Exploiting Global Connectivity Constraints for Reconstruction of 3D Line Segments from Images

Arjun Jain, Christian Kurz, Thorsten Thormählen, and Hans-Peter Seidel
Max-Planck-Institut Informatik[1], Saarbrücken, Germany
ajain@mpi-inf.mpg.de

Abstract

Given a set of 2D images, we propose a novel approach for the reconstruction of straight 3D line segments that represent the underlying geometry of static 3D objects in the scene. Such an algorithm is especially useful for the automatic 3D reconstruction of man-made environments. The main contribution of our approach is the generation of an improved reconstruction by imposing global topological constraints given by connections between neighbouring lines. Additionally, our approach does not employ explicit line matching between views, thus making it more robust against image noise and partial occlusion. Furthermore, we suggest a technique to merge independent reconstructions, that are generated from different base images, which also helps to remove outliers. The proposed algorithm is evaluated on synthetic and real scenes by comparison with ground truth.

1. Introduction

There is an increasing need for geometric 3D models for movie production, games, and other virtual environments. Unfortunately, manual modelling of 3D objects is tedious and 3D scanners are usually costly and cumbersome, and thus not accessible to everybody. Good alternatives are approaches for automatic 3D reconstruction from image sequences or video.

In the attempt to reconstruct 3D models from images, most approaches apply traditional structure-from-motion algorithms (e.g., [9]) to the set of images to estimate the camara parameters and simultaneously generate a 3D point cloud of the scene. Once such an initial scene reconstruction is available, more detailed 3D models can be estimated (e.g., [4, 14]).

In this paper we consider the problem of reconstructing straight 3D line segments from images. This is of particular interest because of its application to man-made objects, like indoor environments, building exteriors, or urban 3D models. Many algorithms for planar reconstruction perform a 3D line estimation and afterwards sweep the 3D space to find the best fitting plane (e.g., [11]).

Given a set of images and corresponding cameras, the problem of 3D reconstruction of straight lines from images has been studied by several research groups in recent years. In general, it can be said that line matching is a difficult task, because of weaker geometric constraint compared to point matching. In the approach by Baillard et al. [1] lines are reconstructed by first finding line correspondences within the epipolar beam in different views.This is done by evaluating the normalized cross correlation scores over the line patches and then calculating the 3D line formed as the intersection of the two half planes defined by the lines of sight through the end points of the corresponding lines. Moons et al. [8] concentrate on aerial footage and therefore have an easier problem of matching lines only within small regions, which are determined using epipolar geometry and flight path information. They admit on having problems when a longer line must be matched to more than one shorter line segment in different views. Heuel et al. [5] attempt to reconstruct 3D lines by using geometric constraints in a probabilistic framework to model uncertainty due to measurement noise. Since only geometric information is used, their results are not satisfying. Woo et al. [15] propose a hybrid method for line matching, where also the elevation maps generated by a stereo approach are employed to reduce the space of matching candidates. Taylor et al. [12] propose a method where they formulate an objective function which measures the total squared distance in the image plane between the observed edge segments and the projections of the reconstructed lines. The 3D line reconstruction algorithm by Schindler et al. [10] additionally takes vanishing point information into account. Finally, Martinec et al. [6] propose a linear method to reconstruct 3D lines from 2D views by factorisation of a matrix containing line correspondences using SVD.

All the above mentioned algorithms for 3D line recon-

[1]This work was also partially funded by the MPC-VCC (BMBF-FKZ01IMC01) and the DAAD (D/08/13983)

struction are solely local, and do not take into account any global topology of the lines. Further, because of cluttered or noisy images or partial occlusion, corresponding line segments for matching may not be detected in all views.

In contrast, in our approach we express the unknown depth parameter of end points of detected line segments as random variables and use a sweeping based approach to define a discrete probability distribution on the different states (depth values) the end points can take. Then we consider connections with neighbouring lines. We assume that two end points of neighbouring line segments are connected, if they share the same depth. Based on this, we consider the joint distribution of the depth values of end points of all lines in the image, conditioning them with respect to another line if they are connected. This joint distribution can be factorised as a graphical model. Finally, we estimate the depth value for each end point, for which the joint probability of all the line end points is maximised. This can be determined for all end points by loopy belief propagation. Thus, we get the global optimum for the 3D reconstruction of line segments. Because only a subset of lines is visible in each image, we repeat the process for different base images. Afterwards, we merge the partial reconstructions from different base images. The redundancy between the partial reconstructions can be used to perform outlier elimination. In summary, our contributions are:

- We introduce a 3D line estimation algorithm that accounts for the global topology of line connections, and thus provide additional constraints for the 3D reconstruction. As we show in our evaluation, this approach performs better than local approaches, which provide less geometric constraints.
- A sweeping based approach is employed that does not need explicit line correspondences and therefore avoids the problem with corresponding lines that are not detected in other views (e.g., due to image noise or partial occlusion).
- An algorithm for merging partial reconstructions from different base images into a global reconstruction is suggested, which does also help to reduce the number of outliers.

