

Current Work in Multimedia Imaging at UoA's Tamaki Campus

Reinhard Klette, Je Ahn, Haokun Geng, Ralf Haeusler, Simon Herman, Waqar Khan, Sathiamoorthy Manoharan, Sandino Morales, John Morris, Radu Nicolescu, FeiXiang Ren, James Russell, Konstantin Schauwecker, Crystal Valente, Xi Yang
CITR and School of Biological Sciences, Tamaki Campus, The University of Auckland, Auckland, New Zealand

Abstract—This report informs about current activities and results in the *.enpeda..* (short for ‘environment perception and driver assistance’) project and related performance evaluation studies, in panoramic visualization, in environmental surveillance based on scanned footprints of small species, in artistic filters, and in the design of efficient geometric algorithms for areas related to 2D or 3D imaging or robotics. The report summarizes some of the current work in multimedia imaging at Tamaki campus; see [52] for a previous report and further areas of research.

Keywords: Multimedia imaging, Tamaki campus, vision-based driver assistance, performance evaluation, panoramic visualization, footprints, artistic filters, efficient geometric algorithms

I. INTRODUCTION

The report summarizes some of the current work in multimedia imaging at Tamaki campus. This report informs about current activities and results in the *.enpeda..* project (with its current focus on vision-based driver assistance), in performance evaluation, in panoramic visualization, in environmental surveillance based on scanned footprints of small species, in artistic filters, and in the design of efficient geometric algorithms, the latter subject for areas related to 2D or 3D imaging, gene research or robotics. See [52] for a 2009 report and further areas of research, not mentioned in this report.

II. VISION-BASED DRIVER ASSISTANCE

The Environment Perception and Driver Assistance (*.enpeda..*) project searches for solutions for *driver assistance systems* (DAS), with a particular focus on the use of various sensors in the *ego-vehicle* (i.e., the vehicle where the considered system operates in) such as cameras or laser-range finders, or modern navigation technologies [35], [39]. We review a few of the current activities in this project. Our main partner in this research is the group led by *Uwe Franke* at Daimler A.G., Germany.

A. Vehicle Tracking and Location Prediction

GPS navigators track vehicle location in real-time [47]. A related problem is tracking a vehicle’s path after the travel. This often is a requirement for fleet management and also for ensuring correct payment of road user

charges. Such path prediction uses data stored during the vehicle’s travel. The problem then is to ensure storing a minimum amount of data, while maximizing the path prediction accuracy: storing large volumes of data will enable better path prediction accuracy while sparse data collection and storage may reduce the prediction accuracy. In addition, processing large volumes of data and/or taking into account a large number of factors may reduce the performance of path prediction. Consequently, an approach that minimizes the storage and data requirements while not compromising path prediction accuracy is required.

There are three steps to location tracking: data recording, matching a stored location to a road, and linking the matched points to form the vehicle track.

Recording location data can be time-driven (i.e., taken at regular intervals) or event-driven (i.e., taken at key events such as ‘turn’ or ‘speed change’). A time-driven approach can lead to large data storage if the interval is too small; it could miss important location information if the interval is too large. Data recording can also be prediction-based: if the next position of the vehicle can be predicted (with some predefined accuracy) using the data already in store, then the data for the next position need not be stored [8], [9].

A map-matching algorithm is used to map a stored location onto a road. It contains two steps: firstly finding the road where the vehicle is, secondly finding the position of the vehicle in the road. There are three types of map-matching algorithms: geometrical, statistical and fuzzy logic [7], [41]. Often due to GPS inaccuracies the stored location can be incorrect. The map matching algorithms take the inaccuracies into account when mapping the location data onto a road.

The map-matched locations (or points) are then linked to form the final vehicle track.

As part of the *.enpeda..* project, we use improved techniques for data storage, map-matching, and linking [19]. We use a hybrid location storage method which uses both time and vehicle direction change as thresholds to decide when to store vehicle information; and a bearing-distance hybrid map-matching algorithm, which maps the GPS location data onto the road network using both heading difference and the distance between the

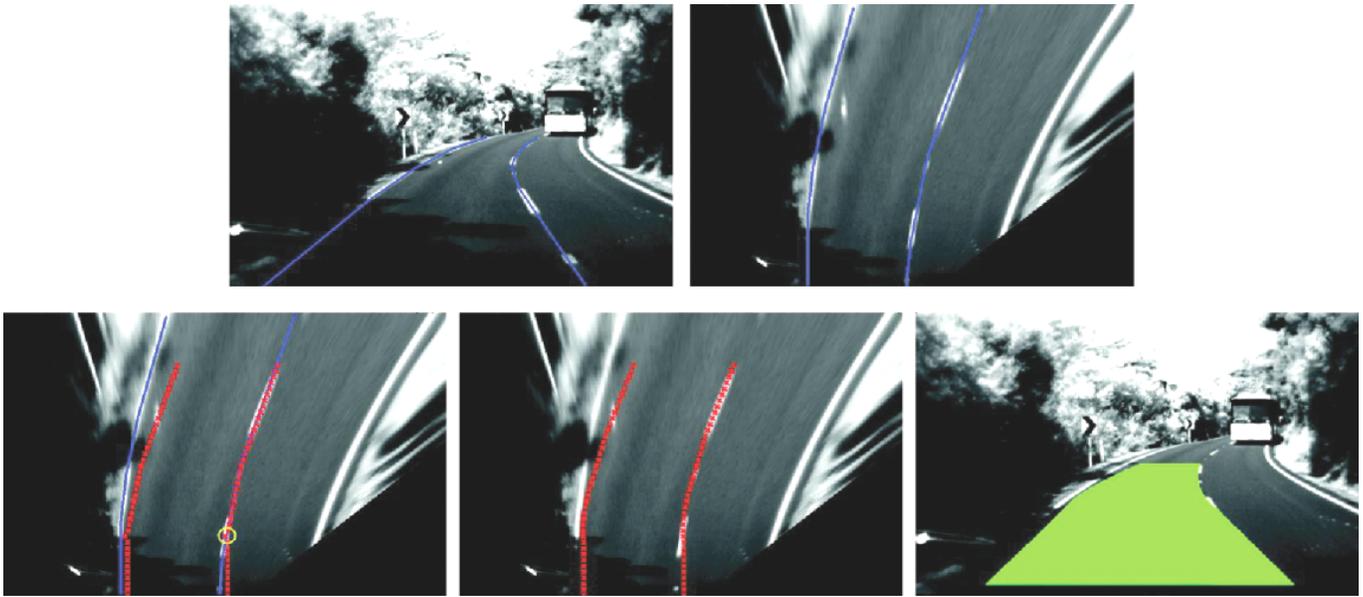


Fig. 1. Lane border and corridor detection. (a) Input image with marked lane borders (in blue). (b) Bird's-eye view. (c) Corridor search in the bird's-eye view (in red). Feature points of the corridor are labeled by a circle (in yellow). (d) Corridor smoothing using a sliding mean. (e) Corridor detection result as a road patch in the input image (in green). See [31].

recorded vehicle location and its nearby roads. Our linking map-matched point algorithm finds all the vehicle paths between the start and the end points, and validates the most likely vehicle path by using total traveled distance, total direction change, vehicle speed and road network topology.

Initial experimental evidence clearly shows that these techniques improve the vehicle tracking system accuracy and efficiency. Furthermore, these methods give reasonable estimations of the most likely vehicle path while using minimal data storage.

Computer vision technologies (e.g., automated lane detection and positioning of the ego-vehicle on the road; see [31], [58]) define a new possibility. Figure 1 illustrates automatically detected lane borders and also the detected *corridor*, which also takes the expected trajectory of the ego-vehicle into account.

B. Mobile DAS

A robust and efficient lane detection system is an essential component of vision-based DAS. Various computation platforms have been proposed in the past few years for the implementation of DAS (e.g., PC, laptop, integrated chips, PlayStation, and so on). In [59] the authors proposed a new platform for the implementation of lane detection, which is based on a mobile phone (in this paper actually the iPhone, but current work [58] was also aiming at various mobile phones).

Mobile DAS is a system installed in mobile phone which aims at helping drivers and improving road safety. It uses the single camera on a cell phone as the main sensor for its mobility and low cost. Our concept

is to develop a light-weight computer vision-based DAS on the mobile phone. Applying real-time sampling and processing, *Mobile DAS* can help to detect lane marks, front vehicle distances and traffic signs on the road in real driving situations. This kind of information can be used to assist a driver, such as giving alarm when approaching an obstacle (e.g., a pedestrian, other vehicles, or other objects on the road), or simply providing information (e.g., the current speed limit zone) for the driver's attention. The three most useful subsystems of DAS are currently: Collision Avoidance System (CAS), Lane Departure Warning System (LDWS) and Traffic Sign Recognition System (TSRS). CAS helps to prevent a collision of the ego-vehicle, LDWS is designed to warn a driver when the ego-vehicle begins to move out of its current lane on the road, and TSRS is used to detect and recognize a traffic sign and interpret its meaning.

Due to physical limitations of mobile phones w.r.t. memory and computing power, a simple and efficient lane detection algorithm using a Hough transform was developed and implemented on the iPhone, as existing algorithms developed based on the PC platform are not suitable for mobile phone devices (currently). This algorithm is based on the road-camera model as discussed in [46]. Experiments in Auckland (see Fig. 2) showed that the rate of successful lane detection is above 90 % with clearly detected straight lane borders.

C. Accuracy of Stereo Matching

Current stereo algorithms are capable of generating dense disparity maps for large sized images (larger than the common 640×480 pixel resolution) at more



Fig. 2. Experimental results of lane detection on an iPhone at night. See [59].

