

Binocular Camera Calibration Using Rectification Error

Derek Bradley
University of British Columbia
Vancouver, Canada
`bradleyd@cs.ubc.ca`

Wolfgang Heidrich
University of British Columbia
Vancouver, Canada
`heidrich@cs.ubc.ca`

Abstract

Reprojection error is a commonly used measure for comparing the quality of different camera calibrations, for example when choosing the best calibration from a set. While this measure is suitable for single cameras, we show that we can improve calibrations in a binocular or multi-camera setup by calibrating the cameras in pairs using a rectification error. The rectification error determines the mismatch in epipolar constraints between a pair of cameras, and it can be used to calibrate binocular camera setups more accurately than using the reprojection error. We provide a quantitative comparison of the reprojection and rectification errors, and also demonstrate our result with examples of binocular stereo reconstruction.

1 Introduction

One of the most common problems in computer vision is camera calibration. The process of calibration is to determine the intrinsic and extrinsic parameters of a camera from a number of correspondences between 3D points and their projections onto one or multiple images [17, 21]. Most often this is accomplished using a calibration plane with a checkerboard or other known marker pattern [5]. In this paper we focus on the problem of multiple camera calibration, where the relative projection matrices between cameras must be very accurate, for example in binocular stereo [11], multi-view stereo [12], human pose reconstruction [18, 4], and novel view interpolation [22, 14]. For these applications, the quality of the camera calibration has a direct impact on the quality of the results. We will specifically explore binocular camera calibration, and later discuss how our approach can be used in a many-camera setup.

The two most common techniques for camera calibration are those of Tsai [17] and Zhang [21]. While the method of Tsai has the advantage that it can handle both coplanar and non-coplanar input points, the easiest and most practical approach is to use a calibration grid or checkerboard

of coplanar points. When the input points lie on a single plane, it is wise to have multiple input images containing different planar grid orientations in order to ensure a robust calibration, even though Tsai's method can operate on a single input image. Zhang's calibration method strictly enforces these conditions, requiring multiple images of a planar calibration grid. In this paper we will employ the method of Zhang, although our algorithm is applicable to other calibration methods, including that of Tsai.

Once multiple calibration images are collected, the calibration process proceeds by finding the projection of known grid points in the images and then solving for the camera parameters that minimize the reprojection error of the detected points. The result is a single set of intrinsic parameters for the entire image sequence and multiple sets of extrinsic parameters, one for each calibration grid location. All of the images are used to compute the intrinsic parameters, however each set of extrinsic parameters is computed from a single image and the corresponding grid location. The problem, as pointed out by Zaharescu et al. [20], is to determine which of the extrinsic parameters to use. When collecting the images, a small number of precisely oriented grid locations could be recorded to ensure that the entire capture space is sampled by points. Alternatively, a short video could be captured where the calibration grid is rotated more or less at random, resulting in hundreds of input images but also covering the entire space. The latter approach is more practical, however there will be many more calibrations to choose from. Each different location of the calibration grid produces different extrinsic parameters for the camera, with varying accuracy depending on the grid orientation, visibility, illumination, noise and a variety of other parameters. The standard approach to determine which grid location to use is to keep the extrinsic parameters that give the lowest reprojection error in the *single* image from which the extrinsics are calculated. This is the only image for which 3D points are known, and thus the best we can hope for in single camera calibration. However, the reprojection error for a single grid location is only guaranteed to be accurate for points that lie on the plane of that grid, and other points off

the plane can have a much higher reprojection error. Although multiple grid locations are captured in the sequence, the 3D location of one grid relative to another is usually not available, and so we are unable to compute reprojection errors for points off the plane of each grid.

