Vanishing Points Detection and Line Grouping for Complex Building Façade Identification

Wenting Duan
The University of Sheffield
Department of Electronic & Electrical Engineering
Mappin Street
Sheffield, S1 3JD
elq06wd@sheffield.ac.uk

Nigel M Allinson
The University of Sheffield
Department of Electronic & Electrical Engineering
Mappin Street
Sheffield, S1 3JD
n.allinson@sheffield.ac.uk

ABSTRACT

We describe a new method for automatically detecting vanishing points and group lines associated with building facades from a single uncalibrated image. Accurate vanishing points detection for buildings is of importance for building façade rectification and 3D scene reconstruction. The challenges arise from confusing scene clutter, occlusions, unusual building shapes and non-Manhattan street layout. Buildings usually possess many straight-line features, e.g., window and door openings, as well as their overall outline. We exploited these features and propose a robust line grouping technique. The method was evaluated on images from the Zubud-Zurich building database. The experiment shows the proposed method works well with varying building structures and image conditions, as well as filtering out "non-building" vanishing points (e.g. vanishing directions detected from road boundaries).

Keywords

Vanishing points, façade identification, line grouping, line segments detection.

1. INTRODUCTION

The rectification of building facades to their fronto-parallel view is an essential step in photogrammetry, 3D scene reconstruction and potential building recognition applications. Robust and accurate rectification requires the precise location of the extrapolated vanishing points obtained directly from, in our case, a single uncalibrated image. Buildings most commonly possess many straight-line features, e.g., window and door openings, as well as their overall outline. Exploiting these features and outlines is the basis of all published approaches. The challenges arise from confusing scene clutter, occlusions, unusual building shapes and non-Manhattan street layout. This paper presents an enhanced method for vanishing point detection that addresses these challenges.

General approaches for estimation of vanishing points are the Gaussian sphere, Hough transform, Expectation Maximisation (EM) and RANSAC-based methods. The Gaussian sphere method was first suggested by Barnard in 1983 [Bar83a]. A line segment from the image intersecting the Gaussian sphere forms a great circle, the vanishing points are found at the intersecting points of these great circles on the Gaussian sphere. The sphere surface is divided into accumulation cells, of which the ones with votes above a specified threshold are selected. This method transforms lines from unbounded to bounded space, so eliminating the singularity caused by a vanishing point located at infinity. Several processes part of the approach have since been improved; the tessellation of the Gaussian sphere [Qua89a], accumulator space [Bri91a] and projection of the intersection points for all line pairs instead of just line segments [Mag84a].

Another technique that maps line information to a bounded space is the Hough Transform [Lut94a]. Vanishing points can be obtained by applying the Hough transformation three times (known as the Cascaded Hough Transform), and was demonstrated by Tuytelaars et al. [Tuy98a]. Rother [Rot02a] pointed out that a drawback of mapping lines onto a bounded space can miss important geometric information such as the original distances between lines and points.

Kosecka and Zhang introduced a method using an EM algorithm to calculate the location of vanishing points for an uncalibrated camera [Kos02a]. This approach simultaneously groups lines parallel to real world axes and estimate dominant vanishing directions by weighting each individual line segment with a probability corresponding to every candidate vanishing direction. The problems of viewpoint change, occlusion, scene clutter and non-orthogonal building facades were addressed by our previous paper [Dua09a]. A refined approach was developed, which was based on an EM algorithm which gradually eliminated low probabilities and assigned lines to their associated group. Eildenauer and Vincze ran RANSAC several times on detected lines to obtain potential vanishing points. The EM algorithm was then used to refine the results [Wil07a]. Their method has been shown to work on images with structure not strictly meeting the Manhattan assumption [Cou03a].

Our work is most closely related to David's technique where RANSAC was used to detect vanishing points and to group lines, and line intersection was exploited to identify vanishing points corresponding to each building façade [Dav08a]. The images we are dealing with are uncalibrated, not conforming to the assumption of a Manhattan world. The aim is to detect all vanishing points and associated building façades appearing in a single image. Within an image of urban areas, detected lines can also arise from road boundaries and markings, vehicles, aerial cables, etc. Therefore, the EM algorithm, which requires an initialised line grouping, and RANSAC's randomly selected line segments, could erroneously pick up many non-building lines. Developing David's work, we suggest that selecting potential vanishing points and initial groups of lines based on a line orientation curve could reduce this type of error. A RANSAC based split and merge scheme was used to refine the vanishing point location. Finally, line intersection was used to find building façades and filter non-building vanishing points. In addition we show how geometric constraints can be utilised to reduce the exhaustive intersection process. In the Discussion, the line detection scheme used in our work was compared to the method used by David [Dav08a]. Parameters involved in our method were analysed to see how the overall grouping performance was affected by the value of the parameters. Example images demonstrate the robustness of our technique compared with previous work.

