CENTER FOR
MACHINE PERCEPTION

CZECH TECHNICAL
UNIVERSITY

REPRINT

3D Metric Reconstruction from Uncalibrated Omnidirectional Images

Branislav Mičušík, Daniel Martinec
and Tomáš Pajdla

micusb1@cmp.felk.cvut.cz,
martid1@cmp.felk.cvut.cz,
pajdla@cmp.felk.cvut.cz

Branislav Mičušík, Daniel Martinec and Tomáš Pajdla, 3D Metric Reconstruction from Uncalibrated Omnidirectional Images, Asian Conference on Computer Vision, Jeju Island, Korea, January 28-30, 2004

Available at
ftp://cmp.felk.cvut.cz/pub/cmp/articles/micusik/Micusik-ACCV2004.pdf

Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Technická 2, 166 27 Prague 6, Czech Republic
fax +420 2 2435 7385, phone +420 2 2435 7637, www: http://cmp.felk.cvut.cz

3D Metric Reconstruction from Uncalibrated Omnidirectional Images

*Branislav Mičušík, Daniel Martinec and Tomáš Pajdla**

Center for Machine Perception, Dpt. of Cybernetics, Faculty of Elect. Engineering,
Czech Technical University in Prague, 121 35 Prague, Karlovo nam. 13, Czech Republic
{micusb1,martid1,pajdla}@cmp.felk.cvut.cz

ABSTRACT

We show that it is possible to obtain a very complete 3D metric reconstruction of the surrounding scene from two or more uncalibrated omnidirectional images. In particular, we demonstrate that omnidirectional images with angle of view above 180° can be reliably autocalibrated. We also show that wide angle images provide reliable information about their camera positions and orientations. We link together a method for simultaneous omnidirectional camera model and epipolar geometry estimation and a method for factorization-based 3D reconstruction in order to obtain metric reconstruction of unknown scene observed by uncalibrated omnidirectional images. The 3D reconstruction is done from automatically established image correspondences only. We demonstrate our method in experiments with Nikon FC–E8 and Sigma 8mm-f4-EX fish-eye lenses. Nevertheless, the proposed method can be used for a large class of non-perspective central omnidirectional cameras.

1. INTRODUCTION

In comparison to standard cameras with narrow view angle, omnidirectional cameras capture larger part of a surrounding scene. Large angle of view often allows to establish more spacious point correspondences which leads to a more complete 3D reconstruction from fewer images. Notice in Fig. 1 how large part of a scene can be reconstructed from *two* omnidirectional images only. An occurrence of degenerate scenes (e.g. when only a single plane is observed in the image) is less probable with omnidirectional images and therefore more stable ego-motion estimation is often achieved.

We show that 3D metric reconstruction of the surrounding scene from two or more uncalibrated omnidirectional images can be performed very similarly as with standard perspective cameras. First, the omnidirectional camera is calibrated using image correspondences and epipolar constraint [15]. Secondly, a projective factorization-based reconstruction from many images handling occlusions [12] is used. Finally, an upgrade to a metric reconstruction is performed. The proposed linear estimation techniques give a good starting point for a non-linear bundle adjustment enforcing metric constraints on the reconstruction and including lens nonlinearity.

Fig. 1. 3D metric reconstruction of a Venice yard from *two* uncalibrated omnidirectional images with automatically detected point correspondences.

In the next, we assume rigid scene, elliptical (usually circular) view field, and approximately known corresponding view angle. We concentrate on fish-eye lenses, namely Nikon FC–E8 and Sigma 8mm-f4-EX fish-eye. However, the proposed method for 3D reconstruction can be applied to a large class of non-perspective central omnidirectional cameras based on lenses as well as on mirrors. Fish-eye lenses with view angle larger than 180° are, at least in

*This research was supported by the following projects: CTU 8306813, CTU 8306713, GAČR 102/01/0971, MSM 212300013, MŠMT Kontakt 22-2003-04, BENOGO–IST–2001–39184, STINT Dnr IG2003-2 062.

terms of image point representation and image formation non-linearity, same as central panoramic catadioptric cameras [17].