In this paper we show that in a binocular camera setup we can use *all* the grid locations in all images to evaluate *each* potential set of extrinsic parameters and more accurately determine the calibration for a pair of cameras. We do this by estimating a reprojection error for the entire volume spanned by the calibration grid over the whole sequence of images, rather than simply a reprojection error for only the points that lie on the calibration plane in one image, as with the single-camera approach described above. Our new approach is partly inspired by the *normalized stereo calibration error* (NSCE) of Weng et al. [19], which evaluates multi-calibration by measuring the triangulation error of a pair of cameras using known 3D points. However, for any given calibration, we do not know the 3D point locations on any calibration plane except the one that defines the extrinsic parameters, so the NSCE is not applicable. Even though we do not know the 3D locations of points on the other grids, the projection of common points onto the two cameras can be found and we can measure the accuracy of the epipolar geometry for each potential set of extrinsic parameters. This accuracy measure, which we call the *rectification error*, measures the subpixel scanline difference between the projection of common points onto the rectified versions of the two camera images. If the calibration of the two cameras is accurate then any scene point visible by both cameras will project onto the same scanline in the rectified images. Any discrepancy indicates an error in the calibration. Combining epipolar geometry and calibration has proven useful in past research. A similar measure to our rectification error was used by Furukawa and Ponce to evaluate their calibration algorithm [6], and Sinha et al. also use epipolar geometry to calibrate camera networks from the silhouettes of objects [13].

A related technique for finding the parameters of a multi-camera setup is bundle adjustment [8]. The goal of bundle adjustment is to simultaneously solve for optimal scene structure and camera parameters given a set of 2D interest points that are common among subsets of views. However, it is well known that bundle adjustment fails if there is insufficient overlap between many camera views [16, 9]. Typically, each feature point must be visible in at least four views in order to produce reliable camera parameter estimates. It is not always possible to provide such redundancy, for example in 360° sparse multi-view reconstruction, where only a small number of cameras are available [18, 4]. Since we only require pairs of cameras to have overlapping views, our rectification error can be used to find very accurate camera calibrations in these cases, more ac-

curate than using the standard reprojection error. This result will be demonstrated quantitatively as well as qualitatively using multi-view stereo reconstruction.

The remainder of the paper is organized as follows. In Section 2 we provide an overview of the camera model and camera calibration. We introduce the rectification error in Section 3, and show experimental results comparing our measure to the reprojection error in Section 4. Finally, we conclude in Section 5.

2 Camera Calibration Overview

We begin by describing the camera model, a typical planar-based calibration method, and the standard rectification error used to evaluate calibration.

2.1 Camera Model

A camera model consists of a set of intrinsic parameters, which define how the camera forms an image, and a set of extrinsic parameters, which define the position and orientation of the camera in the world. The intrinsic parameters include the focal length in pixels (fx, fy), the principal point (px, py), and a skew factor, s (which is often ignored and set to zero [15]). These parameters form the camera matrix K , defined as

$$K = \begin{bmatrix} fx & s & px \\ 0 & fy & py \\ 0 & 0 & 1 \end{bmatrix}. \quad (1)$$

The extrinsic parameters include a rotation matrix $R \in \mathbb{R}^{3 \times 3}$ and a translation vector $t = [tx \ ty \ tz]^\top$, which relate the world coordinate frame to the camera coordinate frame. The full calibration forms a projection matrix $P \in \mathbb{R}^{3 \times 4}$, defined as

$$P = K \cdot [R|t], \quad (2)$$

which maps world points to camera pixels.

The calibration parameters also include radial distortion coefficients to correct lens aberration. From this point on we will assume that radial distortion has been corrected for all camera images.

2.2 Calibration

We will focus on the common calibration technique of Zhang [21]. This method is widely used and an implementation is readily available in the OpenCV library [1]. The input is a set of several views of a known calibration grid, where every view is described by several world-to-2D point correspondences. Typically only 20-30 views are required, however we prefer to capture a video sequence of hundreds

