

# The Cascaded Hough Transform as Support for Grouping and Finding Vanishing Points and Lines

Tinne Tuytelaars, Marc Proesmans and Luc Van Gool

Katholieke Universiteit Leuven, ESAT-MI2  
Kardinaal Mercierlaan 94, B-3001 Leuven, BELGIUM

**Abstract.** In the companion paper [7] a grouping strategy with a firm geometrical underpinning and without the problem of combinatorics is proposed. It is based on the exploitation of structures that remain fixed under the transformations that relate corresponding contour segments in regular patterns. In this paper we present a solution for the complementary task of extracting these fixed structures in an efficient and non-combinatorial way, based on the iterated application of the Hough transform. Apart from grouping, this ‘Cascaded Hough Transform’ or CHT for short can also be used for the detection of straight lines, vanishing points and vanishing lines.

## 1 Introduction

Grouping is the process of combining bits and pieces of visual information into perceptually salient structures. It is an important stepping stone towards a deeper understanding of observed shapes structure and scene organisation.

Recently, a more systematic geometrical basis for grouping has been propounded on the basis of so-called fixed structures [7, 8]. These are the structures in the image that remain fixed under the grouping transformation. As an example, under the transformation that maps one half of a symmetric pattern onto the other the symmetry axis and all its points remain fixed. This also holds when looking at such pattern from an oblique viewpoint.

A problem that has been left open is that exploiting the grouping-specific invariants is only possible when the fixed structures are known. Therefore, methods to find the fixed structures are called for. This paper proposes a way of doing just that, again avoiding combinatorics.

Therefore, a modified version of the Hough transform is introduced — coined the ‘Cascaded Hough Transform’ or CHT for short — that finds structures at the different hierarchical levels by iterating one single kind of Hough transform. It directly supports the quest for fixed points, fixed lines, lines of fixed points, and pencils of fixed lines.

Apart from this framework of efficient grouping, the CHT can also serve to find vanishing points and lines, which are special types of fixed structures that play an increasingly important role in automatic navigation and three-dimensional scene reconstruction. Previously, Hough transforms were already

used for line extraction or vanishing point detection, but these steps were not combined into a single scheme. Moreover, the Hough transforms involved were of a different nature. This is discussed in more detail later. Having a single scheme for the extraction of these different features greatly simplifies the interaction between the different levels at which they are found.

The paper is organised as follows. Section 2 introduces the Cascaded Hough Transform, the central tool introduced in the paper to find good fixed structure candidates. Section 3 discusses some implementation issues of the CHT. The CHT is illustrated in section 4 and section 5 concludes the paper.

## 2 The Cascaded Hough Transform

The Hough transform's history stretches a long way back, as P.V.C. Hough ([2]) patented it in 1962. Yet, now it probably has more applications than ever before and it still is the subject of research and a favourite approach to feature extraction for many [3, 4]. Attempts to replace it by alternative schemes such as SLIDE have not been very successful [6].

The Hough transform is a global, robust technique for the detection of pre-defined shapes in images, esp. straight lines. It is based on the transformation of the edge points to a parameter space. In fact, in its most efficient forms, the Hough transform exploits the symmetry of simple parametric shapes (translational for lines, rotational for circles) in order to reduce the dimensionality of the search [5].

Here, only the extraction of straight lines is considered. In the original version of the Hough transform, straight lines were detected using a slope-intercept parametric representation. In order to detect edge pixels  $(x, y)$  that lie on a straight line  $ax + b + y = 0$ , a voting mechanism in the  $(a, b)$  parameter space is used. The important thing is that the Hough transform finds structures in a non-combinatorial fashion.

Every point  $(x_0, y_0)$  of an edge in the image is transformed into a line  $x_0a + y_0 + b = 0$  in parameter space. Following this parameterisation, there is a degree of symmetry between image space and parameter space. Indeed, the parameters  $a$  and  $b$  are to image space  $(x, y)$  what  $x$  and  $y$  are to parameter space  $(a, b)$ . Lines in one space can be detected as points in the other space and, vice versa, for every point there is also a corresponding line.

