# A new Approach to Vanishing Point Detection in Architectural Environments

Carsten Rother

*Computational Vision and Active Perception Laboratory (CVAP), Royal Institute of Technology (KTH), S–10044 Stockholm, Sweden*

**Abstract**

A man-made environment is characterized by many parallel lines and orthogonal edges. In this article, a new method for detecting the three mutually orthogonal directions of such an environment is presented. Since real-time performance is not necessary for architectural applications, such as building reconstruction, a computationally intensive approach was chosen. However, this enables us to avoid one fundamental error of most other existing techniques. Compared to theirs, our approach is furthermore more rigorous, since all conditions given by three mutually orthogonal directions are identified and utilized. We assume a partly calibrated camera with unknown focal length and unknown principal point. By examine these camera parameters, which can be determined from orthogonal directions, falsely detected vanishing points may be rejected.

*Key words:* Vanishing points, vanishing lines, geometric constraints, architecture, camera calibration

## 1 Introduction

The analysis of vanishing points provides strong cues for inferring information about the 3D structure of a scene. With the assumption of perfect projection, e.g. with a pin-hole camera, a set of parallel lines in the scene is projected onto a set of lines in the image that meet in a common point. This point of intersection, perhaps at infinity, is called the *vanishing point*. Vanishing points which lie on the same plane in the scene define a line in the image, the so-called the *vanishing line*. Figure 1 shows the three vanishing points and vanishing lines of a cube, where a finite vanishing point is defined by a point on the image plane and a vanishing point at infinity is defined by a direction on the image plane. If the camera geometry is known, each vanishing point corresponds to an orientation in the scene and vice versa.

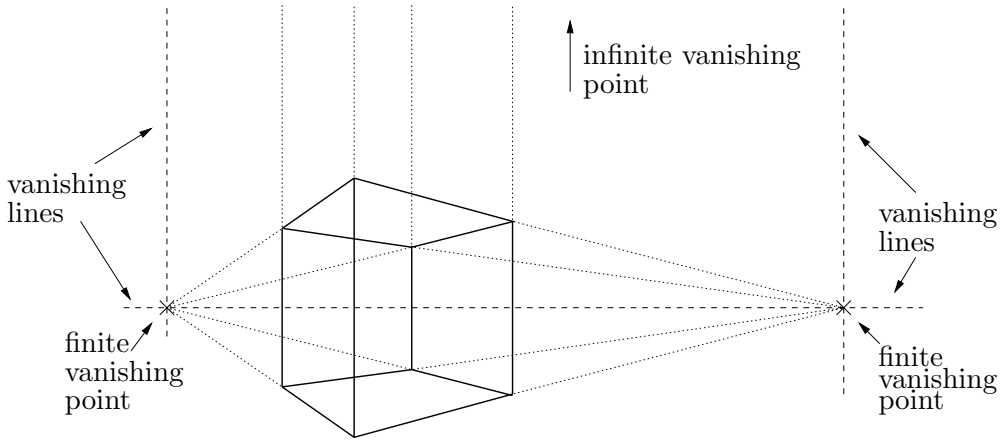



Fig. 1. The three vanishing points and vanishing lines of a cube.

The understanding and interpretation of a man-made environment can be significantly simplified by the detection of vanishing points. For instance, this has been done in the field of navigation of robots and autonomous vehicles [16], in the field of object reconstruction [5,15] or for the calibration of cameras [4,9,17]. A man-made environment has two characteristic properties: Many lines in the scene are parallel and various edges in the scene are orthogonal. In an indoor environment this is true for e.g. shelves, doors, windows and corridor boundaries. In an outdoor environment e.g. streets, buildings and pavements satisfy this assumption. On the basis of these properties the task of detecting the three mutually orthogonal directions of a man-made environment has raised considerable interest [6,8,16].

After a discussion of existing vanishing point detection methods in Section 2, our method is presented in Section 3. Section 4 demonstrates the performance of our method on real image data.

## 2 Previous Work

The majority of vanishing point detection methods rely on line segments detected in the image. A different approach is to consider the intensity gradients of the pixel units in the image directly [6,18]. Since we base our method on line segments, these approaches will be considered in more detail in the following.

The task of detecting those vanishing points that correspond to the dominant directions of a scene is traditionally solved in two steps. Firstly, line segments are clustered together on the condition that a cluster of line segments shares a common vanishing point. We denote this step as the *accumulation step*. In the second step, the dominant clusters of line segments are searched for. We refer to this step as the *search step*.

Let us consider the accumulation step first. In order to reduce the computational complexity of the clustering process, the unbounded image $R^2$ is mapped onto a bounded space. This has the additional advantage that infinite and finite vanishing points can be treated in the same way. The bounded space, also denoted as *accumulator space*, can then be partitioned into a finite number of cells, so-called *accumulator cells*. Barnard [2] suggested the Gaussian sphere centred on the optical centre of the camera as an accumulator space (see fig. 2). A great circle on the Gaussian sphere represents a line segment in the

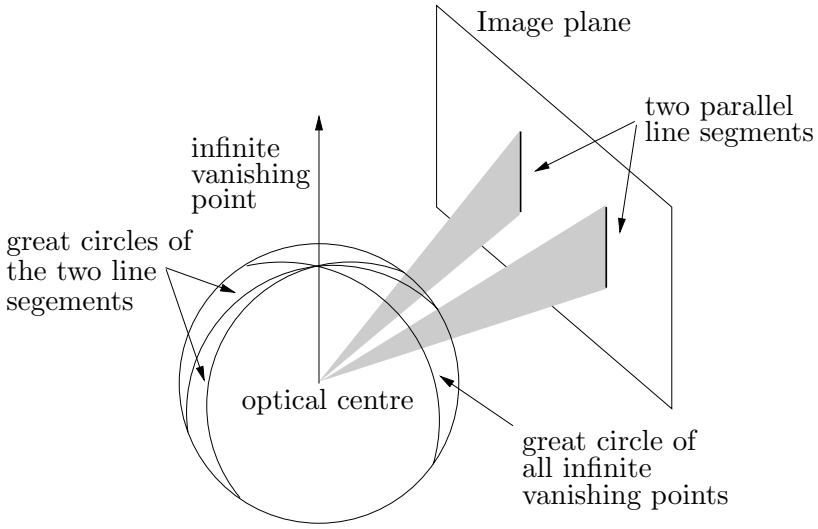



Fig. 2. The Gaussian sphere as an accumulator space.