The paper is organised as follows. Section 2 describes the problem mathematically and introduces some notations. Section 3 starts with an overview of our approach for 3D line reconstruction and afterwards gives details about each step. Section 4 presents the results and the paper ends with a conclusion.

2. Problem Statement

Given K images I_k with $k = 1, \dots, K$ of a scene from different viewing locations, we want to estimate the 3D coordinates of the lines detected in these views.

Let the real world scene be made of J 3D line segments, and let the set containing all these lines be

$$\mathcal{L} := \{\mathbf{L}_1, \mathbf{L}_2, \cdots, \mathbf{L}_J\} \tag{1}$$

where $\mathbf{L}_j$ represents a 3D line segment. A line $\mathbf{L}_j$ is described by its 3D start and end points $\mathbf{P}_s^j, \mathbf{P}_e^j \in \mathcal{P}^3$ given in homogeneous coordinates $\mathbf{P} = (X, Y, Z, 1)^\top$. Let $\mathcal{L}^{(k)} \subseteq \mathcal{L}$ be the 3D line segments visible in a camera image I_k and let us denote this set of 2D lines by

$$\mathcal{E}^{(k)} := \{\mathbf{l}_1^{(k)}, \mathbf{l}_2^{(k)}, \cdots, \mathbf{l}_J^{(k)}\} \quad . \tag{2}$$

The 2D line segment $\mathbf{l}_j^{(k)}$ is described by its start and end points $\mathbf{p}_s^{j,(k)}, \mathbf{p}_e^{j,(k)} \in \mathcal{P}^2$ (see Fig. 1).

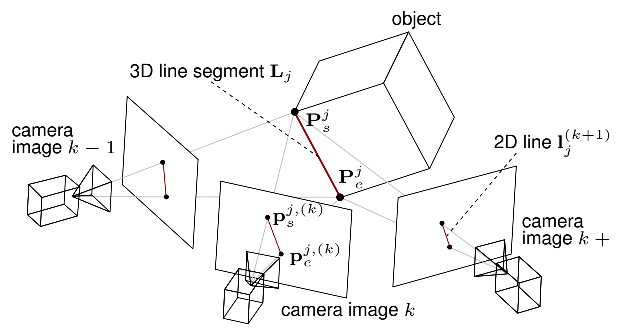

Figure 1. The projection of a 3D line segment $\mathbf{L}_j$ with start point $\mathbf{P}_s^j$ and end point $\mathbf{P}_e^j$ in the camera image I_k gives a 2D line segment $\mathbf{l}_j^{(k)}$ with start point $\mathbf{p}_s^{j,(k)}$ and end point $\mathbf{p}_e^{j,(k)}$.

Given the 3×4 camera matix $\mathtt{A}^{(k)}$ for each camera image k, we have a mapping $\mathcal{L}^{(k)} \mapsto \mathcal{E}^{(k)}$

$$\mathbf{l}_j^{(k)} = \mathtt{A}^{(k)}(\mathbf{L}_j) \quad \forall\, \mathbf{L}_j \in \mathcal{L}^{(k)}, \quad \mathbf{l}_j^{(k)} \in \mathcal{E}^{(k)}\,, \tag{3}$$

which is given by the projection of the start and end points

$$\mathbf{p}_s^{j,(k)} = \mathtt{A}^{(k)}\mathbf{P}_s^j \quad \text{and} \quad \mathbf{p}_e^{j,(k)} = \mathtt{A}^{(k)}\mathbf{P}_e^j \quad . \tag{4}$$

These back-projected 2D points in Eq. (4) and are virtual points and thus are not expected to be visible in all views (e.g., due to occlusions or because the back-projection lies outside the image).

Given the set $\mathcal{E}^{(k)}$ and camera matices $\mathtt{A}^{(k)}$, the objective of this work is to estimate the set of 3D line segments

$$\mathcal{L} = \bigcup_{k=1}^{K} \mathcal{L}^{(k)}. \tag{5}$$

In our approach the 2D lines in the set $\mathcal{E}^{(k)}$ are determined in the images I_k with a straight line detector, which establishes straight lines along image gradients obtained with the Canny edge detector [2]. The camera matrices $\mathtt{A}^{(k)}$ for each image I_k are automatically estimated with a camera tracking software [13].

3. Reconstruction of 3D Line Segments

Let us, for a moment, only consider a single detected 2D line $\mathbf{l}_j$. As illustrated in Fig. 2, the start point $\mathbf{P}_s^j = (X_s^j, Y_s^j, Z_s^j, 1)$ of the corresponding 3D line $\mathbf{L}_j$ can only be located somewhere on the line of sight through the 2D start point $\mathbf{p}_s^j$ of $\mathbf{l}_j$. Similarly, the end point $\mathbf{P}_e^j = (X_e^j, Y_e^j, Z_e^j, 1)$ must lie on the line of sight through the 2D end point $\mathbf{p}_e^j$ of $\mathbf{l}_j$. Therefore, these 3D points have only one degree of freedom, which can be parameterised by their Z-coordinate, Z_s^i and Z_e^i, given in the local camera coordinate system (cp. Fig. 2).