Previous works on the calibration of omnidirectional catadioptric cameras assumed presence of lines in scene [6], planar motion of the camera [7], or used some information about the shape of mirrors [4]. In [10], the calibration was performed from point correspondences and epipolar constraint through minimizing of an objective function. Our approach is similar in its goal but introduces a new method that provides a closed-form solution of camera model parameters and epipolar geometry as a solution of a quadratic eigenvalue problem. It allows to incorporate our method in RANSAC robust estimation technique handling mismatches in automatically established point correspondences.

Previous work related to 3D reconstruction from omnidirectional images assumed uncalibrated [5] or usually *calibrated* catadioptric sensors. In [5], the para-catadioptric camera calibration is performed from an image of the absolute conic. It is shown that Euclidean reconstruction is feasible from two views with constant parameters. Relations that exist between multiple views of a static scene, where the views can be taken by any mixture of para-catadioptric, perspective or affine cameras, were described in [18] and usage of this theory for motion estimation, 3D reconstruction or (self-) calibration was indicated. The 3D reconstruction from large sets of calibrated omnidirectional images with help of GPS was introduced in [14]. Similarly, in [1], a multi-baseline stereo algorithm for 3D reconstruction of an environment from a set of panoramic images with known camera positions was described. There were other works, different in principle from our method, assuming calibrated omnidirectional cameras for 3D reconstruction [3, 11] and structure from motion [2].

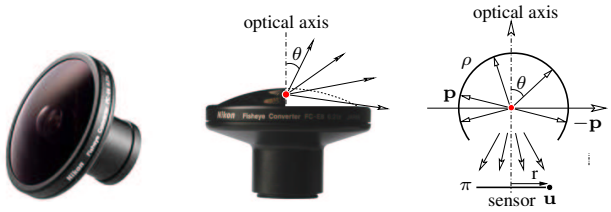
The main contribution of our method insists in the extension of multiple view metric 3D reconstruction from many uncalibrated images to *omnidirectional* cameras with large field of view and *highly nonlinear* image projection.

2. OMNIDIRECTIONAL CAMERA

Rays of the image will be represented as a set of unit vectors in $\mathbb{R}^3$ such that *one* vector corresponds just to *one* image of a scene point, see Fig. 2.

Let us assume that $\mathbf{u} = (u, v)^\top$ are coordinates of a point in a pre-calibrated (will be explained later) image with the origin of the coordinate system in the center of the view field circle $(u_0, v_0)^\top$. The radius r of an image point is transformed to the angle θ of a corresponding 3D vector, see Fig. 2, by a nonlinear function. For Nikon and Sigma fish-eye lenses, respectively, we use the models

$$\theta = \frac{ar}{1+br^2}, \quad \text{resp.} \quad \theta = \frac{1}{b}\arcsin\left(\frac{br}{a}\right), \qquad (1)$$



(a) (b) (c)

Fig. 2. (a) Nikon FC–E8 fish-eye converter. (b) The optical axis of the lens is marked by the dash dotted line and the optical center from which rays emanate is shown as the red dot. The angle between a ray and the optical axis is denoted by θ. (c) The image taken by the lens to the planar sensor π can be represented by intersecting a spherical retina ρ with camera half-rays.

where θ is the angle between a ray and the optical axis, and $r = \sqrt{u^2+v^2}$ is the radius of a point in the image plane w.r.t. $(u_0, v_0)^\top$, and a, b are parameters of the models. The models may have various forms determined by the lens or the mirror construction.

