Bernd Jähne
Rudolf Mester
Erhardt Barth
Hanno Scharr (Eds.)

LNCS 3417

Complex Motion

First International Workshop, IWCM 2004
Günzburg, Germany, October 2004
Revised Papers

Springer

Bernd Jähne Rudolf Mester
Erhardt Barth Hanno Scharr (Eds.)

Complex Motion

First International Workshop, IWCM 2004
Günzburg, Germany, October 12-14, 2004
Revised Papers

Springer

Volume Editors

Bernd Jähne
Universität Heidelberg
IWR und Institut für Umweltphysik
Im Neuenheimer Feld 368, 69210 Heidelberg, Germany
E-mail: Bernd.Jaehne@iwr.uni.heidelberg.de

Rudolf Mester
J.W. Goethe Universität
Institut für Informatik, vsi Arbeitsgruppe
Robert-Mayer-Str. 10, 60054 Frankfurt, Germany
E-mail: mester@iap.uni-frankfurt.de

Erhardt Barth
Universität Lübeck
Institut für Neuro- und Bioinformatik
Ratzeburger Allee 160, 23538 Lübeck, Germany
E-mail: barth@inb.uni-luebeck.de

Hanno Scharr
Forschungszentrum Jülich GmbH
Institut für Chemie und Dynamik der Geosphäre
Institut III: Phytosphäre, 52425 Jülich, Germany
E-mail: H.Scharr@fz-juelich.de

Library of Congress Control Number: 2006939854

CR Subject Classification (1998): I.5, I.4, I.3.5, I.2.10, I.2.6, F.2.2

LNCS Sublibrary: SL 6 – Image Processing, Computer Vision, Pattern Recognition, and Graphics

ISSN 0302-9743
ISBN-10 3-540-69864-7 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-69864-7 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11976837 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 3417

Lecture Notes in Computer Science

For information about Vols. 1–4277

please contact your bookseller or Springer

Vol. 4377: M. Abe (Ed.), Topics in Cryptology – CT-RSA 2007. XI, 403 pages. 2006.

Vol. 4371: K. Inoue, K. Satoh, F. Toni (Eds.), Computational Logic in Multi-Agent Systems. X, 315 pages. 2007. (Sublibrary LNAI).

Vol. 4369: M. Umeda, A. Wolf, O. Bartenstein, U. Geske, D. Seipel, O. Takata (Eds.), Declarative Programming for Knowledge Management. X, 229 pages. 2006. (Sublibrary LNAI).

Vol. 4368: T. Erlebach, C. Kaklamanis (Eds.), Approximation and Online Algorithms. X, 345 pages. 2006.

Vol. 4367: K. De Bosschere, D. Kaeli, P. Stenström, D. Whalley, T. Ungerer (Eds.), High Performance Embedded Architectures and Compilers. XI, 307 pages. 2007.

Vol. 4364: T. Kühne (Ed.), Models in Software Engineering. XI, 332 pages. 2007.

Vol. 4362: J. van Leeuwen, G.F. Italiano, W. van der Hoek, C. Meinel, H. Sack, F. Plášil (Eds.), SOFSEM 2007: Theory and Practice of Computer Science. XXI, 937 pages. 2006.

Vol. 4361: H.J. Hoogeboom, G. Păun, G. Rozenberg, A. Salomaa (Eds.), Membrane Computing. IX, 555 pages. 2006.

Vol. 4360: W. Dubitzky, A. Schuster, P.M.A. Sloot, M. Schroeder, M. Romberg (Eds.), Distributed, High-Performance, and Grid Computing in Computational Biology. X, 192 pages. 2007. (Sublibrary LNBI).

Vol. 4357: L. Buttyán, V. Gligor, D. Westhoff (Eds.), Security and Privacy in Ad-Hoc and Sensor Networks. X, 193 pages. 2006.

Vol. 4355: J. Julliand, O. Kouchnarenko (Eds.), B 2007: Formal Specification and Development in B. XIII, 293 pages. 2006.

Vol. 4354: M. Hanus (Ed.), Practical Aspects of Declarative Languages. X, 335 pages. 2006.

Vol. 4353: T. Schwentick, D. Suciu (Eds.), Database Theory – ICDT 2007. XI, 419 pages. 2006.

Vol. 4352: T.-J. Cham, J. Cai, C. Dorai, D. Rajan, T.-S. Chua, L.-T. Chia (Eds.), Advances in Multimedia Modeling, Part II. XVIII, 743 pages. 2006.

Vol. 4351: T.-J. Cham, J. Cai, C. Dorai, D. Rajan, T.-S. Chua, L.-T. Chia (Eds.), Advances in Multimedia Modeling, Part I. XIX, 797 pages. 2006.