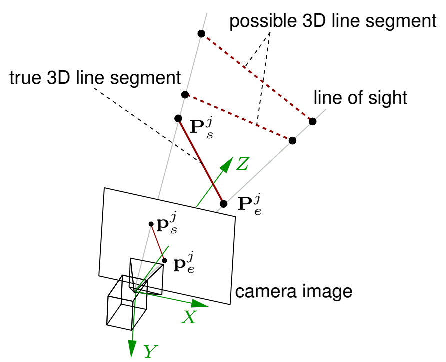

Figure 2. The start and end points $\{\mathbf{P}_s^j, \mathbf{P}_e^j\}$ of the searched 3D line can only be located somewhere on the lines of sight through the start and end points $\{\mathbf{p}_s^j, \mathbf{p}_e^j\}$ of the 2D line.

We define a probability distribution $p(\mathbf{L}_j)$ over the space of possible orientations for a line $\mathbf{L}_j$, described by discrete random variables for Z_s^j, Z_e^j. Thus $p(\mathbf{L}_j) = p(Z_s^j, Z_e^j)$. We describe how we obtain this distribution in subsection 3.1. Thus for a single line $\mathbf{L}_j$,

$$\underset{Z_s^j, Z_e^j}{\arg\max}\; p(Z_s^j, Z_e^j) \tag{6}$$

gives us the optimal 3D position for this line.

Let us now consider multiple lines and their connections. We define another set $\mathcal{J}$, which contains the connections between all 3D line segments, and thus describes the global 3D line segment topology of the scene. If two line segments $\mathbf{L}_p = \{\mathbf{P}_s^p, \mathbf{P}_e^p\}$ and $\mathbf{L}_q = \{\mathbf{P}_s^q, \mathbf{P}_e^q\}$ are connected, the set $\mathcal{J}$ would indicate the equivalence relation for the connected points by

$$\mathcal{J} := \begin{cases} \alpha_{p,q} = 1 & \text{if} \quad \mathbf{P}_a^p = \mathbf{P}_b^q \text{ where } a, b \in \{s, e\} \\ \alpha_{p,q} = 0 & \text{else} \end{cases} \tag{7}$$

As start points can be connected to end points and vice versa, the set $\mathcal{J}$ has a size of $4J(J-1)$ with J the total number of 3D lines in the scene. For image I_k, the set corresponding to the topology of $\mathcal{L}^{(k)}$ is $\mathcal{J}^{(k)}$. We describe how we determine the initial set $\mathcal{J}^{(k)}$ in subsection 3.2.

In subsection 3.3, we look at the joint probability distribution of 3D lines, given the set of connections $\mathcal{J}^{(k)}$. The states at which the random variables Z_s^j, Z_e^j attain a maximum in the joint distribution $p(\mathcal{L}^{(k)} \mid \mathcal{J}^{(k)})$ will give us the globally best position of all 3D line segments $\mathcal{L}^{(k)}$ seen in image k constrained by the topology of the global line connectivity. So, we have to solve

$$\underset{L_1, L_2, \cdots, L_J}{\arg\max}\; p(\mathcal{L}^{(k)} \mid \mathcal{J}^{(k)}) \tag{8}$$

for all lines visible in image k. Since solving Eq. (8) directly is np-hard, we will describe in Section 3.3 how the initial line connectivity given by the set $\mathcal{J}^{(k)}$ helps us to factorise this joint distribution using a graphical model, and then lets us find the max-product for the distribution on the graph, using belief propagation in the loopy setting. Once we have found the best line positions for the initial line connectivity, we alternate between refining the line connectivity $\mathcal{J}^{(k)}$ and estimating the best line positions $\mathcal{L}^{(k)}$ with Eq. (8) until convergence.

We repeat this whole procedure using other images k as base images. For each base image we can only reconstruct the visible subset of lines $\mathcal{L}^{(k)} \subseteq \mathcal{L}$. Thus we need a strategy to group these lines together to generate the conjoined set $\mathcal{L}$ (see Eq. 5). This merging strategy will be described in subsection 3.4, where we also explain, how we use lines from more than one base image to remove outliers.

3.1. Line sweeping

In this subsection, we describe how we define the probability distribution $p(Z_s^j, Z_e^j)$ used in Eq. (6) for all possible 3D positions of line L_j, given by the Z-coordinates Z_s^j, Z_e^j of its start and end point. This probability distribution is estabilished using a sweeping approach [3]. Thereby, the Z-coordinates are given in the coordinate system of the current base camera view k where the corresponding 2D line was detected.