The relationship between the 3D vector $\mathbf{p}$ emanating from the optical center towards a scene point and the image point $\mathbf{u}$ can be expressed up to scale as

$$\mathbf{p} \simeq g(\mathbf{u}) = \begin{pmatrix} \mathbf{u} \\ f(\mathbf{u}, a, b) \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \frac{r}{\tan\theta} \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \frac{r}{\tan\frac{ar}{1+br^2}} \end{pmatrix}, \qquad (2)$$

where $f(\mathbf{u})$ is a rotationally symmetric function w.r.t. the point $(u_0, v_0)^\top$. See [15] for more detail.

2.1. Calibration from the epipolar geometry

The epipolar geometry can be formulated for the omnidirectional central panoramic cameras [17] as well as for omnidirectional cameras with fish-eye converters, which have a single center of projection.

By the calibration of the omnidirectional camera we understand the determination of the affine transformation $\mathtt{A}$ from a view field ellipse to a circle, the symmetry center (u_0, v_0), and camera model parameters a, b. The point (u_0, v_0) is estimated as the center of the elliptical view field. After applying the calibration matrix $\mathtt{A}$, the precalibrated image with square pixel and radial symmetry of non-linear mapping is obtained. Parameters a, b remain unknown and will be estimated from the epipolar geometry.

Function $f(\mathbf{u}, a, b)$ in Eq.(2) can be linearized and the ray direction vector $\mathbf{p}$ can be then written, using Eq.(2), as

$$\begin{aligned} \mathbf{p} &\simeq \left[\begin{pmatrix} \mathbf{u} \\ f(.) - a_0 f_a(.) - b_0 f_b(.) \end{pmatrix} + a\begin{pmatrix} \mathbf{0} \\ f_a(.) \end{pmatrix} + b\begin{pmatrix} \mathbf{0} \\ f_b(.) \end{pmatrix}\right] \\ &\simeq \mathbf{x} + a\mathbf{s} + b\mathbf{t}, \end{aligned}$$

where $\mathbf{x}$, $\mathbf{s}$, and $\mathbf{t}$ are the known vectors computed from image coordinates, a and b are the unknown parameters, and f_a, f_b are partial derivatives of $f(.)$ w.r.t. a and b.

Using the epipolar constraint for vectors $\mathbf{p}_l$ in the left and $\mathbf{p}_r$ in the right image

$$\mathbf{p}_l^\top \mathtt{F} \mathbf{p}_r = 0$$

$$(\mathbf{x}_l + a\mathbf{s}_l + b\mathbf{t}_l)^\top \mathtt{F}(\mathbf{x}_r + a\mathbf{s}_r + b\mathbf{t}_r) = 0$$

leads after arranging of unknown parameters to Quadratic Eigenvalue Problem (QEP) [20]:

$$(\mathtt{D}_1 + a\mathtt{D}_2 + a^2\mathtt{D}_3)\mathbf{h} = 0, \tag{3}$$

which can be solved, e.g., by MATLAB using the function `polyeig`. Parameters a, b, and matrix F can be thus computed simultaneously. A robust technique based on RANSAC with bucketing introduced in [16, 21] can be applied.

Parameters of the camera model described in Eq.(2) and matrix F for an image pair are recovered. Angles between rays and the camera optical axis become known, F is therefore an essential matrix and a *calibrated* camera is obtained. Reader is referred to [15] for more detailed explanation.

3. PROJECTIVE RECONSTRUCTION

Since parameters a, b in Eq. (2) have been obtained from the calibration, a vector $\mathbf{p}$ satisfying multi-view constraints can be constructed for every image point. It allows to perform projective reconstruction of a surrounding scene.

Suppose a set of n 3D points is observed by m central omnidirectional cameras. Not all points are visible in all views. There may be *outliers*, i.e. mismatches in correspondences. The goal is to reject outliers and to recover 3D structure (point locations) and motion (camera locations) from the remaining image measurements.

Let $\mathbf{X}_p$ be the unknown homogeneous coordinate vectors of 3D points, $\mathtt{P}^i$ the unknown 3×4 projection matrices, and $\mathbf{p}_p^i$ the corresponding coordinate vectors of measured image points, where $i = 1, .., m$ labels images and $p = 1, .., n$ labels points. Due to occlusions, some $\mathbf{p}_p^i$ are unknown.

