

Ville Ojansivu

BLUR INVARIANT PATTERN
RECOGNITION AND
REGISTRATION IN
THE FOURIER DOMAIN

FACULTY OF TECHNOLOGY,
DEPARTMENT OF ELECTRICAL AND INFORMATION ENGINEERING,
UNIVERSITY OF OULU;
INFOTECH OULU,
UNIVERSITY OF OULU



ACTA UNIVERSITATIS OULUENSIS
C Technica 339

VILLE OJANSIVU

**BLUR INVARIANT PATTERN
RECOGNITION AND REGISTRATION
IN THE FOURIER DOMAIN**

Academic dissertation to be presented with the assent of
the Faculty of Technology of the University of Oulu for
public defence in Auditorium TS101, Linnanmaa, on 23
October 2009, at 12 noon

OULUN YLIOPISTO, OULU 2009

Copyright © 2009
Acta Univ. Oul. C 339, 2009

Supervised by
Professor Janne Heikkilä

Reviewed by
Professor Karen Egiazarian
Professor Jan Flusser

ISBN 978-951-42-9254-5 (Paperback)
ISBN 978-951-42-9255-2 (PDF)
<http://herkules oulu.fi/isbn9789514292552/>
ISSN 0355-3213 (Printed)
ISSN 1796-2226 (Online)
<http://herkules oulu.fi/issn03553213/>

Cover design
Raimo Ahonen

OULU UNIVERSITY PRESS
OULU 2009

Ojansivu, Ville, Blur invariant pattern recognition and registration in the Fourier domain.

Faculty of Technology, Department of Electrical and Information Engineering, University of Oulu, P.O.Box 4500, FI-90014 University of Oulu, Finland; Infotech Oulu, University of Oulu, P.O.Box 4500, FI-90014 University of Oulu, Finland

Acta Univ. Oul. C 339, 2009

Oulu, Finland

Abstract

Pattern recognition and registration are integral elements of computer vision, which considers image patterns. This thesis presents novel blur, and combined blur and geometric invariant features for pattern recognition and registration related to images. These global or local features are based on the Fourier transform phase, and are invariant or insensitive to image blurring with a centrally symmetric point spread function which can result, for example, from linear motion or out of focus.

The global features are based on the even powers of the phase-only discrete Fourier spectrum or bispectrum of an image and are invariant to centrally symmetric blur. These global features are used for object recognition and image registration. The features are extended for geometrical invariances up to similarity transformation: shift invariance is obtained using bispectrum, and rotation-scale invariance using log-polar mapping of bispectrum slices. Affine invariance can be achieved as well using rotated sets of the log-log mapped bispectrum slices. The novel invariants are shown to be more robust to additive noise than the earlier blur, and combined blur and geometric invariants based on image moments.

The local features are computed using the short term Fourier transform in local windows around the points of interest. Only the lowest horizontal, vertical, and diagonal frequency coefficients are used, the phase of which is insensitive to centrally symmetric blur. The phases of these four frequency coefficients are quantized and used to form a descriptor code for the local region. When these local descriptors are used for texture classification, they are computed for every pixel, and added up to a histogram which describes the local pattern. There are no earlier textures features which have been claimed to be invariant to blur. The proposed descriptors were superior in the classification of blurred textures compared to a few non-blur invariant state of the art texture classification methods.

Keywords: bispectrum, blur invariant features, computer vision, image alignment, image blur, image registration, object recognition, pattern classification, texture analysis

Preface

The research for this thesis was carried out in the Machine Vision Group of the Department of Electrical and Information Engineering at the University of Oulu, Finland during the years 2005-2009.

I would like to express my gratitude to Professor Janne Heikkilä for supervising the thesis. Without his invaluable ideas and guidance the work would have never been possible. I am also grateful to Dr. Esa Rahtu who contributed valuable ideas to the research. I am indebted to my former colleague Dr. Tuukka Toivonen for advices on countless matters related to the research, and engineering in general. I would also like to thank Professor Matti Pietikäinen for allowing me to work in his research group.

I am grateful to Professor Jan Flusser and Professor Karen Egiazarian for reviewing the thesis manuscript. Their findings resulted in clarification of some important theoretical issues. Thanks also to Gordon Roberts for the language revision.

I would like to thank my colleagues in the Machine Vision Group for creating a pleasant atmosphere. Especially, I would like to mention the refreshing coffee room discussions about things related to research and beyond.

The generous financial support provided for this thesis by the Infotech Oulu Graduate School, the Academy of Finland, the Nokia Foundation, the Emil Aaltonen Foundation, the Tauno Tönning Foundation, Seppo Säynäjäkangas Science Foundation, and the Riitta and Jorma J. Takanen Foundation is gratefully acknowledged.

Finally, I would like to thank my mother Riitta and father Tapio for their unconditional support and encouragement over the years. I also wish to thank the rest of my family and friends for the counterbalancing moments outside the research. My warmest thanks belong to Outi for all the support during my studies.