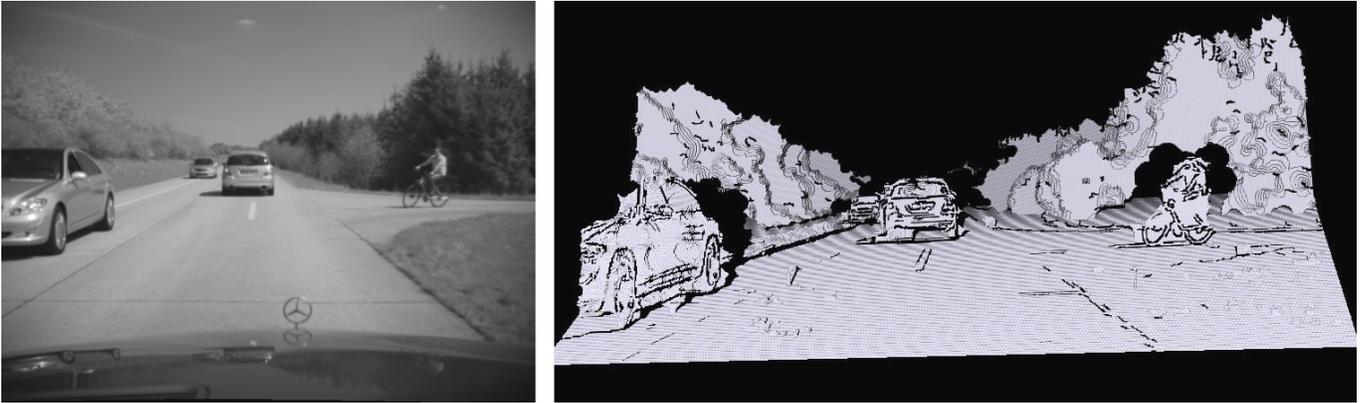


Fig. 3. *Left:* Original input image. *Right:* Disparity depth map image.

than 20 frames per second [17], [32]. The estimation of distances to objects, present in a given traffic scene, together with the detected motion, is of basic importance for understanding dynamic environments that surround the ego-vehicle. For vision-based DAS, the accuracy of stereo algorithms has a crucial influence on the overall performance.

Most stereo algorithms aim at the minimization of an energy E that consists of a *data* and a *smoothness* term. The data term incorporates a matching cost into the energy E , that usually relies entirely on a correspondence measure based on pixel information within a certain neighborhood of tested pixels. The smoothness term adds additional cost in order to solve classic stereo matching problems such as the correct labeling of pixels in homogenous areas, at depth discontinuities or at occlusions. In those areas a data term is not sufficient, due to its locality. Optimization strategies that tackle those problems by incorporating a smoothness term can be categorized into three groups: belief propagation [15], graph-cuts [5], and dynamic programming [53], the latter one also extended to multiple scanlines, and then known as semi-global matching [26]. (For a modification of semi-global matching, see [25].)

There are also further problems that affect the data term and are usually solved by low-level processing [74],

[73]. Illumination differences between images in a stereo pair, or noise obtained during image acquisition have a major influence on the data term itself. Recent studies suggest to decompose input images into a structure and a texture component. The texture component tends to be robust against illumination changes. Additionally, bilateral filtering is used to reduce the noise in an image, since it applies a Gaussian smoothing along with a spatial constraint that preserves image discontinuities.

The major requirements for a stereo analysis algorithm for DAS can be summarized as follows:

- It needs to be robust against sudden changes in the input data (illumination changes, noise, and so forth). This means that a stereo algorithm needs to incorporate a reliable and robust data term to support any downstream optimization step.
- It needs to be depth discontinuity preserving (see Fig. 3). This becomes evident if you think for example of a car passing closely another car (or bicycle). DAS are designed to be able to decide whether a collision is likely to occur. With cars usually passing within 1- 2 meters distance, even marginal blurs at depth discontinuities can affect the decision of an assistance system.
- It needs to run in real-time. This means for current computer hardware that a designed algorithm po-

tentially allows for parallelization in order to run on multiple processors, or to employ the SIMD concept.

D. Belief Propagation Adaptations

The belief propagation (BP) stereo matching algorithm is based on minimizing an energy function that measures the goodness of a disparity assignment. The energy function is composed of two terms, which are a *data term* and a *discontinuity term*,

$$E(f) = \underbrace{\sum_{p \in P} D_p(f_p)(f)}_{\text{Data Term}} + \lambda \underbrace{\sum_{(p,q) \in A} V(f_p - f_q)}_{\text{Discontinuity Term}} \quad (1)$$

where P is the set of all pixels in the left image, D_p a dissimilarity measure, f a disparity assignment, λ a weighting factor which controls the influence of both terms, A the set of adjacent pixels and V a discontinuity-cost function.

The data term penalizes the matching of pixels that are dissimilar, while the discontinuity term penalizes the assignment of different disparities to adjacent pixels. We have evaluated possible improvements of the BP algorithm, which are based on modifications to the data and discontinuity term [64].

The intensity difference is commonly used as data term for BP, which works well for synthetic images, where corresponding pixels always have similar intensities. In real-world stereo images, significant intensity differences can occur, caused by different gain and bias settings of the cameras (other noise sources also contribute). To overcome this problem, a Sobel operator was used in [20] to calculate edge images before processing a stereo pair with BP, which lead to improved results. The usage of other edge detection algorithms has been studied in [21], but the Sobel operator was found to be the most suitable there and also in [79].

A Sobel edge image displays the gradient magnitude, but disregards the gradient direction. To utilize this additional information we created a new data term for BP. The first step in our algorithm is to calculate the gradient maps G_L and G_R for the left and right image. We then define a new dissimilarity measure D_e as the *end point error* of the left and right gradient vector, which is the absolute value of the difference of both vectors:

$$D_e(f_p) = \|G_L(x, y) - G_R(x - l, y)\| \quad (2)$$

With this measure, vectors pointing in opposite directions will have a large cost, as will vectors with very different lengths. We tested a BP algorithm with the new data term on the Middlebury dataset [49] and received an average percentage of bad pixels of 10.8%, which is better than the version relying on intensity differences, for which we obtain 12.1%. If we use the intensity difference on edge images, as done in [21], we receive 11.6% of bad pixels, which is still more than with

the method we propose. For all tested algorithms, we automatically tuned the available parameters and chose the combination with the best performance.

The above results were all gained on synthetic stereo pairs. However, the real strength of the new data term is that it can cope better with large intensity differences, which we expect to observe in real world sequences. We have performed tests with the new algorithm on real world sequences and found that it does provide better results. However, without ground truths it is hard to quantify the results and tune the algorithms in a way that ensures a fair comparison.

We further evaluated the usage of a new discontinuity model. BP can be improved if additional knowledge about the likelihood of a discontinuity at an image pixel is available. In this case, an adaptive discontinuity model can be created that assigns lower costs to pixels with a high discontinuity-likelihood.

We have evaluated two measures for estimating this likelihood. The first one is based on the absolute value of the image gradient G , while the second one uses the local variance V . From both measures we compute a scaling factor, which we then multiply with the previous discontinuity model. We calculate this scaling factor with the equations

$$s_G = \frac{1}{1 + w_G \cdot \|G\|} \quad (3)$$

$$s_V = \frac{1}{1 + w_V \cdot V} \quad (4)$$

The quantities $\|G\|$ and V appear in the denominator of Equation (3) and (4), as they are both anti-proportional to the likeliness of a discontinuity. We multiply both with a weight w_G and w_V to control the adaptiveness of the new model, and add a value of +1 to limit the factor.

We have tuned and tested both approaches in conjunction with our previous improvements on the Middlebury dataset [49]. Our results show that both methods lead to an increased quality of the calculated disparity map. The gradient-based approach performed slightly better with an average bad pixel percentage of 9.6%, while the local variance based measure achieved an average percentage of 9.8%. Figure 4 shows an example of the results gained with the new method.

The performance of the new algorithm on synthetic driving sequences from the EISATS website [14] was, however, worse than with the non-adaptive discontinuity model. We assume that this contradicting behavior is caused by the much fewer discontinuities in the driving sequences, which causes the new discontinuity model to be of little help.

III. PERFORMANCE EVALUATION

It becomes more and more necessary to evaluate the performance of stereo algorithms for various *situations* of on-road traffic, not on short sequences just showing

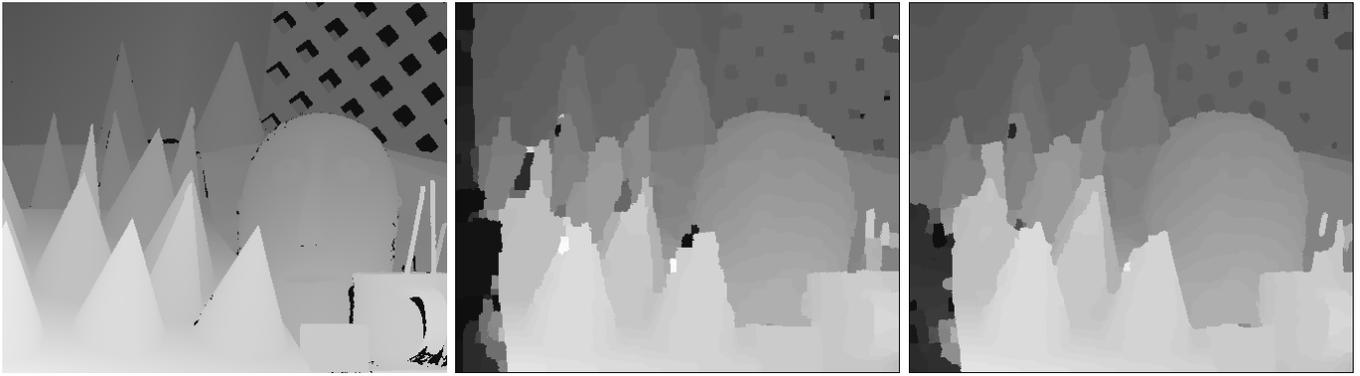


Fig. 4. Comparison of BP algorithms. *Left*: Ground truth. *Middle*: BP with intensity based data term. *Right*: Gradient-based data term and adaptive discontinuity model.