The basic image projection equation says that $\mathbf{p}_p^i$ are projections of $\mathbf{X}_p$ up to unknown scale factors λ_p^i, which will be called (*projective*) *depths*: $\lambda_p^i \mathbf{p}_p^i = \mathtt{P}^i \mathbf{X}_p$. All projections can be gathered into a matrix equation

$$\begin{bmatrix} \lambda_1^1\mathbf{p}_1^1 & \lambda_2^1\mathbf{p}_2^1 & \dots & \lambda_n^1\mathbf{p}_n^1 \\ \times & \lambda_2^2\mathbf{p}_2^2 & \dots & \times \\ \vdots & & \ddots & \vdots \\ \lambda_1^m\mathbf{p}_1^m & \times & \dots & \lambda_n^m\mathbf{p}_n^m \end{bmatrix} = \underbrace{\begin{bmatrix} \mathtt{P}^1 \\ \vdots \\ \mathtt{P}^m \end{bmatrix}}_{\mathtt{P}} \underbrace{[\mathbf{X}_1 \dots \mathbf{X}_n]}_{\mathtt{X}}$$

where marks $\times$ stand for unknown elements which could not be measured due to occlusions, X and P stand for structure and motion, respectively. The $3m \times n$ matrix $[\mathbf{p}_p^i]_{i=1..m,p=1..n}$ will be called the *measurement matrix* (MM). The MM may have missing elements and outliers.

As a result of Sec. 2.1, matches between pairs of images satisfying the epipolar constraint have been obtained. They are not guaranteed to be all the true correspondences. Therefore, an outlier detection technique may still be needed to reject remaining outliers that can only be found using more than two images. Image pairs were read in a sequence. The matches between image pairs were placed into the MM so that the overlapping matches were joined.

The main idea for outlier detection is that minimal configurations of points in triples of images are sufficient to validate inliers reliably. The RANSAC paradigm is used. Trifocal tensors are computed from randomly selected minimal six-tuples of points in triples of images. Ray orientations had to be estimated so that rays projected to the proper image points. Quasiaffine [8] upgrade was computed for the six points in the three images only. Then, the three camera matrices had the right signs w.r.t. the six vectors. Moreover, it turned out in our experiments that most[1] of the points reconstructed using such cameras were properly signed (see more in Sec. 4). If there are sufficiently enough points consistent with the three cameras[2], the consistent points not used to estimate the trifocal tensor are validated as inliers. Sampling is repeated until the set of inliers is sufficient to make the projective reconstruction described below.

Projective reconstruction is done by factorization of the MM with occlusions into structure and motion [12]. This method can handle perspective and also any central views and occlusions jointly. The projective depths of image points are estimated by the method of Sturm & Triggs [19] using the epipolar geometry. Occlusions are solved by an extension of the method by Jacobs [9] for filling the missing data, which can exploit also points with unknown depths.

A set of inliers consistent with the projective reconstruction is found in the following way. In a column of the MM, a pair of image points consistent with the corresponding cameras is searched for. Such a pair forms a *track*. An image point is joined into a track if the point and the track are consistent with cameras as a whole. Columns may also contain more or no track.

It turned out that it is possible to correctly *merge tracks* with close reprojections even if they are *from different columns* of the MM. If there are more candidates, the longest one is chosen because of stability. If they are still more than one, the one from the same column is preferred because there may be some correct pair-wise match by which the tracks got into the same column during building the MM. Also single image points (tracks of length one) are merged this way.

[1] We have not examined how this differed for the right tensors and the ones contaminated by an outlier. Nevertheless, the quasiaffine upgrade existed for the projective reconstruction from all validated points in our experiments.

[2] Image points are consistent if all reprojection errors of a reconstructed world point are below a given threshold.

