In Proceedings "Aktives Sehen in technischen und biologischen Systemen" Workshop der GI-Fachgruppe 1.0.4 Bildverstehen, Hamburg, December 3-4, 1996, edited by Bärbel Mertsching, pp. 194-202, Sankt Augustin: infix
doi: 10.5281/zenodo.16447

The Role of Active Vision in Exploring Growth, Transport, and Exchange Processes

B. Jähne[1,2,4], H. Haußecker[1], F. Hering[1], G. Balschbach[2], J. Klinke[4], M. Lell[1,3], D. Schmundt[1,3], M. Schultz[1], U. Schurr[3], M. Stitt[3], and U. Platt[2]

Abstract

The active vision paradigm plays a significant role in a recently established interdisciplinary research initiative (DFG Forschergruppe) at Heidelberg University. The goal of the Forschergruppe is to develop and apply advanced image sequence processing techniques to study key scientific questions in environmental sciences and biology that cannot be approached with conventional techniques. The Forschergruppe includes researchers from the Interdisciplinary Center for Scientific Computing, the Botany Institute, and the Institute for Environmental Physics. This paper surveys the research strategy of the Forschergruppe with respect to the active vision paradigm and discusses active imaging techniques, camera control and calibration, tracking, motion analysis, 3-D reconstruction, and multichannel image fusion.

1 Introduction

The classical application areas of active vision are in technical disciplines focusing on robotics and autonomous vehicles [*Blake and Yuille*, 1992]. We will report here that the paradigm of active vision plays a key role for scientific applications.

In December 1995, the interdisciplinary research initiative (DFG Forschergruppe) "Image Sequence Analysis to Investigate Dynamic Processes" was founded at Heidelberg University. The goal of the Forschergruppe is to develop and apply advanced visualization techniques and image sequence processing techniques to study key scientific questions in environmental sciences and biology that are out of the reach of conventional techniques. This includes

- the mechanisms of wind-driven small-scale air-sea interaction processes at the ocean surface, especially the transfer of climate-relevant trace gases (e. g., CO_2 and CH_4),
- the nonlinear dynamics of wind generated ocean surface waves,
- the genetic control and external parameters influencing growth processes in plants, and
- the contribution of tropical forest fires to the global NO_x balance using the GOME instrument on the ERS2 satellite.

Although these applications are quite different, they have much in common concerning the image processing tasks. Consequently, a common research strategy has been developed for the acquisition, processing and analysis of multichannel image sequences. The focus of our research is on fast, accurate, and mathematically well-founded algorithms to be used on standard PCs and workstations. We believe that this is the best way to ensure that the results from this project can also be applied to the many and rapidly expanding applications throughout technical and natural sciences.

This paper surveys the role of active vision in this strategy. Discussed are active imaging techniques (Sect. 2), camera control, tracking, and calibration (Sect. 3), motion estimation (Sect. 4), 3-D reconstruction (Sect. 5), and multichannel image fusion (Sect. 6).

2 Active Imaging

Traditionally, computer vision disregarded the image formation process to a large extend. In contrast, the Forschergruppe focuses on an approach that regards image formation and image analysis as an *integral process*. Thus we

[1]Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg
[2]Institute for Environmental Physics, University of Heidelberg, Im Neuenheimer Feld 366, 69120 Heidelberg
[3]Institute for Botany, University of Heidelberg, Im Neuenheimer Feld 360, 69120 Heidelberg
[4]Physical Oceanography Research Division, Scripps Institution of Oceanography, University of California, San Diego

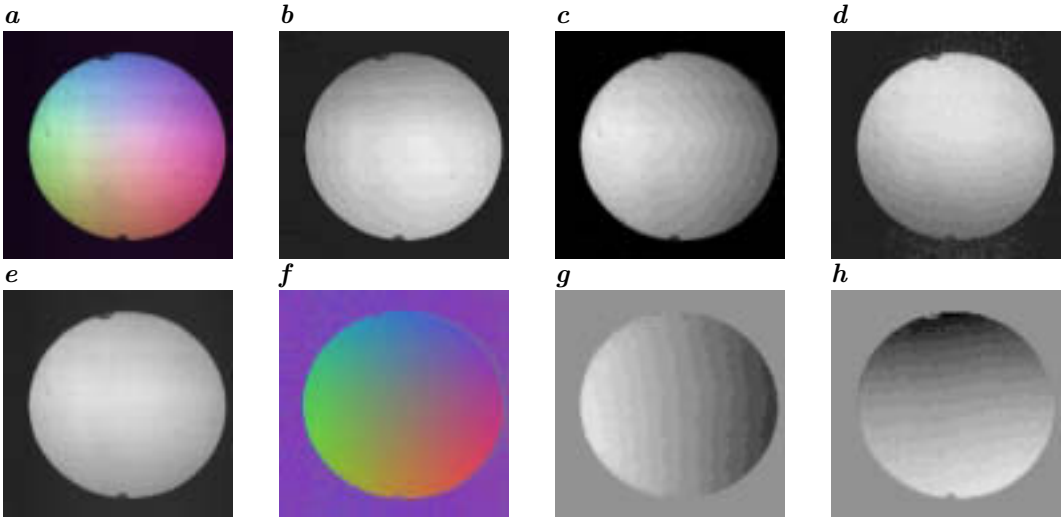

Figure 1: Demonstration of color ratio imaging used with the shape from reflection technique to measure the slope of water surfaces with a spherical lens as a calibration target. **a** original color image. **b - d** : Red, green, and blue channels of the original color image shown in a coarse quantization to highlight contour lines. **e** Intensity of the original image; watch the intensity fall off towards the edge of the lens (higher slopes). **f** Color image normalized by **E**. **g** and **h** : x and y components of the surface slope computed by ratio imaging in a coarse quantization to highlight contour lines.