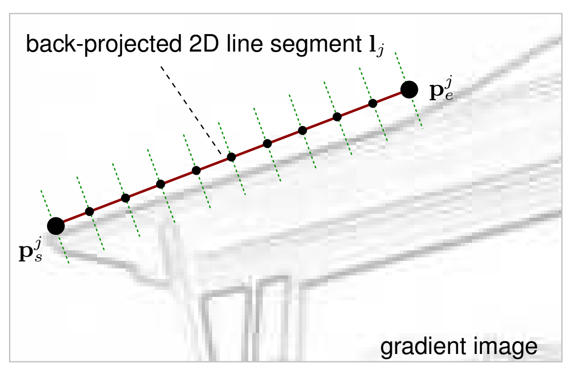

Figure 3. Calculating a scoring function for a back-projected line in the gradient images k'. The small green dots represent measurement points where the gradient is evaluated.

To define $p(\mathbf{L}_j) = p(Z_s^j, Z_e^j)$, we say that the probability of the line $\mathbf{L}_j$ taking a certain position in space

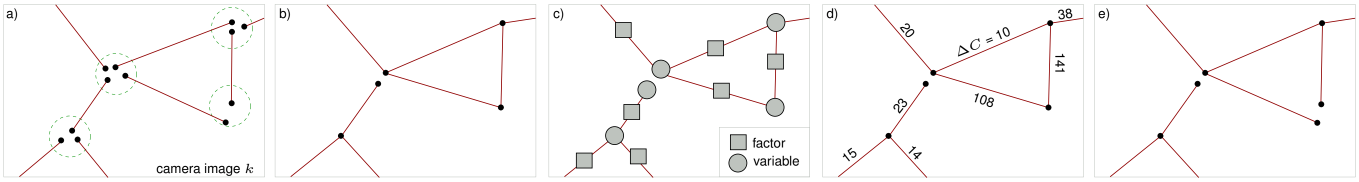

Figure 4. a) Connection candidates are established by searching end points of lines in the proximity of other end points in the camera image. b) The initial line connectivity is found by pairwise evaluation. Lines are connected if the connected cost $C_{p,q}$ is smaller than the unconnected cost $U_{p,q}$. c) The line connectivity can be transformed into a factor graph for loopy belief propagation. d) For each line the additional cost caused by the global connections is calculated. e) The connection that causes the largest cost is erased, if it is larger than a threshold. Afterwards the process is repeated starting from c) until convergence.

is proportional to its cumulative gradient overlap. Therefore, we back-project the 3D line into the other camera views $k' \neq I_k$ using Eq. (4). In order to avoid occlusion, we only back-project into neighbouring camera views k' in the proximity of the base view k with up to 45 degree difference in the cameras' principal axes. For each image k' we also calculate the gradient image $||\nabla I_{k'}||$.

As shown in Fig. 3, the backprojected line is divided between $\mathbf{p}_s(\mathbf{L}_j(Z_s, Z_e)), \mathbf{p}_e(\mathbf{L}_j(Z_s, Z_e))$ into g_p equispaced points. At each such point $\bar{\mathbf{p}}_1, \bar{\mathbf{p}}_2, \cdots, \bar{\mathbf{p}}_{g_p}$, we look at g_l measurement points perpendicular to the line on both sides, and we call this set of points $\mathcal{G}$. With these measurements the probability of the line $\mathbf{L}_j(Z_s, Z_e)$ is given by

$$p(\mathbf{L}_j(Z_s, Z_e)) \propto \sum_{k'} \sum_{i=1}^{g_p} \sum_{\mathbf{x} \in \mathcal{G}} ||\nabla \mathcal{I}_{k'}(\mathbf{x})|| e^{\frac{-\lambda ||\mathbf{x} - \bar{\mathbf{p}}_j||^2}{||g_l||^2}} \quad (9)$$

To get the whole distribution of $\mathbf{L}_j$, we evaluate Eq. (9) for all values of Z_s^i, Z_e^i the line can take. In our experiments we restricted the Z-coordinates between 2 and 9 times the focal length of the camera view k. The distribution is calculated for all lines in $\mathcal{L}^{(k)}$ and for all base images k that we are interested in. Note that without considering connectivity constraints, we could now determine the optimal 3D line segment with Eq. (6).

