Snowcat Track Detection in Snowy Environments*

Alberto Broggi, Virginio Cantoni, Ugo Vallone
Dipartimento di Informatica e Sistemistica
Università di Pavia
Pavia, I-27100, Italy
{broggi,cantoni,vallone}@vision.unipv.it

Alessandra Fascioli
Dipartimento di Ingegneria dell'Informazione
Università di Parma
Parma, I-43100, Italy
fascal@ce.unipr.it

Abstract

This paper describes the image processing techniques designed to localize tracks of snowcats for the Italian mission in Antarctica within the ENEA R.A.S. (Surface Antarctic Robot) project for the autonomous driving of intelligent snowcats. The final goal is to enable snowcats to automatically follow preceding ones.

A camera is used to acquire images of the scene in real-time; the image sequence is analyzed by a computer vision system which identifies the tracks and produces a high level description of the scene. This data is forwarded to a further software module in charge of the control of the snowcat movement. A further optional representation, in which markers highlighting the tracks are superimposed onto the acquired image, is transmitted to a human supervisor located off-board.

1 Introduction

In this paper we present the results of a preliminary study for the automatic driving of a snowcat. The main goal of this system is to automate the following of a manually driven vehicle, during goods transportation between two sites in the South Pole; it will be used in the next Italian scientific missions. The first vehicle will be manually driven by an expert driver, while all the others will follow in a train-like fashion. Moreover, since cracks in the ice can put in serious danger both the driver and the snowcat itself, it is imperative that the following vehicles ride on the same precise path defined by the first vehicle. Since even small drifts from the original driving path defined by the human driver can be extremely dangerous, a precise detection of the tracks left by the previous vehicle, a correct measurement of their position, and a smooth control of the actuators must be carefully designed, tested, and evaluated.

*This work was supported by ENEA within the RAS Project.

A preliminary test phase showed that the most promising sensor that should be able to deliver sufficiently precise measurements is a vision sensor (camera). Many other devices have been considered, even active ones since the specific working site would not present any problem due to interference or to environmental pollution [3]. Anyway, vision seems the only sensing capability that may deliver the highest performance in terms of precision of the localization. Data are acquired from a monocular camera installed inside the driving cabin (see figure 1).

Fig. 1. The snowcat and its sensing capabilities.

Due to the extreme conditions of the working environment, where temperatures can reach even -50 degrees Celsius, the terrain is completely covered by snow or ice, strong sun lighting and reflections may be present, and no specific ground references are available nor assumptions can be made on the soil slope, this application is extremely challenging and presents many additional problems with respect to the driving of unmanned vehicles on traditional (un)structured roads [4].

For this reason an extremely careful analysis and design of the processing techniques is mandatory.