This high symmetry not only emphasizes the duality between lines and points, but also is the very basis of the proposed 'Cascaded Hough Transform', that amounts to iterating the same procedure. Indeed, we can apply a second Hough transform to the (filtered) result of the first one, transforming parameter space back into image space. This is reminiscent of the Fourier transform. There also one goes to the other domain (i.e. the frequency domain) to perform operations that can be carried out more efficiently there, in order then to go back to the original (spatial) domain. In contrast to the Fourier transform, the 'forward' and 'backward' Hough transforms are identical.

The first Hough transform can be used to detect lines in the image, where it is usual to only keep the dominant peaks in the resulting  $(a, b)$ -parameter space. A second Hough transform could then detect lines of collinear peaks in parameter space. This second Hough goes back to image space where the positions of the peaks that it generates correspond to vertices where several straight lines in the original image intersect. Vanishing points are an important case of such vertices. Similarly, a third Hough transform can be applied to the peaks of the second one to detect collinear vertices. An important case of this kind of feature is given by vanishing lines (e.g. the horizon line), containing several of the vanishing points detected by the second Hough transform. These observations are summarised in the flowchart of Fig. 1.

layer	meaning of detected features
layer 0 ↓ Hough 1 ↓	(the original image)
layer 1 ↓ Hough 2 ↓	points ~ lines lines ~ convergent lines
layer 2 ↓ Hough 3 ↓	points ~ intersection points lines ~ collinear intersection points
layer 3	points ~ lines of intersection points

**Fig. 1.** Flowchart summarising the meaning of the features detected at the different layers of the CHT.

The structures found at the second and third layers of the CHT correspond well to the fixed structures we are looking for. Certainly structures that receive many votes are strong candidates to consider as possible fixed structures. In a sense the non-accidentalness principle is applied, but to the interpretation of fixed structures rather than to the interpretation of image contours.

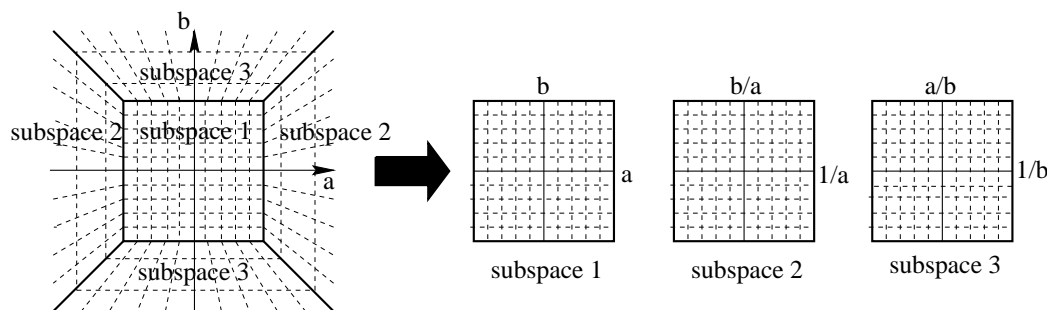
Having this selection process run on straight lines in the original image exclusively, would lower the chance of finding fixed structures except for indoor and building type of scenes, where straight lines abound. Therefore, it is useful to add well-chosen structures at the different levels, as auxiliary input for the next Hough transform. One could for instance add points that correspond to tangent lines at inflections at the level of the first  $(a, b)$ -space. Such lines are bound to reemerge at the next  $(a, b)$ -level, but also structures like symmetry axes can be expected to emerge. The original lines can easily be suppressed.

### 3 Implementation of the CHT

#### 3.1 Parameterisation of the Spaces

The main drawback of the slope-intercept parameterisation of the lines is the unbounded parameter space. Indeed, both  $a$  and  $b$  can take on infinite values. Therefore, another parameterisation was introduced by Duda and Hart [1]. It is based on the orthogonal distance  $\rho$  of the line to the origin and the direction  $\theta$  of the normal to the line. This parameterisation is the one most often used. It yields a bounded parameter space. However, it also breaks the symmetry between the image and parameter spaces. A point in the image is transformed to a trigonometric function in parameter space, and a line to a point. So the basis for the CHT is gone.