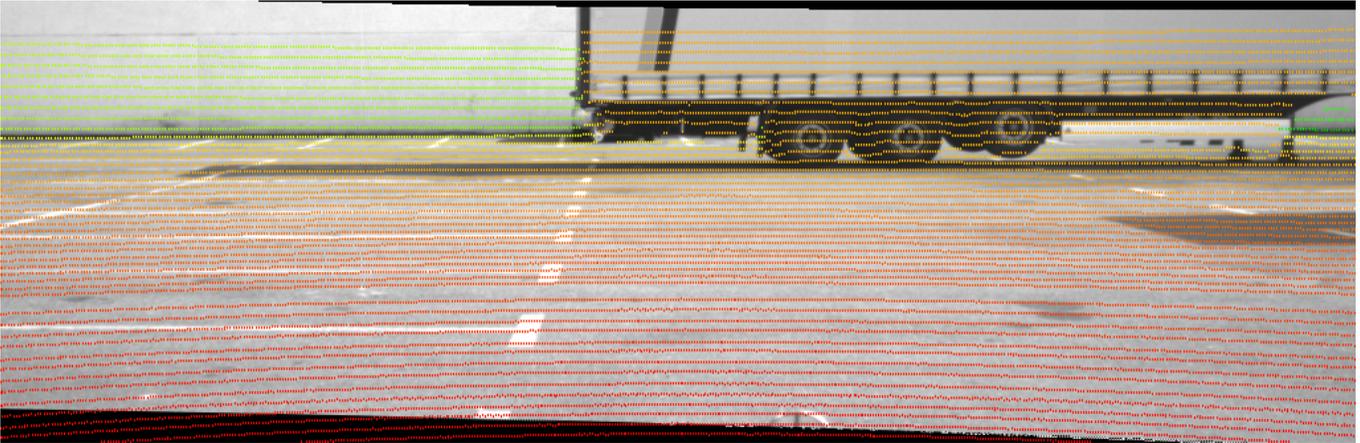


Fig. 5. Laser range-finder data shown as overlay in corresponding image data. Ground truth points (i.e., points acquired with the laser range-finder) are color encoded from red (for close) to green (for further away).

a few seconds of driving, but in tests going for hours or days. This requires to use extensive data sets that are generated under the same (or at least, very similar) conditions as the stereo algorithm will face in day by day routine application.

A. Stereo Ground Truth by a Laser Range-Finder

One option for organizing performance evaluation on long image sequences in the real world is to use a laser range-finder to generate ground truth distance data that can be used to evaluate the performance of stereo algorithms in real-world scenes.¹

From the distance data acquired with a laser range-finder, we generate a (sparse) *ground truth depth map* that is used to evaluate the obtained disparity maps; see Figure 5. By calibrating the coordinate systems of range-finder and stereo camera, both used in the same

ego-vehicle, we are able to obtain “fairly accurate” range readings at about 10% (for a 1320×334 pixel image) of points visible in the reference camera of the stereo set up. For those points we perform a direct comparison between those distance values and the values in a calculated disparity map. For evaluating the remaining points in the disparity map, we use the three nearest points in the ground truth image for generating a patch, whose geometry is compared with the corresponding depth data in that triangular region in the disparity map by using some *confidence measure*.

For generating ground truth data, we calibrate the range-finder and the stereo camera coordinates systems [28]. Then, we generate for the range data an image of the same dimensions as that for the reference camera by projecting the three-dimensional (3D) points acquired by the range-finder that are in the field of view of the reference camera system. For example, for a 1320×334 pixel image, there are range measures available for around 20,000 pixels (e.g., the horizontal resolution of the range-finder is of 11 points per degree, while there are 22 pixels per degree in the cameras used for the experiments). For

¹The used high definition laser range-finder is the Velodyne HDL64E S2 range-finder [75], a rotating device that gets measures from the entire 360° surrounding environment. With an accuracy of 2 cm in a distance range between 1 to 120 meters, and a vertical and horizontal resolution of 0.4° and 0.09° , respectively, it is an adequate tool for evaluating stereo algorithms (and possible other vision-based procedures) in the context of DAS.

all the other points in the ground truth image, we assign a distinctive negative value identifying unavailable data. Now we apply for evaluation one of the two following approaches:

Direct Comparison: Where range data is available we use the percentage of badly calculated pixels (BCP) and the root mean squared error (RMS) as quality metrics [51], [63].

Confidence Measure: To complement the direct comparison (i.e., to evaluate points also “within the triangles” where no range data is available), we decided so far for a confidence measure which follows the approach presented in [12].

Further versions of confidence measures need to be studied, and we are also in particularly interested to extend the current focus on static scenes also on dynamic scenes.

B. Estimated Ground Truth for Simple Road Geometries

Regarding the evaluation of correspondence algorithms for stereo or motion analysis, [46] proposed a way to estimate ground truth when the ego-vehicle is driving in a “geometrically simple” environment.

For example, assume that the ego-vehicle is driving on a planar surface towards a wall. The optical flow $\mathbf{u} = (u, v)$ is then approximately defined by

$$u = \left(\frac{S_t}{S_{t+\delta t}} - 1\right)\left(i - \frac{W}{2}\right) \quad \text{and} \quad (5)$$

$$v = \left(\frac{S_t}{S_{t+\delta t}} - 1\right)\left(\frac{H}{2} - j\right) \quad (6)$$

where S_t and $S_{t+\delta t}$ are the distances between camera (in the ego-vehicle) and wall at time slots t and $t + \delta t$. These two distances can not be measured accurately for δt equals 1/30 of a second because we do not know the exact speed of the vehicle at the required level of accuracy.

However, for estimating the value of $S_t/S_{t+\delta t}$, we can run one of the optic flow algorithms (say, an improved TV- L_1 [78]) on two consecutive image frames first. Using all the calculated values u and v , we estimate $S_t/S_{t+\delta t}$. Finally, we use the estimated ratio $S_t/S_{t+\delta t}$ to have estimated ground truth; AE and EPE are then calculated with respect to those vectors.

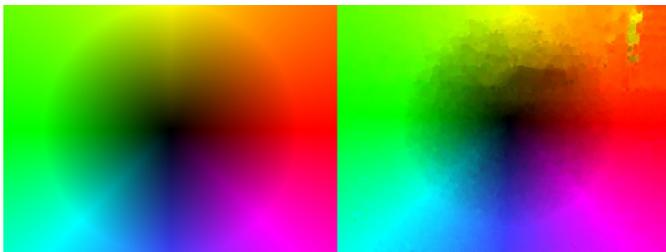


Fig. 6. *Left:* estimated ground truth when driving towards to a wall. *Right:* calculated optical flow using TV- L_1 . See [79].

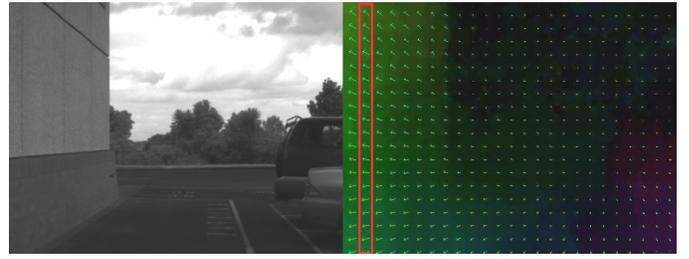


Fig. 7. *Left:* a recorded frame while driving parallel to a wall. *Right:* calculated optical flow using TV- L_1 . See [79].

Figure 6 shows a sample of a TV- L_1 result together with the estimated ground truth flow. Driving parallel to a wall (see Fig. 7), through a tunnel, or on a planar surface are further examples where ground truth may be estimated.

C. Safety of Stereo Driver Assistance Systems

Due to the safety reasons, it is currently not possible to get ground truth data for a collision scenario. In order to validate the performance of a stereo configuration for a safety system like a driver assistance system, we form a model which replicates its constraints like object’s trajectory estimation time, driver response time [10], and the braking time [16] - before the collision. The accuracy of trajectory estimate would have improved significantly if stereo could detect a change in approaching object’s position at each processed frame, but due to the integral disparities stereo could only detect these changes at discrete steps. So basically there is always an error affiliated with the distance being measured by stereo [33]. This error reduces with improved depth resolution with the object coming closer to the ego-vehicle, but then it might be too close for the system to warn the driver about it. Depth resolution can be improved at farther distances, but then the extent of stereo common field of view reduces. This leads us to an interesting study about analyzing the effectiveness of a stereo parameters on the accuracy of trajectory estimation of object,

- located anywhere in front of us,
- moving with some speed to collide with us.

An example of a collision scenario is shown in Fig. 8. Initially there is large error in the estimated trajectory, and with each processed frame the error reduces but when the object is perceived at a different depth, the error reduces significantly.