Several approaches have been considered due to the low

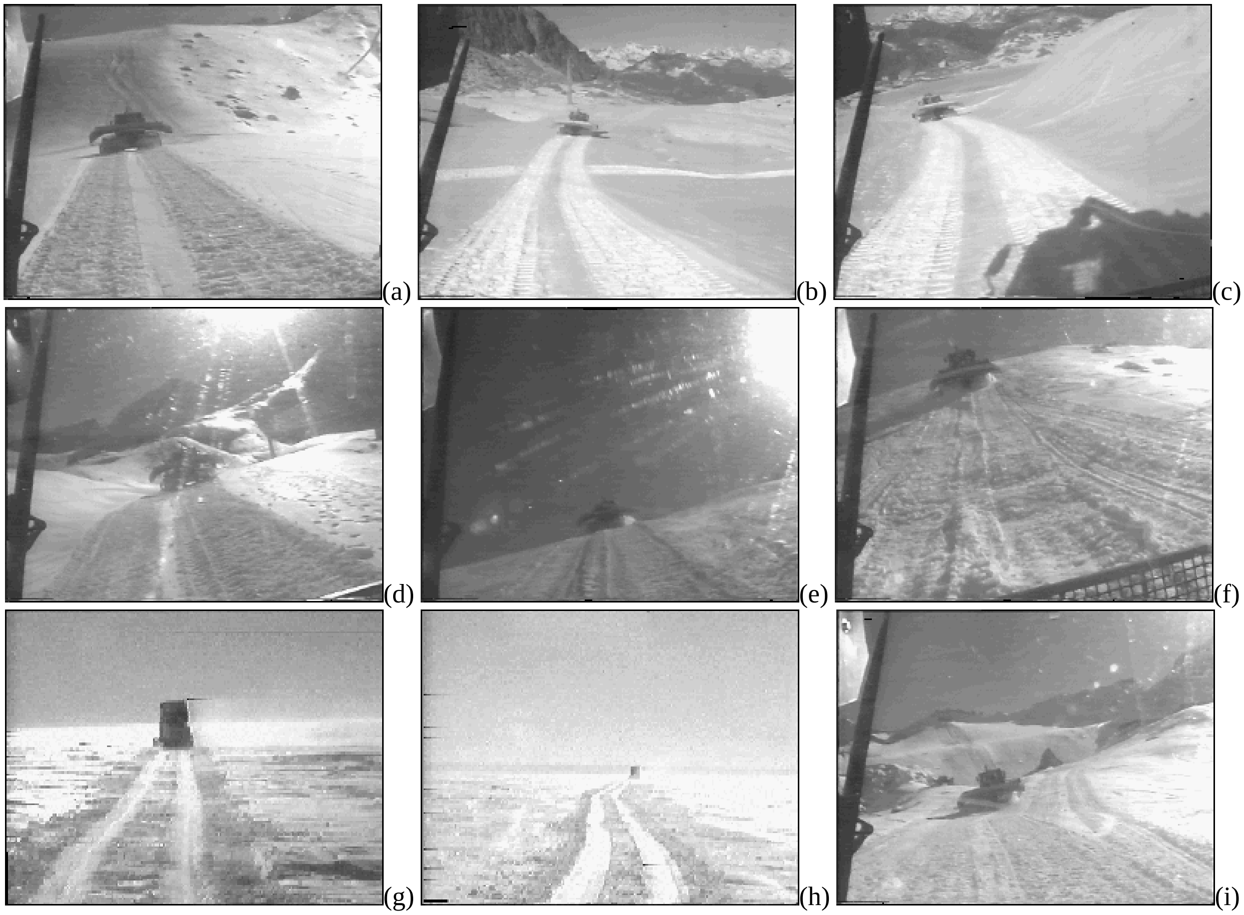

Fig. 2. The many different conditions that must be considered; images come from either the Antarctica region or the Italian Alps

visibility of white tracks on a white background, and specific filters have been developed in order to cope with the typical problems of this environment. The high problem complexity is slightly reduced by the low speed of the vehicle, which permits to focus on the localization of tracks in a reduced close area only.

Moreover, in the automatic driving of road vehicles [1] a special emphasis is generally given to the exploitation of a-priori knowledge in order both to speed-up the computation and make the detection robust. In our case, only a little knowledge about the environmental conditions can be exploited: generally no other vehicle or building is seen by the camera, and the only markings on the ice are due to the preceding vehicle. On the other hand, no assumptions can be made with respect to a possible flatness of the area ahead of the vehicle, nor to a given range of illumination of the scene. In other words, hilly conditions must be considered as well, and therefore the camera orientation generally used in road environments (low towards the road ahead) cannot be replicated here. Besides the acquisition of a large amount of insignificant data during driving in flat areas, the framing of a large portion of the sky can raise another important problem: since in the working site the sun may be very low on the horizon, no specific camera orientation can overcome the problem of direct sun-light into the vision system. This is an extremely difficult issue that must be carefully considered in the development of vision algorithms.

This paper is organized as follows: section 2 discusses the characteristics of the working environment which make the application particularly challenging, section 3 describes the details of the vision algorithm, and finally section 4 illustrates some results and presents future project developments.

2 Environmental Characteristics

The environmental characteristics of the Antarctica region are very challenging and the automatic driving of a vehicle in these conditions is extremely different from traditional