image and a point on the Gaussian sphere corresponds to a vanishing point in the image. Figure 2 shows that the great circles of two line segments in the image plane always intersect in one point, their vanishing point. For the accumulation of line segments, the Gaussian sphere is tessellated into accumulator cells, and each cell is increased by the number of great circles which pass through it. This approach was then enhanced in other works. Since Barnard chose an irregular and quite ad hoc tessellation of the Gaussian sphere, this was improved by Quan and Mohr [14]. Lutton et al. [12] investigated the influence of different error types, e.g. error due to the finite extension of the image, in the accumulation process on the Gaussian sphere. Magee and Aggarwal [13] accumulated the projection of the intersection points of all pairs of line segments in the image onto the Gaussian sphere. This approach is computationally more intensive but on the other hand more accurate. Alternative accumulator spaces were introduced in [16,3]. Brillault [3] established an uncertainty model for a line segment. According to this model an accumulator space is introduced, in which the expected uncertainty of a line segment remains constant in the accumulator space. A different approach to reducing the computational complexity of the accumulation step is to apply the Hough transformation by mapping the parameters of the line segments into a bounded Hough space [1,18]. Tuytelaars et al. [18] applied the Hough transformation three times (Cascade Hough transformation). At different levels of

the Cascade Hough transformation a peak in the Hough space corresponds to a vanishing point and a vanishing line respectively.

A fundamental drawback of all techniques [1–3,12–14,16,18] that transfer information from the image into a bounded space is that the original distances between lines and points are not preserved. Let us consider the two great circles of the two line segments in fig. 2. Due to the perspective effect of the projection from the image plane onto the Gaussian sphere the distance between these two great circles differs when the two line segments undergo the same movement on the image plane. Therefore, the distance between a line segment and a vanishing point depends on their location on the image plane, i.e. the distances between points and line segments on the image plane are not translationally and rotationally invariant. From an abstract point of view, the transformation simplifies the space of information, i.e. the detected line segments, which effects the relation between information, i.e. relative location of geometrical entities. This drawback can be avoided if no transformation is carried out, i.e. the image plane itself is chosen as the accumulator space. Apart from this drawback the advantages of a bounded space in contrast to an unbounded space are: A finite partition of the space is simpler and a distance function between geometrical entities is easier to formulate.

In the past, more effort has been spent on the accumulation step than on the search step. One of the reasons for this is that the directions in the scene of the searched dominant vanishing points do not have to be orthogonal. This means that the orthogonality of the direction of vanishing points was not treated as an additional criterion of the search step. In [13,14] the search step was designed in a straight forward manner. Firstly, the dominant vanishing point, which corresponds to the accumulator cell with most line segments, is detected. After removing the line segments which correspond to this vanishing point the search for a maximum in the accumulator space is repeated. The iterating process stops when the number of line segments of a dominant vanishing point is below a certain threshold. This approach is characterized by a minimal computational effort. Recently, van den Heuvel [8] developed a method for detecting the three mutually orthogonal directions in the scene. The orthogonality criterion was explicitly used, which means that all combinations of three possible vanishing points have to be considered. This method requires higher computational effort than the approach mentioned above. However, in contrast to this work van den Heuvel assumes a calibrated camera. Coughlan and Yuille [6] searched for two orthogonal directions in the scene using Bayesian inference based on statistics which have been learnt in a specific domain.

## 3 Detection of Three Mutually Orthogonal Directions

With increased computing power and without the condition of real-time performance an approach with higher computational effort is reasonable and will here be pursued. As accumulator space we choose the unbounded image plane itself. This has in contrast to [1–3,12–14,16,18] the advantages which was described in the previous section. We will show that, despite the fact that the image plane is unbounded, infinite and finite vanishing points can be treated in the same way. As accumulator cells we choose (like [13]) the intersection points of all pairs of line segments. For the search step we will establish all criteria, which vanishing points with mutually orthogonal directions have to satisfy. We assume, in contrast to [8], a partly calibrated camera with unknown focal length and unknown principal point. It is known that such a camera can be fully calibrated on the basis of three orthogonal vanishing points [4]. By restricting the range of the focal length and the location of the principal point, falsely detected vanishing points can be rejected. Since this implies that each triplet of potential mutually orthogonal vanishing points have to be considered separately, this method requires more computations than other approaches, such as [13,14].

### *3.1 The Accumulation Step*

Due to various reasons, e.g. noise and lens imperfections, the perspective projection of a line segment from the 3D scene onto the 2D image is not congruent with the line segment detected in the image. We denote this perfect projection of a line segment as the *projected line segment*. Hence, all vanishing point detection methods have to formulate either implicitly or explicitly a distance function between a vanishing point and a detected line segment. In this context the basic question is: *How close is a projected line segment $s'$ with the vanishing point $vp$ to its corresponding, detected line segment $s$?* In order to answer the question we represent a line segment with the midpoint representation $(m_x, m_y, l, \alpha_s)$ (see fig. 3 (b)). Compared to other representations, e.g. endpoint representation, it has the advantage that the length of a line segment is treated explicitly. We define: The perfect line segment $s'$ of a line segment $s$ has the same midpoint as $s$ and has $vp$ as vanishing point. On the basis of this definition a distance function $d(vp, s)$ between a vanishing point $vp$ and a line segment $s$ can be defined as the angle $\alpha$ between the corresponding line segments $s'$ and $s$. Figure 3 (a) gives an example for a finite vanishing point. Since we need as well a distance function $d(l, s)$ between a line $l$ and a line segment $s$ in the search step, we define this distance as the tuple $(d, \alpha)$, where $d$ is the distance between $l$ and the midpoint of $s$ and $\alpha$ is the angle between $s'$ and $l$ (see fig. 3 (b)).