3.2. Finding initial line connections

To factorise the joint probability in Eq. (8), we need to find an initial set $\mathcal{J}^{(k)}$ of 3D line connections. This initial set can be found by looking at pairwise connections between lines. As illustrated in Fig. 4a for each start or end point of a 2D line we evaluate, if within a certain radius in the base image I_k, there lies any other start or end point of another 2D line. For all these connection candidates $\{\mathbf{P}^p, \mathbf{P}^q\}$ we evaluate the unconnected cost $U_{p,q}$

$$U_{p,q} = \underset{\mathbf{L}^p, \mathbf{L}^q}{\arg\min} \left(-\log\left(p(\mathbf{L}^p)\ p(\mathbf{L}^q)\right)\right) \quad , \quad (10)$$

where we assume that $p(\mathbf{L}^p)$ and $p(\mathbf{L}^q)$ are statistically independent and each $p(\mathbf{L})$ is given by Eq. (9), and the connected cost $C_{p,q}$

$$C_{p,q} = \underset{\bar{\mathbf{L}}^p, \bar{\mathbf{L}}^q}{\arg\min} \left(-\log\left(p(\bar{\mathbf{L}}^p)\ p(\bar{\mathbf{L}}^q)\right)\right) - B \quad , \quad (11)$$

where $p(\bar{\mathbf{L}})$ is given by Eq. (9) with the additional constaint that the Z-coordinates of the connection candidates $\{\mathbf{P}^p, \mathbf{P}^q\}$ are equal. The user-defined scalar value B is a constant bonus term, which we substract from the negative log likelihood to encourage line connections. Without this bonus term the connected cost $C_{p,q}$ would be always larger than the unconnected cost $U_{p,q}$. However, with this bonus term we often have the situation that $C_{p,q} < U_{p,q}$. In this case we connect the lines by setting $\alpha_{p,q} = 1$ in the set $\mathcal{J}^{(k)}$ of possible connections. Otherwise, we leave the candidate unconnected by setting $\alpha_{p,q} = 0$. Fig. 4b shows a possible connected topology for the example given in Fig. 4a.

3.3. Belief propagation and line connectivity update

Once we have the initial connectivity of lines given by $\mathcal{J}^{(k)}$, this can be transformed into a factor graph for loopy belief propagation (see. Fig. 4c). Each factor vertex is only connected to two variable vertices, where the variables are the unknown Z-coordinates of the 3D line points. Thus, the joint probability from Eq. (8) can be written as:

$$p(\mathcal{L}^{(k)} \mid \mathcal{J}^{(k)}) = \prod_j p(\hat{\mathbf{L}}_j(Z_s^j, Z_e^j)) \quad , \quad (12)$$

and thus we have to solve

$$\underset{L_1, L_2, \cdots, L_J}{\arg\max} \prod_j p(\hat{\mathbf{L}}_j(Z_s^j, Z_e^j)) \quad , \quad (13)$$

where $p(\hat{\mathbf{L}})$ is given by Eq. (9). If two line points are connected, they must be represented by the same random variable Z^j. Loopy belief propagation can be employed on the resulting factor graph to estimate the best 3D positions for the 3D lines taking the global connectivity into account (we used the implementation by Mooij [7]). Once a solution is obtained, we calculate the additional cost ΔC_j for each 3D line $\mathbf{L}_j$ caused by the global connections

$$\Delta C_j = \left(-\log\left(p(\hat{\mathbf{L}}_j)\right)\right) - \left(\arg\min_{\mathbf{L}_j}\left(-\log\left(p(\mathbf{L}_j)\right)\right)\right), \quad (14)$$

where $p(\hat{\mathbf{L}})$ is evaluated at the optimal global Z-coordinates obtained by the belief propagation algorithm and the subtrahend is the cost for an unconnected 3D line (see. Eq. (6)). We sort the resulting values ΔC_j and check if the highest value is above a user-defined threshold H. If this is the case, the connection that causes the highest cost is erased and a new belief propagation is performed with an updated factor graph. This procedure is repeated until all ΔC_j are smaller than the threshold H. This results in a 3D reconstruction where the positions of 3D lines $\mathcal{L}^{(k)}$ and the line connectivity $\mathcal{J}^{(k)}$ is conjointly optimized.

3.4. Outlier elimination by line grouping

We repeat the above process for different base images I_k and thus obtain sets of 3D lines $\mathcal{L}^{(k)}$ for each base image. Then, we group these 3D lines using spatial proximity.

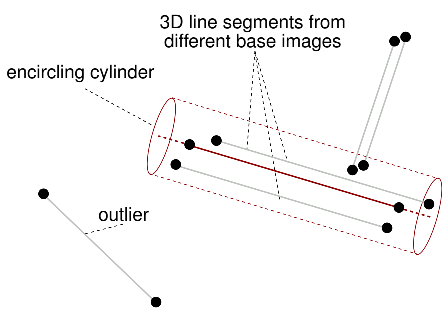

Figure 5. 3D lines segments from different base images I_k are grouped with other 3D lines if these are located within an encircling cylinder. A representative line is estimated for each such group. Lines that do not form a group with at least one other line are considered as outliers.

As show in Fig. 5 we define an cylinder around each 3D line, and check if both end points of a 3D line from another base image fall within this cylinder. If this is the case, these lines form a group. Thereby the cylinder is extended at both sides by 10 percent along the 3D line. Once the groups are established, each group is replaced by a single line. This is done by generating a new line along the principal component direction, which is the eigenvector corresponding to the largest eigenvalue of the scatter matrix of all line points of a group. The new extent of the line segment is defined by projecting all group points onto the principal component direction. The maximal and minimal values in principal component direction define the new start and end point of the segment. All 3D lines that do not form a group with at least one other line are considered as outliers and are removed from the final reconstruction.