Vol. 4349: B. Cook, A. Podelski (Eds.), Verification, Model Checking, and Abstract Interpretation. XI, 395 pages. 2006.

Vol. 4348: S.T. Taft, R.A. Duff, R.L. Brukardt, E. Ploedereder, P. Leroy (Eds.), Ada 2005 Reference Manual. XXII, 765 pages. 2006.

Vol. 4347: J. Lopez (Ed.), Critical Information Infrastructures Security. X, 286 pages. 2006.

Vol. 4345: N. Maglaveras, I. Chouvarda, V. Koutkias, R. Brause (Eds.), Biological and Medical Data Analysis. XIII, 496 pages. 2006. (Sublibrary LNBI).

Vol. 4344: V. Gruhn, F. Oquendo (Eds.), Software Architecture. X, 245 pages. 2006.

Vol. 4342: H. de Swart, E. Orłowska, G. Schmidt, M. Roubens (Eds.), Theory and Applications of Relational Structures as Knowledge Instruments II. X, 373 pages. 2006. (Sublibrary LNAI).

Vol. 4341: P.Q. Nguyen (Ed.), Progress in Cryptology - VIETCRYPT 2006. XI, 385 pages. 2006.

Vol. 4340: R. Prodan, T. Fahringer, Grid Computing. XXIII, 317 pages. 2007.

Vol. 4339: E. Ayguadé, G. Baumgartner, J. Ramanujam, P. Sadayappan (Eds.), Languages and Compilers for Parallel Computing. XI, 476 pages. 2006.

Vol. 4338: P. Kalra, S. Peleg (Eds.), Computer Vision, Graphics and Image Processing. XV, 965 pages. 2006.

Vol. 4337: S. Arun-Kumar, N. Garg (Eds.), FSTTCS 2006: Foundations of Software Technology and Theoretical Computer Science. XIII, 430 pages. 2006.

Vol. 4334: B. Beckert, R. Hähnle, P.H. Schmitt (Eds.), Verification of Object-Oriented Software. XXIX, 658 pages. 2007. (Sublibrary LNAI).

Vol. 4333: U. Reimer, D. Karagiannis (Eds.), Practical Aspects of Knowledge Management. XII, 338 pages. 2006. (Sublibrary LNAI).

Vol. 4332: A. Bagchi, V. Atluri (Eds.), Information Systems Security. XV, 382 pages. 2006.

Vol. 4331: G. Min, B. Di Martino, L.T. Yang, M. Guo, G. Ruenger (Eds.), Frontiers of High Performance Computing and Networking – ISPA 2006 Workshops. XXXVII, 1141 pages. 2006.

Vol. 4330: M. Guo, L.T. Yang, B. Di Martino, H.P. Zima, J. Dongarra, F. Tang (Eds.), Parallel and Distributed Processing and Applications. XVIII, 953 pages. 2006.

Vol. 4329: R. Barua, T. Lange (Eds.), Progress in Cryptology - INDOCRYPT 2006. X, 454 pages. 2006.

Vol. 4328: D. Penkler, M. Reitenspiess, F. Tam (Eds.), Service Availability. X, 289 pages. 2006.

Vol. 4327: M. Baldoni, U. Endriss (Eds.), Declarative Agent Languages and Technologies IV. VIII, 257 pages. 2006. (Sublibrary LNAI).

Vol. 4326: S. Göbel, R. Malkewitz, I. Iurgel (Eds.), Technologies for Interactive Digital Storytelling and Entertainment. X, 384 pages. 2006.

Vol. 4325: J. Cao, I. Stojmenovic, X. Jia, S.K. Das (Eds.), Mobile Ad-hoc and Sensor Networks. XIX, 887 pages. 2006.

Vol. 4323: G. Doherty, A. Blandford (Eds.), Interactive Systems. XI, 269 pages. 2006.

Vol. 4320: R. Gotzhein, R. Reed (Eds.), System Analysis and Modeling: Language Profiles. X, 229 pages. 2006.

Vol. 4319: L.-W. Chang, W.-N. Lie (Eds.), Advances in Image and Video Technology. XXVI, 1347 pages. 2006.

Vol. 4318: H. Lipmaa, M. Yung, D. Lin (Eds.), Information Security and Cryptology. XI, 305 pages. 2006.

Vol. 4317: S.K. Madria, K.T. Claypool, R. Kannan, P. Uppuluri, M.M. Gore (Eds.), Distributed Computing and Internet Technology. XIX, 466 pages. 2006.

Vol. 4316: M.M. Dalkilic, S. Kim, J. Yang (Eds.), Data Mining and Bioinformatics. VIII, 197 pages. 2006. (Sublibrary LNBI).