D. Evaluation of Test Data

Current research in stereo image analysis focuses on improving matching algorithms in terms of accuracy, computational costs, and robustness towards real-time applicability for complex image data and 3D scenes. Interestingly, performance testing takes place for a huge number of algorithms, but, typically, on very small sets

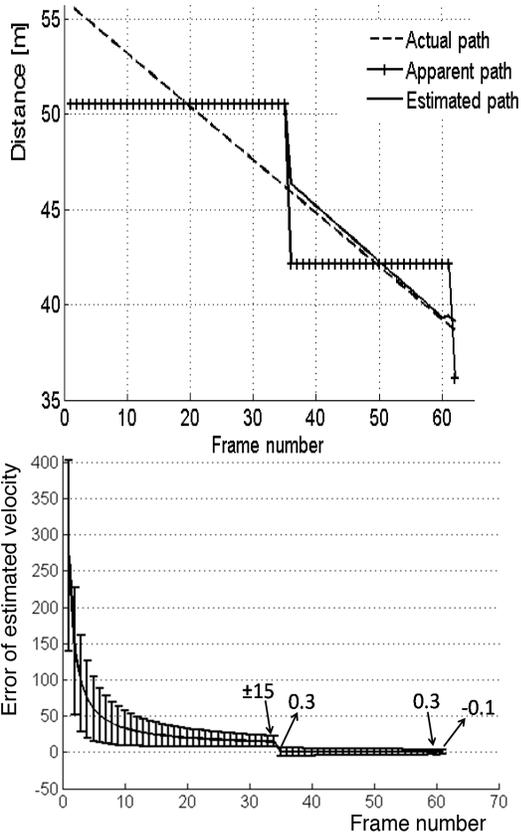


Fig. 8. Accuracy of the trajectory estimate over time. Object is initially located at $(X, Z) = (19, 56)$ meters and moving with a speed -2 m.s^{-1} perpendicular to our vehicle's path, which is also moving with a speed 7 m.s^{-1} . Stereo is assumed to be in a canonical configuration with parameters as: focal length = 5×10^{-3} meters, baseline length = 308×10^{-3} meters, camera resolution $W \times H = 640 \times 480$, and an object detection rate of 25 frames per second. *Top*: Comparison of the actual object path versus the apparent path perceived by stereo and the one estimated by our model. *Bottom*: Error in the estimated speed over time.

of image data only. Even worse, there is little reasoning whether data as commonly applied is actually suitable to prove robustness or even correctness of a particular algorithm.

Performance evaluation of stereo algorithms became increasingly popular since the availability of various test sites such as [63] at Middlebury College. Such evaluations were speeding up progress in the design of stereo matching algorithms. Ranking is typically done by comparing a few error measures, calculated with respect to given ground truth and a relatively small number of images. Evaluations lead to particular insights, for example about the role of used cost functions [26], or of image preprocessing methods [74].

Stereo image data, depending on recorded scenes, sensor quality and so forth, can be of very different characteristics and origin (e.g., synthetic, controlled indoor, real-world outdoor). The question arises: Given a stereo image pair, what is the minimum error we may expect?

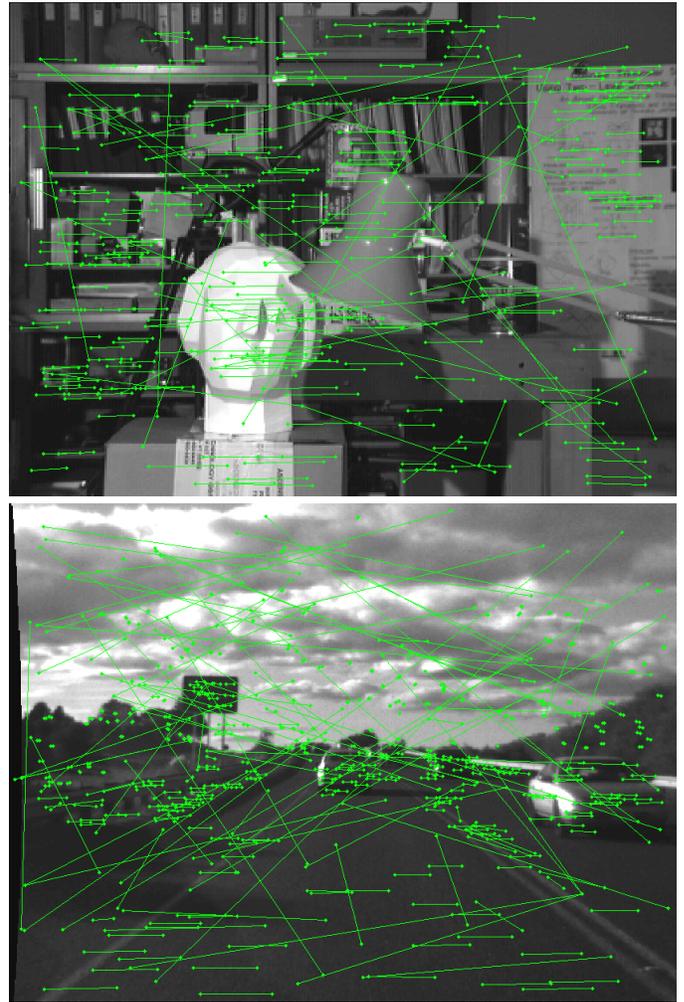


Fig. 9. Illustration of sparse stereo matching with SIFT-features (not constrained by epipolar geometry, but on rectified images) applied to stereo pairs of different characteristics.

We try to answer this question for a wide range of different types of stereo image data, ultimately allowing to quantify this material in terms of quality. However, for the most interesting scenarios – outdoor real-world, highly dynamic and complex scenes with potentially very poor image quality – the common evaluation approach of stereo matching techniques is not feasible due to the lack of ground truth.

Previous work [42], [50] that does not require ground truth needs at least three time-synchronous views of a scene. We develop alternative approaches that only need binocular imagery.

To a human viewer, it is in general obvious whether a stereo pair is of good quality for extracting depth information, or not. For example, stereo photos taken under insufficient lighting conditions (such as outdoors during the night), very high contrast images with poor texturing or stereo pairs with contrast differing with a factor of more than two between left and right image

cannot be matched properly by the human visual system. Similarly, semi-occluded objects or vertical parallax lead to retinal rivalry and therefore to strong eyestrain.

The construction of SIFT descriptors is inspired by the functioning of the primate V1 cortical neurons. Such biological models have been successfully applied to the task of object recognition [13]. We performed experiments which compared our proposed SIFT-based complexity measures [22] with the prediction error analysis for stereo matching on trinocular sequences [51].

We envision four major benefits of assessing stereo image data independently from geometric ground truth. First, it can guide the selection of applied methods as already mentioned above. Second, it may make processing of real-world stereo images more tractable by providing an additional measure of confidence. Third, we can identify “problematic” situations in real-time; this gives a chance to identify unexpected problems when doing an on-line stereo analysis of real-world stereo image sequences, and to be aware of those when further improving stereo matching. Fourth, it may advance theoretical knowledge about stereo matching by implementing performance evaluation on sophisticated synthetic scenes (i.e., using progress in physics-based rendering) and showing its conclusiveness regarding relevance to real-world scenarios.

IV. WEB-SITE FOR MULTIMEDIA IMAGING

The multimedia imaging group at Tamaki campus runs its own website at www.mi.auckland.ac.nz.

A. EISATS Database

Testing computer vision techniques on extensive and varying data sets helps to avoid a bias which occurs when using only selective (e.g., ‘small’) sets of data.

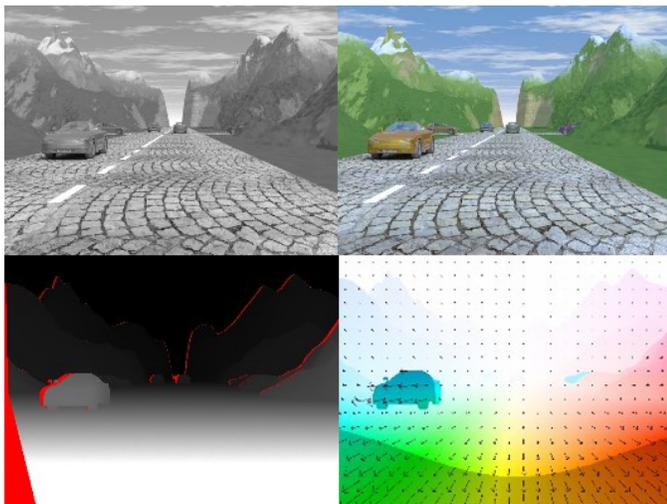


Fig. 10. *Top*: frames of Sequence 1 of Set 2 on [14]. *Bottom*: stereo and motion ground truth.

According to our experience, recorded or synthesized video data may be segmented into subsequences of about 100 to 200 frames (or 4 to 8 seconds of recording, assuming 25 Hz as a current standard) for representing one particular *situation*, defined by a co-occurrence of some *events* in traffic scenes. Examples of events are activities of adjacent traffic (overtaking, oncoming traffic, crossing pedestrians, and so forth), weather and lighting conditions (rain, sun strike, patterns of shadow while driving below trees, and so forth), road geometries (flat or curved, narrow lane, entering a tunnel, driving on a bridge, a speed bump, and so forth), or particular events such as traffic signs, a wet road surface or strong light reflections at night. A situation is defined by a concurrent appearance of such events, such as “driving in daylight on a planar road while overtaking a truck”. We believe that these examples are sufficient for defining the concept of a *situation*. Situations typically change every few seconds in normal traffic, and we consider 4 to 8 seconds as the standard length of a recorded video sequence (also called a *basic sequence*) which may be identified with one particular situation.

Data relevant for DAS applications have basically an unlimited range of variations (‘expect the unexpected’), due to the potential range of events, and thus of their combinations into situations. Selective (‘small’) sets of test data, say with a focus on rendered or engineered (good lighting, indoor) scenes, are insufficient for serious testing. The *.enpeda..* (environment perception and driver assistance) Image Sequence Analysis Test Site (EISATS), see [14], is not (!) focused on one particular set of data or one particular evaluation strategy, but open to researchers in vision-based DAS for applying those data in their evaluations, as well as also for contributing more (best: verified) data. The website contains recently six different sets of test sequences, provided by different research groups in vision-based DAS, and of relevance for particular evaluation strategies.

We see the EISATS data base as a dynamic forum for relevant data and benchmarks for vision-based DAS. See Tab. I for a brief characterization of available sets of image sequences. Figure 11 illustrates an application for one of the trinocular sequences. Each of the EISATS image sequences represents a few seconds of driving (i.e., typically showing one situation). Of course, these sequences are still only representing a very small segment of possible situations in vision-based DAS.

B. Web-Based Visualization of Videos

The *.enpeda..* research domain has special needs for video playback and delivery. This is because researchers need to compare several videos, and the ability to play videos concurrently will provide valuable visual comparison. In addition to the original videos as captured by cameras, one needs to play processed videos that correspond to the originals. For instance, a processed

TABLE I
DATA SETS OFFERED ON THE EISATS WEBSITE IN JULY 2010. SEE [36].