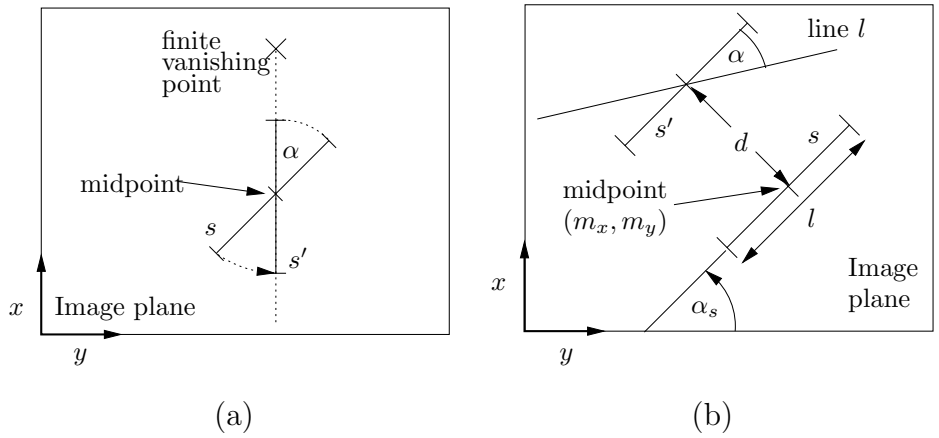



Fig. 3. Explanation for the distance function $d(vp, s)$ between a line segment $s$ and a finite vanishing point $vp$ (a) and for the distance function $d(l, s)$ between a line $l$ and a line segment $s$ (b).

These distance functions fulfill the requirements we state above: Finite and infinite vanishing points are treated in the same way and the distances between points, lines and line segments are independent of their location on the image plane. Note, with this simple approach we disregard the error of a detected line segment and of a potential vanishing point. The modeling of these errors would lead towards a more complex and probabilistic framework. In [10,11] it is assumed that the error in the detected line segment can be modeled by isotropic mean zero Gaussian noise on the endpoints. On the basis of this assumption the maximum likelihood estimate of the corresponding line segment can be explicitly computed by minimizing the sum of squared orthogonal distances from the endpoints of the detected line segment. However, it is questionable if this noise model represents the "true" noise of a line segment in a good way.

On the basis of this framework, we can both formulate and fill the accumulator space. The intersection points, perhaps at infinity, of all pairs of non-collinear line segments are considered as accumulator cells, i.e. potential vanishing points [1]. Since a vanishing point in the 3D scene is a point at infinity, the corresponding vanishing point in the 2D image cannot lie on a line segment, i.e. between the two endpoints of a line segment, with this vanishing point. Therefore, all potential vanishing points which do not satisfy this condition are removed. In order to fill the accumulator space, we state that a line segment $s$ votes for an accumulator cell $a$ if the distance $d(a, s)$ is below a certain threshold $t_a$. Since we have to compare different accumulator cells in the search step, we are interested in the total vote of an accumulator cell. This vote depends on the length of a line segment as well as on the distance between accepted line segments and the accumulator cell. Thereby, we

[1] In the following we do not distinguish between an accumulator cell and a vanishing point.

assume that longer line segments are more reliable than shorter line segments. Therefore, we define

$$vote(a) = \sum_{\text{all } s \text{ vote for } a} w_1 \left(1 - \frac{d(a,s)}{t_a}\right) + w_2 \left(\frac{\text{length of } s}{\text{max length of } s}\right) \quad (1)$$

as the total vote of an accumulator cell $a$, where the weights $w_1$ and $w_2$ establish this trade off.

A brute force version of the accumulation algorithm checks the acceptance of each line segment for each accumulator cell. Afterwards the total vote of each accumulator cell is determined. The computational effort of this algorithm is $O(an) = O(n^3)$, where $a = O(n^2)$ is the amount of accumulator cells and $n$ the amount of line segments.

A similar voting scheme, based on line segments, has been suggested in [10,11]. The vote consists of the sum of all squared orthogonal distances between the endpoints of the detected line segments and the corresponding lines. On the basis of this vote, i.e. cost function, a maximum likelihood estimation of a vanishing point can be established from the line segments voting for this vanishing point.

### *3.2 The Search Step*

The task of the search step is to determine the vanishing points, which correspond to the three mutually orthogonal directions of the scene. Due to this constraint on the vanishing points, three different criteria for the vanishing points can be identified: *The orthogonality criterion, the camera criterion and the vanishing line criterion*. The first two criteria are related to each other and we consider them first.

In [4,9,17] was shown that knowledge of the camera geometry can be deduced from the vanishing points of three mutually orthogonal directions. This knowledge differs in the three different cases, in which none, one or two vanishing points are at infinity. Since these different cases are algebraically discussed in terms of the imaged absolute conic in [9], we illustrate and summarize them here. In order to formulate the orthogonality criterion we have to establish the geometry between the image plane and the camera. We use the same camera model as in [9,17], which deviates from the general perspective camera model in the respect that both image axes are assumed to be orthogonal with the same scale factor. Most real world cameras approximately satisfy this condition. Therefore, the only unknowns of the camera geometry are the focal length $f$ and the principal point $u_0$ (see fig. 4).