By track merging, tracks significantly prolong, as shown in Fig. 3c. In fact, by this technique, new matches are generated. This is significant for closed sequences whose structure of the MM is strictly diagonal at the beginning but the first and last views join after track merging, see Fig. 3d.

4. UPGRADE TO METRIC RECONSTRUCTION

The vector $\mathbf{p}$ in Eq.(2) is determined up to a rotation $\mathtt{R}$ and a scalar λ with respect to the directional vector of the corresponding ray in a Cartesian world coordinate system

$$\mathbf{p}'' = \lambda \begin{pmatrix} \mathtt{R}^{-1} & \\ & 1 \end{pmatrix} \mathbf{p} \simeq (\mathtt{R}''| - \mathtt{R}''\mathtt{T}'')\,\mathbf{X}.$$

Estimated vector $\mathbf{p}$ is thus related to scene point $\mathbf{X}$, expressed in a metric coordinate system, by $\mathbf{p} \simeq \mathtt{R}_T\,(\mathtt{I}| - \mathtt{T}'')\,\mathbf{X}$. The matrix $\mathtt{R}_T$ represents a rotation of the estimated coordinates of the camera in the world coordinate frame. It is clear that calibration matrix [8] $\mathtt{K} = diag(1, 1, 1)$.

Orientation of rays in omnidirectional cameras must be taken into account because the rays with the opposite orientations project into different image points. Therefore, to obtain a metric upgrade after the projective reconstruction, orientation of rays has to be properly estimated. This was done using the so-called oriented-projective or quasi-affine reconstruction [8]. Ray orientations had to be estimated also in the outlier detection stage after the tensor estimation so that rays projected to the proper image points.

The final step of the metric upgrade was done by finding a transformation into a Euclidean basis. Some arbitrarily chosen image pair $[ij]$ was used to estimate the essential matrix [8]. However, there can be too few correspondences between the images to estimate the essential matrix. Therefore, the essential matrix was estimated from the reprojections of all reconstructed points into the image pair. New projection matrices, $\mathtt{A}$, $\mathtt{B}$, were directly estimated from the essential matrix up to a two fold ambiguity [8, page 240][3]. The quasi-affine reconstruction was transformed to a Euclidean one by a projective transformation, $\mathtt{H}$, so that the corresponding camera matrices became aligned with $[\mathtt{A}^\top|\mathtt{B}^\top]^\top$ in some four rows. Nevertheless, due to noise in the data and hence also in the quasi-affine reconstruction, the remaining two rows were not aligned. Consequently, the internal parameter matrices, $\mathtt{K}$s, could not be identity. To put $\mathtt{K}$s as close to identity as possible, $\mathtt{H}$ was iteratively improved by the linear projection in which the new projection matrices are obtained as $\mathtt{R}^i[\mathtt{I}|-\mathbf{t}^i]$ where $\mathtt{P}^i = \mathtt{K}^i\mathtt{R}^i[\mathtt{I}|-\mathbf{t}^i]$[4]. It turned out that this process converges in five or ten iterations for the data used in the paper. After that, $\mathtt{K}$s were very close to identity so setting $\mathtt{P}^i = \mathtt{R}^i[\mathtt{I}| - \mathbf{t}^i]$ increased the reprojection error only slightly even without bundle adjustment. There is still remaining the two fold ambiguity of camera matrices. The two solutions differ exactly by the sign of the last column of $\mathtt{P}$ i.e. they are related by projective transformation $diag(1, 1, 1, -1)$. The one leading to a higher number of scene points with the positive fourth coordinate is chosen.

[3]The four fold ambiguity from [8, page 240] was reduced to two fold one by imploying positivity constraint on 3×3 determinants of the first three columns of the projection matrices.

[4]For stability even in (hypothetical) case of huge number of cameras, the iterative process was first applied on the image pair used to estimate the essential matrix and right afterwards on all the images.