In order to obtain a bounded parameter space, while keeping the symmetry intact, the original, unbounded  $(a, b)$ -space is split into three bounded subspaces. This is shown in Fig. 2. The first subspace also has coordinates  $a$  and  $b$ , but only applies to lines with  $|a| \leq 1$  and  $|b| \leq 1$ . If  $|a| > 1$  and  $|b| \leq |a|$ , the point  $(a, b)$  turns up in the second subspace, with coordinates  $1/a$  and  $b/a$ . If, finally,  $|b| > 1$  and  $|a| < |b|$ , we use a third subspace with coordinates  $1/b$  and  $a/b$ . In this way, the unbounded  $(a, b)$ -space is split into three bounded subspaces. Their coordinates are restricted to the interval  $[-1, 1]$ , while a point  $(x, y)$  in the original space is still transformed into a line in each of the three subspaces. As can be seen at the left side of Fig. 2, this subdivision of the parameter space can also be interpreted as an inhomogenous discretisation of the unbounded parameter space, with cells growing larger as they get further away from the origin.



**Fig. 2.** The original, unbounded space is split into three, bounded subspaces

In order to keep both spaces as symmetrical as possible, the same structure is given to  $(x, y)$ -image space, subdividing it into three subspaces, with coordinates  $(x, y)$ ,  $(1/x, y/x)$  and  $(1/y, x/y)$ . But symmetry is not the only reason for this subdivision.  $(x, y)$ -space is in fact also an unbounded space, not restricted to the size of the image itself. Vanishing points for instance can lie outside the image

boundaries, or even at infinity. With the proposed parameterisation, the points lying at infinity are included in a natural way. The fact that cells grow larger when going out of the field of view reflects the fact that points lying further away are normally determined less accurately anyhow.

In order to give the  $(x, y)$ -parameterisation an additional, practical meaning, the original image is rescaled, such that it fits entirely within the first subspace and is discretized in a homogeneous way. Hence the first subspace actually corresponds to the image itself. Subspace 2 captures information left and right from the field of view, and subspace 3 does the same for information from above or below.

It might seem that setting up and clustering three two-dimensional spaces at each layer causes an enormous computational burden. However, it is not the number of spaces that is important, but their total size. For instance, the computational cost of the clustering stage mainly depends on the number of clusters, which is the same for any kind of Hough transform, whatever parameterisation has been used. On the other hand, our approach does require a lot of memory, especially if a high resolution is needed for the Hough spaces, since each layer needs its own three spaces.

### 3.2 Filtering Stages

Before applying a new Hough transform to the result of a previous one, appropriate data handling may be called for. In general, this will include deciding on

1. which data to eliminate from further consideration by subsequent layers,
2. which data to read out as important information emerging at each layer,
3. which data to add to the input of the next Hough Transform.

A first aspect is that only peaks, i.e. relatively high concentrations of votes, are considered for further processing. In order to better emphasise the initial peaks (at the output of the first Hough transform), locally around edge pixels an estimate of the tangential line is made, and only a small blob around the corresponding  $(a, b)$  values is introduced, rather than a complete line. In this way, the strength of the peak will still reflect the length of the straight line in the image.

Only truly ‘non-accidental’ structures are read out at the different layers. What that means is layer dependent: for the subsequent layers these are straight lines of a minimum length, points where at least three straight lines intersect, and lines that contain at least three line intersections. Note that in the latter case, these can be three intersections of each time two lines.

There also are a number of additional possibilities, that have not yet been implemented completely, but are under study. These include:

- As was mentioned before, one can add structures at the different levels, such as  $(a, b)$  peaks that correspond to tangent lines at inflections. This generalises the approach to scenes and shapes containing few straight lines.

- It can help to suppress structures found at earlier layers. For instance, a line will intersect with all other lines. As a consequence, it is bound to reappear as a line of collinear line intersections. Such lines clearly are not very promising candidates, unless more lines intersect it in the same point.
- Structures found can be fed back into earlier layers. This allows additional structures to be found. As an example, a vanishing line added to the image may yield additional vanishing points as non-accidental read-outs.

