feature correspondences in order to detect outliers. Our method differs from classical robust estimation and outlier detection in so far that we exploit not only the feature point itself but also the spatial color neighborhood around the feature points in the *reference* and the *test* view. In contrast to RANSAC, we do not use any geometric assumption.

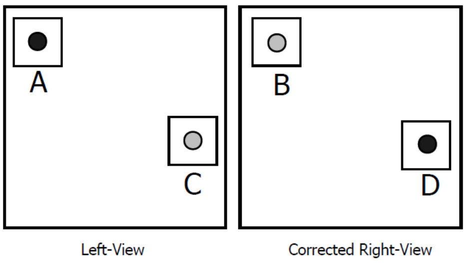


Figure 4: Sample situation where existing classical outlier removal will fail

We prove in the paper that our outlier removal method will improve classical approach in two aspects. First, we expect to reduce the number of false positives in outlier detection. This means we want to reduce the number of cases where a color correspondence is detected as outlier but truly it is not. As a result of reducing the *false positives*, the color mapping model will receive more valid information and will have better precision. Secondly, we also expect to reduce the number of *false negatives*. This means we want to reduce the number of cases where a color correspondence is detected as valid but truly it is an outlier.

Let us analyze a particular example shown in Figure 4. Here, a dark color A of a feature point in the left view corresponds to a light color B in the right view. Note that, this estimation considers all color correspondences and the result of estimation is shown in Figure 5.

Let us further assume that, these corresponding colors A and B are surrounded by lighter colors and an initial color mapping model M is available. If the initial model M cannot map the color A into color B, i.e. M (A) is far away from B, this color correspondence is declared as outlier by classical methods. This decision is a false positive. Due to an erroneous model M, M (A) does not give B. At least in dark colors such as A, the initial model M is erroneous.

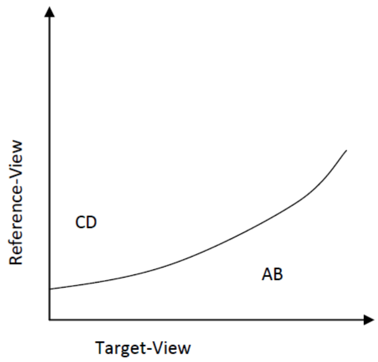


Figure 5: As AB or CD is far from initial estimation, classical method will consider those as outliers whereas in reality they are not outliers.

To solve this issue, we propose to look into the feature points' neighborhoods in the left and the right views to decide about the outlier removal. Why can we expect better outlier detection when considering the neighborhoods? It is quite probable that the neighborhood contains other colors than A in the left image and B in the right image. Let us assume that the neighborhood of color A contains some lighter colors. Let's further assume that the initial model is better for lighter colors than for darker colors. In this case, the model M would work better in the neighborhoods of the colors A and B than for the colors A and B themselves. We may expect that the false positive decision is corrected and the feature point is not any longer detected as outlier.

3.2 Overview of the new method

Our main objective of outlier detection does not depend on the type of color mapping model. For simplicity, we choose a global, Gamma, Offset Gain (GOG) model for each color channel [4].

Let's consider the left view of the stereo pair as the *reference* view and the right view of the stereo pair as the *test* view. Figure 7 shows the flowchart of the proposed outlier detection method. Geometric feature correspondences are computed first. Then, color correspondences are extracted from the two images at the positions of the geometric feature correspondences. In the next step, the GOG parameters of an initial color mapping model are estimated and are used to achieve the *initial color corrected test view*. It's important to note that this initial color compensation uses all corresponding colors including potential outliers. According to our knowledge, the next step is novel. We will compare the colors of spatial neighborhood of the corresponding colors in the *initial color corrected test view* and the *reference view*. As the *test view* is already initially color compensated, the remaining difference of colors in the neighborhood should be small. If the color difference is beyond a chosen *threshold*, the pair of corresponding colors is an outlier and vice-versa. The *threshold* is set to σ where σ^2 is the variance of estimation error, C_{error} :

$$C_{error} = C_{ref} - C_{corrected} \tag{2}$$

Here, for each color channel, C_{ref} is the reference color and $C_{corrected}$ is the iteratively corrected color.