Oulu, September 2009

Ville Ojansivu

Abbreviations

$\mathcal{B}(\mathbf{u})$	Proposed blur invariant feature
c_{pq}	Complex image moment of order (p, q)
$C(p, q)$	Central moment based blur invariant of order (p, q)
$f(\mathbf{n})$	Image intensity function
\mathbf{f}	Vector of stacked rows or columns of image $f(\mathbf{n})$
$F(\mathbf{u})$	Discrete Fourier transform of $f(\mathbf{n})$
F^*	Complex conjugate of F
F^I	Imaginary part of F
F^R	Real part of F
$g(\mathbf{n})$	Blurred image intensity function
\mathbf{g}	Vector of stacked rows or columns of image $g(\mathbf{n})$
$G(\mathbf{u})$	Discrete Fourier transform of $g(\mathbf{n})$
$h(\mathbf{n})$	Point spread function of blur
$h(\mathbf{x})$	Continuous point spread function of blur
$H(\mathbf{u})$	Frequency response of $h(\mathbf{n})$
\mathbf{H}	Superposition blur point spread function matrix
$J(\mathbf{f})$	Criterion function
n	Positive integer
\mathbf{n}	2-D discrete coordinate vector
$T(\mathbf{u})$	Blur invariant feature based on phase-tangent
$\mathcal{T}(\mathbf{u})$	Proposed blur-translation invariant feature
\mathbf{u}	2-D discrete frequency vector
$w(\mathbf{n})$	Image noise
\mathbf{w}	Vector of stacked rows or columns of noise $w(\mathbf{n})$
$W(\mathbf{u})$	Discrete Fourier transform of $w(\mathbf{n})$
\mathbf{x}	2-D continuous coordinate vector
$\phi_f(\mathbf{u})$	Phase spectrum of image $f(\mathbf{n})$
μ_{pq}	Central image moment of order (p, q)
1-D	One-dimensional
2-D	Two-dimensional
3-D	Three-dimensional

BIPC	Blur invariant phase correlation
CLS	Constrained least squares
DFT	Discrete Fourier transform
FFT	Fast Fourier transform
LBP	Local binary pattern
LSI	Local shift invariant
LPQ	Local phase quantization
PSF	Point spread function
SIFT	Scale invariant feature transform
STFT	Short term Fourier transform

List of original articles

- I Ojansivu V & Heikkilä J (2006) Motion blur concealment of digital video using invariant features. Proc. 8th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 2006), Antwerp, Belgium. Lecture Notes in Computer Science 4179: 35-45.
- II Ojansivu V & Heikkilä J (2007) Object recognition using frequency domain blur invariant features. Proc. 15th Scandinavian Conference on Image Analysis (SCIA 2007), Aalborg, Denmark. Lecture Notes in Computer Science 4522: 243-252.
- III Ojansivu V & Heikkilä J (2007) Image registration using blur-invariant phase correlation. IEEE Signal Processing Letters 14(7): 449-452.
- IV Ojansivu V & Heikkilä J (2007) Blur invariant registration of rotated, scaled and shifted images. Proc. 15th European Signal Processing Conference (EUSIPCO 2007), Poznań, Poland: 1755-1759.
- V Ojansivu V & Heikkilä J (2007) A method for blur and similarity transform invariant object recognition. Proc. 14th International Conference on Image Analysis and Processing (ICIAP 2007), Modena, Italy: 583-588.
- VI Ojansivu V & Heikkilä J (2008) A method for blur and affine invariant object recognition using phase-only bispectrum. Proc. 3rd International Conference on Image Analysis and Recognition (ICIAR 2008), Póvoa de Varzim, Portugal. Lecture Notes in Computer Science 5112: 527-536.
- VII Ojansivu V & Heikkilä J (2008) Blur insensitive texture classification using local phase quantization. Proc. 3rd International Conference on Image and Signal Processing (ICISP 2008), Cherbourg-Octeville, France. Lecture Notes in Computer Science 5099: 236-243.
- VIII Ojansivu V & Rahtu E & Heikkilä J (2008) Rotation Invariant Local Phase Quantization for Blur Insensitive Texture Analysis. Proc. 19th International Conference on Pattern Recognition (ICPR 2008), Tampa Bay, Florida, USA: 1-4.
- IX Ojansivu V & Heikkilä J (2009) Weighted DFT based blur invariants for pattern recognition. Proc. 16th Scandinavian Conference on Image Analysis (SCIA 2009), Oslo, Norway. Lecture Notes in Computer Science 5575: 71-80.

The writing of Papers I-VI and IX was mainly carried by the author, who was also in charge of the experiments in these papers. The ideas were processed together with Prof. Heikkilä. The writing of Paper VII was carried by the author with Prof. Heikkilä. The methods were developed together, while the author was in charge of the experiments. The writing of Paper VIII was carried by the author and Dr. Rahtu, who also provided most of the ideas. The author was in charge of the experiments of the paper.

Contents

Abstract	
Preface	5
Abbreviations	7
List of original articles	9
Contents	11
1 Introduction	13
1.1 Background	13
1.2 Motivation	17
1.3 The contribution of the thesis	18
1.4 Summary of original papers	19
2 Image blur and deblurring	21
2.1 Blur modeling	21
2.2 Deblurring	25
3 Global blur and geometric invariants	31
3.1 Global geometric invariants	31
3.1.1 Moment invariants	32
3.1.2 Invariance in the Fourier domain	33
3.1.3 Other invariants	36
3.2 Blur invariant features in the spatial domain	36
3.2.1 Blur invariants based on image moments	36
3.2.2 Other blur invariants in the spatial domain	40
3.3 Fourier domain blur invariants	41
3.3.1 Blur invariants based on phase tangent	41
3.3.2 Weighted blur invariants	43
3.3.3 Novel Fourier domain blur invariants	43
3.3.4 Blur invariant phase correlation	44
3.3.5 Blur and shift invariants	45
3.3.6 Blur and similarity transformation invariants	46
3.3.7 Blur and affine invariant object recognition	47
3.4 Discussion	50
	11