actively influence the image formation process to gain an optimal mapping of the object parameters onto the irradiance at the image plane. Active imaging techniques have been used with tremendous success in the remote sensing community. A recent example is the US/French TOPEX satellite. From the signal of a radar altimeter instrument, such different parameters as the dynamic height of the ocean surface, the significant height of the wind waves, and the wind speed could be derived [*Jähne*, 1996].

Active imaging techniques range from the careful selection and setup of the illumination system to the active probing of the objects. This will be demonstrated with three examples.

Figure 1 demonstrates a novel shape-from-refraction technique that has been developed to measure the shape of the short ocean wind waves. As a review of optical techniques for wave measurements [*Jaehne et al.*, 1994] showed that these techniques are most suitable to handle specular surfaces. In contrast, stereo techniques do not work under this condition. The essential trick is to use ratio imaging. In this way, the surface normal can be derived independent of the intensity from three color illuminations. This is demonstrated with a test object, a spherical lens with constant curvature and thus a linearly changing surface slope. While the original color channels show significant intensity drops at the edge (Fig. 1b–d), the intensity in ratio images is directly proportional to the x and y component of the surface slope, respectively, as it becomes visible in the parallel contour lines in Fig. 1g, h. The same type of technique can be used for more conventional shape from shading techniques. If a Lambertian surface is illuminated by three different light sources, the surface normal vector independent of the reflection coefficient and can be determined [*Jähne*, 1996].

One of the most prominent examples of active imaging are laser induced fluorescence as illustrated in Fig. 2. One step further go techniques, where the illumination is used to probe certain properties of an object. The heat transfer and micro-scale turbulence at the ocean surface, for example, can be studied by heating a spot at the water surface with a pulse from a CO_2 laser and by observing the decay of the heated spot in infrared image sequences [*Haußecker*, 1996]. This technique has become a key instrument to investigate the mechanisms of small-scale exchange processes between the atmosphere and the oceans in the field.

3 Camera Control, Calibration, and Tracking

Many of the optical setups used in the Forschergruppe, require an adoption of the exposure to the available illumination condition yet need to preserve an absolute radiometric calibration is still required. The most demanding application in this respect is the around-the-clock observation of leaf growth under natural illumination conditions.

This problem will be resolved with an automatic control of the exposure time and a continuous observation of radiometric calibration targets based on LEDs that are integrated into the observed scene. We prefer not to change the aperture of the optics in order to avoid another variable that influences the geometrical calibration.

Likewise, a continuous geometrical calibration is required since the imaged sectors need to be adapted to the

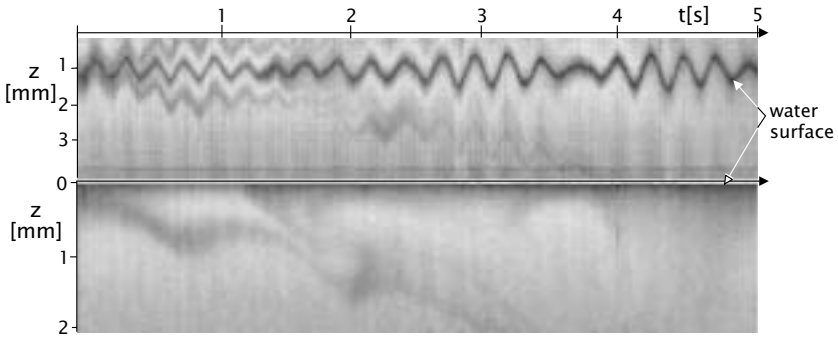

Figure 2: A 5 second time series of vertical concentration profiles of gases dissolved in water right at the interface as measured in the circular wind wave flume at the Institute for Environmental Physics at Heidelberg University. Traces of HCl gas were injected into the air space of the facility. The absorption of HCl at the water surface induces the transport of a fluorescent pH indicator across the aqueous boundary layer giving rise to lower concentration levels of the fluorescent dye towards the water surface. The wind speed was 2.4 m/s. The lower image shows the same sequence of vertical profiles but in a coordinate system moving with the water surface.

changing size of the leaves either by changing the distance of the camera to the object or by using a motorized zoom lens. In order to perform a continuous calibration, calibration targets will be positioned in the scene. It is planned to extend classical bundle triangulation techniques from photogrammetry that do not require calibrated cameras and thus make the experimental setup significantly easier to image sequences. In this way, a much more accurate calibration can be gained and it is possible to include temporal drifts into the calibration procedure. The solution of the nonlinear equation systems will be sped up by the use of calibration patterns that yield not only a position estimate at the image plane but also initial estimates of the distance to the camera and orientation in space.