## 4 Examples of the CHT

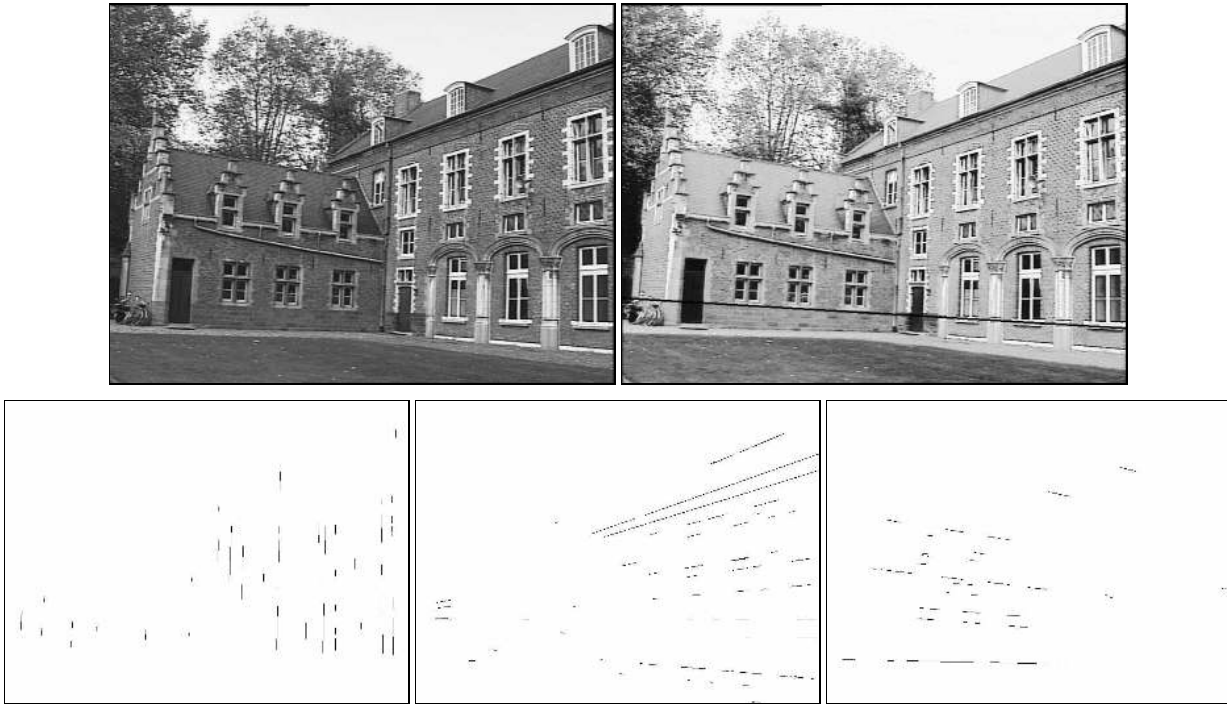
As a first illustration of the CHT, consider Fig. 3. It shows a historic building. In such scenes, straight lines abound, so this example was run without the addition of any extra information (e.g. without adding tangents at inflections). It is illustrated how the CHT yields the vanishing points and lines for the principal directions in the scene. Figure 4 shows the different layers of the CHT in its different rows, with the result of the first Hough transform at the first row, and so on. The different subspaces correspond to the columns, following the order of Fig. 2.

One can clearly see that the peaks in the first layer of the CHT, representing the lines in the original image, are arranged in lines themselves. These lines of peaks correspond to lines in the image that are convergent. The point where they intersect can be detected by applying a second Hough-transform. After filtering, three main peaks can be found in this second layer. They correspond to the three main vanishing points, as can be seen in the bottom part of Fig. 3, where all lines contributing to the detected intersection points are shown.

By applying a third Hough transform we can detect lines of vanishing points, i.e. the vanishing lines in the image. Since we only have three vanishing points, we can find three vanishing lines, all going through two points only. One of them is the horizon line, also shown in Fig. 3.

A similar experiment was carried out for the aerial image shown in Fig. 5. Again, this aerial image contains enough straight lines, allowing to apply the CHT directly, without the addition of any extra features such as tangent lines. Figure 6 shows the different subspaces and layers of the CHT. Peaks in the first row correspond to straight lines in the original image. Some collinear configurations are already salient and correspond to lines intersecting in a single point. These intersections are picked up by the second Hough and hence show up as peaks in the second row. There are two dominant peaks, that correspond to the vanishing points of the two dominant line orientations in Fig. 5. Finally, the third row has local maxima at lines that contain at least two intersections. The two dominant peaks of the second row show up as two dominant lines here and their intersection yields the strongest peak of all. This is the horizon line.