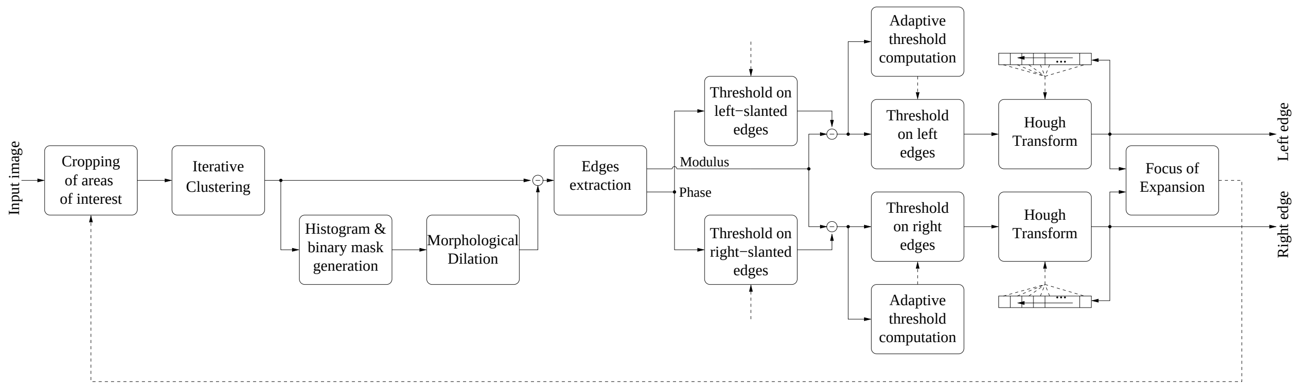

Fig. 3. Block diagram for the complete processing.

highway or urban applications.

The main differences are due to the coverage of the driving area with snow or ice, and the localization of other vehicles' tracks on different kinds of snow or ice requires specific algorithms able to adapt to different scenarios.

As shown in figure 2.a and 2.b, the tracks' characteristics can vary considerably: due to different sun positions, in the first image the tracks are darker than the background, while in the second image the tracks are brighter than the surrounding area. Besides a weak brightness gradient, another invariant that could be exploited is the brightness variance, or –in other words– the texture. Unfortunately, due to the high brightness of the environment, the snow texture provides very weak information. As can be seen in all the images of figure 2 the difference of texture between tracks and background is generally small.

In some cases the shadow of the vehicle itself or of the mountains are captured by the camera (see figure 2.c). Due to the very high contrast of these details, it is impossible to detect weak brightness gradients in the region inside the shadow, which therefore must be eliminated from the analysis. In particular, it is necessary to remove the high brightness gradient generated by shadows, and keep and enhance the weak tracks' edges.

As mentioned, strong sun or light reflections can cause the appearance of reflections patterns in the image, as shown in figure 2.d and figure 2.e. This disturbing effect is also caused by the inevitable presence of small icy particles on the windshield in the region in front of the camera.

No assumptions on terrain slope can be made: in this application domain, no a-priori knowledge on the flatness of the region in front of the vehicle can be used to simplify the localization algorithm. As can be clearly seen from figure 2.e and figure 2.f –acquired with only a few seconds of distance,– the slope can change abruptly, making it difficult even to define an area of interest in the image.

Furthermore, the change in terrain slope can also affect the camera orientation with respect to the sun, and thus can modify the quantity of light acquired from the sensor. This is also visible in figure 2.e and figure 2.f, in which in the former –due to strong sunlight– the snow brightness is lower than in the latter.

The specific traveling conditions may also affect the tracks shape and appearance: in case the ahead vehicle is towing a sledge, the tracks will appear as two compact and uniform stripes surrounded by background with a higher brightness variance, as in figure 2.g and figure 2.h. On the contrary, the tracks shape when no sledge is used are characterized by a higher brightness variance than the background (see all the other images of figure 2).

Finally, no ground references at all can be exploited, as shown by figure 2.g and figure 2.h.

In this first version of the system, problems of divergence from previous tracks as visible in figure 2.i are not considered. Furthermore, in some of the sequences acquired for the first tests, the snowcat was equipped with a shovel –visible in the bottom of figure 2.f– and a windscreen wiper is present in almost all images. Both objects have been filtered out through a specific filter, as discussed in the following section.

3 Tracks Detection

This section presents the complete processing steps for tracks detection; figure 3 sketches the corresponding block diagram.

In order to reduce the complexity of the detection of snowcat tracks in a snowy environment, some assumptions are taken. In the first place, thanks to the low speed of the vehicle, the localization of the tracks in a nearby area suffices for the automatic driving of the vehicle. Secondly, focusing on a close region ahead of the vehicle, the nearest portion