Tracking is another important aspect because all to be investigated objects move. The most demanding applications in this respect are all high-resolution measurement close to the water surface. Concentration fields of dissolved gases, for instance (Fig. 2), require spatial resolution of down to 20 μm on a water surface moving up and down by wave motions by up to several meters. An elegant solution to this problem is a multistage approach. First the observation platform is made to move with the large scale motion. In the ocean, a freely drifting shallow buoy is used for that purpose (Fig. 3). The buoy follows the motion of the large waves. Thus small scale-structures stay longer in the observation sector and the water surface height varies by at most 20 cm. This low amplitude makes it feasible to use an optical wave follower with a galvanometer scanner as a second-stage tracking system.

4 Motion Analysis

The central task to be solved for all applications in the Forschergruppe is an accurate and robust estimate of the motion field. The mainstream of research in image sequence processing still works with only two or a few consecutive images of a sequence. This approach, however, yields only a "snapshot" of the motion field. Unfortunately, a snapshot of a visual motion field does not contain sufficient information to apply a regularization scheme to compute dense motion fields. We cannot infer, for example, whether a brightness edge coincides with an occlusion boundary or is just a discontinuity in the optical properties within the surface of an object moving with constant speed. The tremendous difficulties with the quantitative extraction of motion from consecutive images let many researchers to conclude that only qualitative properties of the motion field can reliably be extracted from consecutive images of a sequence [*Verri and Poggio*, 1989].

While this is useful and sufficient for many aspects of human and computer vision, it is not appropriate for the type of applications encounter in our research work. The demands for accuracy are significantly higher than in many other applications, since key parameters are not contained in the motion field itself but in its first derivatives. For instance, the divergence of the field is related to the growth rate and the curl to the vorticity of a flow field.

Consequently, we turned to the analysis of motion in spatio-temporal images, where the images are sampled densely enough also in the time axis to meet the sampling theorem. Then the correspondence problem is eliminated, regions of constant motion appear as oriented structures in the space-time images and the extraction of motion is therefore much more reliable.

Early roots of this approach can be found in biological vision. In a series of papers, *Adelson and Bergen* [1983, 1985, 1986] discussed spatio-temporal energy models for the perception of motion and pointed out the analogy between motion and orientation in the space-time domain. Tensor techniques for efficient and accurate computation of

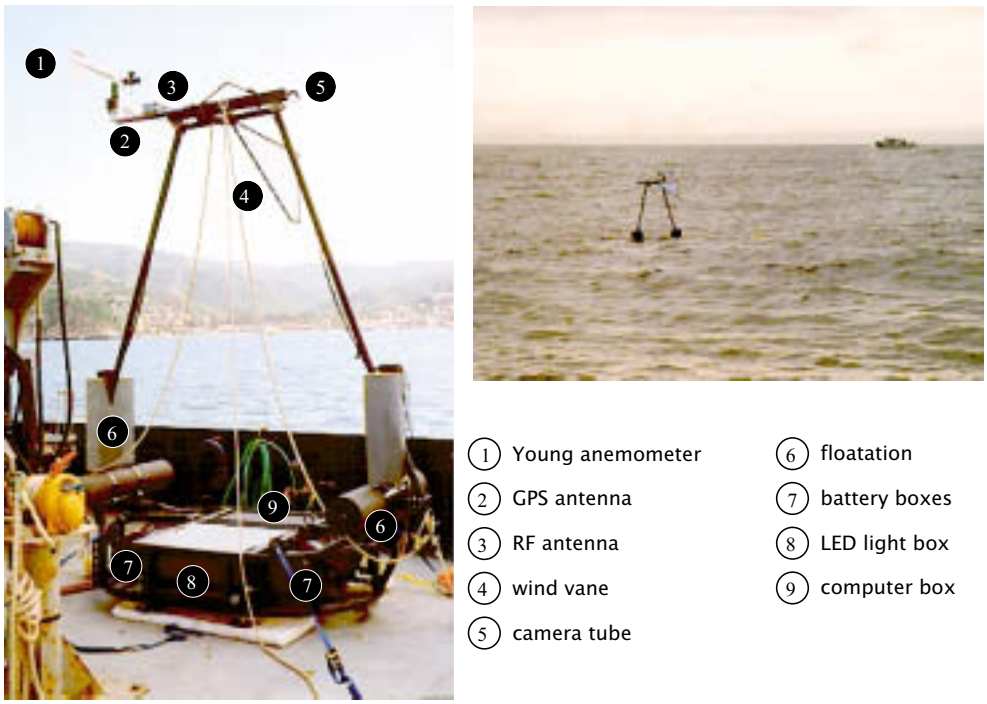

Figure 3: Instrumentation to take image sequences of the slope of short wind waves at the ocean interface as an example of a passive tracking system. Left: Wave-riding buoy at the stern of the RV New Horizon. Right: deployed buoy. Almost all parts of the buoy including the light source are submerged.

orientation origin from the work of *Granlund* [1978] and *Knutsson* [1982] who devised a tensor technique based on directional quadrature filters. Later, *Bigün and Granlund* [1987] showed that the tensor could be computed directly without using a set of quadrature filters and [*Jähne*, 1990] applied the tensor approach to motion analysis.