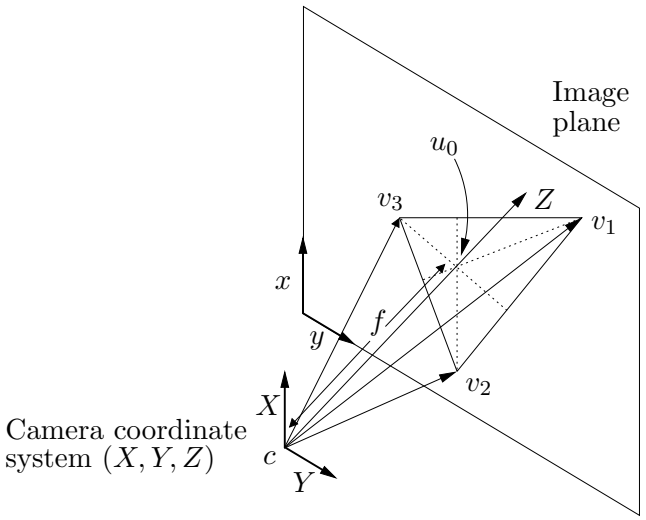



Fig. 4. Explanation for the camera geometry and the orthogonality criterion.

We denote the three vanishing points on the image plane by $v_1, v_2$ and $v_3$. Since the vector $cv_i$ from the camera centre $c$ to the vanishing point $v_i$ has the vanishing point $v_i$, we can formulate the *orthogonality criterion* as:

$$\langle cv_1, cv_2 \rangle = 0, \langle cv_1, cv_3 \rangle = 0 \text{ and } \langle cv_2, cv_3 \rangle = 0, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product. The question for the orthogonality criterion is: Do the vanishing points $v_1, v_2$ and $v_3$ satisfy these three equations (2) with reasonable values for $u_0$ and $f$, i.e. $f \in [0, \infty)$. We discuss the three different cases with none, one or two vanishing points at infinity separately:

(1) *Three finite vanishing points $v_1, v_2$ and $v_3$:*
The triangle $(v_1, v_2, v_3)$ forms together with the principal point an orthocentric system (see fig. 4). Therefore, the intersection point of the heights of this triangle defines the principal point. The size of this triangle defines the focal length uniquely. The orthogonality criterion can be defined as the condition that each angle of this triangle is smaller than $90^o$.

(2) *Two finite vanishing points $v_1, v_2$ and one infinite vanishing point $v_3$:*
The principal point lies on the line segment, which is defined by the two endpoints $v_1$ and $v_2$. For real world cameras the principal point is more likely positioned in the centre of the image. Therefore, we choose the principal point as the point which lies on the line segment and is closest to the midpoint of the image. By determining the principal point, the focal length is uniquely defined. In this case the orthogonality criterion is defined by the condition that the direction of the infinite vanishing point $v_3$ is orthogonal to the line defined by $v_1$ and $v_2$.

(3) *One finite vanishing point $v_1$ and two infinite vanishing points $v_2, v_3$:*
In this case the principal point is identical with the vanishing point $v_1$. The focal length cannot be determined. The orthogonality criterion is

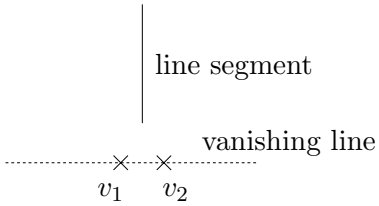



Fig. 5. Two close located vanishing points $v_1, v_2$ which do not fulfill the vanishing line criterion.

defined by the condition that the directions of $v_2$ and $v_3$ are orthogonal.

We can now specify the *camera criterion*. This criterion is fulfilled if the principal point and the focal length are inside a certain range, in case they are calculable.

Let us consider the *vanishing line criterion*. Two different vanishing points $v_1$ and $v_2$ have a finite vanishing line if not both vanishing points are at infinity (see fig. 1). Therefore, a line segment which lies on the vanishing line does vote for the two accumulator cells which correspond to $v_1$ and $v_2$. Hence, we define that two accumulator cells fulfill the vanishing line criterion if each line segment which votes for both accumulator cells lies on the corresponding vanishing line. Figure 5 shows a case where a line segment supports two close located, finite vanishing points but does not lie on the vanishing line. In case both vanishing points are at infinity, the two sets of line segments of the corresponding accumulator cells have to be disjoint. With the distance function $d(l, s)$ we can check if a line segment $s$ lies on a vanishing line $l$. Since $d(l, s)$ returns a tuple $(d, \alpha)$, we check if $d$ and $\alpha$ are below certain thresholds $t_d$ and $t_\alpha$ respectively. The two vanishing points in fig. 5 do not fulfill the criterion with the threshold $t_\alpha < 90^o$ since $\alpha = 90^o$. Note, the lower the threshold $t_a$ (see sec. 3.1) is chosen, i.e. the closer a line segment has to be to an accumulator cell in order to vote for it, the more pairs of accumulator cells fulfill the criterion. Furthermore, if the orthogonality and the camera criterion restrict the camera geometry "considerably", the vanishing line criterion could be omitted since vanishing points which fulfill the vanishing line criterion should fulfill the other two criteria as well.

With the criteria developed above we can define an algorithm for the search step:

Take the accumulator cell $a_1$ with the highest vote $vote(a_1)$ (see equation 1)
Go through all pairs of accumulator cells $(a_i, a_j)$
  If the vanishing line criterion is fulfilled for $(a_1, a_i)$, $(a_1, a_j)$ and $(a_i, a_j)$
    If the orthogonality and the camera criterion are fulfilled for $(a_1, a_i, a_j)$
      Calculate $vote = vote(a_1) + vote(a_i) + vote(a_j)$
Take the accumulator cells $a_i, a_j$, with the highest vote $vote$

The vanishing points which correspond to the accumulator cells $a_1, a_i$ and $a_j$ represent the three mutually orthogonal directions of the scene. The computational effort of the search step is $O(a^2 n) = O(n^5)$, where $a = O(n^2)$ is the amount of accumulator cells and $n$ the amount of line segments.