Set	Comments
1	Night vision stereo sequences (Daimler AG) These seven stereo night vision sequences (12 bit, between 220 and 300 pairs of frames each) have been provided by Daimler AG, Germany, in June 2007 (group of Dr. Uwe Franke). These sequences come with ego-motion data and time stamps for each frame.
2	Synthesized stereo sequences (.enpeda.. & Daimler AG) These synthesized stereo sequences (with ground truth) have been provided by Tobi Vaudrey (.enpeda..) and Clemens Rabe (Daimler AG).
3	IMOs in color stereo sequences (Drivisco) These three day-time, color stereo sequences have been provided by the European Drivisco project. Independent moving objects ground truth and gaze data is now available.
4	"Normal camera" binocular stereo sequences (Hella Aglaia Mobile Vision & .enpeda..) A few of those day- or night-time, gray-level stereo sequences have been provided by Hella Aglaia Mobile Vision GmbH, Germany; most of them have been recorded by students in the .enpeda.. project.
5	"Normal camera" trinocular stereo sequences (.enpeda..) Three-camera stereo sequences (rectified by pairs) captured with HAKA1.
6	Grey-level stereo sequences with range scans (HU Berlin and others) So far three stereo vision sequences where the test vehicle drives through a car park; ground truth from a laser scanner; SGM, block and cross matcher disparity maps are also included. More to come here soon.



Fig. 11. An application for a trinocular sequence in Set 5 of EISATS. *Left*: third view. *Middle*: virtual view for the disparity map shown on the right. The applied matching algorithm was belief propagation stereo analysis. The specularities (see image on the left), apparent both in the recorded left and right image, causes a 'defect' in the calculated disparity data.

video could depict the optical flow as seen on the original video. To this end, a specialist web-application is developed to support concurrent video playback. Figure 12 illustrates the architecture of the designed system.

Delivering video over the web requires careful attention to bandwidth usage. When the videos are in high-definition, they consume a large bandwidth. The web-application therefore delivers low-definition videos with the ability to seamlessly zoom into it until the highest definition is reached.

The application is developed using Silverlight [40], [68] and Deep Zoom [57]. The application design and architecture are described in [77].

C. ScanT.NET

The .enpeda.. research domain has special interests for developing extensible online image databases, capable of handling complex pattern analysis and image processing for selected tasks, which can support both an expert community in a given research area and novice users from the public.

In the general context of environmental surveillance (see also Section VI-A), we are currently developing a web-based application called *ScanT.NET*, which allows ecology experts to upload, share, analyze, and compare their image data (e.g., scanned images of footprints of small species) with other existing image samples stored in database, as well as match and answer queries from the general public.

Such footprints give information of what kind of species they might be. However, the "puzzle" is that the images of a footprint may have very different appearances, including normal footprints, missing toe footprints, overlapped footprints; finding the sex of mice by analyzing their footprints is another possible point of interest. Our project aims to make the "puzzle solving" step much easier and quicker, including for people in the field.

The overall design of *ScanT.NET* follows the scheme shown in Fig. 13. Users should be able to access it via Silverlight applications running in browsers and mobile phones.

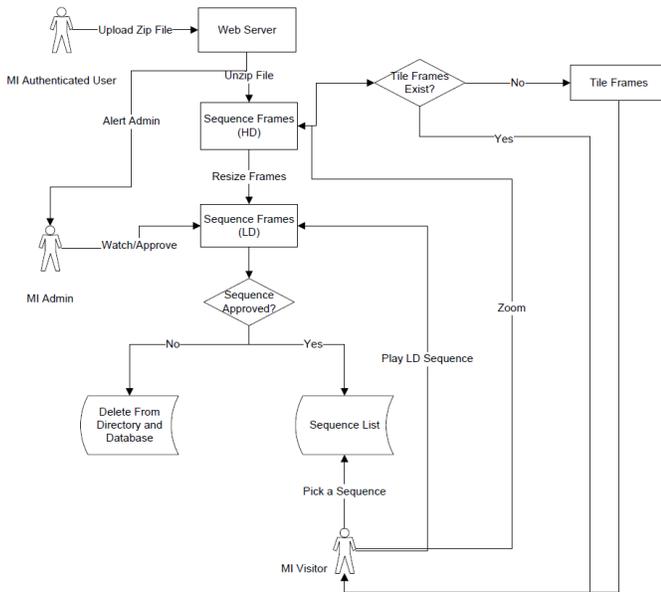


Fig. 12. System architecture.

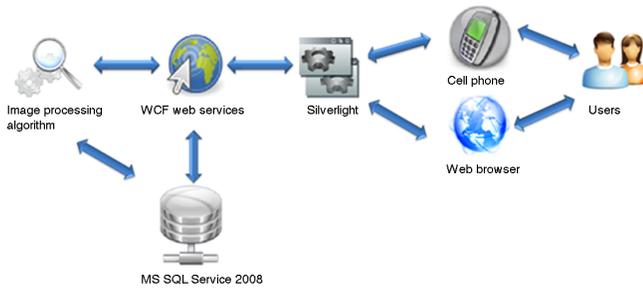


Fig. 13. Design overview of the ScanT.NET project.

V. PANORAMIC VISUALIZATION

We review a few of the current activities in this project. Our main partner in this research is the group led by *Fay Huang* at Ilan University, Taiwan. There is still ongoing collaboration with the department of *Anko Börner* at the German Air and Space Institute at Berlin. The 2008 monograph [28] and the paper [66] are results of those collaborations. The general interest is in wide-angle image and range data recording and visualization. [23] discussed how to reconstruct a 3D scene from a single cylindrical panorama.

A. Calibration of Fish-Eye Cameras

Fish-eye cameras are of interest for vision-based DAS because they allow to record wide-angle views, thus approaching the viewing abilities of the human visual system [1]. The geometry of fish-eye lenses may be mapped into the geometries of other panoramic sensors; see [70].

However, using fish-eye lenses in the DAS context poses several challenges such as a robust and precise



Fig. 14. Test vehicle HAKA1 of the .enpeda.. project with mounted fish-eye cameras on the roof.

synchronization and calibration of the cameras while mounted in a car and providing an unobstructed view for the cameras. The calibration must be done with the cameras mounted as they are setup for recording (see Fig. 14) to ensure accuracy, which poses additional challenges compared to calibrating in a lab, for example.

Altogether, the calibration method should not only be accurate and robust but also practical and appropriate for outdoors and cameras mounted on a car. Three fish-eye calibration methods were tested and compared within the DAS context: Bouguet's fish-eye calibration [4], Mei's omnidirectional method [48], and Scaramuzza's toolbox for omnidirectional cameras [62].

In the calibration experiments of stereo cameras with fish eye lenses, series of either 5,10,15, or 20 images of a planar checkerboard (90 cm \times 90 cm) were used. (The checkerboard proved to be appropriate and practical.) The calibration methods were compared by analyzing the differences in reported principal points, using back-projected images of vertices of checkerboard squares (into the recorded checkerboard images; see [38] for this method), as well as comparing stereo matching results on the rectified images using one selected stereo matching technique.

The results of the experiments show that Bouguet's calibration method is precise and converges nicely when taking at least 10 "good" calibration images. Mei's method follows closely with results that are nearly consistent with those of Bouguet's method, but less precise. Finally, Scaramuzza's method was not that easy to use since the automatic corner detection did not perform correctly always, and moreover, the principal point was off by up to 13 pixels compared to the results obtained by the other methods; it was also less precise in general.

In conclusion, Bouguet's fish-eye calibration method seems to be sufficiently precise and practical for the



Fig. 15. Anaglyphic stereo panorama of a classroom at Ilan University captured in 2009 with a rotating sensor matrix camera. See [29].

calibration of fish-eye lenses in a DAS context.

B. Calibration of Rotating Sensor Matrix Cameras

The calibration of rotating sensor-line cameras (see Fig. 16) has been a major subject in [28], [65]. However, sensor-line cameras are still in general not widely available. Thus, instead of rotating a sensor-line camera with a fixed viewing angle ω (as illustrated in Fig. 16), it is also possible to rotate a common matrix camera, and to compose the panoramic image by merging data recorded with one fixed sensor column of the matrix, thus implementing angle ω with an alternative method. Of course, the effective focal length of the selected sensor column will change with its position in the array.

As discussed in [28], two panoramas, recorded with a symmetric pair ω and $-\omega$ of viewing angles, support stereo visualization (thus also stereo analysis). The symmetric pair can now be recorded with the rotating matrix camera by selecting a symmetric (with respect to the center) pair of sensor columns of the sensor array. Three sensor parameters, off-axis distance R , principle angle ω , and corresponding focal length f , are critical to achieve high-quality stereo visualization. The actual focal length may be calibrated using some common method for sensor matrix cameras. The paper [29] presents a new camera calibration method for estimating accurately R and ω . To our best knowledge, there was no other camera

calibration method published prior to this paper for such a purpose.

Figure 15 illustrates an example of a stereo panorama captured by a rotating sensor-matrix camera, using the calibration method described in [29].

C. Panoramic Stereo Visualization

The precision of depth perception is limited by (digital) image resolution and constrained by the resolution of the displaying medium. A sensible goal is to optimize the stereo quality while viewing a high-resolution panoramic image by increasing the total number of potentially possible disparity values, including both crossed (i.e., negative, in front of the viewing medium) and uncrossed (i.e., positive, behind the viewing medium) disparity values. An insufficient number of image disparities produces a “cardboard effect”, where the 3D scene is perceived as a set of parallel cardboards, sorted in depth, one sitting in front of the other.