4	Local blur insensitive descriptors	53
4.1	Local descriptors	53
4.2	Blur insensitive descriptors	57
4.2.1	Local phase quantization	57
4.2.2	Rotation invariant local phase quantization.....	62
4.3	Discussion	63
5	Applications	65
5.1	Preprocessing for multichannel deconvolution	65
5.2	Image fraud detection	67
5.3	Face recognition	68
5.4	Stereo matching	70
6	Conclusions	73
	References	74
	Original articles	81

1 Introduction

1.1 Background

According to Jain *et al.* (2000), "pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and make sound and reasonable decisions about the categories of the patterns". In this context, the word pattern may refer to many different concepts such as a handwritten letter, fingerprint, human face, speech signal, or the structure of a surface.

Humans deal with various patterns such as text, images, and objects everyday, and humans by far are the best pattern recognizers in most situations. The more relevant the patterns are the better humans can recognize them (Jain *et al.* 2000). Humans can easily recognize very dissimilar and partially occluded instances of the same pattern. In a limited setting, computers can also recognize patterns, but despite about half a century of research, a general purpose pattern recognizer remains a dream (Jain *et al.* 2000). Computers are better suited for tasks that are labor intensive or handle patterns that are very hard to distinguish through our psychophysical senses. The use of a computer may be justified also with economical reasons. Pattern recognition can be applied for example in medicine, biology, economics, signal processing, and computer vision. For an optimal solution to a particular problem, it may be necessary to combine several sensing modalities (Jain *et al.* 2000). In computer vision, the patterns are image intensity functions and cameras act as sensors. Typical pattern recognition application areas for computer vision are medical imaging, biometric recognition, remote sensing, data mining, and industrial automation. Nowadays it is possible to recognize very challenging patterns such as images of human faces or expressions (Li & Jain 2005).

Two fundamental tasks in computer vision are recognition and registration of images. In the following, recognition is discussed first. In image recognition, the objective is to classify a pattern into one of the possible classes. The pattern to be recognized is typically an object or texture. By texture we mean the repetitive structure in the scene, such as the surface structure of grass or a crop field in an aerial image. Depending on the task, the goal may be to recognize a given pattern to be a particular pattern, as in object recognition, or a representative of a class of patterns, as in classification of texture or categorization of an object into one of the classes of, for example, "bicycles"

and "cars".

There are several approaches to pattern recognition, and their suitability often depends on the type of the application. Jain *et al.* (2000) divide pattern recognition methods into four groups, which are not completely independent: template matching, syntactic or structured matching, artificial neural networks, and statistical classification. Template matching is the simplest of the methods. A pattern is matched against a stored template or prototype of the pattern taking into account different poses and scales, for example, by using correlation. Generally, template matching is computationally demanding.

The syntactic approach to recognition assumes that patterns are built from small sub-patterns called primitives. The patterns are then represented by the primitives and their interrelations. This is analogous to languages which are formed from words, letters of an alphabet, and a grammar. The approach often leads to difficulties in segmenting the primitives and inference of the grammar (Jain *et al.* 2000). In the case of texture, the primitives are called textons. Tuceryan & Jain (1998) divide the methods of the syntactic approach into those that use textons as features utilizing their statistical distributions, and into others that use relative placement rules of the textons.

Artificial neural networks are weighted directed graphs in which the nodes are artificial neurons, and the weighted edges connect the outputs and inputs of neurons of different levels. The networks can learn complex non-linear input-output relationships for a specific classification task and may require relatively less domain specific knowledge from the user. Also features of statistical methods can be used with neural networks (Jain *et al.* 2000).

The pattern recognition methods covered in this thesis extract statistical features from patterns, and therefore belong to the statistical approach. Statistical recognition is the most studied and applied approach in practice (Jain *et al.* 2000). In the statistical approach, the pattern is represented by d features or measurements as a point in the d -dimensional space. The idea is to extract features that describe well the pattern of a specific class and separate the patterns of different classes in the feature space. Decision boundaries are established in the feature space to bound the different classes.

Statistical pattern recognition has typically three steps: preprocessing, feature extraction, and classification. The purpose of the preprocessing is to segment or separate the pattern of interest from the surrounding pattern or back ground, and possibly perform some operations to represent the data compactly. The feature extraction step creates the representation of the object using the features. Finally, the classification step assigns the

input pattern to a specific object class. The most straightforward approach is to use the nearest neighbor classification. In this method, the object is assigned to the closest class in the feature space. The closest class can be found, for example, by using a Euclidean distance measure between the feature vectors.

In addition, the statistical recognition system has two operating modes: training and classification, the latter being the normal operating mode. In the training phase, the features to be used in classification are selected, and the classifier learns the decision boundaries between the classes in the feature space. The training can be done in a supervised or unsupervised manner. In supervised learning, the training samples are labeled according to their class. In unsupervised learning, the object classes are deduced from the data.

This thesis concentrates on the feature extraction step in the case of gray scale images. It is assumed that the patterns are already segmented from the background. The treatment is further divided into global features computed from the whole image, and local features. The local features can be computed in every image location, or only around interest points that contain salient details. The given application often dictates the type of features used.