2. METHOD

2.1 Line Segments Detection

Due to the presence of geometric regularity in images containing man-made structures, lines can be easily derived from the image edges. The majority associated with building façades are also mutually parallel or orthogonal. As parallel lines of buildings intersect at vanishing points in the image under perspective transformation, line segments detected from building image are a reliable source to locate vanishing points. Firstly, we up-sampled the image if its size was less than 1M pixels in order to improve the quality of edges [Wil07a]. A Canny edge detector followed by non-maximum suppression, was used to find local edges. A binary image of the strongest edges is obtained after applying hysteresis thresholding. The line fitting stage is similar to Kosecka's method [Kos02a] originated in [Kah90a]. The normal edge orientation span [0 π] was equally divided into k bins. Take k = 18, (the set value 18 gives a bin size of 10°. This is chosen to give a reasonable bin size, enabling connected components to form within the set orientation range), hence the edge location are assigned with the same label if their orientation is within certain bin range (e.g. 10°<edge_orientation<20°). After getting 18 sets of edge pixels, every three sets with consecutive labels are grouped together for connected components analysis. To reduce repeated detection of lines and avoid boundary cut-off, the end label of the last grouping is used as the beginning label of the next. For example, (1,2,3), (3,4,5) and (5,6,7), etc. Edge lists of detected connected components with supporting points less than 15 pixels are removed. For each connected component, their supporting edge pixels are denoted as (x_i, y_i). The variance $\tilde{x} = x_i - \bar{x}$ and $\tilde{y} = y_i - \bar{y}$, where $\bar{x} = \frac{1}{n}\sum_i x_i$ and $\bar{y} = \frac{1}{n}\sum_i y_i$, was used to form the matrix D. The eigenvector $\boldsymbol{e_1}$ computed from matrix D, which corresponds to the largest eigenvalue, was used to derive the angle θ for the line we are trying to fit.

$$D = \begin{bmatrix} \sum_i \tilde{x}_i^2 & \sum_i \tilde{x}_i\tilde{y}_i \\ \sum_i \tilde{x}_i\tilde{y}_i & \sum_i \tilde{y}_i^2 \end{bmatrix} \quad (1)$$

$$\theta = \tan^{-1}(\boldsymbol{e_1}(2), \boldsymbol{e_1}(1)) \quad (2)$$

The line length is defined as the major axis length from its supporting edge pixels. The mid-point is also $(\bar{x}, \bar{y})$. Together with θ, the endpoints can be easily derived. In Figure 1, two example images and their detected lines are shown.

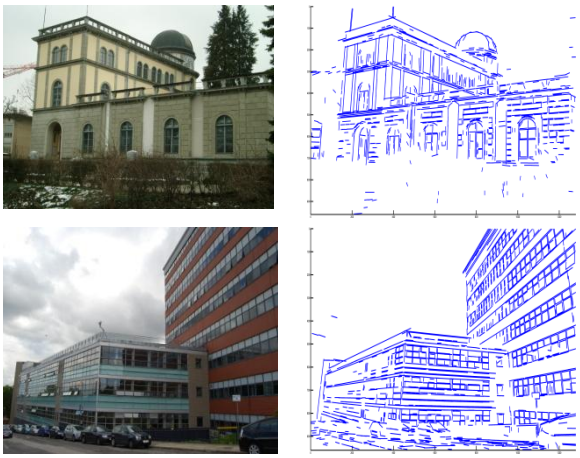


Figure 1. Two examples of line segments detected from building images.

2.2 Initial Vanishing Points Estimation and Line Grouping

According to the RANSAC based approach described by David [Dav08a] and Wildenauer and Vincze [Wil07a], two line segments were randomly selected to compute an estimation of vanishing point. Lines belonging to this direction were then grouped. The disadvantage of this method is that it will iterate exhaustively until: 1).a predefined *K* vanishing points candidate are found; or 2). Insufficient inliers are assigned. Instead, we suggest using a simple initial line grouping according to the line direction in order to create a guided selection of intersecting lines. Prior to the initial estimation scheme, each line segments are weighted by their length and colinearity. The weight range is [1, 2, 3], with every segment being equally weighted as 1 at the start. If a length is larger than 12.5% of the image size or it has more than two collinear lines, the value of 1 is added to the weighting respectively. The orientation histogram computed from the detected line segments with their associated weightings, i.e. segments with weightings of 2 are counted twice. This significantly boosts the peaks after fitting a curve to the Gaussian smoothed histogram. The curvature is computed as [Kos02a]:

$$\boldsymbol{C}(k) = h_\theta(k) - \frac{1}{s}\sum_{i=k-\frac{s}{2}}^{k+\frac{s}{2}+1} h_\theta(i) \qquad (3)$$

where k = 60 and s = 9. The starting side of the histogram is also padded with the last 5 bins at the ending side and so did the ending side before fitting the curve so that the peaks at the boundary can emerge (Figure 2(b)). The two minima located nearest to each unique peak are located and chosen as the orientation range for the initial line grouping. Hence, lines with directions falling into the defined range are grouped together. Those not assigned are grouped as the outliers as shown in black in Figure 2(c).

2.3 Split and Merge Scheme

As shown in Figure 2, this type of building structure can result in a fine initial grouping and its vanishing points can be easily detected using existing techniques. However, there are many buildings not so regular in appearance, such as the second example in Figure 1. This section shows how a split and merge scheme can appropriately group line segments.

For a homogeneous representation, $\boldsymbol{x_1} = (x_1, y_1, 1)$ and $\boldsymbol{x_2} = (x_2, y_2, 1)$. The lines are expressed in their normalised representation of the coincident infinite line $\boldsymbol{l_n}$ following the formula (4) for efficient computing of vanishing points:

$$\boldsymbol{l} = \boldsymbol{x_1} \times \boldsymbol{x_2} \qquad (3)$$

$$\boldsymbol{l_n} = \boldsymbol{l}/\sqrt{l_1^2 + l_2^2} \qquad (4)$$

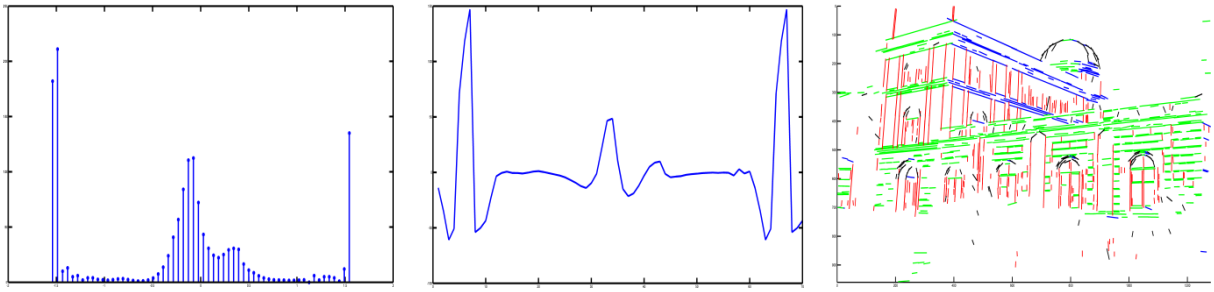
where $\boldsymbol{l} = (l_1, l_2, l_3)$. At the split stage, we assess each initially grouped line set. The line segments are sorted from highest to lowest weighting. As the lines with the higher weightings have a higher probability of giving a better estimation of vanishing point locations, two segments at the beginning of the line queue are intersected to produce the prototype vanishing point $\boldsymbol{v} = \boldsymbol{l_1} \times \boldsymbol{l_2}$. The angle θ_v and distance $\boldsymbol{d_v}$ to the prototype vanishing point from the rest of lines were calculated according to the formulas below:

$$\boldsymbol{m} = (x_m, y_m, 1) \qquad (5)$$

$$\boldsymbol{d_v} = \frac{|\boldsymbol{l_{nv}}.\boldsymbol{x_1}|}{\sqrt{\boldsymbol{l_{nv}}(1)^2 + \boldsymbol{l_{nv}}(2)^2}} \qquad (6)$$

$$\theta_v = \sin^{-1}(\boldsymbol{d}/_{0.5 * L}) \qquad (7)$$

The line $\boldsymbol{l_v}$ crosses the prototype vanishing point $\boldsymbol{v}$ and one of the assessing lines' midpoint $\boldsymbol{m}$. Its normalised version $\boldsymbol{l_{nv}}$ is calculated by (4) and was used to obtain $\boldsymbol{d_v}$. In (6), the distance between one of the prototype line's endpoints $\boldsymbol{x_1}$ and $\boldsymbol{l_{nv}}$ is obtained. Equation (7) measures the angle diversion of the prototype line from prototype vanishing point. The lines with θ_v smaller than σ_θ, and $\boldsymbol{d_v}$ smaller than σ_d are taken as inliers. If the number of inliers is larger than ψ_{min}, then the set of inliers is split from the testing line group forming a new line group. If one prototype line intersecting every other line within the group does not manage to form an inliers set, it is moved to the outlier group. The iteration will stop when there are less than ψ_{min} lines left in the prototype group, these leftover lines are also assigned to the outlier group.