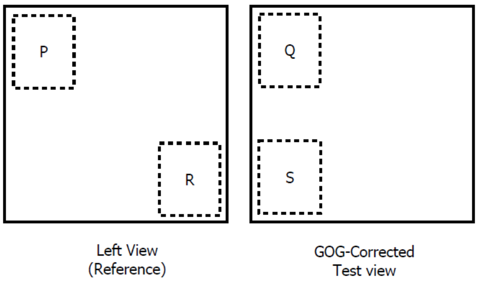


Figure 6: Neighborhood comparison

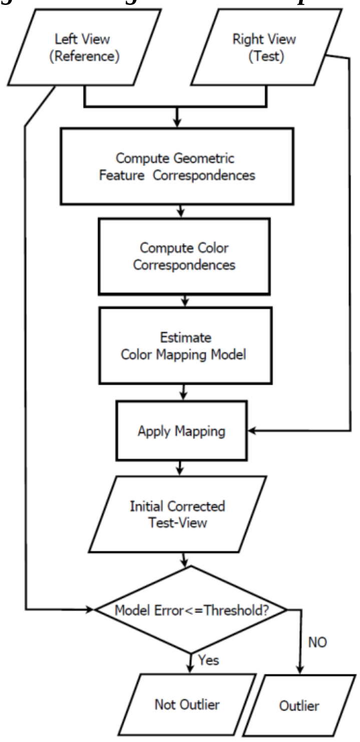


Figure 7 Flowchart of robust outlier detection

3.3 Neighborhood comparison methods

We first want to define the problem of neighborhood comparison and then present two possible methods. Let's analyze a simple scenario to understand how neighborhood comparison works. Figure 6 shows true geometric feature correspondence at position <P, Q>. The image content at position P in the *reference* view is the same as at position Q in the *test* view. In contrast let's assume <R, S> is a wrong correspondence. The dotted boxes around a geometric feature point is called neighborhood.

Let us remind that the neighborhood comparison of a geometric feature correspondence is carried out between the *reference view* and the iteratively updated *corrected test* view. Note that comparison process is done channel wise. For the sake of presentation let's assume the neighborhood size is 3x3. We will show some possible ways to compare the neighborhood and their advantages and disadvantages.

First of all, for each channel, we may compute the *absolute difference of means* between corresponding neighborhoods such as shown by equation 3. Here, $diff_{p\times p}$ refers to the difference of $p \times p$ size windows around the correspondences. C_{ref} and $C_{corrected}$ refer to the *reference* view colors and the *corrected test* view colors respectively.

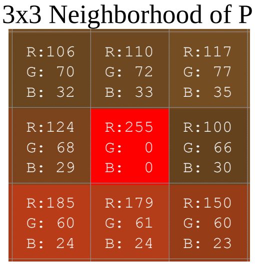
$$diff_{p\times p} = abs\left(\frac{1}{p^2}\sum_{i=0}^{p^2-1} C_{ref}\ (i) - \frac{1}{p^2}\sum_{i=0}^{p^2-1} C_{corrected}\ (i)\right) \tag{3}$$

If the *absolute difference of means* is larger than the threshold, the correspondence is an outlier:

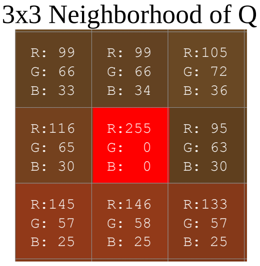
$$Is\ outlier = \begin{cases} no, & diff_{p\times p} \le threshold \\ yes, & diff_{p\times p} > threshold \end{cases} \tag{4}$$

The main disadvantage of the *absolute difference of means* is the assumption of linear color mapping. Therefore, as different colors in the neighborhood increases, the error increases as well.

3x3 Neighborhood of P



3x3 Neighborhood of Q



Blue Channel

M	M	M
M		M
M	M	M

Green Channel

M	M	M
M		M
M	M	M

Red Channel

M	NM	NM
M		M
NM	NM	NM

Assume Threshold=8
M=Match, NM=Non-Match
Total Match=19>Total Non-Match=5; So It is not Outlier

Figure 8: An example of 3x3 neighborhood comparison of a feature correspondence and voting in three channels for outlier decision.