In image registration, the objective is to find the transformation parameters between the given images representing the same scene at different times, from different view-points, or using different sensors, and then overlaying the images. The transformation is often assumed to be a shift, rotation, scaling, or affine transformation. The applications of image registration include, for example, fusion of images of different sensors, image mosaicing, and super-resolution image formation.

Zitová & Flusser (2003) divide the image registration methods into area based global methods and methods based on local interest points. The area based methods can be divided into template matching type methods, and methods that use global features. Template matching is used in a similar way to that used in object recognition, and the registration parameters can be deduced from the best matching pose of the template. In the case of global features, the transformation parameters are estimated in the feature domain. The advantage of feature based methods is that they can be made to be invariant to specific image distortions which alter the pure image intensity values.

Interest point based registration methods detect salient details in the images which are subsequently matched between the images, typically by using suitable local features. The transformation parameters can be inferred from the matching interest points after false matches are discarded. The interest points can be high-contrast closed boundary

regions, lines, or points which are scattered around the image (Zitová & Flusser 2003).

After computation of the transformation parameters, the images are aligned using appropriate interpolation of the pixel values. Global registration methods are more suitable for images which do not contain much salient details. This is often the case with medical images (Zitová & Flusser 2003). On the other hand, local interest point based methods can better handle non-global transformations and occlusions of the images. This thesis concentrates on global feature based registration of gray scale images.

Image recognition and registration are not trivial tasks in practice, since the captured images are degraded due to various factors, such as varying illumination, view angle, occlusions, noise, and blur. An ideal pattern recognition system should be invariant to the deformations of the objects, but still discriminate between different kinds of objects. Obviously, there is a trade-off between these two. Providing a general solution to the problem which is feasible for every situation is difficult. Typically, it is assumed that image degradations are limited. In this thesis, we are specifically interested in recognition and registration of blurred gray scale images using features that are invariant to blur, i.e., blur invariants for short. However, the features can be simultaneously invariant also to geometrical transformations and illumination variations. Image blurring may be caused, for example, by motion between the camera and the scene during the exposure, an out of focus lens, or atmospheric turbulence (Banham & Katsaggelos 1997). Figure 1(a) shows an example image which is motion blurred in Figure 1(b).



Fig 1. (a) Original image and (b) a motion blurred version of the same image.

1.2 Motivation

As mentioned in the previous section, an ideal image pattern recognition system should be invariant to image blur among other image degradations. There are basically two options to achieve blur independent analysis of images. One option is first to deblur the images, and after that use ordinary image analysis methods. Unfortunately, image deblurring is a very difficult problem which in practice introduces new artifacts to the image altering them even more (Banham & Katsaggelos 1997). For this reason, the second option, direct blur invariant analysis of the images is justified. However, despite the vast amount of research on invariant pattern recognition (Wood 1996), there are not many known methods for blur invariant image analysis.

Most of the earlier methods for blur invariant pattern recognition use the blur invariant features based on image moments proposed by Flusser & Suk (1998). These global features are invariant to centrally symmetric blurring of images. Also combined moment invariants, which, in addition to blur invariance, are invariant to translation, scaling, rotation, or even affine transformation of the images have been proposed (Flusser & Zitová 1999, Suk & Flusser 2003). The moment invariants are quite efficient to compute for image recognition applications. A shortcoming of the moment invariants is their sensitivity to noise and clutter of the images, which is due to the inherent properties of the geometric image moments. The moment invariants can be used also for template-matching type registration, but in this case the computational load is more burdensome as the invariants have to be evaluated in every image location.

Also Fourier domain blur invariant features have been proposed, but their practical utilization is limited due to stability problems, lack of geometrical invariances (Flusser & Suk 1998), or problems related to practical implementation (Lu & Yoshida 1999).

All the above mentioned image moment and Fourier transform based invariants, having various deficiencies, are computed globally from the image as a whole. To the author's best knowledge, blur insensitive local features or blur insensitive texture analysis in general have not been mentioned in the literature. In principle, the moment invariants could be used also for local image analysis, but they are better suited to global pattern recognition.

1.3 The contribution of the thesis

Previously, blur invariant pattern analysis carried out in the Fourier domain has been difficult due to practical deficiencies of the methods. This is unfortunate since Fourier analysis has some desirable properties, such as robustness to noise, compared to the methods based on the geometric moments. The contributions of this thesis are related to the development of new blur invariant features in the Fourier domain. Based on novel ideas, we have developed methods for blur invariant object and texture recognition as well as registration.

A shortcoming of the earlier blur invariants based on the Fourier transform phase was their instability (Flusser & Suk 1998). To overcome this problem, we developed a method to weight the invariants so that the effect of unstable coefficients is minimized. Later, we derived the invariants in a different way to overcome the instability problem completely. Similar to the earlier blur invariants, also our new features are computed globally and are invariant to centrally symmetric blur, which includes motion, out of focus, and atmospheric turbulence blur as special cases (Banham & Katsaggelos 1997). The novel Fourier domain invariants were extended for translation, scale, and rotation invariance. A similar global approach can be utilized for blur invariant phase correlation registration. The registration method was extended also for scaled and rotated images. In addition, we developed a method for blur and affine invariant object recognition in the Fourier domain. In extensive experiments, all of these methods were compared with good results to other non-blur invariant and blur invariant pattern recognition and registration approaches.