Paper [27] discusses the optimization of stereo viewing of panoramic images, especially with respect to zooming in or out. When zooming into a stereo panoramic image, disparity values will be “enlarged”. As a result, the perceived depth perception also increases. This implies that the depth cues (observed from the image parallaxes) lead to the opposite conclusion than actually wanted:

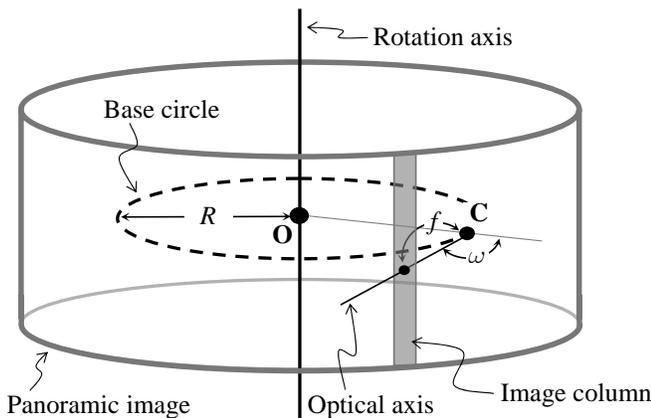


Fig. 16. The shaded area stands for the rotating sensor-line camera. It is in distance f to the focal center (on a base circle), which rotates at distance R around the rotation axis. The sensor line is tilted by an angle ω away from the normal direction of the cylinder defined by the base circle. See [28].



Fig. 17. The back-projection screen and polarized system for stereo panorama visualization at Ilan University, as used in the experiments reported in [27].



Fig. 18. *Left*: tracking tunnel with replaceable scan card. *Right*: an inked scan card with mouse footprints.

objects appear further away from the viewer. Obviously, this establishes a contradiction to our brain during a virtual navigation process, and may cause discomfort during stereo-viewing.

Also note that stereo viewing requires that disparities stay below the maximum disparity limit for human stereo fusion. This maximum disparity (or *fusibility limit*) depends on the viewing distance, and is suggested in [72] as being approximately equal to 0.03 times the viewing distance. If image disparity exceeds the upper disparity limit of human vision, then this causes double images (or *dipodia*). This would result in uncomfortable stereo viewing as well as eyestrain [76].

Paper [27] provides a solution for zooming in and out of stereo panoramic images, also using experiments in the Ilan University Virtual Reality Lab, shown in Fig. 17.

VI. PATTERN ANALYSIS AND IMAGE PROCESSING

Our group has been active in various applications of 2D or 3D image analysis and image processing in the past, such as 3D image analysis of human brain tissue [34]. Here we list three current activities.

A. Environmental Surveillance

Our main partners in this research are *James Russel* in the School of Biological Sciences at Tamaki campus, and the group led by *Young Woon Woo* at Busan University, Korea.

Track analysis plays an important role in environmental surveillance [54], [2] but is generally an un-automated labour-intensive process relying upon expert ability. Research by the group has demonstrated that much of this track analysis process (see Fig. 18) can be successfully automated, from digitizing the original animal prints, to creating reference databases, and finally testing unknown prints against the database using image-matching algorithms [60]. Focal taxa for this research include insects [24], [67], reptiles and small

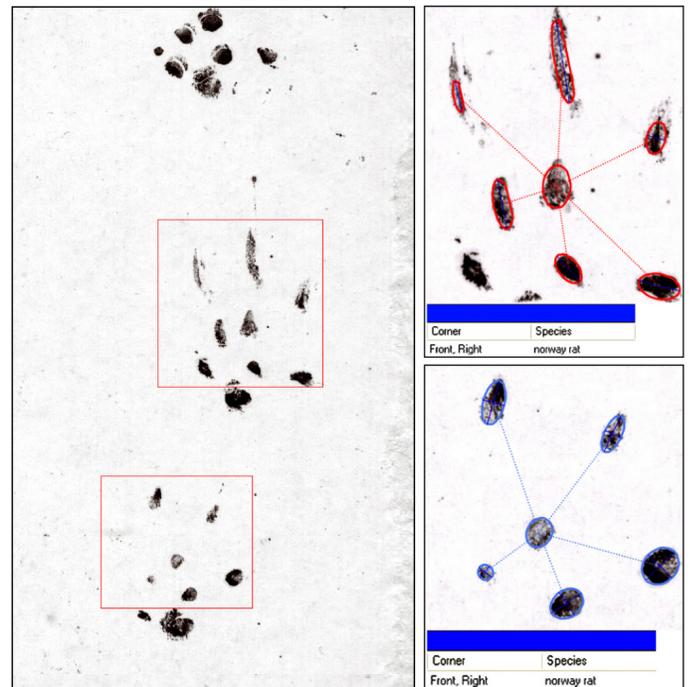


Fig. 19. Results of automated footprint labeling.

mammals. For introduced rodents of New Zealand (see Fig. 19) success rates around 80% are to be expected in the image-matching [60].

The benefits of this work include removing track analysis reliance on the few taxonomic experts currently available, as well as leveraging further analytical capacity for differentiating among and within species (e.g., age and sex), which experts are not currently able to do reliably. Application of the work includes to large-scale community restoration projects where the number of tracks collected outweighs labour ability to analyze, and also to border biosecurity where new unwanted

species arrivals may be detected.

B. Small Artists

As a side-project to automated track analysis, a large number of varying animal prints have been collected. Given the random behavioral component of many animals, these prints can themselves have an intrinsic artistic value [61]. By applying certain subsets of transformations and color filters it is possible to explore the artistic elements of these prints (e.g., Figs. 20 and 21). These artistic prints provide a novel method for environmental education, via their linkage to our track analysis research, but also through raising awareness of the diversity of small animals found throughout New Zealand.

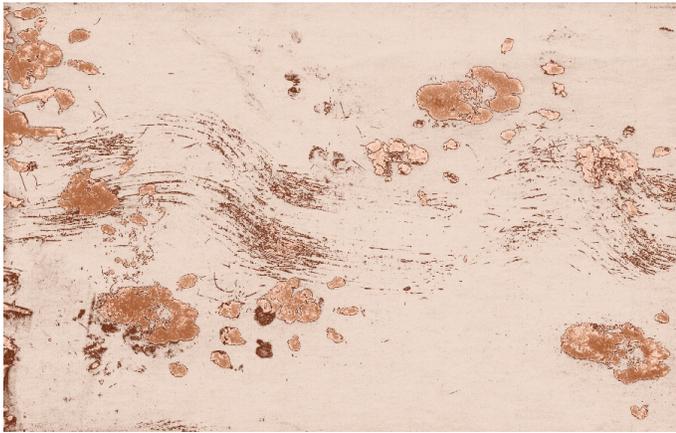


Fig. 20. The print of a large Wistar-strain male laboratory rat *R. norvegicus*. Pseudo-colored in earthy tones [61].

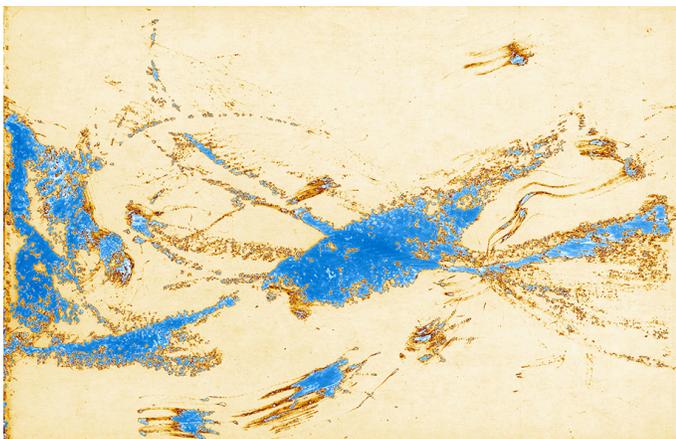


Fig. 21. Processed print of a robust skink *Oligosoma alani* using contrast enhancement and subsequent pseudo-coloring [61].

C. Artistic Filters

Painterly rendering defines a challenging subject for image processing. Captured images are processed in a way to simulate the brush strokes of painters. Different

authors have created their particular way to render photographs in the way of impressionism, pointillism, or of some specifically selected painters; see [71]. For example, so-called *Glass patterns* [18] may be used to simulated curved strokes as known from paintings of van Gogh [56].



Fig. 22. Original photographs by Angela Palmer. See [71].

However, artistic filters were aiming at simulating one particular style so far, not leaving much freedom for creating a new artistic expression. The techniques described in the report [71] allow us to blend different styles, thus selecting, for example, a blend of “60% pointillism and 40% impressionism”. This defines a new way to support creativity when using the offered interface for filtering photographs.

Figure 22 shows results, in particular highlighting the opportunity to apply the implemented blending procedures for creating ‘instant’ portrait paintings. These could be used, for example, as a commercial product or as a welcome gift for visitors to a city or University.

VII. EFFICIENT GEOMETRIC ALGORITHMS

We review a few of the current activities in this project. Our main partner in this research is the group led by *Fajie Li* at Huaqiao University, Xiamen, China. This work is closely related to open problems as identified in the monograph [37]. (A current journal publication, [30], contributed to the digital topology subject as discussed in this monograph.)

Basically, the main interest in this project is about shortest paths in Euclidean spaces (and not on finite graphs). The “turning points” of those shortest paths may freely move within a Euclidean space, just restricted by a set of polygonal or polyhedral objects. Such Euclidean shortest path problems come in various specifications, as defined by application areas such as robotics [69], driver assistance [45], gene research, 3D digital image analysis, and so forth.

In this project we are focussing on exploring the possibilities of applying rubberband algorithms. The initiating paper for this research was [6]. A summarizing report about related results was given in [43].

A. Watchman-Route in a Simple Polygon

Let Π be a planar, simple, topologically closed polygon with n vertices, and $\partial\Pi$ be its frontier. A point $p \in \Pi$ is *visible* from point $q \in \Pi$ iff $pq \subset \Pi$. The (*floating*) *watchman route problem* (WRP) of computational geometry, as discussed in [3], is defined as follows: *Calculate a shortest route $\rho \subset \Pi$ such that any point $p \in \Pi$ is visible from at least one point on ρ .* See Fig. 23 for an example. If a start point of the route is given on $\partial\Pi$ then this refined problem is known as the *fixed* WRP.