(a) (b) (c)

Figure 2. Initial processing of Figure 1 (left): (a) line direction histogram; (b) curvature computation; (c) line grouping with outlier group marked black.

The initial outlier group is assessed last, because the assigned outliers from the previous groups are included. Another constraint added to the processing of this group is that the angle between each intersecting line pair needed to be smaller than 45° to continue. 45° is a practical limit angle for the parallel lines intersection to occur under perspective transformation. The vanishing points are recalculated for each line group formed. The parameters σ_θ and σ_d are 1° and 1, respectively in our experiment. This seems a rather strict threshold but the advantage is that accurate candidate parallel sets can emerge. ψ_{min} was set as 50 empirically for a reliable parallel line set to form. The point $\boldsymbol{v}$ lies on the line $\boldsymbol{l_n}$ if $\boldsymbol{v^T l_n} = 0$. This leads us to the linear least square estimation problem:

$$min_v \sum_{i=1}^{n} \left(\boldsymbol{l_i^T v}\right)^2 \qquad (8)$$

where n is the number of lines. We can rewrite this as

$$min_v \|\boldsymbol{Av}\|^2 \qquad (9)$$

The estimation of $\boldsymbol{v}$ is the eigenvector associated with the smallest eigenvalue of $\boldsymbol{A^T A}$. After the vanishing points are relocated, RANSAC was applied again to the outlier group for each vanishing direction. In The split stage has resulted in some over-grouped line sets, i.e. it is sensitive to line deviation from prototype vanishing points. The benefit is that some proper candidates of vanishing points start to emerge.

At the merge stage, a pair of line sets with similar vanishing directions is combined to calculate a new vanishing point based on (9). The inliers and outliers of the combined line group belonging to new vanishing point are found. If more than, say, 70% of lines from each merging set are inliers, the inliers are allowed to merge setting a new vanishing direction. The outliers are also moved to their associated set. The whole candidate line sets are also refreshed to eliminate the tested pair and to add the new set. Otherwise, nothing is changed and the next pair of sets is picked. The outlier group is again assessed at the end to find any inliers for the new estimated vanishing points. The resultant groups are shown in Figure 3.

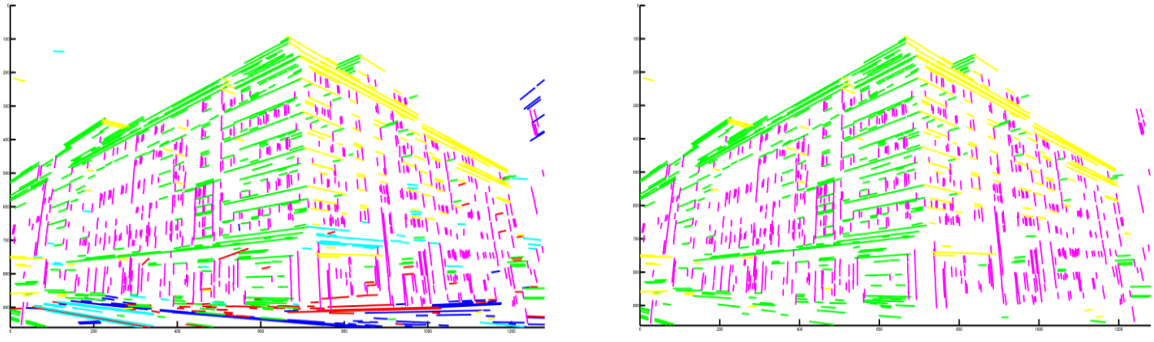
2.4 Filtering of Vanishing Points Associated with Building

Since the aim is to find vanishing points associated with building facades, and that there can be confounding line sets belonging to other regular objects such as roads or vehicles, we used a method similar to David [Dav08a] but less complicated to remove these unwanted groups. The example image in Figure 3 is used to demonstrate. In Figure 3(a) it is shown that at least two sets of lines are not associated with the building. The vanishing points corresponding to the vertical direction of the image are first selected. As the image is taken at ground level, vanishing points with negative value of y-axis are defined as vertical ones. Then, the intersection points between each of the vertical line set and every other line set are computed. The method is illustrated in Figure 4. Line segments in each set are indexed and extended by e_d pixels from their endpoints. A blank image of the same size as the original image is created for each participating line set. The lines in each group are indexed, and on their associated blank image, points consisting the extended line are labelled at their pixel location by line index *i*. The locations of labelled images with two or more recorded index are searched. This way, lines intersecting the vertical lines are found and the others filtered out. The line sets with lines intersecting infrequently are non-building sets.