The second approach is pixel to pixel comparison and voting. Let us assume that we are comparing the top-left pixel of neighborhood of P with the top-left pixel of neighborhood of Q as shown in Figure 6. If the difference between two pixels is within a given *threshold*, we will increment the counter of *matches*. Similarly, if the difference is more than the *threshold*, we will increment a counter of *non-matches*. We will repeat this process for all three (R, G, B) color channels. If total number of *matches* is higher than total number of *non-matches*, we will declare <P, Q> as valid. Otherwise, we will declare it as outlier. A toy example is presented in Figure 8 where the central pixels of the two 3x3 patches refers to P and Q. Let us assume that the threshold is 8. Now, for each channel, comparing all eight neighborhood colors as presented above deduce that number of *matches* is higher than number of *non-matches*. Therefore, this particular geometric feature correspondence is not an outlier. We have shown our experimental results based on this method. The main disadvantage with this comparison method is that it is not invariant to geometric deformation.

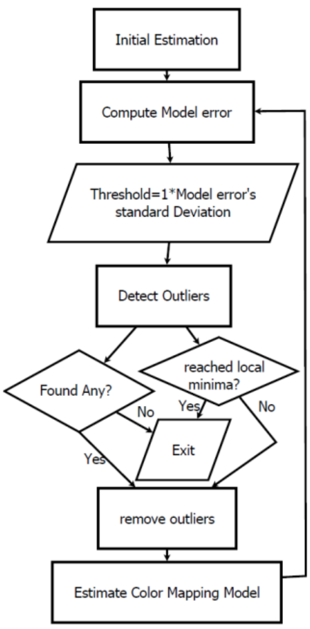


Figure 9: Threshold gets updated as outliers get removed

3.4 Threshold choice and iterative mapping

The *threshold* to compare neighborhood is updated according to the standard deviation (σ) of the model error. The iterative outlier removal works as follows:

1. Do the following until difference of two consecutive model errors is smaller than 0.1σ or no outlier is found.
2. Detect outliers (Section 3.3)
3. Remove outliers from data
4. Estimate the color mapping model.
5. Go to step 1

Details of these steps are shown in Figure 9.

Outlier removal Method	Observation Used
Classical	Feature Point Color
Neighborhood (Proposed)	(Feature+Neighborhood) Color
RANSAC	Image Coordinate

Table 1: Different observations used by three reported outlier detection methods.

4 Experimental results

For our experiments, we have used the Middlebury [9, 8, 3] and BOLD [5] stereo datasets. We will first show the performance of three outlier detection methods. It is important to remind that different outlier removal methods use different types of observations as shown in Table 1.

Now, to compare these three outlier removal methods we have used Middlebury [9, 8, 3] stereo datasets as they provide the ground truth. Our comparison will be based on the following definitions [1]:

- TP: true positives, i.e., number of correct matches;
- FN: false negatives, matches that were not correctly detected;
- FP: false positives, proposed matches that are incorrect;
- TN: true negatives, non-matches that were correctly rejected.

$$Precision = \frac{TP}{TP + FP}; Recall = \frac{TP}{TP + FN};$$
$$Accuracy = \frac{TP + TN}{(TP + FN) + (FP + TN)}$$

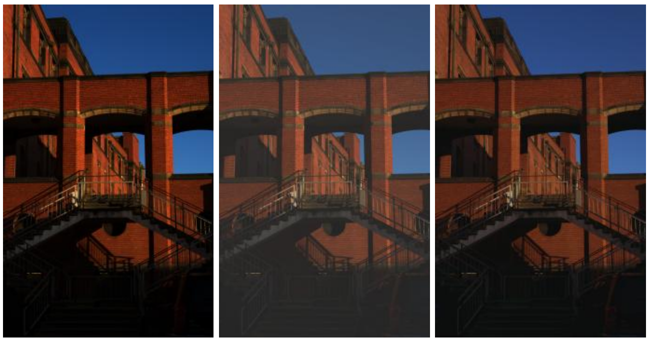
Stereo Set	Method	Recall	Precision	Accuracy
Teddy	Classical	**0.964**	0.102	0.141
	Neighborhood	0.893	**0.129**	**0.379**
	RANSAC	0.75	0.093	0.235
Cones	Classical	0.969	0.10	0.159
	Neighborhood	0.75	**0.134**	**0.512**
	RANSAC	**1.0**	0.106	0.192
Aloe	Classical	0.865	0.173	0.344
	Neighborhood	0.784	**0.216**	**0.531**
	RANSAC	**1.0**	0.179	0.295
Baby1	Classical	0.964	**0.435**	0.503
	Neighborhood	0.25	0.4	**0.566**
	RANSAC	**0.97**	0.401	0.43
Rocks1	Classical	0.816	0.148	0.496
	Neighborhood	0.816	**0.178**	**0.593**
	RANSAC	**0.974**	0.105	0.146
All Images	Classical	0.916	0.192	0.329
	Neighborhood	0.698	**0.212**	**0.516**
	RANSAC	**0.939**	0.177	0.259