Vol. 4313: T. Margaria, B. Steffen (Eds.), Leveraging Applications of Formal Methods. IX, 197 pages. 2006.

Vol. 4312: S. Sugimoto, J. Hunter, A. Rauber, A. Morishima (Eds.), Digital Libraries: Achievements, Challenges and Opportunities. XVIII, 571 pages. 2006.

Vol. 4311: K. Cho, P. Jacquet (Eds.), Technologies for Advanced Heterogeneous Networks II. XI, 253 pages. 2006.

Vol. 4309: P. Inverardi, M. Jazayeri (Eds.), Software Engineering Education in the Modern Age. VIII, 207 pages. 2006.

Vol. 4308: S. Chaudhuri, S.R. Das, H.S. Paul, S. Tirthapura (Eds.), Distributed Computing and Networking. XIX, 608 pages. 2006.

Vol. 4307: P. Ning, S. Qing, N. Li (Eds.), Information and Communications Security. XIV, 558 pages. 2006.

Vol. 4306: Y. Avrithis, Y. Kompatsiaris, S. Staab, N.E. O'Connor (Eds.), Semantic Multimedia. XII, 241 pages. 2006.

Vol. 4305: A.A. Shvartsman (Ed.), Principles of Distributed Systems. XIII, 441 pages. 2006.

Vol. 4304: A. Sattar, B.-H. Kang (Eds.), AI 2006: Advances in Artificial Intelligence. XXVII, 1303 pages. 2006. (Sublibrary LNAI).

Vol. 4303: A. Hoffmann, B.-H. Kang, D. Richards, S. Tsumoto (Eds.), Advances in Knowledge Acquisition and Management. XI, 259 pages. 2006. (Sublibrary LNAI).

Vol. 4302: J. Domingo-Ferrer, L. Franconi (Eds.), Privacy in Statistical Databases. XI, 383 pages. 2006.

Vol. 4301: D. Pointcheval, Y. Mu, K. Chen (Eds.), Cryptology and Network Security. XIII, 381 pages. 2006.

Vol. 4300: Y.Q. Shi (Ed.), Transactions on Data Hiding and Multimedia Security I. IX, 139 pages. 2006.

Vol. 4299: S. Renals, S. Bengio, J.G. Fiscus (Eds.), Machine Learning for Multimodal Interaction. XII, 470 pages. 2006.

Vol. 4297: Y. Robert, M. Parashar, R. Badrinath, V.K. Prasanna (Eds.), High Performance Computing - HiPC 2006. XXIV, 642 pages. 2006.

Vol. 4296: M.S. Rhee, B. Lee (Eds.), Information Security and Cryptology – ICISC 2006. XIII, 358 pages. 2006.

Vol. 4295: J.D. Carswell, T. Tezuka (Eds.), Web and Wireless Geographical Information Systems. XI, 269 pages. 2006.

Vol. 4294: A. Dan, W. Lamersdorf (Eds.), Service-Oriented Computing – ICSOC 2006. XIX, 653 pages. 2006.

Vol. 4293: A. Gelbukh, C.A. Reyes-Garcia (Eds.), MICAI 2006: Advances in Artificial Intelligence. XXVIII, 1232 pages. 2006. (Sublibrary LNAI).

Vol. 4292: G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. Nefian, G. Meenakshisundaram, V. Pascucci, J. Zara, J. Molineros, H. Theisel, T. Malzbender (Eds.), Advances in Visual Computing, Part II. XXXII, 906 pages. 2006.

Vol. 4291: G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. Nefian, G. Meenakshisundaram, V. Pascucci, J. Zara, J. Molineros, H. Theisel, T. Malzbender (Eds.), Advances in Visual Computing, Part I. XXXI, 916 pages. 2006.

Vol. 4290: M. van Steen, M. Henning (Eds.), Middleware 2006. XIII, 425 pages. 2006.

Vol. 4289: M. Ackermann, B. Berendt, M. Grobelnik, A. Hotho, D. Mladenič, G. Semeraro, M. Spiliopoulou, G. Stumme, V. Svátek, M. van Someren (Eds.), Semantics, Web and Mining. X, 197 pages. 2006. (Sublibrary LNAI).

Vol. 4288: T. Asano (Ed.), Algorithms and Computation. XX, 766 pages. 2006.

Vol. 4287: C. Mao, T. Yokomori (Eds.), DNA Computing. XII, 440 pages. 2006.

Vol. 4286: P.G. Spirakis, M. Mavronicolas, S.C. Kontogiannis (Eds.), Internet and Network Economics. XI, 401 pages. 2006.