The grouping may disturb the established connections between 3D lines. Therefore, we need to refine this solution by solving a linear cost function, which imposes that the connections between lines are reenforced. For all remaining connections in the set $\mathcal{J}$ where $\alpha_{p,q} = 1$ we update the current points $\mathbf{P}_p^j, \mathbf{P}_q^j$ to the refined points $\hat{\mathbf{P}}_p^j, \hat{\mathbf{P}}_q^j$ by solving

$$\arg\min_{\hat{\mathbf{P}}_p^j, \hat{\mathbf{P}}_q^j} = \sum_{\mathcal{J}} \alpha_{p,q} ||\hat{\mathbf{P}}_p^j - \mathbf{P}_p^j||^2 + ||\hat{\mathbf{P}}_q^j - \mathbf{P}_q^j||^2 \,. \quad (15)$$

4. Results

In this section our approach is evaluated on 3 datasets: a synthetic data set, a real data set captured in our lab together with a laser scan for ground truth evaluation, and another example taken outside with a consumer camera.

The synthetic image sequence is generated from a 3D CAD model of a timber-frame house. From this scene we rendered 240 images with a resolution of 1280×960 pixels. Examples of the input images are shown in Fig. 6. After we generated a 3D reconstruction of the 3D lines with the presented approach, we compared the result with the known 3D CAD model. In Tab. 1 we compare the root mean square error (RMSE) of our 3D reconstruction with a 3D reconstruction that would be obtained without considering global line connectivity. The RMSE is shown for different cut-off thresholds. If the measured error for a particular point on a 3D line is higher than this cut-off threshold, the error is not included in the RMSE measurement. It can be seen, that the approach with global connectivity constraints outperforms the local approach for all cut-off thresholds. These results show, that the additional geometric constraints introduced by the connections of 3D lines improves the accuracy and helps to reduce the number of outliers in the final reconstruction. In Fig. 6 we also show the color-coded reconstruction error of our approach for visual inspections. It can be seen that a large majority of the lines have a very small reconstruction error.

Note, that we employed the outlier elimination from subsection 3.4 for both methods before we compared them in Tab. 1. Otherwise the improvement provided by our approach would be even more significant. Fig. 7 shows a comparison of the results before and after line grouping and outlier elimination.

The next example was recorded with a HDV video camera in our lab. The input images have a resolution 1440×1080 pixels and show yellow and red building blocks on a planar black and white checker board. Each square of the checker board has an edge length of 50mm. A total of 84 images was recorded. At the same time the scene was reconstructed with a commercial laser scanner, which generated a 3D model for ground truth evaluation. The camera matrices were estimated with a camera tracking software

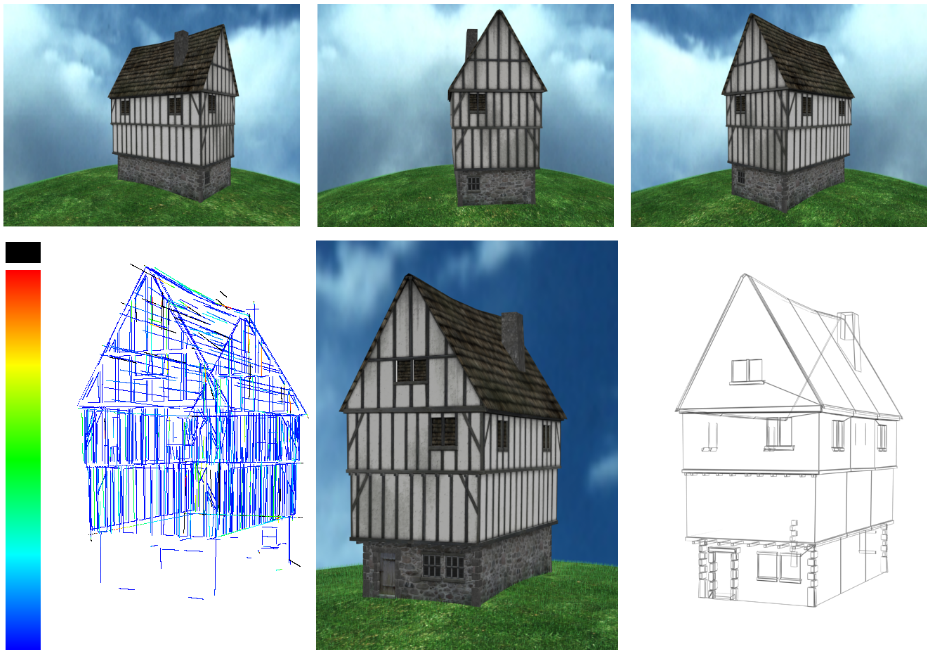

Figure 6. **Timber-frame house** (synthetic scene): Top row: example images from the input sequence. Bottom left: color coded reconstruction error of our approach (blue indicates a low error, red a high error, and black an error larger than 0.5m). Bottom right: ground truth model rendered textured and in wireframe.