5. FINDING CORRESPONDENCES

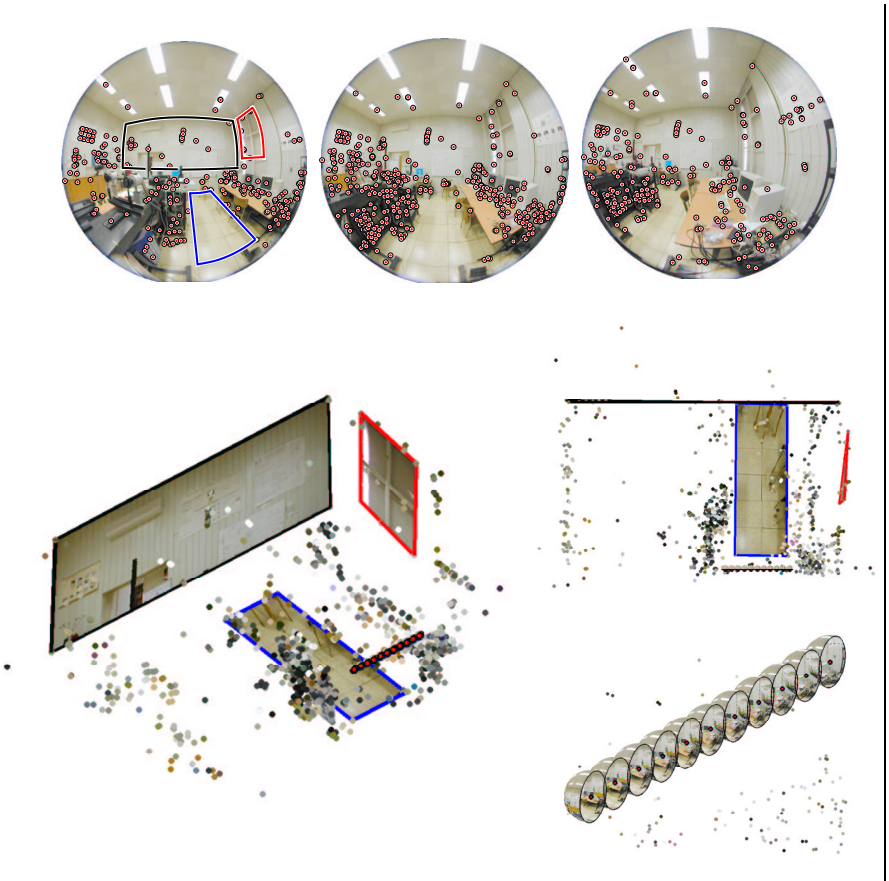
The automatic search for correspondences in omnidirectional images becomes more complicated than in perspective images because the affine invariance of corresponding features, used by most of methods, is preserved only approximately for large motions of the camera.

Nevertheless, in many practical situations, omnidirectional images can be matched by technique [13] that has been developed for conventional wide baseline perspective images. In experiments presented here, only moderate motions of the camera were made, and only smaller regions in images were used to establish tentative correspondences.

6. EXPERIMENTS

In all experiments, the tentative correspondences, i.e. centers of gravity of every region, were obtained by [13]. As a result of applying the calibration method described in Sec. 2.1, the camera model and the essential matrix were obtained and most outliers rejected. Partial correspondences from image pairs were integrated into the measurement matrix as described in Sec. 3. Final 3D reconstructions were improved by a non-linear bundle adjustment tuning all 3D points $\mathbf{X}_i$, camera matrices $\mathtt{P}^i$ and camera model parameters $(u_0, v_0)^\top$, a, and b, and enforcing the same internal parameters for all cameras. To show the quality of the 3D reconstruction, some correspondences, like corners on the walls, have been established manually. The estimated camera matrices were used for reconstructing these points. Finally, textures were mapped on the planes created by the reconstructed 3D points.