Vol. 4285: Y. Matsumoto, R.W. Sproat, K.-F. Wong, M. Zhang (Eds.), Computer Processing of Oriental Languages. XVII, 544 pages. 2006. (Sublibrary LNAI).

Vol. 4284: X. Lai, K. Chen (Eds.), Advances in Cryptology – ASIACRYPT 2006. XIV, 468 pages. 2006.

Vol. 4283: Y.Q. Shi, B. Jeon (Eds.), Digital Watermarking. XII, 474 pages. 2006.

Vol. 4282: Z. Pan, A. Cheok, M. Haller, R.W.H. Lau, H. Saito, R. Liang (Eds.), Advances in Artificial Reality and Tele-Existence. XXIII, 1347 pages. 2006.

Vol. 4281: K. Barkaoui, A. Cavalcanti, A. Cerone (Eds.), Theoretical Aspects of Computing - ICTAC 2006. XV, 371 pages. 2006.

Vol. 4280: A.K. Datta, M. Gradinariu (Eds.), Stabilization, Safety, and Security of Distributed Systems. XVII, 590 pages. 2006.

Vol. 4279: N. Kobayashi (Ed.), Programming Languages and Systems. XI, 423 pages. 2006.

Vol. 4278: R. Meersman, Z. Tari, P. Herrero (Eds.), On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops, Part II. XLV, 1004 pages. 2006.

Preface

The world we live in is a dynamic one: we explore it by moving through it, and many of the objects which we are interested in are also moving. Traffic, for instance, is an example of a domain where detecting and processing visual motion is of vital interest, both in a metaphoric as well as in a purely literal sense. Visual communication is another important example of an area of science which is dominated by the need to measure, understand, and represent visual motion in an efficient way.

Visual motion is a subject of research which forces the investigator to deal with complexity; complexity in the sense of facing effects of motion in a very large diversity of forms, starting from analyzing simple motion in a changing environment (illumination, shadows, ...), under adverse observation conditions, such as bad signal-to-noise ratio (low illumination, small-scale processes, low-dose x-ray, etc.), covering also multiple motions of independent objects, occlusions, and going as far as dealing with objects which are complex in themselves (articulated objects such as bodies of living beings). The spectrum of problems includes, but does not end at, objects which are not 'bodies' at all, e.g., when analyzing fluid motion, cloud motion, and so on. Analyzing the motion of a crowd in a shopping mall or in an airport is a further example that implies the need to struggle against the problems induced by complexity. We cannot be sure that the named or similar application areas already represent the high-end of complexity in motion analysis – actually, there will probably be even harder problems in the analysis of complex visual motion which we have not faced yet, or which we have not yet dared to address.

Based on the observation that the current state of the art in the field of motion analysis is in a rather advanced shape already, but also taking into account that there are so many real-life problems which have not been solved yet, a group of researchers from different German research institutions decided to initiate an international workshop to attract renowned scientists and young researchers from different areas of visual motion analysis.

Therefore, in October 2004, the 1st International Workshop on Complex Motion (IWCM 2004) was held at Schloss Günzburg, a beautiful mansion and scientific convention center administered by the University of Ulm (South Germany). The Steering Committee of IWCM 2004 aimed at inspiring and encouraging the members of the computer vision community to share experiences and exchange opinions on the contemporary development of the field. There were several invited talks given by renowned senior researchers who not only appreciated the historic development of research in visual motion, but demonstrated and discussed the current grand challenges in a vivid and stimulating way.

This workshop was particularly devoted to advancing the repertoire of methods dealing with complex visual motion and to initiating a more intensive and hopefully continuing discussion amongst leading experts in the field. The topics of presentations were optical flow, local motion estimation for image signals affected by strong disturbances, structure from motion, multicamera flow analysis, dynamic stereo, fluid motion analysis, the estimation of multiple motions, motion tracking, and many other areas where complex visual motion patterns have to be evaluated. In fact, there were several plenary discussions which covered open issues, unsolved problems, and also different opinions in a highly constructive manner and which apparently ignited many further discussions adjacent to the official workshop programme, and presumably initiated exchange and cooperation between researchers who had not been in direct contact before.

The workshop was organized by the members of the LOCOMOTOR Project (Nonlinear analysis of multidimensional signals: LOcal adaptive estimation of COmplex MOTion and ORientation patterns), which is part of the priority research program SPP 1114 "Mathematical methods for time series analysis and digital image processing" supported by the *German Research Council (DFG).* We particularly appreciate the generous support that our workshop received from the *German Pattern Recognition Association (DAGM, Deutsche Arbeitsgemeinschaft für Mustererkennung)* and its president, Prof. Dr. Hans Burkhardt (University of Freiburg). Without this generous support this workshop would not have been possible. We also appreciate the kind support of Springer for giving us the opportunity to distribute the scientific essence of this workshop to the computer science and engineering community via a volume in the Springer LNCS series. We hope that the present compilation of research papers presented at IWCM 2004 reflects the diversity of challenges, the prosperity of the research field, and possibly also a bit of the enjoyable atmosphere we shared at Schloss Günzburg.