This exhaustive search for the best three orthogonal vanishing points might be too time consuming for certain applications. It could be speed up by RANSAC (see [7]). A RANSAC version of our algorithm would randomly choose pairs of accumulator cells $(a_i, a_j)$ and check if they fulfill the three criteria. The algorithm terminates if the total vote of the cells $a_1, a_i$ and $a_j$ is above a certain threshold.

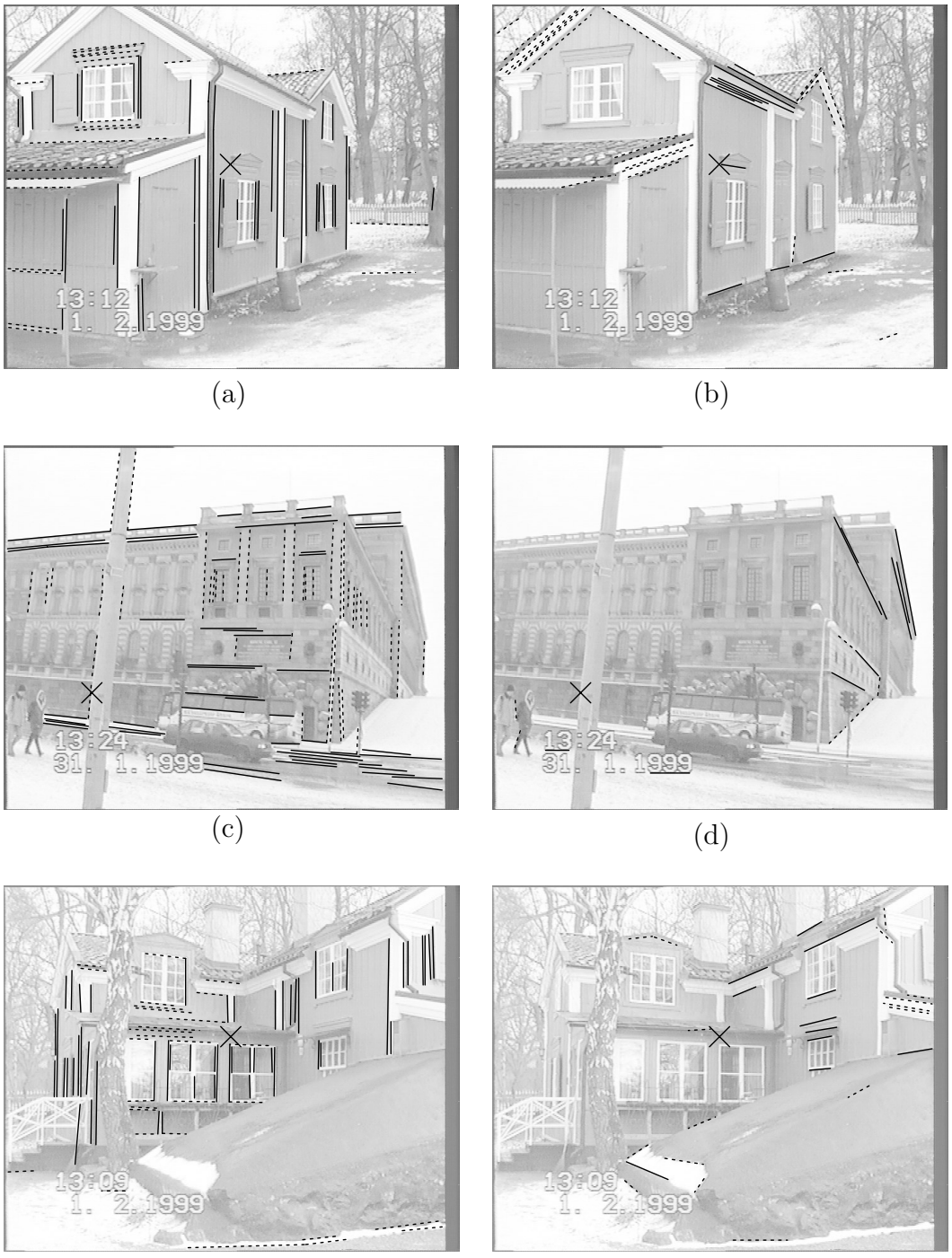
## 4 Experimental Results

For the experiments a standard, auto-focus, hand-held digital camera (KODAK DC 120) was used. 16 images with size of $720 \times 576$ pixels were taken of the Royal Palace of Stockholm (see fig. 6 (c)) and a residential house (see fig. 6 (a)). The camera parameters remained fixed while the pictures were taken. For the process of vanishing point detection the parameters were set as: $t_a = 5^o, t_\alpha = 5^o, t_d = 30$ pixel, $w_1 = 0.3$ and $w_2 = 0.7$. The maximal difference between the principal point and the midpoint of the image was set to 400 pixel, i.e. 43% of the image diagonal. The results of two different experiments are summarized in Table 1.

Table 1
Classification of the results of two different experiments on 16 images.