Table 2: Comparison between three outlier detection methods for some Middlebury stereo datasets [9, 8, 3]

Following parameters are chosen for the experiment:

- Classical: *threshold* =1σ of the model error
- Neighborhood: *threshold* =1σ of the model error and neighborhood size= 3x3
- RANSAC: Distance threshold =1

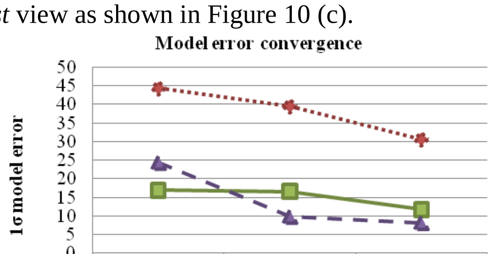
Table 2 reports the comparison between three outlier removal methods for five stereo datasets where the neighborhood method outperforms the other two methods in terms of overall accuracy and precision. Though both classical and neighborhood method uses color information as observations, overall performance of neighborhood method is better than classical method. The third method, RANSAC [2] which uses geometric information as observations, have poor precision and accuracy, but recall is significantly good and it is around 94%.



(a) (b) (c)

Figure 10: (a) Reference view (b) test view (c) color corrected test-view of "Building"[5]

A sample result of color compensation without the removal of outliers is shown in Figure 10. Color difference in *test* view is synthesized by manual change of contrast. Figure 10 shows a *reference*, a *test* and a *corrected test*-view of a scene called "Building". After computing the feature and color correspondences, Figure 12, Figure 13 and Figure 14 show the Red, Green and Blue channel estimation by nonlinear regression using levenberg-marquardt algorithm. So, from this estimation we have the gamma, offset, and gain parameters for all three channels. If we apply these estimated parameters (GOG) on the *test* view then we will get the *corrected test* view as shown in Figure 10 (c).



	Iteration1	Iteration2	Iteration3
Red 1σ model error	44.3344	39.5026	30.6083
Green 1σ model error	16.9027	16.6222	11.7341
Blue 1σ model error	24.4255	9.81435	8.1232

Figure 11: Threshold (1σ model error) eventually converges either by reaching local minimal (<0.1σ change in two consecutive iterations) or no outlier get detected.

Finally, Figure 11 shows how the model error estimation changes in the first three iterations for the "Building" scene (section 3.4)

5 Conclusions

We have proposed a robust outlier detection and thus robust color correction to compensate stereo color differences. The novelty is gained by using spatial color information around the correspondences between the *reference* and the iteratively color compensated *test* view. Experimental results show that our proposed method could bring new information in terms of outlier removal. The new method has two main advantages over the existing methods. The first advantage comes from the fact that it's more robust and effective to analyze the iterative color-corrected image (*initial corrected test-view to corrected view*) than a non color-corrected image. The second advantage is that unlike RANSAC, it does not depend on geometric assumptions such as planar scene or do not use geometric information. However, the proposed method has limitation such as on homogeneous neighborhoods. But we argue from our experimental result that very few correspondences are found in homogeneous neighborhoods and thus the impact of this limitation is not significant for the estimation of color mapping. In our future work, we would like to combine RANSAC and the proposed neighborhood method as they use different types of observations. We also want to analyze the impact of the proposed method on different estimation model. Lastly, we also expect to take into account the temporal aspect for the video applications.