November 2006

Rudolf Mester
Bernd Jähne
Erhardt Barth
Hanno Scharr

Organization

Steering Committee

Rudolf Mester	Frankfurt University (Programme Chair, overall management)
Bernd Jähne	Heidelberg University (Workshop Proceedings)
Erhardt Barth	Lübeck University
Hanno Scharr	Research Center Jülich

Sponsoring Institutions

German Pattern Recognition Association (DAGM)
German Research Council (DFG)

Table of Contents

Optical Flow Estimation from Monogenic Phase

Michael Felsberg*

Linköping University, Computer Vision Laboratory,
SE-58183 Linköping, Sweden
mfe@isy.liu.se
http://www.isy.liu.se/cvl/

Abstract. The optical flow can be estimated by several different methods, some of them require multiple frames some make use of just two frames. One approach to the latter problem is optical flow from phase. However, in contrast to (horizontal) disparity from phase, this method suffers from the phase being oriented, i.e., classical quadrature filter have a predefined orientation in which the phase estimation is correct and the phase error grows with increasing deviation from the local image orientation. Using the approach of the monogenic phase instead, results in correct phase estimates for all orientations if the signal is locally 1D. This allows to estimate the optical flow with sub-pixel accuracy from a multi-resolution analysis with seven filter responses at each scale. The paper gives a short and easy to comprehend overview about the theory of the monogenic phase and the formula for the displacement estimation is derived from a series expansion of the phase. Some basic experiments are presented.

1 Introduction

The aim of this paper is not only to present just another optical flow estimation approach, but also to give a tutorial-like introduction to the topic of the monogenic signal and its phase. We try to strip off all theoretic background which is unnecessary and focus on a simple, concrete, and complete description of the framework. For a more formal treatment, we refer to the earlier publications on the monogenic framework [1,2,3].

Based on the monogenic phase, we then derive a simple formula for measuring oriented displacements between two images, which is then applied in a multi-scale approach for the estimation of the optical flow. This method is quite similar to the one presented in [4], with two differences: We do not know the direction of displacement in advance and we show the pure estimates, i.e., we do not post-process the point estimates with channel smoothing. Combining the point-wise estimator and a non-linear smoothing technique easily compensates errors in the estimates and partly the aperture effect, but it was not our aim to present a complete motion estimation system but rather the signal processing part of it.

2 Why Is Phase-Based Image Processing Preferable?

Before we start to introduce the framework of the monogenic phase, we motivate why we want to use phase-based methods at all. Basically, there are three reasons:

---

* This work has been supported by EC Grant IST-2003-004176 COSPAL and by EC Grant IST-2002-002013 MATRIS.

B. Jähne et al. (Eds.): IWCM 2004, LNCS 3417, pp. 1–13, 2007.

1. There is strong evidence that the human visual system makes use of local phase in V1 [5].
2. Phase-based processing is to a large extend invariant to changes of lighting conditions. The local image intensity can additionally be used to measure the reliability of measurements.
3. Perceptually, the reconstruction of an image from phase information is much better than that from amplitude information.

Reasons 2 and 3 have to be explained in some more detail. Note that we always consider a local region in the spatial-frequency domain, i.e., we look at local image regions at certain frequency ranges (or equivalently: at certain scales). Note in this context that we denote spatial frequency or wavenumber by 'frequency', not a frequency in the temporal sense.

2.1 The Image Model of Local Phase

Our image model that we apply in phase-based processing is

$$I(\mathbf{x}) = \tilde{A}(\mathbf{x}) \cos(\tilde{\varphi}(\mathbf{x})) + \bar{I} \tag{1}$$

where $\mathbf{x} = (x_1, x_2)^T$ indicates the spatial coordinate vector, $I(\mathbf{x})$ the image, $\bar{I}$ the average intensity (DC level), $\tilde{A}(\mathbf{x})$ the instantaneous amplitude (real-valued, non-negative), and $\tilde{\varphi}(\mathbf{x})$ the instantaneous phase [6]. The average intensity is irrelevant for the analysis of the image contents and in the human visual system it is largely compensated already during the image acquisition. In this model the decomposition into $\tilde{A}(\mathbf{x})$ and $\tilde{\varphi}(\mathbf{x})$ seems to be highly ambiguous. This is however not the case, since the amplitude is a non-negative real number. Hence, the zeros of $I(\mathbf{x}) - \bar{I}$ must be covered by zeros of $\cos(\tilde{\varphi}(\mathbf{x}))$. Assuming sufficiently smooth functions, the zero crossings are in direct correspondence to the full phase [7] and the instantaneous phase becomes a uniquely defined feature.