3. EXPERIMENTAL RESULTS

The algorithm was tested on images from the Zubud-Zurich building database [Sha03a]. The images are in PNG format with size 640*480. The results were also compared with the previous paper [Dua09a] using EM algorithm for vanishing points detection and line grouping. First of all, the line

detection algorithm [Kah90a] was compared to Kovesi's method implemented in Matlab code [Kov07a] (used in David's [Dav08a] and Wildenauer's [Wil07a]. The parameters involved in the two methods were set to be the same, e.g. σ=1 in Canny detector, connected edge with length greater than 15 was accepted for line fitting. The result is shown in Fig 5 and Table 1. The advantages of the suggested method are: 1. Greater accuracy in line detection - this is clear on the major structural lines of the left frontal facade of the example building where the lines detected by Kovesi's method weren't parallel. 2. Quicker and computationally less complicated - the two methods were tested on the first eight images of the Zubud-Zurich building database [Sha03a]. From Table 1, we can see that the time required for Kovesi's method varies with the amount of line segments detected, whereas the current method is more stable and a lot faster. 3. A higher proportion of line segments associated with man-made structures were found; line segments associated with scene clutter such as trees were eliminated by the line detection procedure. The parameters ψ_{min}, σ_θ and σ_d were analysed to test how the overall grouping performance varies with different values. The overall grouping performance was measured by the average deviation distance from each final grouped line to their associated vanishing point. When incorrect groupings occurred, i.e. more or less groupings than actual line groupings appearing in the image were detected, a penalty weight was assigned to its measured distance. The results were obtained by running the proposed method on 30 images randomly selected from the Zubud-Zurich building database [Sha03a]. Figure 6 shows how the deviation distance varies with the setting of ψ_{min}. The dramatic changes around 30 and 70 are caused by the fact that in most cases: $\psi_{min} < 30$ leads to over-grouping (actual groups being split further) and $\psi_{min} > 70$ causes under-grouping (some groups not detected because minimum threshold set too high). Parameters σ_θ and σ_d were observed together, since individually they had a trivial effect compared to when they were changed in concert. As shown in Figure 6(b), as both σ_θ and σ_d increase, the average deviation distance from the final grouped lines to their associated VP also increases. In Figures 7 – 10, the results tested on different building image conditions are shown. In general, more line segments are detected with the new line detection method. The line sets are grouped more accurately, i.e. there are very few lines that are not associated with the detected vanishing direction. Figure 7 shows the results computed from an image taken with a sharp viewing angle to the building. The new method has shown to perform well while the EM based method fails to detect enough lines for its grouping. On buildings with multiple facades as shown in Figure 8, existing methods under the Manhattan assumption will fail to group correctly. However, the proposed method cannot handle buildings with a round structure. The lines belonging to the round part in Figure 8(a) are simply assigned to their closest vanishing direction. Figure 9 shows the results on building images with occlusion. The proposed method is actually more sensitive to occlusion where it tends to put the lines from the object causing occlusion as another group but still manages to find the groups associated with the building. A blurred image is shown in Figure 10. The method is weaker in these conditions than the EM based method. However, with fast improving camera and storage space, this may not be a problem in the future.



(a) (b)

Figure 3. Example image for demonstration of filtering out non-building segments: (a) computed group after the merge stage; (b) final line sets obtained after filtering stage.