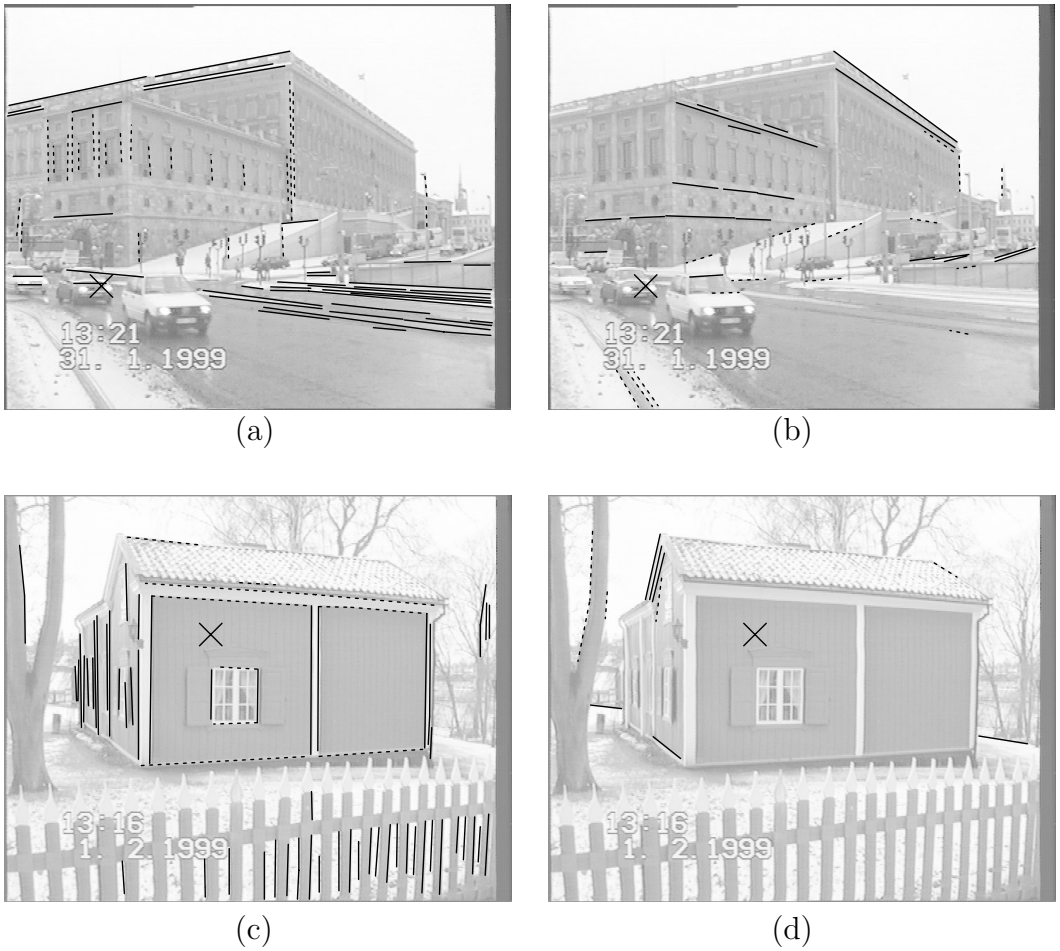
| Result | good | average | poor |
|---|---|---|---|
| $f \in [0, \infty)$ | 10 | 3 | 3 |
| $f \in (700, 1200)$ | 13 | 2 | 1 |

Let us consider the first experiment, in which the acceptable range of the focal length was not limited, i.e. $f \in [0, \infty)$. For 10 (out of 16) images all three mutually orthogonal vanishing points were found correctly, i.e. classified as good. Figures 6 (a-f) demonstrate examples of good results. We see that good results were achieved for images with a cluttered environment (see fig. 6 (e,f)) or with a substantial amount of outliers (see fig. 6 (a,b)). A result was classified as average if at least one of the vanishing points was not detected accurately. This can occur if a considerable amount of line segments which have a different vanishing point are additionally assigned to this vanishing point. Note, since all points of the unbounded image plane represent the projection of vanishing points from the 3D scene onto the image plane, every line segment on the image plane passes through points, i.e. possible vanishing points which are not its vanishing point, e.g. the two solid line segments in the left, down corner in fig.




(a) (b) (c) (d) (e) (f)

Fig. 6. Three results of the first experiment which were classified as good. The two different types of line segments in the left images (a,c,e) and the solid line segments in the right images (b,d,f) represent the three mutually orthogonal directions. The dashed line segments in the right images (b,d,f) show the remaining line segments which were not assigned to one of the three vanishing points. The principal point of the camera is drawn as a cross.