RMSE without [m]	with [m]	Threshold [m]	Improvement [%]
0.3361	0.1970	none	41.1
0.2019	0.1810	3.5	10.3
0.1918	0.1736	2.5	9.4
0.1470	0.1262	1.5	14.1
0.0964	0.0807	0.5	16.2

Table 1. **Timber-frame house**: RMSE of the 3D line reconstruction without and with global connectivity constraints. The RMSE is shown for different cut-off thresholds.

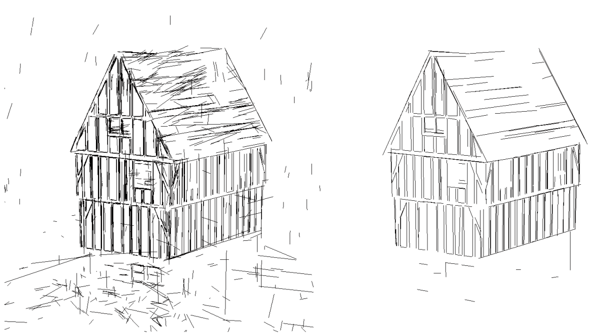

Figure 7. **Timber-frame house**: 3D line reconstruction before (left) and after (right) line grouping and outlier elimination.

and the laser scan data was fitted to the camera images using the features points provided by the checker board [1]. In Fig. 8 as well as in Tab. 2 a comparison between our approach with and without global connectivity constraints is shown. Again our method with global connectivity shows significant improvements of the RMSE. Though for a cut-off threshold of 5mm the improvement is only 3.3%. However, this particular comparison is maybe already affected by the measurement error of the laser scanner.

[1]Ground truth data, and our results can be downloaded from http://www.mpi-inf.mpg.de/resources/LineReconstruction/

RMSE without [mm]	with [mm]	Threshold [mm]	Improvement [%]
10.66	9.65	none	9.4
8.87	6.29	75.0	29.1
7.41	4.30	50.0	42.0
5.48	3.83	25.0	30.1
2.36	2.28	5.0	3.3

Table 2. **Building blocks**: RMSE of the 3D line reconstruction without and with global connectivity constraints. The RMSE is shown for different cut-off thresholds.

The third example is a 3D line reconstruction from a set of 20 photos taken with a consumer SLR camera in a low-light situation, which resulted in images with a high pixel noise. Fig. 9 shows the 3D line reconstruction of this scene. Many details are reconstructed including the tiles on the wall of the rightmost house.

5. Limitations and Conclusion

We have presented a novel approach for 3D line reconstruction from image sequences. In contrast to existing ap-

Figure 9. **Street** (real scene): Top: example images from the input sequence. Bottom: 3D line reconstruction with global connectivity constraints.

proaches, we automatically establish connections between neighboring 3D lines. These additional geometric constraints improves the reconstruction significantly as shown by our evaluation with ground truth data. The root mean squared reconstruction error is reduced by approx. 20 percent.

Our sweeping-based approach does not use explicit 2D line matching and, thus, can often reconstruct a line in situations where matching based approaches fail because the corresponding line is not detected in the neighbouring views. This can often happen due to noise or partial occlusions. However, a disadvantage of the sweeping approach is that evaluating all possible Z-coordinates is computationally more expensive than explicit 2D line matching. For some scenes with many base images we had to run our algorithm over night (8 to 10 hours) to obtain our results. Another limitation of the sweeping approach is that 3D lines that are not in the sweeping range (of 2 to 9 times the focal length in our case) can not be correctly reconstructed.

We have also presented an automatic approach for merging partial reconstructions from different base image, which tries to merge lines using spatial proximity. If a line does not form a group with at least one other line from a different reconstruction, it is rejected as outlier. As shown in our