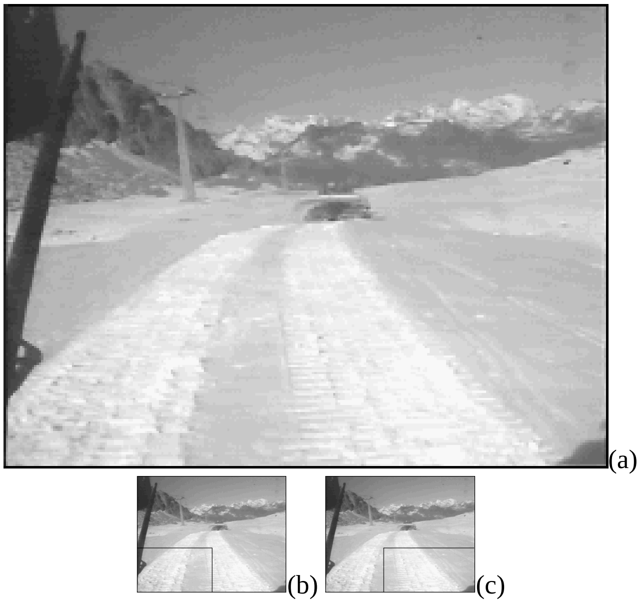

Fig. 4. *(a)* Original image; *(b)* area of interest for left border; *(c)* area of interest for right border.

of the tracks is supposed to be straight and their position is assumed to be slowly varying from frame to frame.

Therefore, for each track border a specific area of interest is defined and analyzed (see figure 4). In these two regions, edges are extracted by means of a classical gradient based approach (Sobel operator), followed by thresholding. In order to reduce the sensitivity to both noise and the threshold value itself, a preliminary clustering is applied as well as a specific filter to mask the presence of shadows and/or dark objects. The two edge images are then used to recover the tracks position by means of the Hough transform.

Since the contrast between the track and the snowy or icy ground is generally low, a clustering algorithm is needed which is able to enhance also weak and isolated intensity discontinuities. An iterative procedure proposed in [2] has been used, which repeatedly substitutes each pixel's brightness with a weighted average computed over its neighborhood. The definition of the weights comprises a function of the neighborhood which enhances sharp edges and preserves weak edges, while averaging uniform areas. Figure 5 shows the result of seven iterations of a 3×3 filter applied to figure 4.a.

Subsequently, an histogram of the pixels intensity is computed in the union of the two search regions, in order to devise a brightness threshold which allows to discriminate between the soil, which is generally bright, and shadows or other dark objects. In this way strong edges deriving from dark objects can be masked out, while leaving weak edges generally representing the tracks position. Figure 6.a shows the gray-level histogram computed in the search areas of the clusterized image: the contribution of dark objects (a small shadow in the right bottom corner and part of the windscreen wiper on the left) can be clearly distinguished from the bright ground. The result of the threshold (a binary image used as a mask) is then dilated with a 5×5 morphological structuring element to enlarge the masked areas. In figure 6.b the result of masking is presented.

To separately detect the two tracks' borders, the gradient based filtering is followed by thresholding the edges' phase so to extract edges belonging to forward slanting oblique

Fig. 5. The result of the iterative clustering. The procedure is actually applied to the areas of interest only; the whole clusterized image is here presented for displaying purposes.

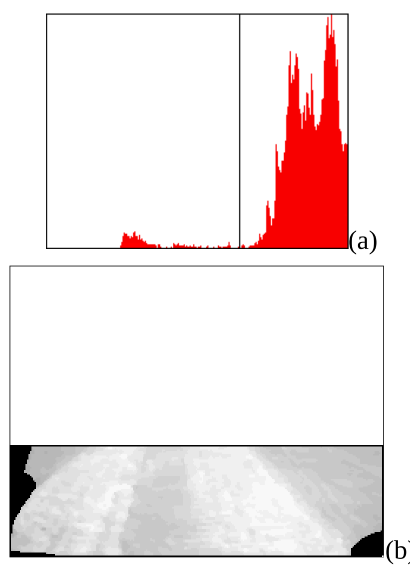

Fig. 6. *(a)* Histogram of gray-level values; *(b)* result of masking.

borders in the left area of interest, and edges belonging to backward slanting oblique borders in the right area of interest. Such filtering has been designed to rely on edges' direction only–and not on a complete 360^o phase– in order to work both with tracks brighter and darker than the ground (see for example figure 2.a and 2.b). Oblique edges point are then filtered with respect to their modulus by means of an adaptive threshold which extracts a fixed number of surviving edges. The value of this parameter was experimentally computed from the analysis of several different sequences. In this way the process is adjusted to the variable contrast between tracks and ground: a constant number of edges is obtained by lowering the threshold when the luminance difference is low and raising the threshold when the contrast is high. The threshold value is easily determined from a cumulative histogram of the gray-level intensity values. Figure 7 shows the edges extracted from figure 6.b in the two different areas of interest.