A unified analysis of all common techniques for motion determination in space-time images was undertaken by *Jähne* [1993]. He showed that regularized terms correlating the derivatives in the different directions in the space-time image are of central importance to a low-level estimate of the optical flow for any type of technique:

$$G_{pq}(\boldsymbol{x},t) = \langle g_p g_q \rangle = \int w(\boldsymbol{x}-\boldsymbol{x}', t-t') \frac{\partial g}{\partial p} \frac{\partial g}{\partial q} \mathrm{d}^2x' \mathrm{d}t' \tag{1}$$

where $g(\boldsymbol{x},t)$ is the spatio-temporal gray value structure. The partial derivative in the direction p is abbreviated by g_p. These terms can easily be approximated in discrete images by a combination of derivative filters, point operations, and smoothing filters. Expressed in an operator notation, we obtain

$$\mathcal{B}(\mathcal{D}_p \cdot \mathcal{D}_q), \tag{2}$$

where $\mathcal{D}_p$ is a suitable discrete first-order derivative operator in the direction p, $\mathcal{B}$ an averaging operator, and the nonlinear operator $\cdot$ denotes pointwise multiplications of two images. The operator expression in (2) includes the following sequence of image processing operations:

1. Apply the convolution operators $\mathcal{D}_p$ and $\mathcal{D}_q$ to the images sequence to obtain images with the first-order derivatives in direction p and q.
2. Multiply the two derivative images point by point.
3. Smooth the resulting images in all directions using the averaging kernels B.

The standard differential technique minimizes the continuity of the optical flow

$$||\boldsymbol{e}||_2^2 = \int w(\boldsymbol{x}-\boldsymbol{x}', t-t')(\boldsymbol{u}^T \boldsymbol{\nabla} g - g_t)^2 \mathrm{d}^2x' \, \mathrm{d}t' \longrightarrow \min \tag{3}$$

and results in the following linear equation system

$$\begin{bmatrix} \langle g_x g_x \rangle & \langle g_x g_y \rangle \\ \langle g_x g_y \rangle & \langle g_y g_y \rangle \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \langle g_x g_t \rangle \\ \langle g_y g_t \rangle \end{bmatrix} \tag{4}$$