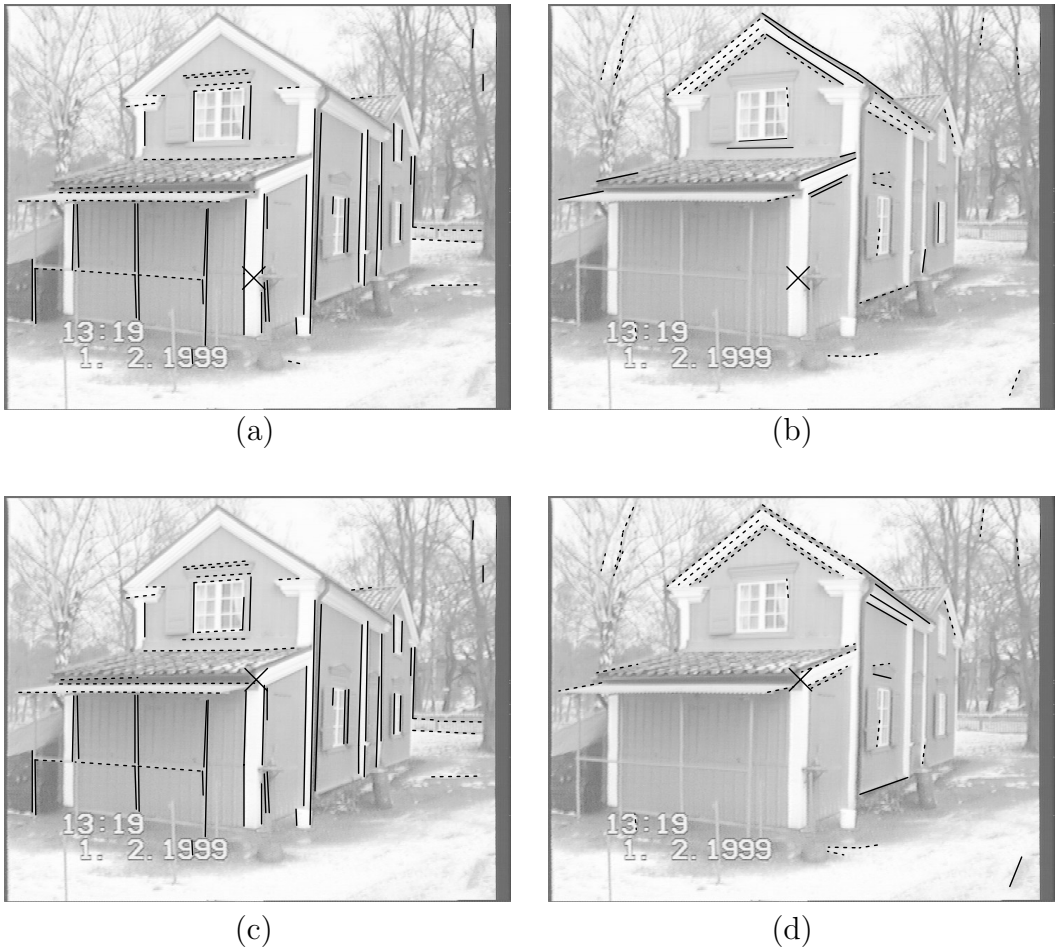



(a) (b) (c) (d)

Fig. 7. An average result (a,b) and a poor result (c,d) of the first experiment. The line markings are as in fig. 6.

6 (d). Figure 7 (a,b) shows one of the three average results. In fig. 7 (a) the solid line segments of the building and of the street in front of the building are assigned to the same vanishing point. Indeed, the building is approximately rotated about $20^o$ relatively to the street. A result was classified as poor when at least one of the vanishing points was detected incorrectly. Figure 7 (c,d) and 8 (a,b) demonstrate poor results. The solid line segments in fig. 7 (d) and 8 (b) do not represent a vanishing point.

Let us consider the distribution of the determined camera parameters for the 16 images (see fig. 9). We see that the principal point of a good result can deviate up to 251 pixel, i.e. 27% of the image diagonal, from the midpoint of the image, which is presumably close to the true principal point. The image which corresponds to this principal point is displayed in fig. 6 (c,d). Furthermore, the principal points of average or poor results are close to the midpoint of the image. Therefore, a further specification of the acceptable range of the principal point in advance would not necessarily improve the results. However, the distribution of the focal lengths (see fig. 9 (b)) shows that the focal lengths of some average or poor results deviate more from the average focal length




(a) (b)

(c) (d)

Fig. 8. A poor result of the first experiment (a,b) improved to a good result (c,d) in the second experiment. The line markings are as in fig. 6.

than the focal lengths of good results. The maximal deviation of a focal length of a good result from the average focal length, which is 953 pixel, is 173 pixel, i.e. 18%.

On the basis of this investigation we carried out a second experiment, in which the possible range of the focal length was limited to $f \in [700, 1200]$. The second line of tabular 1 shows that the number of good results improved. The poor result in fig. 8 (a,b) improved to a good result (see fig. 8 (c,d)), whereat the focal length changed from 658 pixel (first experiment) to 953 pixel (second experiment). The only remaining poor result, with a focal length of 1654 pixel in the first experiment, is displayed in fig. 7 (c,d). In this case it is not possible to obtain a better result, since the third orthogonal direction (solid line segments in fig. 7 (d)) is only "supported" by one solid line segment.

Let us consider the processing time of the accumulation step and the search step on an Ultra Sparc 10. For 77 line segments and 737 accumulator cells the runtime was 3,50 sec for the accumulation step and 3,16 sec for the search

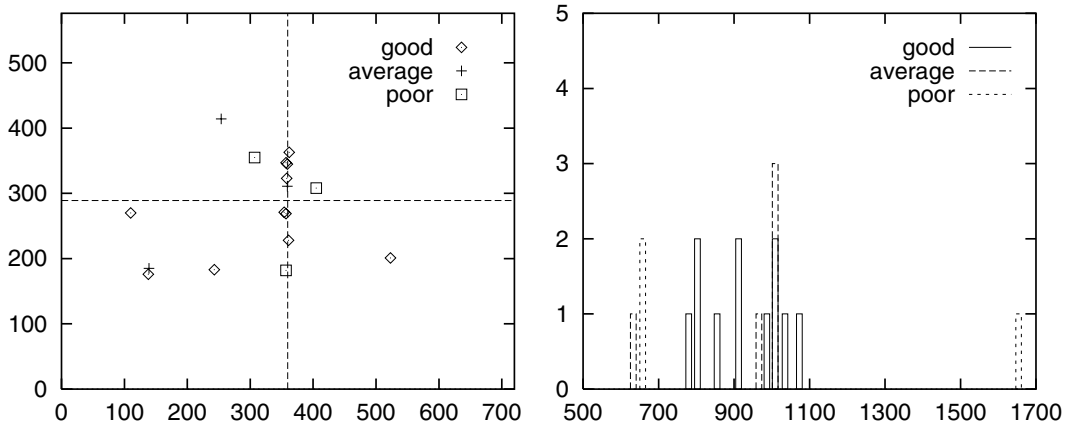



Fig. 9. The determined principal points (a) and focal lengths (b) of the 16 images of the first experiment. The principal points and focal lengths are classified with respect to good, average and poor results. Diagram (a) corresponds to the image size of $720 \times 576$ pixel, where the dashed lines indicate the midpoint of the image. Diagram (b) shows the frequency of focal lengths, where focal lengths with a difference of less than 15 pixel were accumulated. All focal lengths were in the range of $f \in [500, 1700]$.

step. A different run with 105 line segments and 1552 accumulator cells needed 9,17 sec for the accumulation step and 14,53 sec for the search step. Therefore, our method, in its current version, is applicable if the number of line segments is not large and no real-time conditions are required. Note, since low processing time was not our main goal, the method was designed in a simple and straightforward manner.