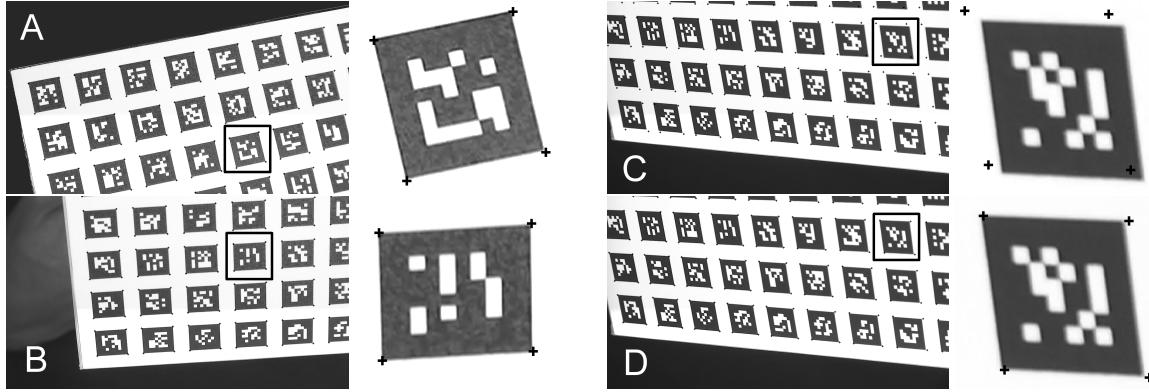


Figure 1. Visualizing reprojection errors for two different calibrations of a single camera. The top row (A & C) is calibrated with extrinsic parameters from grid A, while the bottom row (B & D) is calibrated from grid B (intrinsic parameters are the same for both). The points on the common grid in C and D lie off the planes of both A and B. The calibration from grid B is more accurate than the one from grid A, as shown in the zoom images on the far right.

of views, rather than attempting to choose the small number of views that will provide the best calibration. The world points are defined by the calibration plane ($z = 0$), and an elegant marker pattern and corner detection scheme has been proposed by Fiala and Shu [5] to detect the 2D correspondences. The calibration method then solves for all the camera parameters such that the reprojection error of the points is minimized. The result is a single K matrix and multiple R_i and t_i transformations, one for each view of the calibration grid. Since all grid locations are used to estimate K we assume the intrinsic parameters are computed robustly. This assumption is common for multi-camera setups [20], and in practice we have observed this to be true. However, since an R_i and t_i are determined for each input image, the accuracy of each transformation depends on a single grid location and how well the grid pattern was detected in the single image. Therefore, some transformations can be more accurate than others. We illustrate this effect in Figure 1. Here we choose two different grid locations, A and B , each resulting in a different projection matrix $P_A = K \cdot [R_A|t_A]$ and $P_B = K \cdot [R_B|t_B]$. Both P_A and P_B appear to be accurate when reprojecting the *planar* grid points onto the image, as we see in the first set of zoom images. However, when reprojecting points that are not on the calibration plane used to compute the extrinsic parameters (e.g. the plane in Figure 1, C and D), we see that P_A (top row) is less accurate than P_B (bottom row).

The problem is to determine which of the multiple R_i and t_i transformations is the most accurate. The common approach is to select the transformation with the lowest reprojection error for the single calibration grid used to compute the transformation, as we describe next.

2.3 Reprojection Error

Let $P_i = K \cdot [R_i|t_i]$ be the projection matrix of camera c for calibration grid view i . Assume we have detected k grid points x_j in the image, corresponding to 3D planar points X_j . Then the reprojection error for image i is

$$e_{rep}^c[i] = \frac{1}{k} \sum_{j=1}^k \|P_i(X_j) - x_j\|. \quad (3)$$

The reprojection error has been widely used as the main tool for evaluating camera calibration, either in the form presented above [5, 10, 20], or in a normalized form [15, 19]. This is because a low reprojection error indicates an accurate projection matrix, at least for the points on the plane that were used to compute the projection matrix. The problem is that the reprojection error may increase for 3D points off the plane, as we saw in Figure 1. The worst case is when the planar grid is perpendicular to the optical axis of the camera, and the calibration may only be accurate for essentially a single depth. The reprojection error would be more accurate if there were additional 3D points available, off the plane of the calibration grid, for which we had corresponding detected 2D pixels. The image sequence does in fact contain many different grid locations for which corner points are detected, however the corresponding 3D locations of those points are not known. Figure 2 (a) illustrates the problem. The camera is calibrated with respect to calibration grid i_1 , which defines the world coordinate system. Unfortunately, the 3D location of point Q on grid i_2 is not available, and so Equation 3 cannot be applied. For this reason, the reprojection error can only be evaluated on