Blur insensitive local features were also developed and applied to texture classification. Although the features are based on phase, the approach is different compared to the global features. Here the phase is computed locally and quantized, and the resulting features are presented as a histogram. We also developed a method for estimating the local characteristic orientation of a pattern, which can be used in conjunction with the local features to obtain rotation invariance. This method was utilized for blur and rotation invariant texture analysis. The proposed methods gave good results in comparative experiments with state of the art non-blur invariant texture classification methods.

The methods developed in this thesis are presented in the context of recognition and registration of 2-D patterns; more precisely image intensity functions. However, the methods could possibly also be utilized for other kinds of patterns, for example 1-D signals, color or multispectral images, and volumetric data.

1.4 Summary of original papers

This thesis consists of nine publications. Paper I presents a method of weighting the earlier blur invariant moment and Fourier phase based features according to their estimated reliability. The weighting especially improves the results obtained with Fourier phase invariants, but also the results of the moment invariants are improved. Furthermore, an application of the invariants for blur concealment in the background of an image sequence is presented. The experiments support the use of the weighting scheme, emphasizing its necessity for Fourier domain invariants.

Paper II proposes a new approach to constructing global blur invariant features using the Fourier transform phase. The approach is also extended to include shift invariance using a phase-only bispectrum. The features are applied to object recognition. The experiments show that these invariants are superior to the earlier Fourier domain invariants, and they also outperform the moment invariants in the presence of noise. For practical image sizes, the proposed invariants are also much faster to compute than the moment invariants.

Paper III presents an approach to performing blur invariant phase correlation at subpixel accuracy to be used for image registration. This method uses a similar blur invariant form of the Fourier phase as in Paper II. The experiments show that the proposed method outperforms the traditional phase correlation in the registration of blurred images. The method also gives slightly better registration results compared with moment invariants, and the required computation time is only a fraction.

Paper IV proposes a method for blur invariant registration of rotated, scaled, and shifted images using global features based on the bispectrum phase. The rotation and scaling parameters are solved using log-polar sampling of the blur invariant form of the phase-only bispectrum, after which the translation component is solved using the blur invariant phase correlation proposed in Paper III. According to the experiments, the proposed method results in more accurate registration of blurred images compared with similar methods without blur invariance.

Paper V introduces blur, rotation, scale, and shift invariant global features for object recognition. The method is largely similar to the registration method in Paper IV, but here only a small number of features is used, instead of correlating the blur invariant bispectrum slices. In addition, the shift invariance property requires an additional step. Experimental comparisons of the proposed invariants with moment invariants having similar invariance properties confirm their supremacy for noisy images.

Paper VI proposes a method for object recognition which is invariant to blur and an affine transformation of the image. The method operates globally on a rotated set of log-log sampled blur invariant bispectrum slices. Although this method has the invariance properties, it does not produce invariant features as the methods in Paper II and V did. According to experiments, the method outperforms the blur and affine invariant moment based features when images are noisy or contain a slight perspective projection.

Paper VII introduces a completely new approach to constructing local blur insensitive features. The features are applied for blurred texture classification. This method is insensitive to moderate centrally symmetric blur. The method utilizes low frequency phase coefficients of the short term Fourier transform which is computed locally. The phase coefficients are quantized and coded using 8-bits. For texture classification, these codes are computed for every pixel and presented as a histogram for the whole image. The method outperforms many state of the art texture classification methods even for sharp images, and for blurred images the method is superior.

Paper VIII presents an approach to constructing blur and rotation insensitive local features. First, a method for estimating the local characteristic orientation is described. Then we present an oriented version of the descriptor proposed in Paper VII, which is computed in the direction corresponding to the estimated local orientation. The method was applied for blurred and rotated texture classification. According to the comparisons with other state of the art texture classification methods, the proposed method is superior in the case of blurred textures, and similar or slightly better for sharp textures.

Paper IX presents a method for weighting of the novel Fourier domain blur invariants of Paper II according to their estimated reliability. The reliability is based on the estimated signal-to-noise ratio of the noisy invariants. The method is similar to that presented in Paper I for earlier blur invariants; however, here the presentation is more rigorous. The weighting improved significantly the classification results of the novel Fourier phase invariants.

2 Image blur and deblurring

Blur is generated to images by the non-ideal image capturing process. The effect of blurring is the same as lowpass filtering the image. Typical causes of blurring are motion between the camera and the scene during exposure, and an out of focus optical system. In addition to the blur, the captured image always contains some amount of noise. The noise results from various factors such as non-ideal sensors and quantization of the digital images. The blur invariants discussed later in this thesis are based on the assumption that the blur obeys a certain model. In the following, we discuss how the typical image blurs can be modeled. We also discuss shortly the difficulties of image deblurring, which is an alternative to blur invariant image analysis.