The Hough transform is then applied to localize the straight line that best fits the edge points of each track border. When selecting the line which gains the highest score, a region centered on the average position of the track in the previous few frames is considered, in order to exploit the strong temporal/spatial correlation.

Once two lines approximating the nearest portion of the track borders have been selected (see figure 8), the focus of expansion (FOE) is determined by computing their intersection. The position of the FOE is compared to the previous ones: if it is too distant from previous results or if it exits from a specific area whose size and position have

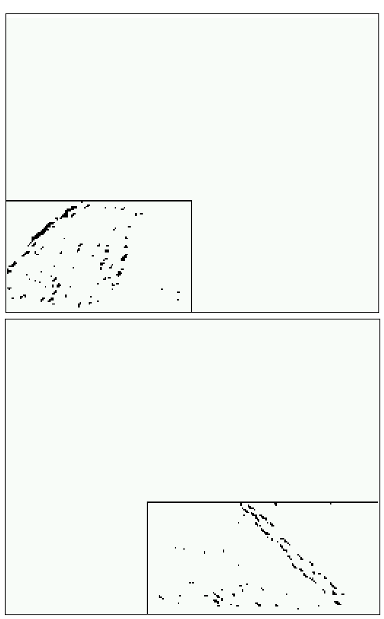

Fig. 7. Edges extracted in the two different areas of interest.

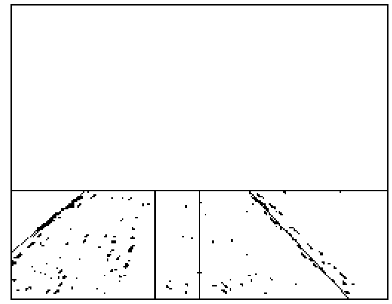

Fig. 8. Straight lines that approximate the track borders.

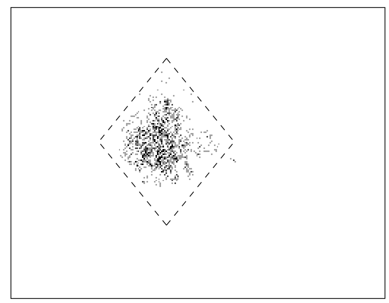

Fig. 9. Image representing the recurrence of the FOE position: the darker the point, the higher the frequency of occurrence in considered sequences; the dashed bounding box represents the area used to validate the final result.

been determined from the analysis of several sequences (see figure 9), the current result is discarded.

Moreover, the two search areas are dynamically resized: their height and width are adapted to an average of the FOE's position in a few previous images; the FOE's position encodes information on the terrain slope and the relative orientation between the vehicle and terrain (see figure 10).

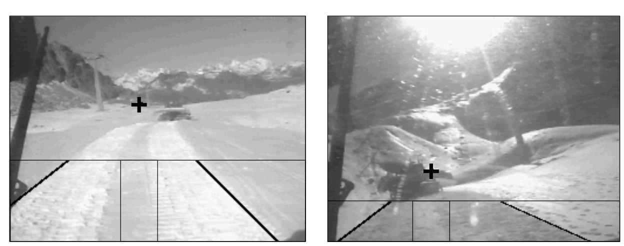

Fig. 10. Different search areas are considered depending on the FOE position, displayed with a black cross.