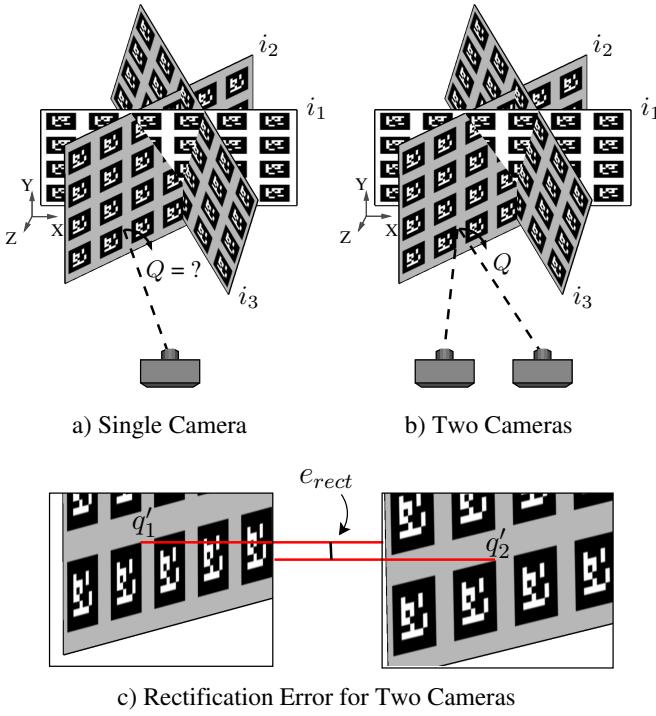


Figure 2. Evaluating the extrinsic parameters computed using calibration grid i_1 . a) With a single camera we cannot compute reprojection errors for points on other planes. b) With two cameras, although the 3D location of Q is not known, its projection onto the two cameras can be found. c) The rectification error is the scanline difference between the projection of Q onto the rectified versions of the two cameras.

plane i_1 . However, if two cameras observe the same calibration grid sequence (such as the case of binocular stereo), then we propose a pair-wise calibration algorithm using the rectification error.

3 Rectification Error

When two cameras observe the same sequence of calibration grid locations, all grids can be used to evaluate the calibration accuracy for each individual set of extrinsic parameters. As we have seen, the standard reprojection error in Equation 3 cannot be applied to points off the main grid (grid i_1 in Figure 2). However, if a point, Q , on some other grid is visible in both cameras then epipolar constraints tells us that the projection of Q onto the rectified versions of the left and right images should lie on the same scanline, if the

calibration of the cameras is accurate (see Figure 2 (b) and (c)). This fact is independent of the 3D location of Q , and thus we are able to use all detected points from all grid locations that are common in both views.

From the above observation, we form a measure of *rectification error* for two cameras c_1 and c_2 , and calibration grid view i as follows. For each calibration grid, let the k^{th} detected grid point on the image plane of c_1 corresponding to unknown 3D point Q^k be $q_1^k = (u_1^k, v_1^k)$, and on the image plane of c_2 be $q_2^k = (u_2^k, v_2^k)$. For $c \in \{1, 2\}$, we denote $q_c^k[0]$ to refer to u_c^k and $q_c^k[1]$ to refer to v_c^k . Then,

$$e_{rect}^{c_1}[i] = \frac{1}{N} \sum_{j=1}^N \left(\frac{1}{M_j} \sum_{k=1}^{M_j} \left| (T_i^{c_1} q_1^k)[1] - (T_i^{c_2} q_2^k)[1] \right| \right) \quad (4)$$

where $T_i^{c_1}$ is the rectifying transformation for camera c_1 using calibration i , and $T_i^{c_2}$ is defined similarly for camera c_2 . N is the total number of grid positions in the sequence, and M_j is the number of grid points that are commonly detected in both camera views for grid position j . Note that the rectification error is symmetric, so $e_{rect}^{c_2} = e_{rect}^{c_1}$. We compute the rectifying transformations using the method of Fusiello et al. [7]. We illustrate the rectification error for a particular point Q in Figure 2 (c), where

$$\begin{aligned} q'_1 &= T_i^{c_1} q_1, \\ q'_2 &= T_i^{c_2} q_2. \end{aligned}$$

This rectification error measure can now be used to determine more accurate binocular camera calibrations than the standard method of using the reprojection error. In fact, the calibration grid A in Figure 1 was the one with the lowest reprojection error, and grid B had the lowest rectification error. As we saw in that figure, the rectification error determined a more accurate calibration.

4 Experimental Results

We demonstrate the quality of the rectification error by calibrating a multi-camera setup of 14 cameras, arranged as seven binocular pairs. We show qualitative results by performing stereo reconstruction of a static object, and quantitatively prove that the rectification error is a better tool for evaluating binocular camera calibrations than the reprojection error.