Note that this aerial image is a nadir rather than an oblique view. Perspective effects have previously been used almost exclusively for oblique, reconnaissance

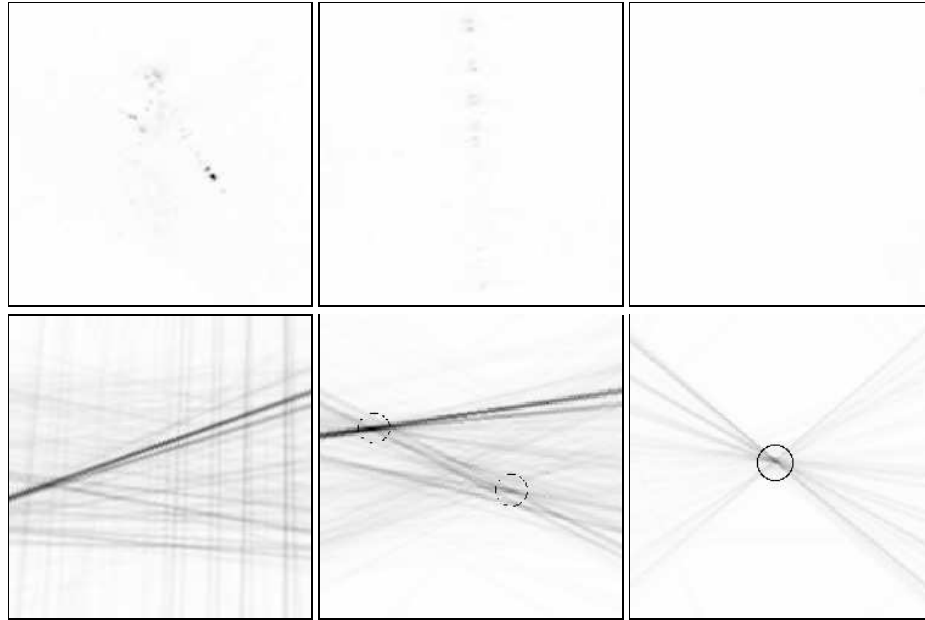


**Fig. 3.** Upper left: the original image. Upper right: idem, with the detected horizon line superimposed. Bottom left to right : lines belonging to one of the three detected vanishing points.

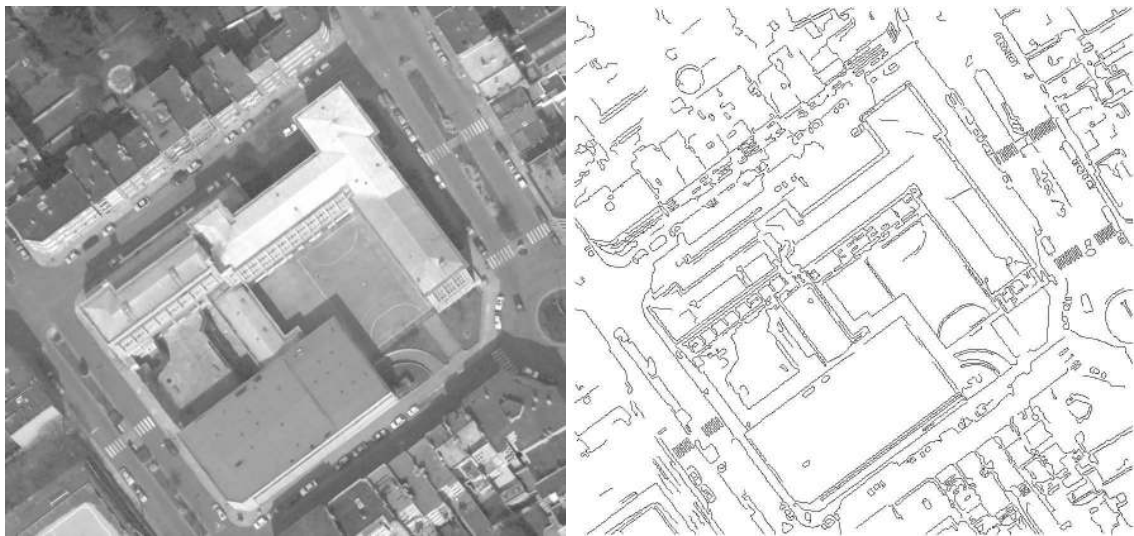
type of views. Aerial imagery for civil applications has the camera oriented almost vertically, however. Nevertheless, even if the perspective effects are much weaker in such case, the results show them to be of sufficient importance to warrant their explicit detection. Thanks to the special parameterisation, vanishing points and lines ‘approaching’ infinity can be handled adequately.