2.1 Blur modeling

In general, the image degradation process caused by blur and noise, which results in observed image $g(\mathbf{n})$, can be described using the following discrete model (Banham & Katsaggelos 1997):

$$g(\mathbf{n}) = \sum_{\mathbf{k}} h(\mathbf{n}, \mathbf{k}) f(\mathbf{k}) + w(\mathbf{n}), \quad (1)$$

where $f(\mathbf{n})$ is the ideal image of the scene, $h(\mathbf{n}, \mathbf{k})$ is the 2-D point spread function (PSF) of the imaging system and $\mathbf{n} = (n_1, n_2)^T$ and $\mathbf{k} = (k_1, k_2)^T$ are 2-D discrete coordinate vectors. The PSF spreads out each intensity point of the ideal image to a larger area blurring the image. Additive noise is represented by $w(\mathbf{n})$. It is very difficult to deal with spatially varying blur like in (1). However in many practical cases, the image blur can be assumed to be spatially invariant. This kind of blur results from relative motion between the camera and the scene, being out of focus, or atmospheric turbulence. In this case, the degradation process can be modeled as a discrete linear shift-invariant (LSI) system given by

$$\begin{aligned} g(\mathbf{n}) &= \sum_{\mathbf{k}} h(\mathbf{n} - \mathbf{k}) f(\mathbf{k}) + w(\mathbf{n}) \\ &= (f * h)(\mathbf{n}) + w(\mathbf{n}), \end{aligned} \quad (2)$$

where $*$ denotes 2-D convolution and $h(\mathbf{n})$ is spatially invariant PSF of the blur. This

blurring process can be presented equivalently in the Fourier domain as suggested by the convolution theorem, namely

$$G(\mathbf{u}) = F(\mathbf{u}) \cdot H(\mathbf{u}) + W(\mathbf{u}), \quad (3)$$

where capitals denote discrete Fourier transforms (DFT) and $\mathbf{u} = (u_1, u_2)^T$ is a 2-D coordinate vector in the discrete Fourier domain.

In the following, we present spatially continuous PSFs $h(\mathbf{x})$ of various blur forms which typically occur in practice (Lagendijk & Biemond 2005, Banham & Katsaggelos 1997). With $\mathbf{x} = (x_1, x_2)^T$ we denote a 2-D continuous coordinate vector. Discrete PSF $h(\mathbf{n})$ is derived by sampling $h(\mathbf{x})$ appropriately.

In an ideal image capturing process without blur, the PSF is the Dirac delta function $h(\mathbf{x}) = \delta(\mathbf{x})$, which equals a unit pulse in a discrete case. In practice, there is always some blurring in images due to non-zero aperture and other non-idealities so that PSF has greater support.

Motion blur results from the relative motion between the camera and the scene during exposure. The motion can be, for example, in the form of translation, rotation, rapid zooming, or some combination of these. Here, we consider only linear translational motion with constant velocity, which is encountered in many practical situations and can be modeled using (2). In this case, the image of any point source is a line and the PSF of the blur is given in the continuous form by

$$h(\mathbf{x}) = \begin{cases} \frac{1}{L}, & \text{if } x_1 \sin \theta - x_2 \cos \theta = 0, (x_1 - \frac{L}{2} \cos \theta)^2 + (x_2 - \frac{L}{2} \sin \theta)^2 \leq (\frac{L}{2})^2 \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where L is the length of the blur, which equals the velocity of the motion multiplied by the exposure time, and θ is the angle of the motion direction (Lagendijk & Biemond 2005). Later on, we assume that $h(\mathbf{x})$ is shifted so that the center of the non-zero values is at the origin. This does not have any effect on the blurring, but is the same as assuming the blurred image shifted by distance $\frac{L}{2}$ in the direction opposite to the motion. In this case, $h(\mathbf{x}) = h(-\mathbf{x})$, which is the requirement for the blur invariant features discussed in Sections 3.2, 3.3, and 4.2. Figure 2(a) illustrates a discrete motion blur PSF $h(\mathbf{n})$ corresponding to $L = 8$ and $\theta = \pi/4$ and Figure 3(a) shows the corresponding magnitude of its DFT $|H(\mathbf{u})|$. Motion blurring corresponds to lowpass filtering in the direction of the motion. The cross-section of $H(\mathbf{u})$ in this direction is an oscillating

sinc-function with zeros in intervals approximately equal to N/L if the dimensions of the image and DFT are N -by- N . Figure 4 shows a simple image which is blurred in Figure 5(a) by motion blur PSF of Figure 2(a).

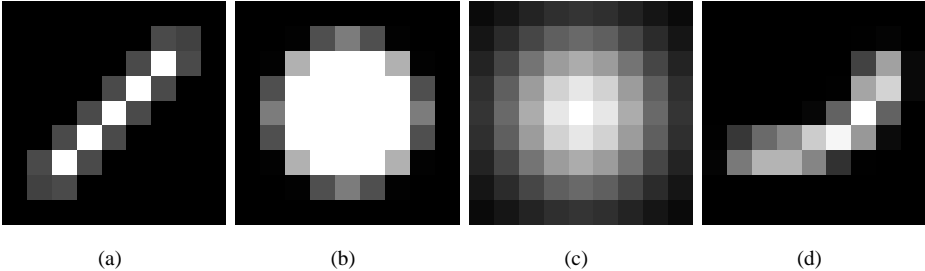


Fig 2. Discrete PSFs: (a) motion blur, (b) out of focus blur, (c) Gaussian blur, and (d) arbitrary blur.