Our static object is a human head model made from styrofoam, which we have painted in order to give it a high-frequency surface texture to aid the stereo reconstruction. We use an ARTag calibration grid [5] and the method of Zhang [21] to compute one set of intrinsic parameters and multiple sets of extrinsic parameters for each camera. In all

of our experiments we record a calibration video for each camera, where the calibration grid is rotated and translated throughout the capture volume, resulting in over one thousand planar grid orientations. As we have discussed, the problem lies in choosing which extrinsic parameters to use. We will show in Section 4.4 that this choice is critical to the quality of the calibration and the resulting stereo reconstruction.

In this paper we are in fact advocating two principles. One, cameras should be calibrated in binocular pairs, and two, cameras should be calibrated using the rectification error. In order to show the importance of combining these principles we have performed three experiments. In the first experiment we find the best extrinsic parameters for all cameras globally using reprojection error. This is what we refer to as the standard approach, which we use as a baseline for comparison. In the second experiment we calibrate the cameras in pairs, but still use the reprojection error. This experiment will show that simply calibrating in pairs is not sufficient, if the rectification error measure is not used. Finally, in the third experiment we calibrate in pairs and use the rectification error. As we will see, the calibration quality achieved in the third experiment is consistently superior to that of the first two experiments. Additionally, we will see that the rectification error is directly proportional to the quality of the stereo reconstruction, no-matter which method is used to choose the extrinsic parameters, unlike the reprojection error which can be misleading.

4.1 Exp. 1 - Global Reprojection Error

Let S be the set of all calibration grids visible in *every* camera view. We choose the single grid location that yields the lowest average reprojection error among all N_c cameras. Specifically, the grid i that minimizes

$$\min_{i \in S} \sum_{c=1}^{N_c} (e_{rep}^c[i]). \quad (5)$$

The benefit of this approach is that all cameras are calibrated to the same world coordinate system. However the drawback is that some cameras will be calibrated better than others (see Figure 3 and Section 4.4). In practice, depending on the camera setup it may also be difficult to find calibration grids that are visible in all camera views.

4.2 Exp. 2 - Pair-wise Reprojection Error

Let S be the set of all calibration grids visible by a specific *pair* of cameras, c_1 and c_2 . We choose the single calibration grid that yields the lowest average reprojection error for those two cameras. Specifically, the grid i that minimizes

$$\min_{i \in S} \sum_{c \in \{c_1, c_2\}} (e_{rep}^c[i]). \quad (6)$$

The benefit of this approach is that there will be more grids to choose from, resulting in lower reprojection errors (see Table 1 and Section 4.4). The drawback is that the low reprojection errors are misleading, as some calibrations are still not very accurate (Figure 3). Also, the cameras will not be calibrated in the same world coordinate system, so combining stereo results from different pairs is no-longer trivial.

4.3 Exp. 3 - Pair-wise Rectification Error

Let S be the set of all calibration grids visible by a specific *pair* of cameras, c_1 and c_2 . We choose the single calibration grid that yields the lowest average rectification error for those two cameras. Specifically, the grid i that minimizes

$$\min_{i \in S} \sum_{c \in \{c_1, c_2\}} (e_{rect}^c[i]). \quad (7)$$

This approach has the same benefit as experiment two, in that there are more grids to choose from than in experiment one, and the same drawback in that the cameras will not be calibrated in the same coordinate system. However, this approach yields the most accurate calibrations (again, see Figure 3).

4.4 Stereo Reconstruction Analysis

We analyze the quality of the calibrations in each experiment by performing stereo reconstruction of our styrofoam head. We expect that the accuracy of each calibration will be reflected in the number of outliers in the corresponding depth map. We use the reconstruction method of Bradley et al. [3], although other methods could equally be employed. In this reconstruction algorithm, depth outliers are automatically rejected by thresholding on the correlation score, and also through a spatial depth filtering post-process. In effect, the quality of the calibration is related to the completeness of the depth map.

Our cameras are Sony HDR-SR7 camcorders, which capture high-definition video, although we only reconstruct a single frame. The cameras are placed very close to the reconstruction object and zoomed in to see the painted surface details. This type of close-range setup is very challenging to calibrate as even the slightest calibration error results in inaccurate reconstruction results.