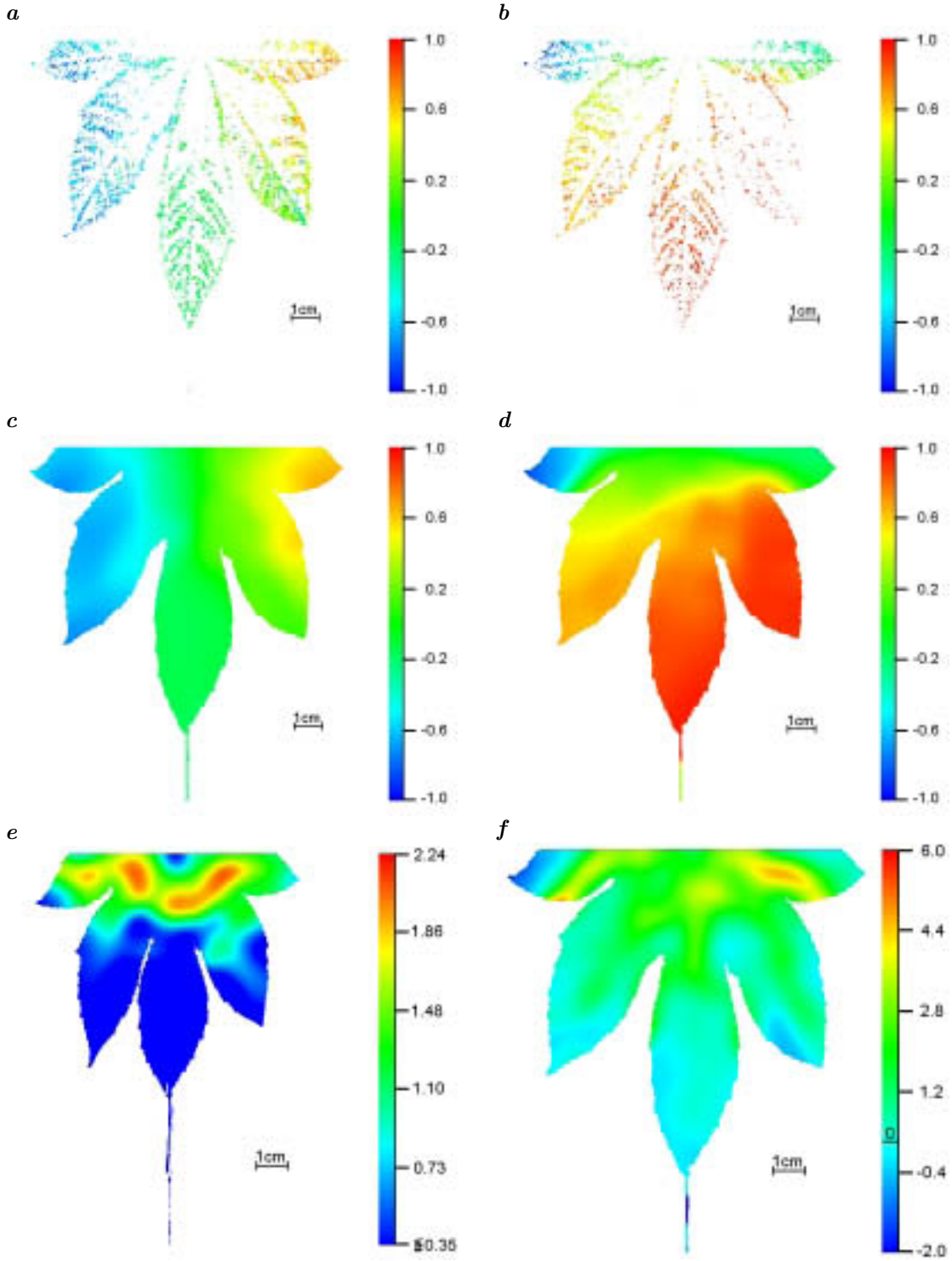

Figure 4: 2-D velocity determination of plant leave growth: **a** and **b** Masked horizontal and vertical velocities, **c** and **d** Horizontal and vertical velocities interpolated by normalized convolution, **e** and **f** Area dilation (divergence of the velocity vector field). Grayscale reproduction of color images from *Lell* [1996].

Although this approach yields exact results for a neighborhood with constant motion, additive noise leads to a systematic error biasing the motion estimates towards lower values.

As shown by *Jähne* [1993], the tensor approach does not show this disadvantage. It computes the eigenvectors and eigenvalues of the structure tensor

$$\begin{bmatrix} \langle g_x g_x \rangle & \langle g_x g_y \rangle & \langle g_x g_t \rangle \\ \langle g_x g_y \rangle & \langle g_y g_y \rangle & \langle g_y g_t \rangle \\ \langle g_x g_t \rangle & \langle g_y g_t \rangle & \langle g_t g_t \rangle \end{bmatrix}. \tag{5}$$

The direction of the eigenvector to the largest eigenvalue gives one or both components of the motion vector. When the rank of the tensor is two, a neighborhood with distributed spatial orientation but constant motion is encountered. When the rank of the tensor is one, the neighborhood is also spatially oriented and only the velocity component perpendicular to the direction of constant gray values can be computed. This is the well-known aperture problem. A tensor of rank zero corresponds to a constant neighborhood in which no velocity information can be obtained, while a tensor of rank three indicates a motion discontinuity. A detailed study of the classification of local neighborhoods by the tensor method can be found in *Haußecker and Jähne* [1996].

The differential and tensor methods can easily be extended to multichannel images by summing the terms with appropriate weighting over all channels

$$G_{pq}(\boldsymbol{x}, t) = \sum_{n=1}^{N} \frac{1}{\sigma_n^2} \langle g[n]_p g[n]_q \rangle, \tag{6}$$

where the channel number is indicated by square brackets.

Some initial results obtained with the tensor method are illustrated in Fig. 4. Fig. 4a,b shows the initial estimates only for the points at which reliable 2-D velocity estimates could be obtained corresponding to a rank of 2 for the structure tensor. Continuous velocity maps were computed using normalized convolution [*Granlund and Knutsson*, 1995] (Fig. 4c,d). Finally, the divergence of the velocity field (area dilation) nicely indicates the growth regions of the leaf at two different points in time (Fig. 4e,f).

5 3-D Reconstruction

The principal of continuous image acquisition can also be applied to 3-D reconstruction. This means that a single moving camera continuously taken images replaces a stereo camera setup. This direction of research was initiated almost a decade ago by the work of *Bolles et al.* [1987], who introduced epipolar-plane image analysis, as a method to extract structure from motion. Under the restricted condition that a camera moves along a linear path with a viewing direction orthogonal to the travel direction, they were able to recover the geometry of a static scene. Later this work was extended to more general camera movements [*Baker and Bolles*, 1989].

This approach has significant advantages over a stereo or multi-camera setup for 3-D reconstruction and is a nice example of the power of active vision systems. First, only a single camera is required. Thus 3-D reconstruction becomes feasible when only one camera is available, for instance an expensive infrared camera.

Second, the correspondence problem is avoided. The disparity indicating depth in a stereo camera setup is replaced by different orientations of gray value structures in space-time images. This is illustrated in Fig. 5 showing a leaf placed over a calibration grid. The camera is mounted on a rotating arm, the center of rotation coinciding with the position of the calibration grid. In the space-time slice (Fig. 5d), the leaf that is slightly closer to the camera shows a different orientation than the grid. Thus at depth jump is marked by a discontinuity in the orientation.

Third, occlusions no longer pose an intractable problem. This is demonstrated in Fig. 6 with a piece of crumbled paper. Occlusions are no longer just regions of missing correspondence between two stereo images but become *events* in space-time images. At a certain time, a pieces of surface disappear or appear Fig. 6c,d. These events can be detected as discontinuities in the orientation.

6 Multichannel Image Fusion

An important and common aspect of all research projects in the Forschergruppe is that each requires simultaneous measurements with multiple cameras. Thus the problem of image fusion arises.