So far, the best result in running time for solving the fixed watchman route problem is $\mathcal{O}(n^3 \log n)$, published in [11]. The paper [44] provides an algorithm with $\kappa(\varepsilon) \cdot \mathcal{O}(kn)$ runtime, where n is the number of vertices of the given simple polygon Π , and k the number of essential cuts; $\kappa(\varepsilon)$ defines the numerical accuracy in dependency of a selected constant $\varepsilon > 0$. Moreover, the proposed algorithm is significantly simpler, easier to understand and implement than previous ones for solving the fixed watchman route problem.

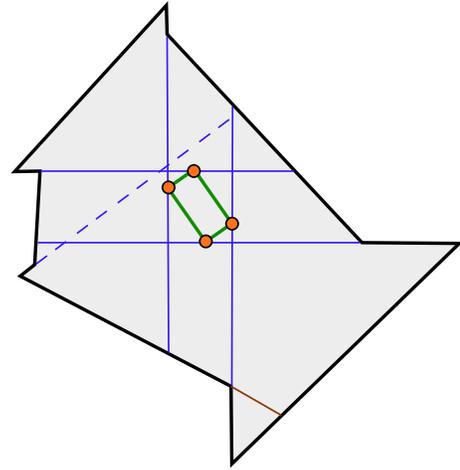


Fig. 23. This simple polygon has five essential cuts. Only four of those carry one of the four vertices of the shown shortest watchman route. This shortest watchman route may actually change its rectangular shape (i.e., there is an infinite number of shortest watchman routes). See the presentation of [44] at CCCG’2010.

B. Touring Polygons Problem

Assume that two points p and q are given and a finite ordered set of simple polygons, all in the same plane; the basic version of a touring-a-sequence-of-polygons problem (TPP) is to find a shortest path such that it starts at p , then visits these polygons in the given order, and ends at q .

According to [11], “one of the most intriguing open problems” identified by their results “is to determine the complexity of the fixed TPP for pairwise disjoint nonconvex simple polygons”. The paper [55] is focussing on the unconstrained *fixed* TPP (i.e., given start and end point of the path) and the *floating* TPP (i.e., no given start

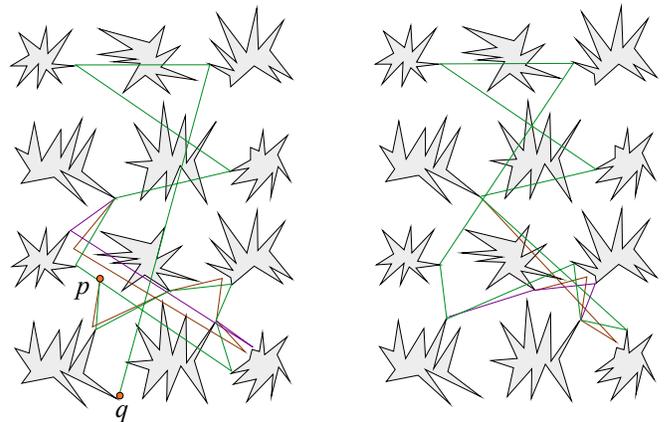


Fig. 24. The **red route** is obtained as the “initial path” for the rubberband algorithm when analyzing the convex hulls of those polygons; the **blue route** is an intermediate result, and the **green route** is the final one (left: fixed TPP; right: floating TPP). See [55].

or end point) under the condition that the convex hulls of the input polygons P_i are pairwise disjoint, but the polygons P_i itself may be nonconvex.

The paper [55] provides $\kappa(\varepsilon)\mathcal{O}(n)$ approximation algorithms for solving the TPP, either for given start and end points p and q , or with allowing to have those variable, where n is the total number of vertices of the given k simple and pairwise disjoint polygons; $\kappa(\varepsilon)$ defines the numerical accuracy in dependency of a selected $\varepsilon > 0$. Figure 24 illustrates a solution to the fixed TPP, and also a solution to the floating TPP.

VIII. CONCLUDING REMARKS

This report is basically covering activities within the previous one to two years (i.e., since the middle of 2009). We have not aimed at presenting all of our research during that period, but to present a few representative subjects. The main (i.e., “driving”) project in our works is certainly currently the *.nepeda.* project with its accompanying activities in stereo and motion analysis, and in creating web-based applications.

ACKNOWLEDGMENT

The authors thank previous and current members of the *.nepeda* group at Tamaki campus who contributed to subjects mentioned in this overview paper. Thanks also go to Daimler A.G., Germany, for making the Velodyne range scanner available for Sandino Morales while he was there recently for six months of internship.

REFERENCES

- [1] Ahn, J.: Calibration of various configurations of multiple cameras for driver assistance systems. Master thesis, Auckland University of Technology, 2010.
- [2] Alibhai, S. K., Z. C. Jewell, P. R. Law: A footprint technique to identify white rhino *Ceratotherium simum* at individual and species levels. *Endangered Species Research*, 4: 205–218, 2008.
- [3] Asano, T., S. K. Ghosh, and T. C. Shermer. Visibility in the plane. In: *Handbook of Computational Geometry* (J.-R. Sack and J. Urrutia, editors), pages 829–876, Elsevier, 2000.
- [4] Bouguet, J.: Camera calibration toolbox for Matlab, see www.vision.caltech.edu/bouguetj/calib_doc/.
- [5] Boykov, Y., O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Analysis Machine Intelligence*, 23:1222–1239, 2001.
- [6] Bülow, T., and R. Klette. Digital curves in 3D space and a linear-time length estimation algorithm. *IEEE Trans. Pattern Analysis Machine Intelligence*, 24:962–970, 2002.
- [7] Chen, Y., L. Gao, Z. Li, and Y. Liu: A new method for urban traffic state estimation based on vehicle tracking algorithm. In: *Intelligent Transportation Systems Conference*, pages 1097–1101, 2007.
- [8] Civilis, A., C. S. Jensen, J. Nenortait, and S. Pakalnis: Efficient tracking of moving objects with precision guarantees. In: *Int. Conf. Mobile Ubiquitous Systems: Networking and Services*, pages 164–173, 2004.
- [9] Civilis, A., C. S. Jensen, S. Pakalnis: Techniques for efficient road-network-based tracking of moving objects. *IEEE Trans. Knowledge Data Engineering*, 17:698–712, 2005.
- [10] Dewar, R., and P. Olson: *Human Factors in Traffic Safety*. Lawyers & Judges Publ., Tucson, 2002.
- [11] Dror, M., A. Efrat, A. Lubiw, and J. Mitchell. Touring a sequence of polygons. In: *STOC*, pages 473–482, 2003.
- [12] Egnal, G., M. Mintz, and R. Wildes: A stereo confidence metric using single view imagery with comparison to five alternative approaches. *Image Vision Computing*, 22:943–957, 2004.
- [13] Edelman, S., N. Intrator, and T. Poggio: Complex cells and object recognition. Unpublished, cogprints.org/561/, 1997.
- [14] EISATS: The *.nepeda.* image sequence analysis test site. <http://www.mi.auckland.ac.nz/EISATS>.
- [15] Felzenszwalb, P. F., and D. Huttenlocher: Efficient belief propagation for early vision. *Int. J. Computer Vision*, 70:41–54, 2006.
- [16] Fumi, D., and I. Sultan. A novel in-vehicle real-time brake-monitoring system. *Proc. Institution Mechanical Engineers, Part D: J. Automobile Engineering*, 223:793–804, 2009.
- [17] Gehrig, S., F. Eberli, and T. Meyer: A real-time low-power stereo vision engine using semi-global matching. In: *Int. Conf. Computer Vision Systems*, pages 134–143, 2009.
- [18] Glass, L.: Looking at Dots. *The Mathematical Intelligencer*, 24(4):37–43, 2002.
- [19] Gong, X: Finding the most likely vehicle path using spatial data. MSc thesis, Computer Science, The University of Auckland, 2010.
- [20] Guan, S., and R. Klette: Belief-propagation on edge images for stereo analysis of image sequences. In: *Robot Vision*, pages 291–302, 2008.
- [21] Guan, S., R. Klette, and Y. W. Woo: Belief propagation for stereo analysis of night-vision sequences. In: *PSIVT*, LNCS 5414, pages 932–943, 2009.
- [22] Haeusler, R., and R. Klette: Testing stereo data - Not the matching algorithms. Technical report, MItch-TR-55, Multimedia Imaging, The University of Auckland, 2010.
- [23] Haeusler, R., R. Klette, and F. Huang. Monocular 3D reconstruction of objects based on cylindrical panoramas. In: *PSIVT*, LNCS 5414, pages 60–70, 2009.
- [24] Heo, G., R. Klette, Y. W. Woo, K.-B. Kim, and N. H. Kim: Fuzzy support vector machine with a fuzzy nearest neighbor classifier for insect footprint classification. In: *IEEE Int. Conf. Fuzzy Systems*, to appear, 2010.
- [25] Hermann, S., R. Klette, and E. Destefanis. Inclusion of a second-order prior into semi-global matching. In: *PSIVT 2009*, LNCS 5414, pages 633–644, 2009.
- [26] Hirschmüller, H., and D. Scharstein: Evaluation of stereo matching costs on images with radiometric differences. *IEEE Trans. Pattern Analysis Machine Intelligence*, 31:1582–1599, 2009.
- [27] Huang, F., and R. Klette: Stereo panorama acquisition and automatic image disparity adjustment for stereoscopic visualization. *Multimedia Tools Applications*, 47:353–377, 2010.
- [28] Huang, F., R. Klette, and K. Scheibe: *Panoramic Imaging: Sensor-Line Cameras and Laser Range-Finders*. John Wiley & Sons Ltd, Chichester, 2008.
- [29] Huang, F., R. Klette, J.-C. Tien, and Y.-W. Chang: Rotating sensor-matrix camera calibration. In: *ICIP Conf.*, to appear, 2010.
- [30] James, T., and R. Klette: The topology of incidence pseudographs. *Pure Mathematics Applications*, 20:77–101, 2009 (published online on 10 March 2010).
- [31] Jiang, R., R. Klette, T. Vaudrey, and S. Wang: Corridor detection for vision-based driver assistance systems. *Int. J. Pattern Recognition Artificial Intelligence*, to appear, 2010.
- [32] Kalarot, R., and J. Morris: Comparison of FPGA and GPU implementations of real-time stereo vision. In: *Workshop Embedded Computer Vision*, computervisioncentral.com/sites/all/files/ecv2010/03.pdf, 2010.
- [33] Khan, W., J. Morris, and R. Klette: Stereo accuracy for collision avoidance. In: *Image Vision Computing New Zealand*, IEEE Catalog Number CFP0967E, 2009.
- [34] Klette, G.: Skeletal curves of 3D astrocyte samples. *Machine Graphics & Vision*, 17:105–129, 2008.
- [35] Klette, R., R. Jiang, S. Morales, and T. Vaudrey: Discrete driver assistance. In: *ISMM*, LNCS 5720, pages 1–12, Springer, Berlin, 2009.
- [36] Klette, R., N. Krüger, T. Vaudrey, K. Pauwels, M. van Hulle, S. Morales, F. I. Kandil, R. Haeusler, N. Pugeault, C. Rabe, and L. Markus: Performance of correspondence algorithms in vision-based driver assistance using EISATS. Technical report, MItch-TR-54, Multimedia Imaging, The University of Auckland, 2010.