Qualitative Analysis. Figure 3 shows the depth maps for each experiment. As we can see, the results from experiment one tend to be rather poor, with only one camera pair

(8-9) producing a valid and mostly complete depth map. We can also see that depth maps for some pairs are only accurate at a single depth (i.e. pairs 2-3 and 10-11), which indicates that the chosen calibration grid was likely nearly-perpendicular to the optical axes of the cameras, resulting in a low reprojection error but also low calibration accuracy. We see that experiment two sometimes produces better depth maps than experiment one, although sometimes they are worse. This indicates that calibrating in pairs alone is not sufficient. The results of experiment three always produce the most complete depth maps, indicating the importance of calibrating in pairs and using the rectification error.

Quantitative Analysis. Table 1 shows reprojection errors (Equation 3) and rectification errors (Equation 4) for the resulting calibration of every camera from each experiment, measured in pixels. This data corresponds to the stereo results in Figure 3. Using this table we can quantitatively analyze the difference between the two measures. As expected, the reprojection errors for experiment two are lower than experiments one and three, although this has no correlation with the quality of the stereo reconstruction and so the reprojection error is not always indicative of calibration accuracy. On the other hand, the rectification error varies between experiments one and two, and a lower error value directly correlates with better stereo reconstruction (refer again to Figure 3). As expected, experiment three produced the lowest rectification errors and also the best reconstruction results. It is clear from comparing Table 1 to the depth images in Figure 3 that our proposed rectification error accurately measures the quality of binocular camera calibrations.

Merging stereo pairs. If we wish to generate a 3D model from the stereo reconstructions, the cameras must be aligned in the same world coordinates. For experiments two and three, the resulting pair-wise calibrations need to be transformed rigidly into a global coordinate system. This can be achieved in a number of ways. First, experiment one could be performed to establish the world coordinate system, and then each pair could be commonly transformed into the global coordinates. In this approach, the same transformation would be applied to each camera in a pair, keeping the relative transformations between the two cameras fixed and therefore very accurate. The drawback is that inter-pair alignment would be less accurate than the pairwise alignment. A second option is to perform pair-wise calibration between non-stereo pairs and *build up* a global calibration by sequentially adding pairs. For example, if the stereo pairs are $[c_1, c_2]$, $[c_3, c_4]$, and $[c_5, c_6]$, then pair-wise calibration could be performed between cameras c_2 and c_3 , thus aligning the first two pairs, and then between c_4 and c_5 , thus aligning the last pair with the first two. Finally, a third option is to complete the binocular stereo reconstruction in

Cam	Reprojection Error			Rectification Error		
	Exp. 1	Exp. 2	Exp. 3	Exp. 1	Exp. 2	Exp. 3
0	1.57	0.66	1.45	1.20	1.21	1.01
1	2.04	1.08	1.84	1.20	1.21	1.01
2	1.77	0.94	2.08	12.64	1.75	1.08
3	1.80	1.02	2.19	12.64	1.75	1.08
4	1.75	0.54	1.40	4.48	1.28	0.75
5	0.83	0.56	1.62	4.48	1.28	0.75
6	1.89	0.85	2.10	2.09	1.81	0.78
7	1.76	1.18	1.19	2.09	1.81	0.78
8	1.96	0.93	1.83	1.43	3.33	1.02
9	2.65	1.51	2.22	1.43	3.33	1.02
10	2.03	0.90	1.33	14.93	7.76	1.22
11	1.64	1.16	1.89	14.93	7.76	1.22
12	1.63	1.04	3.42	3.05	3.59	1.49
13	1.56	1.12	4.60	3.05	3.59	1.49

Table 1. Comparing reprojection and rectification errors for the three experiments. Note that the reprojection errors are quite similar and non-indicative of the stereo results in Figure 3, while the rectification errors vary drastically and they directly reflect the quality of the stereo reconstruction.

pairs and then align the different depth maps or 3D surfaces using a rigid alignment technique such as ICP [2]. In practice, this is how we generate the final 3D model in Figure 4. This result, using calibrations from experiment three, shows the high accuracy of calibrations computed using the rectification error.

5 Conclusion

We propose a new technique for calibrating binocular cameras using a pair-wise rectification error. This technique can be used to significantly improve stereo reconstruction results, as compared to using the standard reprojection error for calibration.