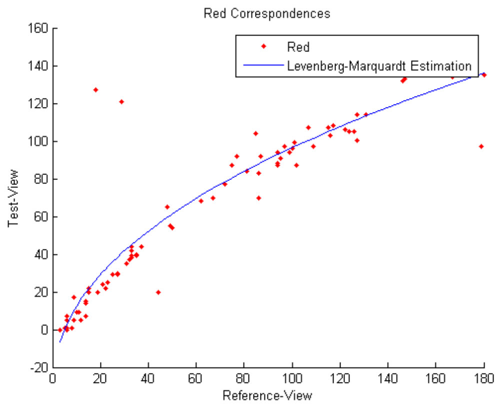


Figure 12: Red channel estimation of scene "Building"

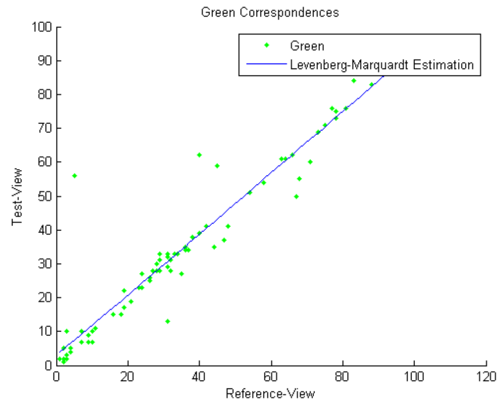


Figure 13: Green channel estimation of scene "Building"

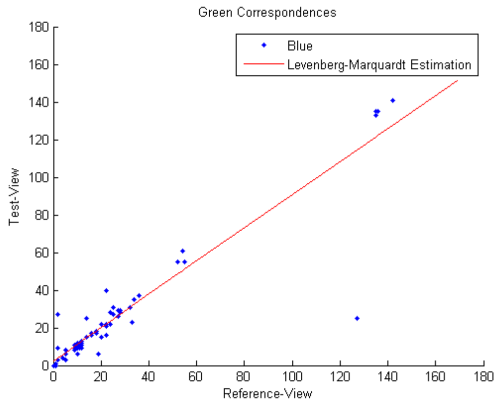


Figure 14: Blue channel estimation of scene "Building"

6 References

[1] T. Fawcett. An introduction to ROC analysis. *Pattern recognition letters*, 27(8):861–874, 2006.

[2] M.A. Fischler and R.C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[3] H. Hirschmuller and D. Scharstein. Evaluation of cost functions for stereo matching. In *Proc. Computer Vision and Pattern Recognition, 2007*, pages 1–8, 2007.

[4] Jung J., Kang M., and Ho Y. Color Corrected Newspaper Sequence for 3D Video Coding. *ISO/IEC JTC1/SC29/WG11*, January 2011.

[5] X. Jiang, A. Schofield, and J. Wyatt. Correlation-based intrinsic image extraction from a single image. In *Proc. ECCV*, pages 58–71, 2010.

[6] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[7] F. Pitié, A.C. Kokaram, and R. Dahyot. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding*, 107(1-2):123–137, 2007.

[8] D. Scharstein and C. Pal. Learning conditional random fields for stereo. In *Proc. Computer Vision and Pattern Recognition 2007*, pages 1–8, 2007.

[9] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proc. Computer Vision and Pattern Recognition, 2003*, volume 1, pages I–195, 2003.

[10] Yu-Wing Tai, Jiaya Jia, and Chi-Keung Tang. Local Color Transfer via Probabilistic Segmentation by Expectation-Maximization. In *Proc. of CVPR*, volume 1, pages 747–754, 2005.

[11] M.P. Tehrani, Ishikawa A., Sakazawa S., and Koike A. Iterative colour correction of multicamera systems using corresponding feature points. *Journal of Visual Communication and Image Representation*, 21(5-6):377–391, 2010.

[12] Q. Wang, X. Sun, and Z. Wang. A Robust Algorithm for Color Correction between Two Stereo Images. In *Proc. Asian Conference on Computer Vision*, volume 5995, pages 405–416. 2010.

[13] K. Yamamoto and R. Oi. Color correction for multi-view video using energy minimization of view networks. *International Journal of Automation and Computing*, 5(3):234–245, 2008.