If we switch to a local region in the spatial-frequency domain, i.e., we consider an image region and a small range of frequencies, the model (1) becomes much simpler. Under the assumption of small magnitude variations of the considered frequency components in the local spectrum, the amplitude becomes approximately constant in the local region. It is therefore referred to as the local amplitude $A_{\mathbf{x}}$, where $\mathbf{x}$ now indicates the origin of the local region, i.e., all estimates with subscript $\mathbf{x}$ refer to the region with origin $\mathbf{x}$. The model (1) becomes for the local region

$$\tilde{I}(\mathbf{x}+\mathbf{y}) = A_{\mathbf{x}} \cos(\varphi_{\mathbf{x}}(\mathbf{y})), \tag{2}$$

where $\tilde{I}(\mathbf{x}+\mathbf{y})$ is the local image patch, $\mathbf{y}$ the local patch coordinate vector, and $\varphi_{\mathbf{x}}(\mathbf{y})$ is the local phase.

Assuming a small range of frequencies, the local phase cannot vary arbitrarily fast, it has a high degree of smoothness. Therefore, it can be well approximated by a first order series in $\mathbf{y} = 0$:

$$\varphi_{\mathbf{x}}(\mathbf{y}) \approx \mathbf{f}_{\mathbf{x}}^T \mathbf{y} + \varphi_{\mathbf{x}}(0) = f_{\mathbf{x}} \mathbf{n}_{\mathbf{x}}^T \mathbf{y} + \varphi_{\mathbf{x}}(0), \tag{3}$$

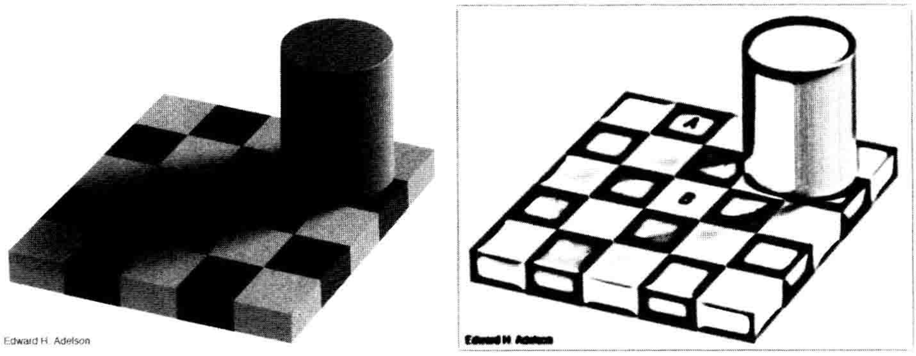

Fig. 1. The checker shadow illusion 'disillusioned'. Left: original image from *http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html*. Right: reconstruction from local phase.

where $f_{\mathbf{x}}$ is the local frequency, $\mathbf{n}_{\mathbf{x}}$ is the local orientation (unit vector), and $\varphi_{\mathbf{x}}(0)$ is some phase offset. That means, our local image model directly led to the assumption of intrinsic dimensionality one [8], or simple signals [6], where $\tilde{I}(\mathbf{x}+\mathbf{y}) = \hat{I}_{\mathbf{x}}(\mathbf{n}_{\mathbf{x}}^T\mathbf{y})$ ($\hat{I}$ being a suitable 1D function).

We can group the series of the local phase (3) in two different ways:

$$\varphi_{\mathbf{x}}(\mathbf{y}) \approx f_{\mathbf{x}}(\mathbf{n}_{\mathbf{x}}^T\mathbf{y}) + \varphi_{\mathbf{x}}(0) = \bar{\varphi}_{\mathbf{x}}(\mathbf{n}_{\mathbf{x}}^T\mathbf{y}) \qquad \text{and} \tag{4}$$

$$\varphi_{\mathbf{x}}(\mathbf{y}) \approx \mathbf{n}_{\mathbf{x}}^T(\mathbf{n}_{\mathbf{x}}\mathbf{n}_{\mathbf{x}}^T\mathbf{y} f_{\mathbf{x}} + \mathbf{n}_{\mathbf{x}}\varphi_{\mathbf{x}}(0)) = \mathbf{n}_{\mathbf{x}}^T\mathbf{r}_{\mathbf{x}}(\mathbf{y}). \tag{5}$$

Whereas the former expression is a 1D function with a scalar product as an argument, the latter expression is a 2D vector field, the local phase vector $\mathbf{r}_{\mathbf{x}}$, which is projected onto the orientation vector. Although the distinction seems to be trivial, the local phase vector is simpler to estimate, because we do not need to know the local orientation in advance. We will return to this issue later.