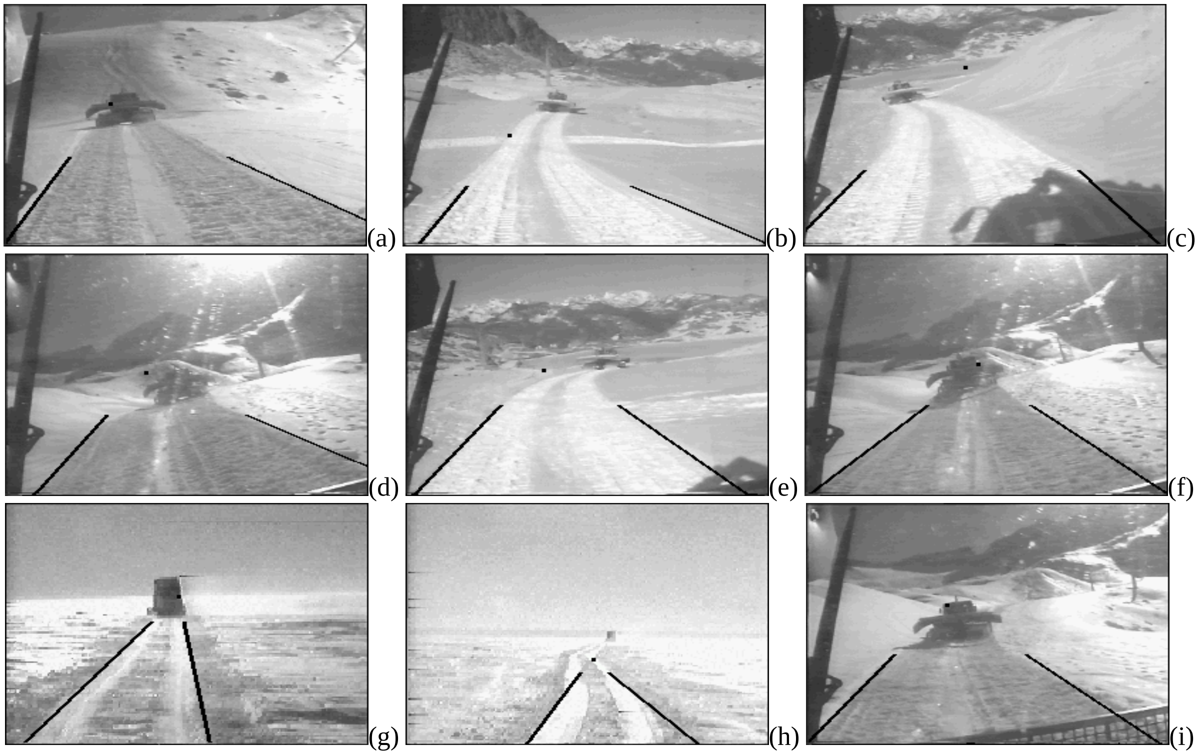

Fig. 11. Results of snowcat track detection in different conditions.

4 Results and Future Research

Figure 11 shows some results of snowcat track detection in different conditions.

Figures 11.a-d and 11.g-h present the result on the corresponding images of figure 2. The remaining images of figure 2 represent very critical situations where the track cannot be distinguished from background with the current algorithm.

Conversely, figures 11.e-f and 11.i illustrate situations where the detection is successful even if noisy or critical conditions such as shadows, sun reflections, unknown terrain slope, and dark objects are present.

Generally the tracks feature a quasi-linear behavior in the region of interest. Anyway it can happen that when approaching a curve, the tracks begin to deviate from this assumption and the Hough transform result is not precise. To cope with this problem, a new method, based on the Hough transform is being developed, which will allow a fine tuning of the current result.

As a further development an extension will be implemented which permits to recover the slope of the path ahead from the FOE's vertical position, when a correct calibration is available. On the other hand, the FOE's horizontal position, together with the measurement of the snowcat's roll, will be used to assess the orientation of the path.

Acknowledgments

We would like to thank Dr. C. Moriconi of ENEA, Italy, and his institution for the support to our research. Our gratitude goes also to the staff of the sky resort of Passo Tonale (Italian Alps) for their help during the tests.

References

[1] M. Bertozzi, A. Broggi, and A. Fascioli. Vision-based Intelligent Vehicles: state of the art and perspectives. *Journal of Robotics and Autonomous Systems*, 32(1):1–16, June 2000.

[2] A. Broggi. Parallel and Local Feature Extraction: a Real-Time Approach to Road Boundary Detection. *IEEE Transactions on Image Processing*, 4(2):217–223, February 1995.

[3] A. Broggi, M. Bertozzi, and A. Fascioli. Requirements for Visual Perception of Automotive Environments. In G. L. Foresti, P. Mähönen, and C. S. Regazzoni, editors, *Multimedia Video-Based Surveillance Systems: Requirements, Issues and Solutions*, volume 573 of *the Kluwer international series in engineering and computer science*, chapter 1.3. Kluwer Academic Publishers, Boston, USA, Oct. 2000. ISBN 0-7923-7927-6.

[4] A. Broggi and E. D. Dickmanns. Applications of computer vision to intelligent vehicles. *Image and Vision Computing Journal*, 18(5):365–366, Apr. 2000.