Especially for the orientation going from top-left to bottom-right, there is a perspective effect that cannot be discarded. As an example, it is interesting to focus on the pedestrian walkway with this orientation, that is visible at the left side of the image. Figure 7a shows the line that is obtained from a single application of the Hough transform. Although it is positioned rather well, this line does not follow the walkway precisely. Part of the problem are the cars and their shadows that have pulled the edge to the right.

One can expect a better result, by pooling information on the corresponding, dominant orientation. Therefore, the average orientation of the different Hough-based lines belonging to this group was calculated. Only sufficiently long straight edge segments were used, as indicated in Fig. 9a. They were found as segments along the lines corresponding to peaks in the first Hough space where contrast over the lines was sufficiently high. This threshold was set quite conservatively.

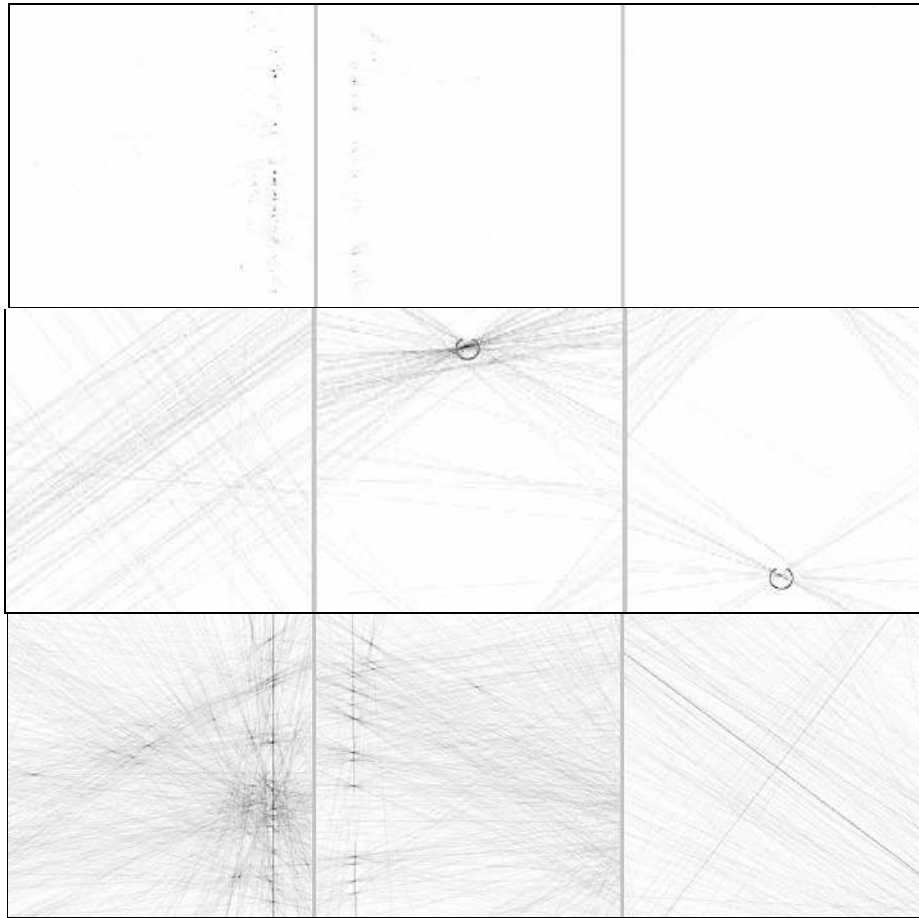


**Fig. 4.** The different layers of the CHT of the image of Fig. 3. The result of the first Hough transform is shown at the first row, and so on. The different subspaces correspond to the columns, following the order of Fig. 2.