- [37] Klette, R., and A. Rosenfeld: *Digital Geometry - Geometric Algorithms for Digital Picture Analysis*. Morgan Kaufmann, San Francisco, 2004.
- [38] Klette, R., K. Schlüns, and A. Koschan: *Computer Vision - Three-Dimensional Data from Digital Images*. Springer, Singapore, 1998.
- [39] Klette, R., T. Vaudrey, J. Wiest, R. Haeusler, R. Jiang, and S. Morales: Current challenges in vision-based driver assistance. Invited chapter in: *Progress in Combinatorial Image Analysis* (P. Wiederhold and R. Barneva, eds.), pages 3-22, Research Publ. Services, Singapore 2010.
- [40] Laurence, M.: Get started building a deeper experience across the web. *MSDN Magazine*, June, 2007.
- [41] Li, X., M. Li, W. Shu, and M. Wu: A practical map-matching algorithm for GPS-based vehicular networks in Shanghai urban area. In: *IET Conf. Wireless Mobile Sensor Networks*, pages 454-457, 2007.
- [42] Leclerc, Y. G., Q.-T. Luong, and P. V. Fua: Self-consistency and MDL: A paradigm for evaluating point-correspondence algorithms, and its application to detecting changes in surface elevation. *Int. J. Computer Vision*, 51:63-83, 2002.
- [43] Li, F., and R. Klette. Rubberband algorithms for solving various 2D or 3D shortest path problems. Invited talk, in: *IEEE Conf. Computing Theory Applications*, The Indian Statistical Institute, Kolkata, pages 9-18, 2007.
- [44] Li, F., and R. Klette. Watchman route in a simple polygon with a rubberband algorithm In *Canadian Conf. Computational Geometry*, Winnipeg, Canada, 2010.
- [45] Li, F., R. Klette, and S. Morales: An approximate algorithm for solving shortest path problems for mobile robots or driver assistance. In: *IEEE Intelligent Vehicles Symp.*, pages 42-47, 2009.
- [46] Liu, Z., and R. Klette: Approximated ground truth for stereo and motion analysis on real-world sequences. In: *PSIVT*, LNCS 5414, pages 874-885, 2009.
- [47] Manoharan, S.: On GPS tracking of mobile devices. In: *IEEE Int. Conf. Networking Services*, pages 415-418, 2009.
- [48] Mei, C., and P. Rives: Single view point omnidirectional camera calibration from planar grids. In: *IEEE Conf. Robotics Automation*, pages 3945-3950, 2007.
- [49] Middlebury Stereo Vision Page: Middlebury College, vision.middlebury.edu/stereo/.
- [50] Morales, S., Klette, R.: A third eye for performance evaluation in stereo sequence analysis. In: *Computer Analysis Images Patterns*, LNCS 5702, pages 1078-1086, 2009.
- [51] Morales, S., T. Vaudrey, and R. Klette: Robustness evaluation of stereo algorithms on long stereo sequences. In: *IEEE Intelligent Vehicles Symp.*, pages 347-352, 2009.
- [52] Morris, J., R. Haeusler, R. Jiang, R. Kalarot, T. Khan, W. Khan, S. Manoharan, S. Morales, T. Vaudrey, J. Wiest, and R. Klette: Current work in the .enpeda.. project. In: *Image Vision Computing New Zealand*, IEEE Catalog Number CFP0967E, 2009.
- [53] Y. Ohta and T. Kanade. Stereo by two-level dynamic programming. In: *Int. Joint Conf. Artificial Intelligence*, volume 2, pages 1120-1126, 1985.
- [54] Palma, A. R. T., and R. Gurgel-Goncalves: Morphometric identification of small mammal footprints from ink tracking tunnels in the Brazilian Cerrado. *Revista Brasileira Zoologia*, 24: 333-343, 2007.
- [55] Pan, X., F. Li, and R. Klette: Approximate shortest path algorithms for sequences of pairwise disjoint simple polygons. In *Canadian Conf. Computational Geometry*, Winnipeg, Canada, 2010.
- [56] Papari, G., and N. Petkov: Continuous glass patterns for painterly rendering, *IEEE Trans. Image Processing*, 18:652-664, 2009.
- [57] Prosis, J.: Taking silverlight deep zoom to the next level. *MSDN Magazine*, July 2009.
- [58] Ren, F., J. Huang, and A. Al Sarraf: Lane detection on a mobile phone. Runner-up in the New Zealand *Imagine Cup* 2010, www.microsoft.com/nz/imaginecup/default.aspx.
- [59] Ren, F., J. Huang, M. Terauchi, R. Jiang, and R. Klette: Lane detection on the iPhone. In: *ArtsIT*, LNICST 30, pages 198-205, Springer, Heidelberg, 2009.
- [60] Russel, J., N. Hasler, R. Klette, and B. Rosenhahn: Automatic track recognition of footprints for identifying cryptic species. *Ecology*, 90(7):2007-2013, 2009.
- [61] Russel, J., R. Klette, and C.-Y. Chen: Tracking small artists. In: *ArtsIT*, LNICST 30, pages 165-172, 2009.
- [62] Scaramuzza, D., and R. Siegwart: A practical toolbox for calibrating omnidirectional cameras. In: *Vision Systems Applications* (G. Obinata and A. Dutta, editors), pages 297-310, I-Tech Education and Publishing, Vienna, 2007.
- [63] Scharstein, D., and R. Szeliski: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Computer Vision*, 47:7-42, 2002.
- [64] Schauwecker, K.: Adaptations of belief propagation stereo. Internal report, Multimedia Imaging, The University of Auckland, 2010.
- [65] Scheibe, K., F. Huang, and R. Klette: Calibration of rotating sensors. In: *CAIP*, LNCS 5702, pages 157-164, 2009.
- [66] Scheibe, K., F. Huang, and R. Klette: Pose estimation of rotating sensors in the context of accurate 3D scene modeling. *JUCS*, 16:1269-1290, 2010.
- [67] Shin, B.-S., E.-Y. Cha, K.-B. Kim, K.-W. Cho, R. Klette, and Y. W. Woo: Effective feature extraction by trace transform for insect footprint classification. *J. Computational Theoretical Nanoscience*, 7:868-875, 2010.
- [68] Smith, J.: WPF apps with the Model-View-ViewModel design pattern. *MSDN Magazine*, February 2009.
- [69] Sommer, G., and R. Klette (editors): *Robot Vision 2008*. LNCS 4931, Springer, Berlin, 2008.
- [70] Torii, A., and R. Klette: Panoramic and 3D computer vision. In: *ArtsIT*, LNICST 30, pages 9-16, Springer, Heidelberg, 2009.
- [71] Valente, C.: Artistic emulation - Filter blending for painterly rendering. Dissertation for COMPSCI 789, Computer Science Department, The University of Auckland, 2010.
- [72] Valyrus, N.: *Stereoscopy*. Focal Press, London, 1966.
- [73] Vaudrey, T., S. Morales, A. Wedel, and R. Klette: Generalized residual images' effect on illumination artifact removal for correspondence algorithms. *Pattern Recognition*, to appear, 2010.
- [74] Vaudrey, T., A. Wedel, and R. Klette: A methodology for evaluating illumination artifact removal for corresponding images. In: *CAIP*, LNCS 5702, pages 1113-1121, Springer, Berlin, 2009.
- [75] Velodyne Lidar Inc.: Velodyne's HDL-64E S2 user manual - www.velodyne.com/laserrangefinder/ManualList.aspx , 2010.
- [76] Viire, E.: Health and safety issues for VR. *Comm. ACM*, 40:40-41, 1997.
- [77] Wang, B., S. Manoharan, and R. Klette: Web-based visualization of media data for driver assistance. In: *Int. Conf. Advances Multimedia*, pages 44-48, 2010.
- [78] Wedel, A., T. Pock, C. Zach, H. Bischof, and D. Cremers: An improved algorithm for TV- L^1 optical flow. In: *Visual Motion Analysis Workshop*, LNCS 5604, pages 23-45, Springer, Berlin, 2009.
- [79] Yang, X., and R. Klette: Evaluation of motion analysis on synthetic and real-world image sequences. Technical report, Mitech-TR-58, Multimedia Imaging, The University of Auckland, 2010.