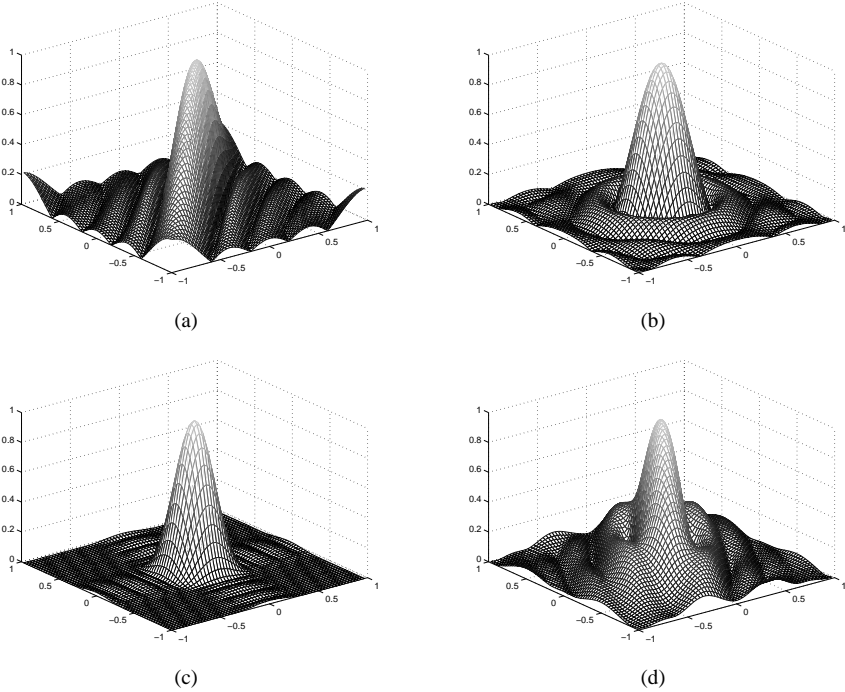


Fig 3. Amplitude responses of the filters (a)-(d) in Figure 2.

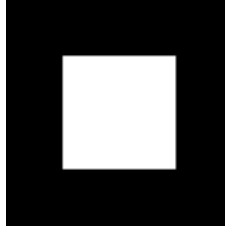


Fig 4. A simple image.

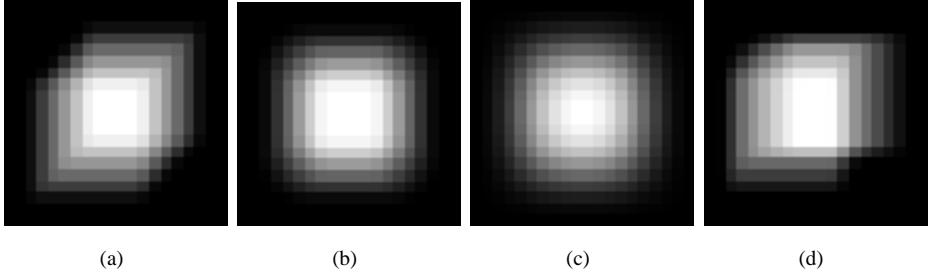


Fig 5. Image of Figure 4 blurred using filters of Figure 2(a-d): (a) motion blur, (b) defocus blur, (c) Gaussian blur, and (d) arbitrary blur.

Out of focus blur appears in the image if the lens system of the camera is not focused for the correct distance to the scene. Because the aperture of the lens is circular, the blurred image of a point source is a small disk. In general, the extent of the blur depends on the focal length, the aperture size, and the distance between the camera and the object, and it is not constant in the whole image. When the scene is relatively planar and orthogonal to the camera, the out of focus blur can be estimated by a uniform and circular PSF, which is spatially invariant according to (2), and is given in the continuous form by

$$h(\mathbf{x}) = \begin{cases} \frac{1}{\pi R^2}, & \text{if } \sqrt{x_1^2 + x_2^2} \leq R \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where R is the radius of the blur (Lagendijk & Biemond 2005). Figure 2(b) illustrates a discrete out of focus blur PSF $h(\mathbf{n})$ corresponding to $R = 3$. Figure 3(b) illustrates the corresponding magnitude response $|H(\mathbf{u})|$ in the Fourier domain, again with an oscillating shape, and lowpass behavior. This time zero-pattern is circular. Figure 5(b) illustrates an image which is out of focus blurred using the PSF in Figure 2(b).

Atmospheric turbulence or Gaussian blur is encountered for example in remote sensing. The PSF of the blur depends on various factors, such as the air temperature and the wind speed, but for relatively long exposure time, it can be estimated using a Gaussian distribution, given by

$$h(\mathbf{x}) = C \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma_G^2}\right), \quad (6)$$

where variance σ_G^2 determines the severity of the blur, and constant C is chosen so that the sum of the coefficients is equal to one (Legendijk & Biemond 2005). Figure 2(c) illustrates a discrete atmospheric turbulence blur PSF $h(\mathbf{n})$ corresponding to $\sigma_G^2 = 9/4$ and Figure 3(c) shows the corresponding magnitude response $|H(\mathbf{u})|$. In the continuous case, the magnitude response would be a pure 2-D Gaussian, but in the discrete case, it contains some oscillation due to the finite support of the PSF. Figure 5(c) illustrates an image which is Gaussian blurred using the PSF in Figure 2(c). Figure 2(d) shows an arbitrary blur PSF, the magnitude response being shown in Figure 3(d). The corresponding blurred image is illustrated in Figure 5(d).