2.2 Lighting Invariance and Perceptual Image Contents

Before we present an estimation procedure for the local phase, we continue the discussion on the reasons 2 and 3 for using phase-based methods. The decomposition of an image into local amplitude and local phase at a particular frequency range means to neglect the high frequencies, to represent the intermediate frequencies by the local phase, and to cover the lower frequencies by (more global) changes of the local amplitude. Changes of lighting conditions are then to a large extend represented by changes of the local amplitude, with the exception of moving shadow boundaries. Hence, most changes of lighting conditions are not visible in the phase representation, cf. Fig. 1.

Since the famous paper [9], it is well known that the reconstruction from the phase spectrum is much better from a perceptional point of view than the one from the magnitude spectrum. Considering the Fourier spectrum means to consider the part of the spatial-frequency domain with maximum localization in frequency and no localization in position. If we move to some point with finite localization in both spaces, the results

Fig. 2. Decomposing an image into its local phase and its local amplitude. From left to right: $\cos(\varphi_{\mathbf{x}_0})$, $A_{\mathbf{x}_0}$, $\tilde{I}(\mathbf{x}_0)$, where the intensities are adjusted to obtain similar intensity ranges. Grey means zero, white means positive values, and black means negative values. Top row: full size images. Bottom row: image detail. $\tilde{I}(\mathbf{x}_0)$ is obtained from $I(\mathbf{x}_0)$ applying the filters from [10] using scales $\{3,4,5\}$.

of the experiment from [9] still remain valid, cf. Fig. 2, although we now consider the local phase.

If the image is decomposed into its amplitude and phase information, it becomes evident that the local amplitude is basically just a measure for the confidence of the extracted phase, i.e., in technical terms it represents the signal-to-noise ratio (SNR), cf. Fig. 2, center column. The local phase represents most of the image structure, cf. Fig. 2, left column. In the areas where the amplitude is close to zero, thus meaning 'no confidence', the local phase contains mainly noise. In the regions of non-zero confidence, the cosine of the local phase results in a visual impression which comes very close to the bandpass filtered image, cf. Fig. 2, right column.

3 The Monogenic Signal: A Survey

The monogenic signal provides a framework to estimate the local phase, the local orientation, and the local amplitude of an image [3]. It can be considered as a 2D generalization of the analytic signal. Former 2D generalizations tried to estimate the local phase according to (4), which was only partly successful since the local orientation has to be known in advance to steer the filters [11,12]. In the monogenic framework, however, we

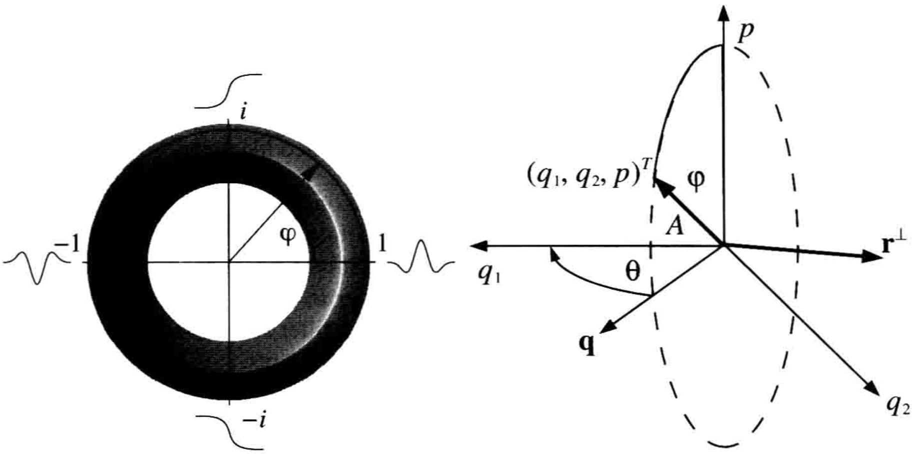

Fig. 3. Local phase models. Left: the 1D phase, the corresponding filter shapes at 1, i, -1, and $-i$, and the continuously changing signal profile (grey values in the background). Right: the local phase of the monogenic signal. The 3D vector $(q_1, q_2, p)^T$ together with the p-axis define a plane at orientation θ in which the rotation takes place. The normal of this plane multiplied by the rotation angle φ results in the rotation vector $\mathbf{r}^\perp$.

try to estimate the local phase vector (5) instead, where the local orientation is part of the result and need not be known in advance. The estimated vector has a natural upper bound for its error [1].