**Fig. 5.** Aerial image of high resolution with a school building and surrounding streets (left), and the corresponding edges (right).





**Fig. 6.** Nine CHT subspaces for the image of Fig. 5. The first, second, and third rows give the results for the first, second, and third Hough, resp. The columns give the first, second, and third subspaces, ordered from left to right according to Fig. 2.

Figure 7b shows a second estimate for the walkway edge. It has the average orientation and runs through the center of gravity of the walkway segments in Fig. 9a. This line follows the walkway edge more closely. A deviation from the real edge is still visible, however. The edge should be rotated slightly clockwise. This is a deviation that is in agreement with what we would expect from the perspective distortion in the image. Finally, Fig. 7c shows another line, through the same center of gravity and the vanishing point corresponding to this dominant orientation. The vanishing point was found after the second step of the CHT. As the system keeps track of the lines that contribute to the different vanishing points, they are easy to match. As can be seen in the figure, the line now follows the walkway edge more closely.

The relative quality of the three edge estimations is better illustrated on a detail near the bottom of the image, because there the differences show up most clearly. Figure 8 shows that part of the walkway. The three lines are shown together with, from top to bottom, the line found on the basis of a single Hough, based on the averaged orientation, and based on the vanishing point. Only the latter line remains within a 1-pixel distance from the visual edge throughout the image. A similar correction can be applied to the other lines of the same direction. When looking again for support along these lines, we now find longer edge segments (Fig. 9b).

## 5 Conclusion

A non-combinatorial approach for finding the fixed structures in an image was propounded based on the Cascaded Hough Transform or CHT. Applying several Hough transforms in a cascaded way was shown to yield an efficient method for finding fixed point and line candidates. Such fixed structures are the key element for the geometric grouping proposed in [7].

Of course, the structures found at the different layers of the CHT can also be useful as such. For instance, they can be used as cues for three dimensional reconstructions, or they can serve as a starting point towards robot navigation by e.g. detecting the horizon and the vanishing point of the street which a robot is on.