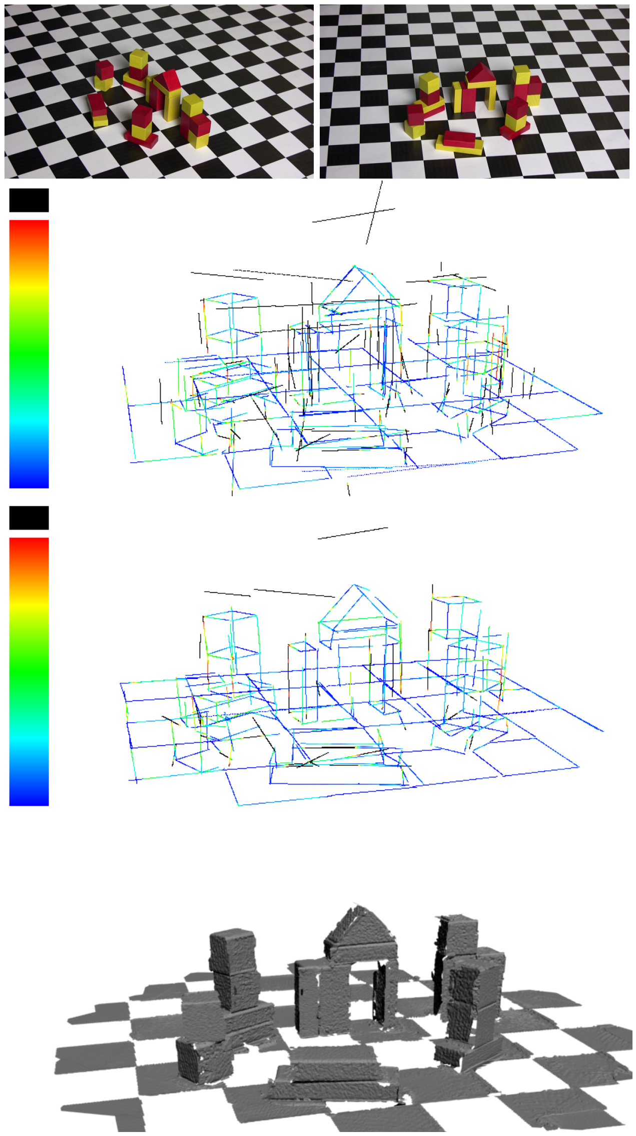

Figure 8. **Building blocks** (lab scene): Top to bottom: example images from the input sequence; color coded reconstruction error without global connectivity constraints (blue indicates a low error, red a high error, and black an error larger than 5mm); color coded reconstruction error with global connectivity constraints; laser scan used for ground truth evaluation.

results, only very few outliers remain. A limitation of this approach is that we sometimes also merge lines that are in fact no outliers but different 3D lines in close proximity.

To further increase the reconstruction quality, in future work, we want to consider additional geometric constraints, like the perpendicularity of 3D lines often present in man-made environments. Furthermore, the achieved 3D line reconstruction can be a perfect starting point for algorithms that try to extract a more complete surface reconstruction from images.

References

[1] C. Baillard, C. Schmid, A. Zisserman, A. Fitzgibbon, and O. O. England. Automatic line matching and 3d reconstruction of buildings from multiple views. In *In ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, IAPRS vol.32, Part 3-2W5*, pages 69–80, 1999. 1

[2] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, November 1986. 2

[3] R. T. Collins. A space-sweep approach to true multi-image matching. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:358, 1996. 3

[4] S. Gibson, R. J. Hubbold, J. Cook, and T. Howard. Interactive reconstruction of virtual environments from video sequences. *Computers & Graphics*, 27(2):293–301, 2003. 1

[5] S. Heuel and W. Förstner. Matching, reconstructing and grouping 3d lines from multiple views using uncertain projective geometry. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:517, 2001. 1

[6] D. Martinec and T. Pajdla. Line reconstruction from many perspective images by factorization. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 1:497, 2003. 1

[7] J. M. Mooij. libDAI: A free/open source C++ library for Discrete Approximate Inference. http://www.libdai.org/, 2009. 4

[8] T. Moons, D. Frère, J. Vandekerckhove, and L. J. V. Gool. Automatic modeling and 3d reconstruction of urban house roofs from high resolution aerial imagery. In *ECCV '98: Proceedings of the 5th European Conference on Computer Vision-Volume I*, pages 410–425, London, UK, 1998. Springer-Verlag. 1

[9] M. Pollefeys, L. V. Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch. Visual modeling with a hand-held camera. *IJCV*, 59(3):207–232, 2004. 1

[10] G. Schindler, P. Krishnamurthy, and F. Dellaert. Line-based structure from motion for urban environments. *3D Data Processing Visualization and Transmission, International Symposium on*, 0:846–853, 2006. 1

[11] S. Scholze, T. Moons, and L. V. Gool. A probabilistic approach to roof extraction and reconstruction. In *Proceedings of the ISPRS Commision III Symposium*, volume 34, pages 231–236, September 2002. 1

[12] C. J. Taylor and D. J. Kriegman. Structure and motion from line segments in multiple images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(11):1021–1032, 1995. 1

[13] T. Thormählen and H. Broszio. Voodoo camera tracker. Freely avaiable for download at http://www.digilab.uni-hannover.de. 2

[14] G. Vogiatzis, P. H. S. Torr, and R. Cipolla. Multi-view stereo via volumetric graph-cuts. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2*, pages 391–398, Washington, DC, USA, 2005. IEEE Computer Society. 1

[15] D.-M. Woo, D.-C. Park, and S.-S. Han. Extraction of 3d line segment using disparity map. *Digital Image Processing, International Conference on*, 0:127–131, 2009. 1