In order for a camera pair to be considered as a binocular pair for calibration, the only requirement is that the two cameras observe the same sequence of calibration grid images. This means that the cameras should be placed fairly close together with significant overlap in their views. However, this condition is already met in most binocular camera setups, if the cameras are to be used for applications such as stereo reconstruction or novel view interpolation.

We validate our approach by calibrating seven binocular pairs using our rectification error measure, and further demonstrate how the cameras can be combined into an

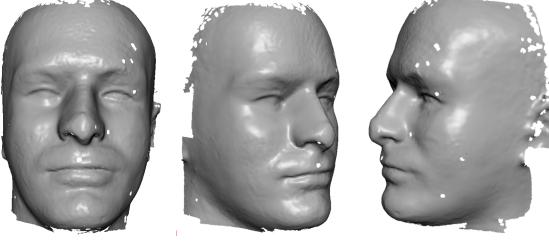


Figure 4. Combined multi-view 3D result from Experiment three.

accurate multi-view reconstruction setup. We believe our method can benefit a number of applications that make use of camera arrays, where binocular pairs are available.

References

- [1] Open source computer vision library.
- [2] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Trans. on PAMI*, 14(2):239–256, 1992.
- [3] D. Bradley, T. Boubekeur, and W. Heidrich. Accurate multi-view reconstruction using robust binocular stereo and surface meshing. In *Proc. CVPR*, 2008.
- [4] E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel, and S. Thrun. Performance capture from sparse multi-view video. *ACM Trans. Graphics (Proc. SIGGRAPH)*, page 98, 2008.
- [5] M. Fiala and C. Shu. Self-identifying patterns for plane-based camera calibration. *Machine Vision and Applications*, 19(4):209–216, 2008.
- [6] Y. Furukawa and J. Ponce. Accurate camera calibration from multi-view stereo and bundle adjustment. *Int. J. Comput. Vision*, 84(3):257–268, 2009.
- [7] A. Fusiello, E. Trucco, and A. Verri. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, 12(1):16–22, 2000.
- [8] R. I. Hartley. Euclidean reconstruction from uncalibrated views. In *Proceedings of the Second Joint European - US Workshop on Applications of Invariance in Computer Vision*, pages 237–256, 1994.
- [9] Y. S. Hung and W. K. Tang. Projective reconstruction from multiple views with minimization of 2d reprojection error. *Int. J. Comput. Vision*, 66(3):305–317, 2006.
- [10] C. Ouyang, G. Wang, Q. Zhang, W. Kang, and H. Ding. Evaluating harris method in camera calibration. In *IEEE Eng. in Medicine & Biology*, pages 6383–6386, 2005.
- [11] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. Journal of Computer Vision*, 47(1-3):7–42, 2002.
- [12] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *CVPR*, 2006.
- [13] S. N. Sinha, M. Pollefeys, and L. McMillan. Camera network calibration from dynamic silhouettes. *CVPR*, 1:195–202, 2004.
- [14] A. Smolic, K. Müller, K. Dix, P. Merkle, P. Kauff, and T. Wiegand. Intermediate view interpolation based on multiview video plus depth for advanced 3d video systems. In *Proc. ICIP*, pages 2448–2451, 2008.
- [15] W. Sun and J. R. Cooperstock. An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques. *Machine Vision and Applications*, 17(1):51–67, 2006.
- [16] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon. Bundle adjustment - a modern synthesis. *Lecture Notes in Computer Science*, 1883:298–372, 1999.
- [17] R. Y. Tsai. An efficient and accurate camera calibration technique for 3d machine vision. In *Proc. CVPR*, pages 364–374, 1986.
- [18] D. Vlasic, I. Baran, W. Matusik, and J. Popović. Articulated mesh animation from multi-view silhouettes. *ACM Trans. Graphics (Proc. SIGGRAPH)*, page 97, 2008.
- [19] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on PAMI*, 14(10):965–980, 1992.
- [20] A. Zaharescu, R. Horaud, R. Ronfard, and L. Lefort. Multiple camera calibration using robust perspective factorization. In *3DPVT*, pages 504–511, 2006.
- [21] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. on PAMI*, 22(11):1330–1334, 2000.
- [22] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High-quality video view interpolation using a layered representation. In *ACM SIGGRAPH*, pages 600–608, 2004.