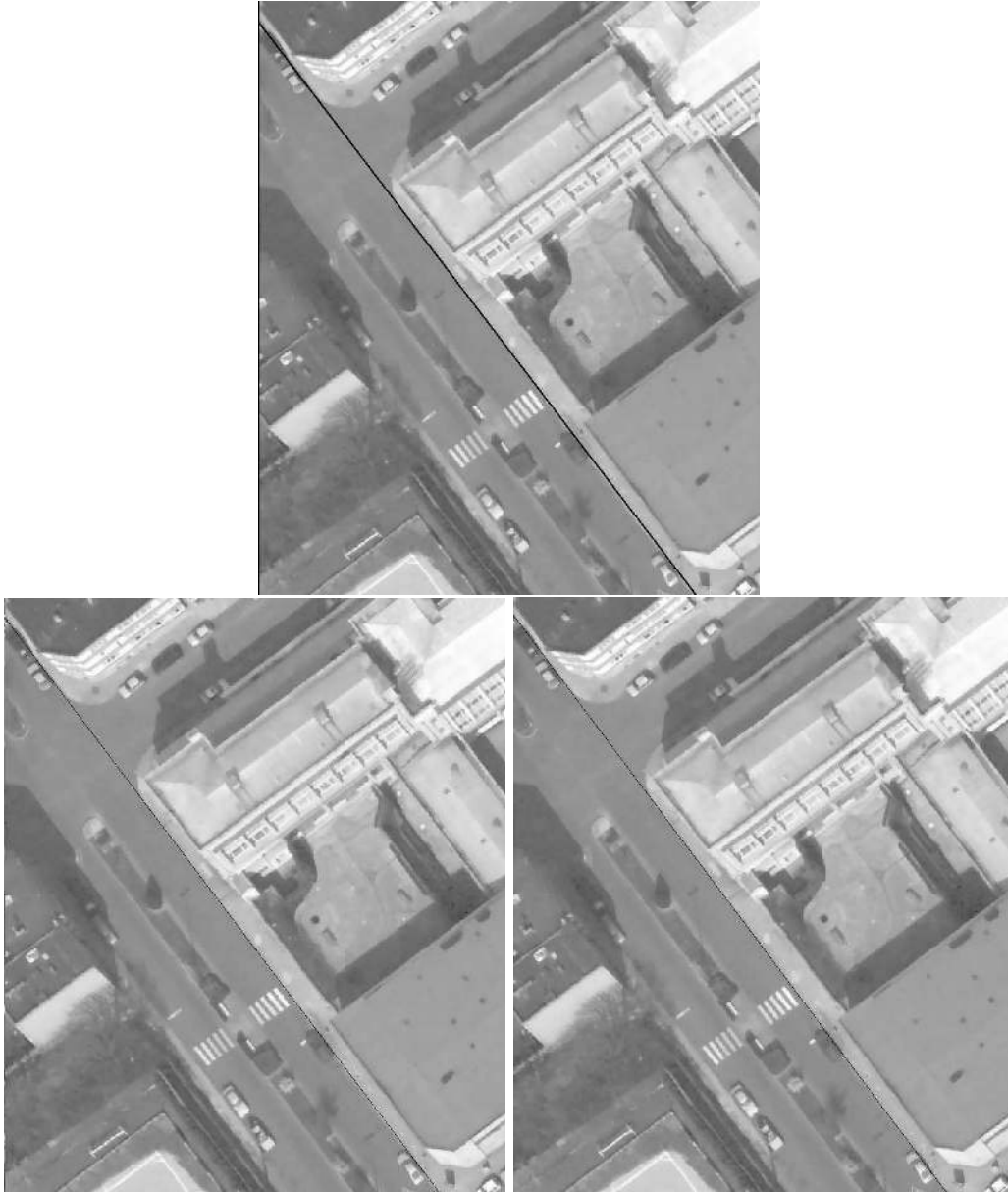
### Acknowledgements

The authors gratefully acknowledge the support by IUAP 4/24 project IMechS, financed by the Belgian OSTC. TT is a Research Assistant of the Fund for Scientific Research – Flanders (FWO-Vlaanderen), MP is a Postdoctoral Fellow of the Flemish Institute for the advancement of Science and Technology in Industry (IWT).

## References

1. Duda, R.O., Hart, P.E.: Use of the Hough transform to detect lines and curves in pictures. *Commun. ACM* **15** (1972) 11–15
2. Hough, P.V.C.: Method and Means for Recognising Complex Patterns. U.S. Pattern No. 3069654 (1962)
3. Illingworth, J., Kittler, J.: A survey of the Hough transform. *CVGIP* **44** (1988) 87–116
4. Leavers, V.F.: Which Hough transform ? *CVGIP* **58** (1993) 250–264
5. Pintsov, D.: Invariant pattern recognition, symmetry, and the Radon transform. *J. of the Optical Soc. of America A*, **6** (1989) 1544–1554
6. Sheinvald, J., Kiryati, N.: On the magic of SLIDE. *Machine Vision and Applications* **9** (1997) 251–261
7. Van Gool, L.: From Fixed Structures to Geometrical Grouping. *Proc. Int. workshop Algebraic Frames for the Perception-Action Cycle* (1997)
8. L. Van Gool, M. Proesmans, T. Moons, *Groups for Grouping*, SPIE Int. Symp. on Optical Science, Appl. of Digital Image Processing XVIII, Vol.2564, pp.402-413, 1995.

This article was processed using the  $\LaTeX$  macro package with LLNCS style



**Fig. 7.** The line detected by the first Hough-transform (top) clearly is not very accurate. One can improve the orientation of this line by giving it the average orientation of that direction (bottom left). However, a better result is obtained if the line is adapted using the vanishing point of that direction (bottom right).



**Fig. 8.** A detail of Fig. 5 with the three lines fitted to the walkway: the Hough line at the right, the average line in the middle and the line through the vanishing point at the left.



**Fig. 9.** a/ Line segments belonging to the same direction as the walkway (top) b/ idem, after adapting the orientation of the lines such that they pass through the vanishing point (bottom).