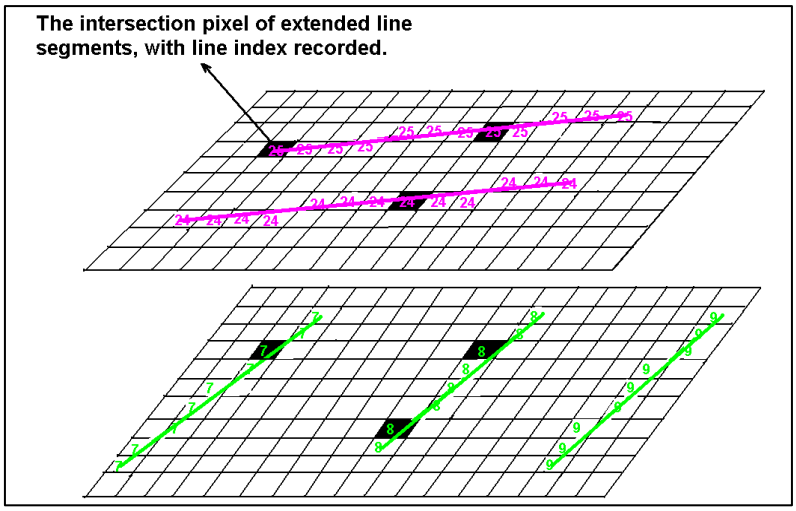
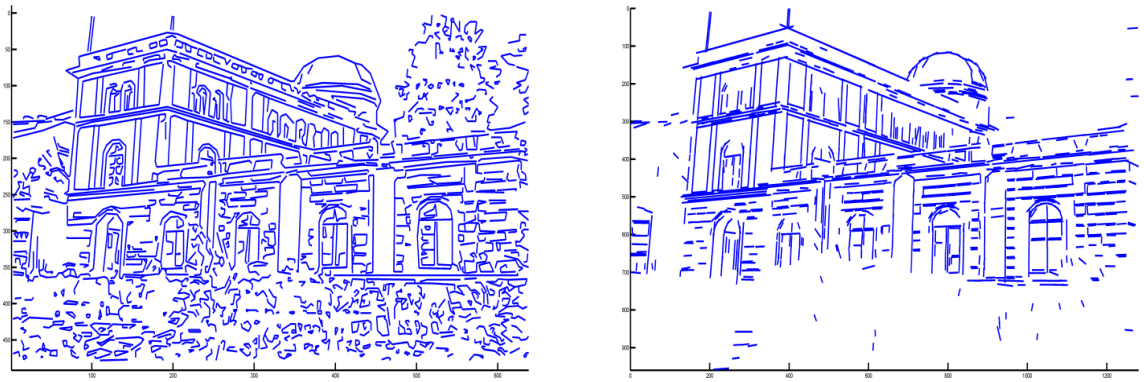


Figure 4. The illustration of how the intersections between lines were found.

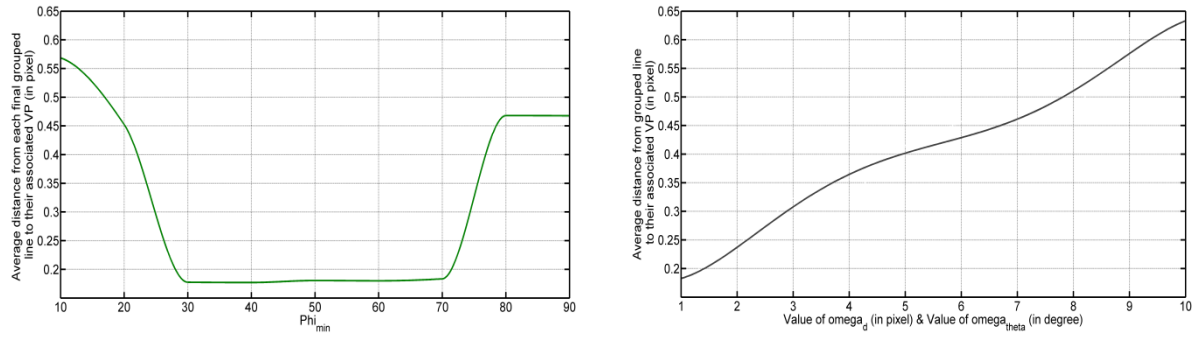


(a) (b)

Figure 5. (a) Line segments detected using Kovesi's method; (b) Line segments detected using Kahn's approach.