In our first experiment, one image pair was selected from the Venice Yard QY dataset, acquired by the Sigma 8mm-f4-EX fish-eye lens with view angle 180° mounted on the Canon EOS–1Ds digital camera with resolution 4064×2704 pxl. The obtained calibrated cameras and validated point matches were used for 3D reconstruction achieved by a linear technique [8]. See Fig. 1 for a result. Notice how precise and complete the 3D reconstruction from only *two* omnidirectional images can be obtained. The RMS of the reprojection error was 0.25 pxl.

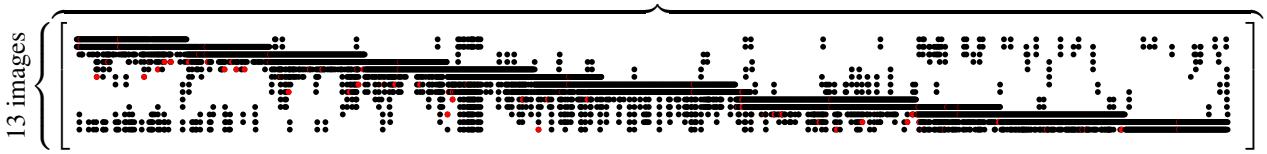


Scene	*Lab side motion*: 11 imgs [1172 × 1172]	*Venice*: 13 imgs [2528 × 2528]
Correspondences / missing data	1119 / 62.62 %	1846 / 77.58 %
Image points consistent with EG	4614 (48.57 % of all regions)	5385 (53.92 % of all regions)
Image points consistent with MVG	**4591** (99.50 % of 4614 image points)	**5301** (98.44 % of 5385 image points)
Outliers	**23** (0.50 % of 4614 image points)	**84** (1.56 % of 5385 image points)
Rec. / partially rec. / not-rec. corr.	1117 / 16 / 2 of 1119	1833 / 54 / 13 of 1846
Mean / maximal reprojection error	**0.23** / 1.00 pxl (inliers @ thr = 1 pxl)	**0.28** / 1.00 pxl (inliers @ thr = 1 pxl)

(a) (b)

Length of reconstructed tracks	11	10	9	8	7	6	5	4	3	2	all
No track merging	0	0	13	6	14	19	44	148	440	1391	2075
After track merging	11	1	4	19	40	52	79	175	357	1112	1850

(c)



(d)

Fig. 3. *Metric reconstruction*. (a) The first, the middle, and the last image from the lab side motion sequence with inliers and three views of reconstructed scene points with camera positions are shown. Notice that all reconstructed points are the points in the view field of the cameras and the cameras lie on a line. (b) Reconstruction of the Venice Yard QY sequence with camera positions. For comparison, only reconstructed points are shown on top whereas only textures from the same view-point are shown in the middle. Summary table is given for both sequences. (c) Lengths of the reconstructed tracks without and with track merging. Notice the amount of longer tracks after merging. (d) Measurement matrix for the Venice sequence. Without track merging, the matrix would be dominantly diagonal, i.e. no correspondences between first and last views would be used.

In our second experiment, all images from the Venice Yard QY dataset were used. The obtained camera model was used in 3D reconstruction by factorization described in Sec. 3, followed by the metric upgrade given in Sec. 4. One pixel was set as the threshold on outlier detection accuracy.

Fig. 3ab shows the metric reconstruction of scene points and camera positions. The table shows number of images and sizes of the cropped images containing the view field circle, number of found correspondences, amount of the missing data, amount of the image points used in the multi-view projective reconstruction (these were already consistent with epipolar geometries), amount of the detected inliers consistent with the multi-view geometry, amount of the reconstructed, partially reconstructed, and not-reconstructed correspondences, and the Euclidean reprojection errors of the reconstruction without outliers. In structure of the MM of the Venice scene, see Fig. 3d, "•" stand for inliers, "•" stand for outliers, and " " stand for the missing data.