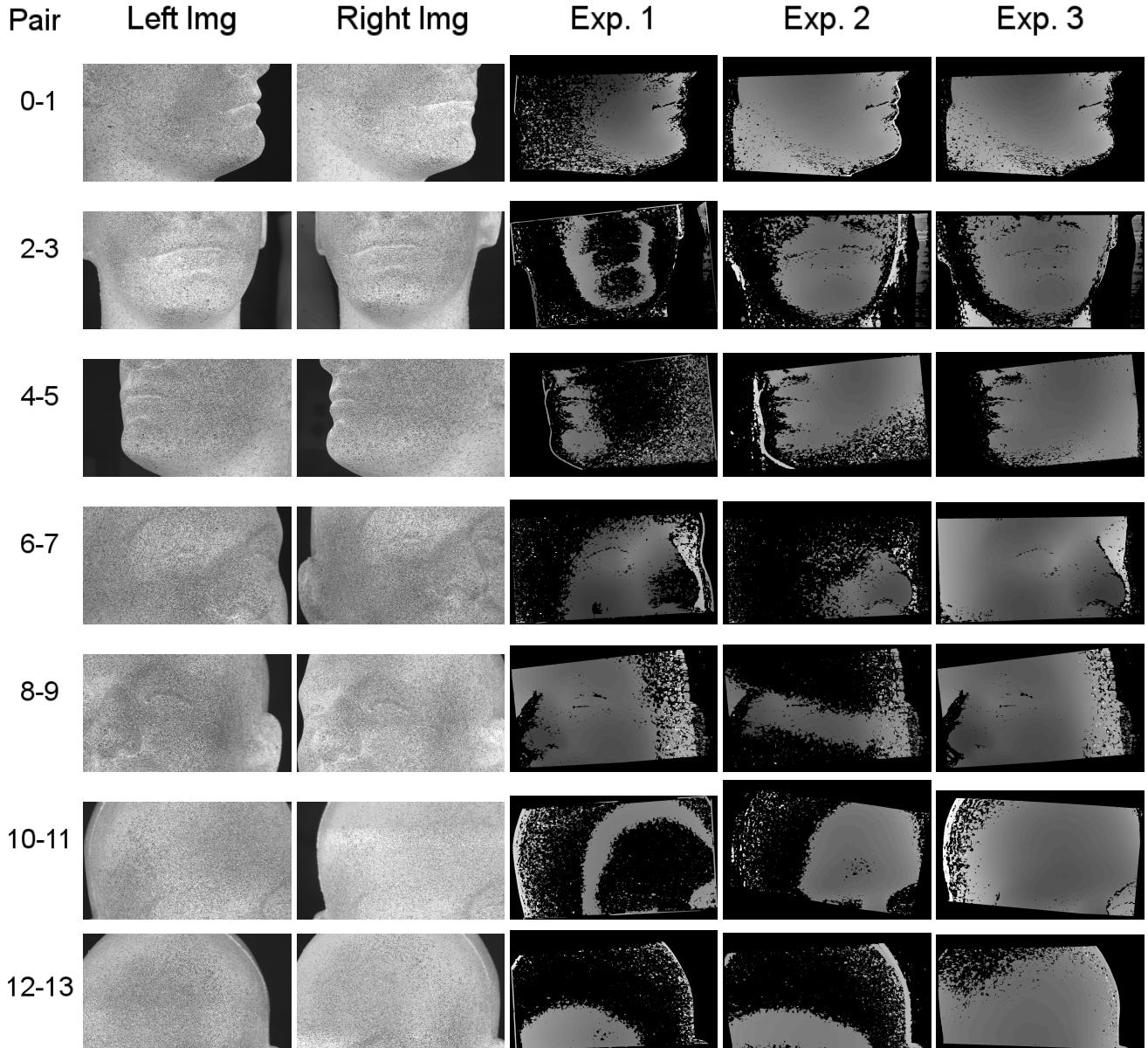


Figure 3. Binocular stereo results for a styrofoam head using the calibrations from the three experiments. Each row is a separate camera pair. From left to right we show the two camera views, then the depth maps from experiments one, two and three respectively. Corresponding quantitative results are shown in Table 1.