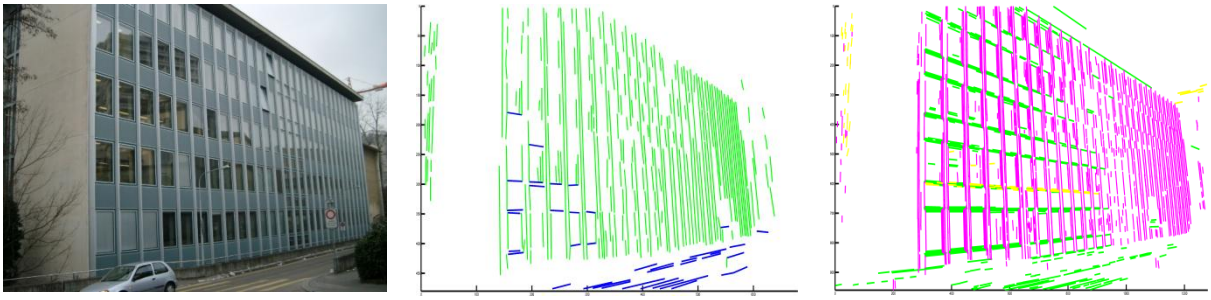
	Image1	Image2	Image3	Image4	Image5	Image6	Image7	Image8
No. lines detected by Kahn's method	1171	2513	1925	1210	1577	1034	2023	1570
No.lines detected by Kovesi's code	1058	801	1324	651	702	838	728	704
Time used for computing Kahn's method	17.96	17.48	19.82	16.48	15.23	15.92	14.97	15.39
Time used for Kovesi's original matlab code	426.31	121.24	688.08	59.09	67.06	190.63	37.04	43.26

Table 1. Comparison of two method on line segments detection in terms of line quantity and time used.



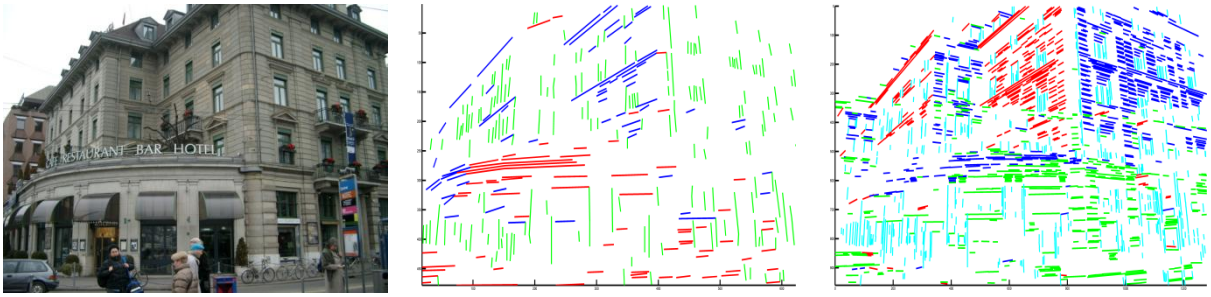
(a) (b)

Figure 6. Average deviation distance from final grouped lines to their estimated vanishing point VS (a) ψ_{min}; (b) σ_0 and σ_d.



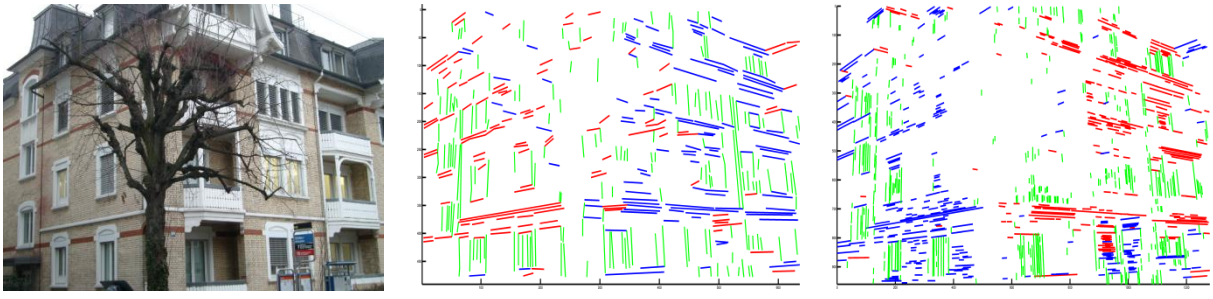
(a) (b) (c)

Figure 7. Line grouping results on a building image with sharp viewpoint. (a) original sample images; (b) line grouping results from previous EM based method; (c) results from our approach.