## 5 Discussion and Future Work

A new method for detecting the three mutually orthogonal directions of a man-made environment has been presented. Since real-time performance is not necessary for architectural application, such as building reconstruction, an approach has been chosen which is computationally more intensive than other methods, e.g. [13,14]. A simple and coherent framework for the accumulation step and the search step has been introduced. By using the unbounded image plane as accumulator space, the original distances between vanishing points and line segments are preserved, in contrast to techniques which transfer the line segments from the image plane into a bounded space [1–3,12–14,16,18]. In the search step, all criteria for vanishing points of three mutually orthogonal directions have been identified and utilized. We assume, in contrast to [8], a partly calibrated camera with unknown focal length and unknown principal point. By examine these camera parameters, which can be determined from

orthogonal directions [4], falsely detected vanishing points may be rejected.

The experiments have shown that the method produces good results, even for images with a cluttered environment and with a substantial amount of outliers. Furthermore, we see that the performance can be improved by providing more information about the camera, e.g. the acceptable range of the focal length.

In the simple framework for the accumulation step, two possible "sources of error" were neglected, which were discussed in the literature. Firstly, we could introduce an error model for a line segment, like in [3,10,11]. With this model a probabilistic distance function between line segments, lines and vanishing points could be derived, like the Mahalanobis distance for line segments. Furthermore, on the basis of this error model, the uncertainty of a possible vanishing point can be propagated, like in [11]. Depending on the location and uncertainty of the three mutually orthogonal vanishing points, the uncertainty of the determined principal point and focal length can finally be propagate. This would improve the specification of the acceptable range for the principal point and the focal length. Secondly, due to image limitation the probability of a line segment passing through a point in the centre of the image or at infinity might be different (see Lutton et al. [12]). On the basis of the distance function between line segments and vanishing points and of the distribution of line segments inside the image borders, this bias could be predicted and integrated in the accumulation step.

In our future task of building reconstruction we will see if the accuracy of the detected vanishing points is sufficient. If not, the transition from our simple framework to a more complex and probabilistic framework might be necessary.

Especially buildings have the property that in certain poses not enough – or even none – of the detected line segments specify a searched for vanishing point (see fig. 7 (d)). Therefore, we currently extend our method towards an approach which detects all "visible", mutually orthogonal directions of a scene.

## Acknowledgment

We thank S. Carlson and J.-O. Eklundh for their careful reading of a draft.

## References

[1] D. H. Ballard and C. M. Brown: *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ, 1982.

[2] S. T. Barnard: *Interpreting perspective images*. Artificial Intelligence 21, 1983, pp. 435-462.

[3] B. Brillault-O'Mahony: *New Method for Vanishing Point Detection*. In Computer Vision, Graphics, and Image Processing 54(2), 1991, pp. 289-300.

[4] B. Caprile and V. Torre: *Using Vanishing Points for Camera Calibration*. International Journal of Computer Vision 4, 1990, pp. 127-139.

[5] A. Criminisi, I. Reid and A. Zisserman: *Single View Metrology*. Proceedings of the 7th International Conference on Computer Vision (ICCV'99), September 1999, Kerkyra, Greece, pp. 434-441.

[6] J. M. Coughlan and A. L. Yuille: *Manhattan World: Compass Direction from Single Image by Bayesian Inference*. Proceedings of the 7th International Conference on Computer Vision (ICCV'99), September 1999, Kerkyra, Greece, pp. 941-947.

[7] M.A.Fishler and R.C.Bolles: *Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography*. Comm. Assoc. Comp. Mach. 24(6), March 1981, pp. 381-395.

[8] F. A. van den Heuvel: *Vanishing point detection for architectural photogrammetry*. International Archives of Photogrammetry and Remote Sensing, Vol. XXXII part 5, 1998, pp. 652-659.

[9] D. Liebowitz and A. Zisserman: *Combining Scene and Auto-calibration Constraints*. Proceedings of the 7th International Conference on Computer Vision (ICCV'99), September 1999, Kerkyra, Greece, pp. 293-300.

[10] D. Liebowitz and A. Zisserman: *Metric rectification for perspective images of planes*. IEEE Conf. Computer Vision and Pattern Recognition, june 1998, Santa Barbara, CA, pp. 482-488.

[11] D. Liebowitz: *Camera Calibration and Reconstruction of Geometry*. D.Phil. thesis, University of Oxford, june 2001.

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

[13] M. J. Magee and J. K. Aggarwal: *Determining Vanishing Points from Perspective Images*. Computer Vision, Graphics and Image Processing 26, 1984, pp. 256-267.

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

[15] C. Rother and S. Carlsson: *Linear Multi View Reconstruction and Camera Recovery*. Proceedings of the 8th International Conference on Computer Vision, 2001, Vancouver, Canada, pp. 42-51.

[16] M. Straforini, C. Coelho and M. Campani: *Extraction of vanishing points from images of indoor and outdoor scenes*. Image and Vision Computing, Vol. 11(2), 1993, pp. 91-99.

[17] D. Svedberg and S. Carlsson: *Calibration, Pose and Novel Views from Single Images of Constrained Scenes*. Proceedings of the 11th Scandinavian Conference on Image Analysis (SCIA'99), June 1999, Kangerlussuaq, Greenland, pp. 111-117.

[18] T. Tuytelaars, L. Van Gool, M. Proesmans and T. Moons: *The Cascaded Hough Transform as an Aid in Aerial Image Interpretation*. Proceedings of the 6th International Conference on Computer Vision (ICCV'98), January 1998, Bombay, India, pp. 67-72.