3.1 Spherical Quadrature Filter

Quadrature filters in 1D are constructed in two steps:

1. Select a suitable bandpass filter which is responsible for the localization in the time-frequency domain, i.e., it responds to signal contributions which are in a certain frequency range (the passband) and in a certain time window. This bandpass filter is an even filter, i.e., it is symmetric.
2. Compute the Hilbert transform of the bandpass filter in order to construct the corresponding odd, i.e., antisymmetric, filter which has the same magnitude spectrum.

Practical problems concerning computing the Hilbert transform are out of the scope of this paper. The quadrature filter pair is mostly applied as a complex filter to the 1D signal. The response is a complex signal, which is divided into magnitude and argument. One can easily show that the argument is an estimate for the local phase. In Fig. 3 left, the 1D phase interpretation is illustrated.

The figure is generated by projecting all possible phase-responses onto the filter. As a result we get those input signals which would generate the in-fed responses. Keeping the amplitude constant and varying the phase from 0 to 2π results in the sketched signal profiles (at 1, i, -1, and $-i$) and the continuously varying intensities in the background. With increasing phase angle, the quadrature filter turns from a purely even filter towards a purely odd filter. Continuing further than $\pi/2$ leads towards an even filter again, but with opposite sign. After π the filter becomes first odd (opposite sign) and finally it turns back into the initial filter. The corresponding signal profile changes from a positive impulse over a positive step (from inside to outside), over a negative impulse, and over a negative step, until it is a positive impulse again.

The 2D spherical quadrature filters (SQF) are constructed likewise:

1. Select a suitable *radial* bandpass filter, i.e., a rotation invariant filter. The passband consists of frequencies of a certain range in absolute value, but with arbitrary direction. This bandpass filter is an even filter, i.e., it is symmetric.
2. Compute the *Riesz* transform of the bandpass filter in order to construct the corresponding odd, i.e., antisymmetric about the origin, filter which has the same magnitude spectrum. The odd filter consists of two components.

The radial bandpass filter is given by some suitable frequency response $B_{\mathrm{e}}(\rho)$, where $(\rho,\ \phi)$ are the polar coordinates of the frequency domain, such that it is rotational symmetric and therefore symmetric (even) about the origin. The corresponding antisymmetric (odd) filters are then given by

$$B_{\mathrm{o}1}(\rho,\phi) = i\cos\phi\, B_{\mathrm{e}}(\rho) \qquad \text{and} \qquad B_{\mathrm{o}2}(\rho,\phi) = i\sin\phi\, B_{\mathrm{e}}(\rho). \tag{6}$$

All together, an SQF provides three responses; the even filter response $p(\mathbf{x}) = (I * b_{\mathrm{e}})(\mathbf{x})$ and the two odd filter responses $\mathbf{q}(\mathbf{x}) = (q_1(\mathbf{x}),\ q_2(\mathbf{x}))^T = ((I * b_{\mathrm{o}1})(\mathbf{x}),\ (I * b_{\mathrm{o}2})(\mathbf{x}))^T$.

3.2 Extracting Local Phase

The local amplitude can be extracted likewise as in the 1D case by calculating the magnitude of the 3D vector:

$$A_{\mathbf{x}} = \sqrt{q_1(\mathbf{x})^2 + q_2(\mathbf{x})^2 + p(\mathbf{x})^2}, \tag{7}$$

cf. Fig. 4 for an example. The phase, however, cannot be extracted as the argument of a complex number, since we need two angles to describe the 3D rotation from a reference point (on the p-axis) into the SQF response. These angles are indicated in Fig. 3, and they have direct interpretations in terms of local orientation and local phase.

It has been shown in [3] that an image patch with intrinsic dimensionality one and local orientation θ (w.r.t. the horizontal axis) results in a response of the form $(q_1(\mathbf{x}),\ q_2(\mathbf{x}),\ p(\mathbf{x}))^T = (\cos\theta\, q(\mathbf{x}),\ \sin\theta\, q(\mathbf{x}),\ p(\mathbf{x}))^T$ for a suitable $q(\mathbf{x})$, i.e., according to Fig. 3, the rotation takes place in a plane which encloses angle θ with the $(q_1,\ p)$-plane. For non-zero $\mathbf{q}$ this angle can hence be estimated as

$$\theta_{\mathbf{x}} = \tan^{-1}\left(\frac{q_2(\mathbf{x})}{q_1(\mathbf{x})}\right) \in (-\pi/2; \pi/2], \tag{8}$$

where an orientation – direction ambiguity occurs, since the directional sense of a 2D signal cannot be extracted from a local signal [6], i.e., $\mathbf{q}$ and $-\mathbf{q}$ map onto the same orientation, cf. Fig. 4.