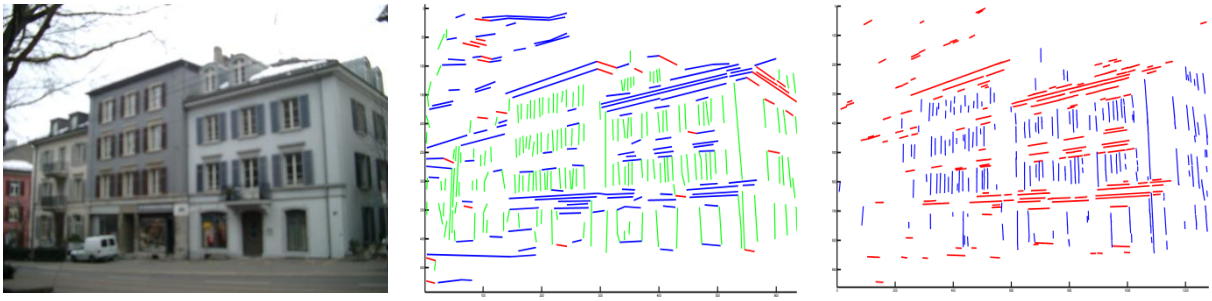
(a) (b) (c)

Figure 8. Line grouping results on buildings with multiple facades. (a) original images; (b) grouping results from previous EM based method; (c) grouping results from proposed approach.



(a) (b) (c)

Figure 9. Line grouping results on buildings with occlusion. (a) original images; (b) grouping results from previous EM based method; (c) grouping results from proposed approach.



(a) (b) (c)

Figure 10. Line grouping results on blurred building image. (a) blurred image sample; (b) line groups computed from previous EM method; (c) line groups computed from proposed approach.

4. CONCLUSIONS

A new approach for building planar surface associated vanishing points detection and line grouping has been presented. Working under uncalibrated camera conditions and with images not conforming to the Manhattan assumption, the experimental results have shown that the method is able to perform well on a wide range of building structures and conditions such as a cluttered environment. It also outperforms existing techniques in working in complex building images. Further improvement to this method would be the introduction of an error model as in [Cou03a]. In the future, we will investigate ways of integrating this method into 3D reconstruction of buildings and building recognition task.

5. REFERENCES

[Bar83a] Barnard, S. Interpreting perspective images. Artificial Intelligence, vol. 21, 1983.

[Bri91a] Brillault-O'Mahony, B. New method for vanishing point detection. Computer Vision, Graphics, and Image Processing, vol. 54(2), pp.289-300, 1991.

[Cou03a] Coughlan, J.M. and Yuille, A.L. Manhattan world: Orientation and outlier detection by Bayesian inference. Neural Computation, 15(5): 1063-1088, 2003.

[Dav08a] David, P. Detection of Building Facades in Urban Environments. Proc. of SPIE conf. on Visual Information Processing XV!!, vol. 6978, pp.9780, 2008.

[Dua09a] Duan, W. and Allinson, N.M. Automatic approach for rectifying building facades from a single uncalibrated image. Proc. of 6th International Conference on Informatics in Control, Automation and Robotics (ICINCO), vol.2, pp.37-43, 2009.

[Kah90a] Kahn, P., Kitchen, L. and Riseman, E.M. A fast line finder for vision-guided robot navigation. IEEE Transactions on PAMI, 12(11), pp.1098-1102, 1990

[Kos02a] Kosecka, J. and Zhang, W. Video compass. ECCV, vol. 2353, pp. 29-32, 2002.

[Kov07a] Kovesi, P.D. MATLAB and Octave functions for computer vision and image processing. School of Computer Science & Software Engineering, The University of Western Australia. Available from: http://www.csse.uwa.edu.au/~pk/research/matlabfns/

[Lut94a] Lutton, E., Maitre, H. and Lopez-Krahe, J. Contribution to the determination of vanishing points using Hough transform. IEEE Transaction on Pattern Analysis and Machine Intelligence, 16(4), pp.430-438, 1994.

[Mag84a] Magee, M.J. and Aggarwal, J.K. Determining vanishing points from perspective images. Computer Vision, Graphics and Image Processing, vol. 26, pp.256-267, 1984.

[Qua89a] Quan, L. and Mohr, R. Determining perspective structures using hierarchical Hough transform. Pattern Recognition Letters, vol. 9, pp.279-286, 1989.

[Rot02a] Rother, C. A new approach to vanishing point detection in architectural environments. Image and Vision Computing, vol.20, pp.647-655, 2002.

[Sha03a] Shao, T.S.H. and Gool, L.V. Zubud-zurich buildings database for image based recognition. *Technical report No. 260*, Swiss Federal Institute of Technology, 2003, http://www.vision.ee.ethz.ch/showroom/zubud/

[Tuy98a] Tuytelaars, T., Gool, L.V., Proesmans, M. and Moons, T. The Cascaded Hough Transform as an Aid in Aerial Image Interpretation. Proceedings of the Sixth International Conference on Computer Vision (ICCV'98), pp.67-72, 1998.

[Wil07a] Wildenauer, H. and Vincze, M. Vanishing point detection in complex man-made worlds. ICIAP, 2007.