Plant leaf growth, for instance is observed with

- mobile monochrome cameras with several channels in the visible and near infrared range to measure leave irradiance, size and reflectivity (Fig. 7a,b),

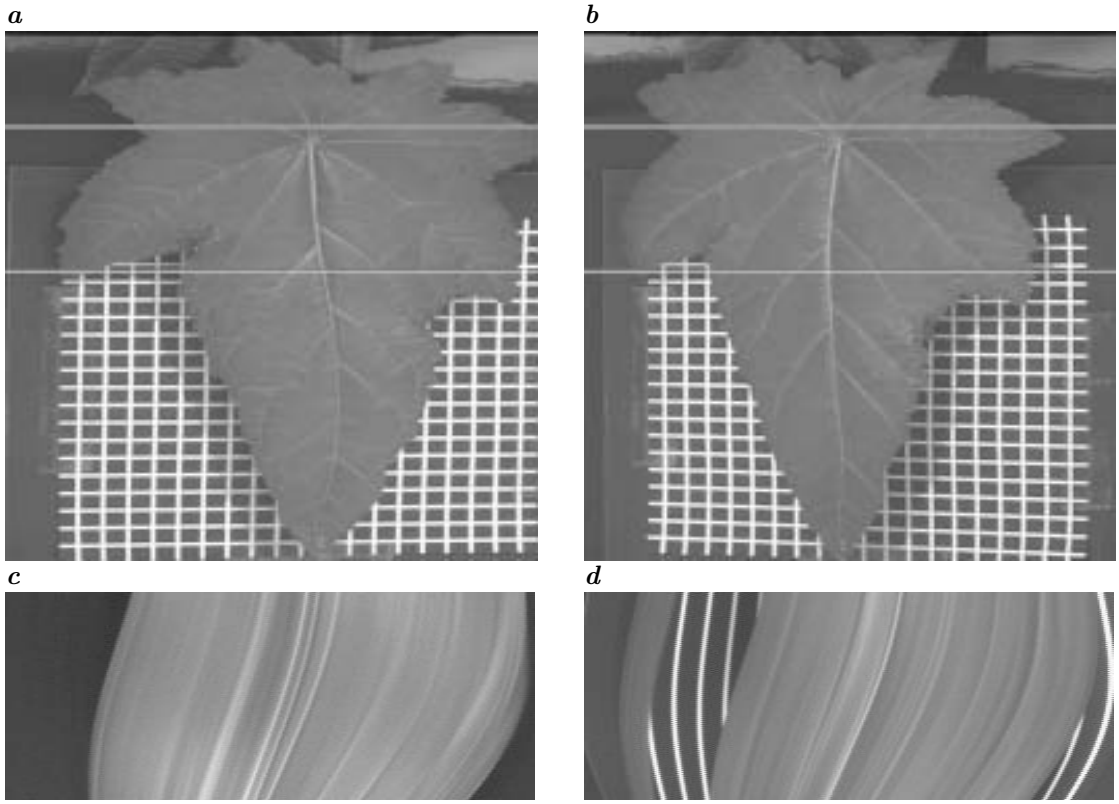

Figure 5: 3-D reconstruction from space-time images of a moving camera. **a** and **b** first and last image of the sequence, **c** and **d** space-time slice at the upper and lower marked lines in **a** and **b**, respectively.