In the third experiment, Nikon FC–E8 fish-eye lens was mounted on the Nikon COOLPIX digital camera with 1600×1200 pxl. The camera was moving along a line in constant steps (15cm) capturing the scene at direction perpendicular to motion. Notice in Fig3a, that the estimated trajectory is really straight and distances between all 11 cameras are equal. It can be seen in the top-view of Fig. 3a that all reconstructed points are the points in the field of view of the cameras. See table in Fig. 3a for results.

7. SUMMARY AND CONCLUSIONS

The paper presented a 3D metric reconstruction technique from uncalibrated omnidirectional images. As the main contribution, the paper shows that omnidirectional cameras with highly non-linear projection can be used for 3D reconstruction in the same manner as the standard perspective cameras with narrow view angle.

Using camera with resolution 1200×1200 pxl and lens with field of view 183° is equivalent to using camera with resolution 300×300 pxl and standard lens with field of view 45° in the sense of ratio pxl/angle. Our method shows that very accurate reconstruction of camera positions and accurate reconstruction of scene points can be obtained with relatively small (4 times lower) resolution in comparison to a standard camera. The accuracy of reconstruction of scene points can be improved by using a camera with a higher resolution, e.g. Canon EOS–1Ds as in first two experiments.

8. REFERENCES

[1] R. Bunschoten and B. Kröse. Robust scene reconstruction from an omnidirectional vision system. *IEEE Transaction on Robotics and Automation*, 2002.

[2] P. Chang and M. Hebert. Omni-dir. structure from motion. *IEEE Workshop on Omnidir. Vision*, pp 127–133, 2000.

[3] P. Doubek and T. Svoboda. Reliable 3D reconstruction from a few catadioptric images. *IEEE Workshop on Omnidir. Vision*, pp 71–78, 2002.

[4] J. Fabrizio, J.-P. Tarel, and R. Benosman. Calibration of panoramic catadioptric sensors made easier. *IEEE Workshop on Omnidirectional Vision*, pp 45–52, 2002.

[5] C. Geyer and K. Daniilidis. Structure and motion from uncalibrated catadioptric views. *CVPR*, pp I: 279–286, 2001.

[6] C. Geyer and K. Daniilidis. Paracatadioptric camera calibration. *PAMI*, 24(5):687–695, May 2002.

[7] J. Gluckman and S. K. Nayar. Ego-motion and omnidirectional cameras. *ICCV*, pp 999–1005, 1998.

[8] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[9] D. Jacobs. Linear fitting with missing data: Applications to structure from motion and to characterizing intensity images. *CVPR*, pp 206–212, 1997.

[10] S. B. Kang. Catadioptric self-calibration. *CVPR*, pp I: 201–207, 2000.

[11] S. B. Kang and R. Szeliski. 3-D scene data recovery using omnidirectional multibaseline stereo. *IJCV*, 25(2), 1997.

[12] D. Martinec and T. Pajdla. Structure from many perspective images with occlusions. *ECCV*, pp II: 355–369, 2002.

[13] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. *BMVC*, pp I: 384–393, UK, 2002.

[14] J. Mellor. Geometry and texture from thousands of images. *IJCV*, 51(1), 2003.

[15] B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. *CVPR*, I: 485–490, 2003.

[16] B. Mičušík and T. Pajdla. Omnidirectional camera model and epipolar geometry estimation by RANSAC with bucketing. *SCIA*, pp I: 83–90, 2003.

[17] T. Pajdla, T. Svoboda, and V. Hlaváč. Epipolar geometry of central panoramic cameras. *Panoramic Vision : Sensors, Theory, and Applications*, pp 85–114. Springer Verlag, 2001.

[18] P. Sturm. Mixing catadioptric and perspective cameras. *IEEE Workshop on Omnidirectional Vision*, pp 60–67, 2002.

[19] P. Sturm and B. Triggs. A factorization based algorithm for multi-image projective structure and motion. *ECCV*, pp II: 709–720, 1996.

[20] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43(2):235–286, 2001.

[21] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2):87–119, 1995.