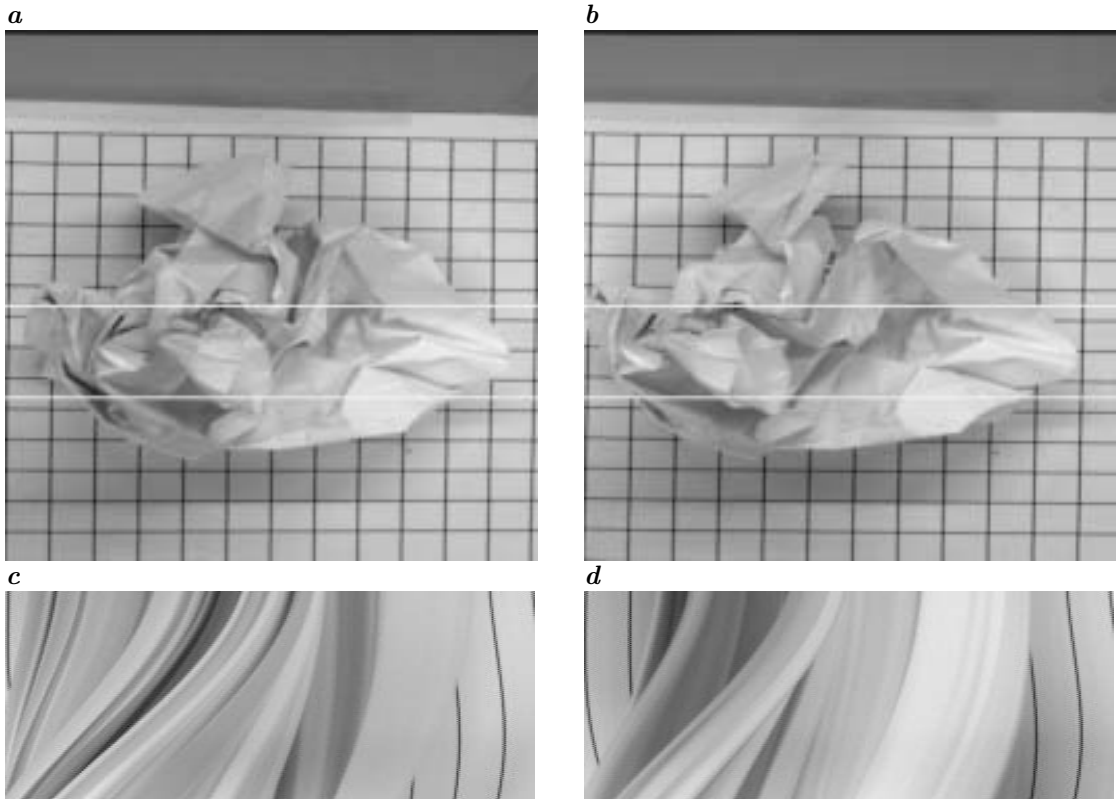

Figure 6: Occlusion detection in 3-D reconstruction from space-time images of a moving camera demonstrated with a crumbled sheet of paper. **a** and **b** first and last image of the sequence, **c** and **d** space-time slice at the upper and lower marked lines in **a** and **b**, respectively.

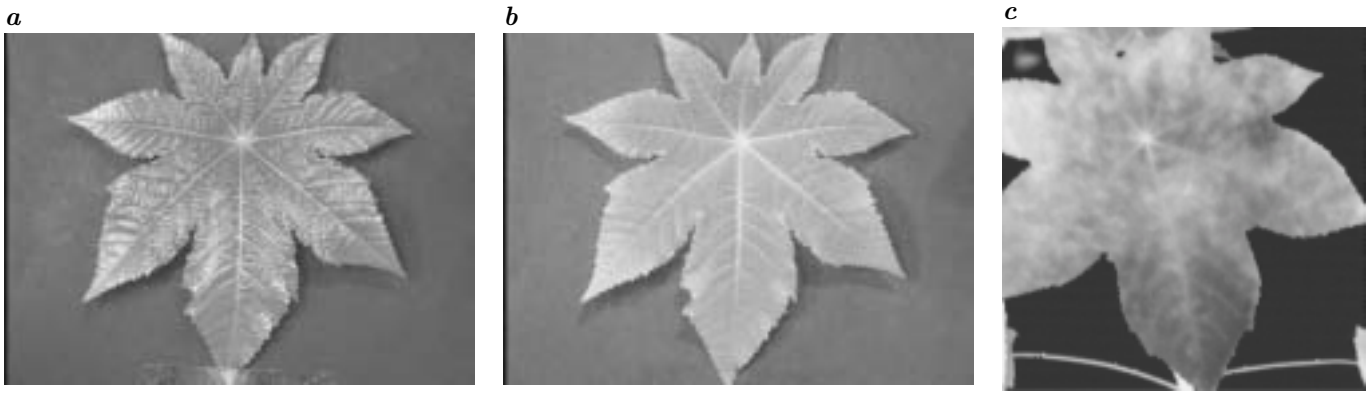

Figure 7: Images of a ricinus leaf taken in different wavelength regions: **a** visible, **b** near infrared (800–900 nm), **c** thermal infrared (3–5 μm). The patchy leaf surface in **c** indicates that evaporation and, therefore, opening of the stomatae is not equally distributed over the surface of the leaf.

- a high sensitive stationary slow-scan camera to measure chlorophyll fluorescence (no images taken yet), and
- an infrared camera to investigate leave transpiration (Fig. 7c).

Likewise, the investigation of the small-scale air sea interaction processes requires simultaneous image sequence acquisition of

- the water surface slope using a shape-from-refraction technique [*Jähne*, 1993], which itself requires a color camera or multiple grayscale cameras as discussed in Sect. 2,
- the 3-D flow field using particle tracking [*Hering et al.*, 1995, 1996] (also itself a multi-camera setup),
- measurements of the water surface temperature to study the heat exchange techniques [*Haussecker et al.*, 1995], and
- laser-induced fluorescence measurements to measure the concentration of dissolved gases close to the water surface [*Münsterer et al.*, 1995].

Again, the fusion of the images taken with different cameras is significantly facilitated by the fact that not single images but whole sequences can be registered. A nice example of this kind is the solution of the stereo correspondence problem for flow visualization [*Netzsch and Jähne*, 1995]. While the solution of the problem with individual particles requires at least three cameras, it can much easier be solved if the particles are first tracked through the sequence and then the particle traces are matched.

Acknowledgments

We gratefully acknowledge financial support for this research by the German Science Foundation (Forschergruppe "Bildfolgenanalyse zum Studium dynamischer Vorgänge", Ja 395/6-1) and in part by the US Office of Naval Research (grants N00014-93-J-0093 and N00014-94-1-0050) and the US National Science Foundation (grant OCE94-09182).

References

[1] Adelson, E. H., and J. R. Bergen, 1983. Spatio-temporal energy models for the perception of motion, *J. Opt. Soc. Am.*, *73*, 1861

[2] Adelson, E. H., and J. R. Bergen, 1985. Spatiotemporal energy models for the perception of motion, *J. Opt. Soc. Am.*, *A 2*, 284–299

[3] Adelson, E. H., and J. R. Bergen, 1986. The extraction of spatio-temporal energy in human and machine vision, In *Proc. on Motion: Representation and Analysis*, Charleston, 1986, pp. 151–155, IEEE Computer Society Press, Washington

[4] Baker, H. H., and R. C. Bolles, 1989. Generalizing epipolar-plane image analysis on the spatiotemporal surface, *Int. J. Comp. Vision*, *3*, 33–49

[5] Bigün, J., and G. H. Granlund, 1987. Optimal orientation detection of linear symmetry, *Proc. 1st Int. Conf. Comp. Vis.*, London 1987, pp. 433–438, IEEE Computer Society Press, Washington.

[6] Bigün, J., G. H. Granlund, and J. Wiklund, 1991. Multidimensional orientation estimation with application to texture analysis and optical flow, *IEEE Trans. PAMI*, *13*, 775–790.

[7] Blake, A., and A. Yuille, eds., 1992. *Active Vision*, MIT Press, Cambridge, MA

[8] Bolles, R. C., H. H. Baker, and D. H. Marimont, 1987. Epipolar-plane image analysis: an approach to determining structure from motion, *Int. J. Comp. Vision*, *1*, 7–55

[9] Granlund, G. H., 1978. In search of a general picture processing operator, *Comp. Graph. Imag. Process.*, *8*, 155–173

[10] Granlund, G., and H. Knutsson 1995. *Signal Processing For Computer Vision*, Kluwer, Dordrecht.

[11] Haußecker, H., B. Jähne, and S. Reinelt, 1995. Heat as a proxy tracer for gas exchange measurements in the field: principles and technical realization, Air-Water Gas Transfer, Selected Papers, 3rd Intern. Symp. on Air-Water Gas Transfer, edited by B. Jähne and E. Monahan, AEON, Hanau, pp. 405-413

[12] Haußecker, H., 1996. Messung und Simulation von kleinskaligen Austauschvorgängen an der Ozeanoberfläche mittels Thermographie, Dissertation, Faculty for Physics and Astronomy, Univ. Heidelberg

[13] Haußecker, H., and B. Jähne, 1996. A tensor approach for local structure analysis in multi-dimensional images, Proceedings Workshop 3D Bildanalyse und -synthese, Graduiertenkolleg 3D Bildanalyse und -synthese, Universität Erlangen-Nürnberg, 18.-19. November 1996, in press

[14] Hering, F., M. Merle, D. Wierzimok, and B. Jähne, 1995. A robust technique for tracking particles over long image sequences, Proc. ISPRS Intercommission Workshop From Pixels to Sequences, Zuerich, March 22 - 24, E. P. Baltsavias ed., Int. Arch. of Photog. and Rem. Sens., Vol 30, Part 5W1, pp. 202-207, RICS Books, Coventry UK, ISSN 0256-1840.

[15] Hering, F., C. Leue, D. Wierzimok, and B. Jähne, B., 1996. Particle tracking velocimetry beneath water waves, Experiments in Fluids, in press.

[16] Jähne, B., 1990. Motion determination in space-time images, *Proc. Computer Vision — ECCV 90, Lecture Notes in Computer Science 427*, edited by O. Faugeras, pp. 161–173, Springer, New York

[17] Jähne, B., 1993. Spatio-Temporal Image Processing - Theory and Scientific Applications, Lecture Notes in Computer Science 751, Springer, Berlin

[18] Jähne, B., 1996. Practical Handbook on Image Processing for Scientfic Applications, CRC-Press, Boca Raton, FL, in press

[19] Jähne, B., J. Klinke, and S. Waas, 1994. Imaging of short ocean wind waves: a critical theoretical review, J. Optical Soc. Amer. A, 11, 2197-2209.

[20] Knutsson, H., 1982. *Filtering and reconstruction in image processing*, Diss., Linköping Univ.

[21] Lell, M., 1996. Ortsaufgelöste Bestimmung von Blattwachstum durch Strukturanalyse von Bildsequenzen aus dem nahen Infrarot, Diplomarbeit, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Univ. Heidelberg.

[22] Münsterer, T., H. J. Mayer, and B. Jähne, 1995. Dual-tracer measurements of concentration profiles in the aqueous mass boundary layer, Air-Water Gas Transfer, Selected Papers, 3rd Intern. Symp. on Air-Water Gas Transfer, edited by B. Jähne and E. Monahan, AEON, Hanau, pp. 637-648

[23] Netzsch, T., and B. Jähne, 1995. A high performance algorithm for 3-dimensional particle tracking velocimetry in turbulent flow research using image sequences, Proc. ISPRS Intercommission Workshop From Pixels to Sequences, Zuerich, March 22-24, E. P. Baltsavias ed., Int. Arch. of Photog. and Rem. Sens., Vol 30, Part 5W1, pp. 74-79, RICS Books, Coventry UK, ISSN 0256-1840.

[24] Verri, A., and T. Poggio, 1989. Motion field and optical flow: qualitative properties, *IEEE Trans. PAMI*, *11*, 490–498