The existing blur invariant image features are based on the assumption that the PSF of the blur is spatially invariant (2) and centrally symmetric with respect to the origin. These conditions are met by the above mentioned forms of motion, out of focus, and atmospheric turbulence blur PSFs.

2.2 Deblurring

Image restoration, or deblurring, is concerned with the reconstruction of the original uncorrupted image from the blurred and noisy image by an inverse operation. The operation is different to artificial image sharpening. For spatially invariant blurring, the deblurring is done by deconvolution. There are many problems in deblurring including the estimation of the PSF of the blur and the difficulties of deconvolution.

It is possible to present the LSI blurring process (2) in a vector-matrix form, as follows (Gonzales & Woods 1992):

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{w}. \quad (7)$$

Here the rows or columns of the N_1 -by- N_2 image $g(\mathbf{n})$ are stacked into the N_1N_2 -by-1 vector \mathbf{g} . N_1N_2 -by-1 vectors \mathbf{f} and \mathbf{w} are similarly formed. The N_1N_2 -by- N_1N_2 matrix \mathbf{H} represents the superposition blur operator. With appropriate zero padding, the linear

convolution can be made equivalent to circular convolution, which enables the application of frequency domain techniques. In this case, \mathbf{H} has a block circulant structure (Banham & Katsaggelos 1997).

If \mathbf{H} is known, a straightforward approach to solving (7) would seem to be a simple matrix inverse. However, the operation is ill-conditioned, and solution existence, uniqueness, and stability become problems (Karl 2005). When null space of \mathbf{H} is empty, a direct approach to solution existence is to find an estimate $\hat{\mathbf{f}}$ that minimizes the criterion function

$$J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2, \quad (8)$$

with respect to \mathbf{f} , where $\|\cdot\|_2$ denotes l^2 norm resulting in least square fit to the data (Banham & Katsaggelos 1997). This leads to the generalized inverse filter solution

$$\hat{\mathbf{f}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{g}. \quad (9)$$

The main problem with the generalized inverse filter is its instability due to noise amplification. This can be noticed by examining the inverse filter solution in the Fourier domain (Gonzales & Woods 1992), namely

$$\hat{F}(\mathbf{u}) = \frac{G(\mathbf{u})}{H(\mathbf{u})} = F(\mathbf{u}) + \frac{W(\mathbf{u})}{H(\mathbf{u})}. \quad (10)$$

The difficulties arise when $H(\mathbf{u})$ becomes very small or zero. In this case, the inverse filtered noise, on the right in the equation, is amplified strongly. This poses a problem especially because the zeros of $H(\mathbf{u})$ exist also at high frequencies (as shown in Figure 3) where the signal energy is relatively small. As a result, the restored image becomes dominated by noise at these frequencies.

If $H(\mathbf{u})$ becomes exactly zero at some frequencies, the inverse filter is not defined. In this case, the null space of \mathbf{H} is not empty and images differing at these frequencies will produce identical observations and there is not a unique solution. At these frequencies, the solution can be set to zero which corresponds to choosing the solution of minimum energy (Karl 2005, Banham & Katsaggelos 1997).

Regularization is used in image restoration to manage the ill-conditioned deconvolution. Regularization can stabilize the solution, producing physically meaningful estimates. The basic idea is to constrain the least squares solutions somehow by inclusion of prior knowledge in the form of a side constrain (Karl 2005).

The constrained least squares (CLS) solution can be formulated by minimizing the Lagrangian

$$J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \alpha \|\mathbf{L}\mathbf{f}\|_2^2, \quad (11)$$

where $\mathbf{L}\mathbf{f}$ represents a high pass filtered version of the image \mathbf{f} (Banham & Katsaggelos 1997). This is essentially a constrain that promotes the smoothness of the image by minimizing the amount of high pass energy in the restored image. It is assumed that in most practical images the energy is concentrated at lower frequencies and high frequency energy is limited. Typically \mathbf{L} is a 2-D Laplacian operator. In (11), α is the Lagrangian multiplier that controls the trade-off between fidelity to the data and smoothness of the solution (Banham & Katsaggelos 1997).

Minimization of (11) leads to the CLS estimate

$$\hat{\mathbf{f}} = (\mathbf{H}^T\mathbf{H} + \alpha\mathbf{L}^T\mathbf{L})^{-1}\mathbf{H}^T\mathbf{g}, \quad (12)$$

which can be presented also in the frequency domain (Gonzales & Woods 2002) by

$$\hat{F}(\mathbf{u}) = \frac{H^*(\mathbf{u})}{H^*(\mathbf{u})H(\mathbf{u}) + \alpha L^*(\mathbf{u})L(\mathbf{u})}G(\mathbf{u}), \quad (13)$$

where $L(\mathbf{u})$ is the DFT of the Laplacian operator. The influence of the term $\alpha L^*(\mathbf{u})L(\mathbf{u})$ becomes stronger when $H(\mathbf{u})$ approaches zero, thus suppressing the noise amplification. CLS filtering can be performed also iteratively.

LSI restoration filters result in two kinds of errors in the restored image. These are image dependent ringing, or a halo around the edges of the image and filtered noise, which can be seen best in the flat regions of the image. There is always a trade-off between these two types of errors, as can be seen in the following equation where \mathbf{G} represents the restoration filter:

$$\begin{aligned} \hat{\mathbf{f}} &= \mathbf{G}\mathbf{g} = \mathbf{G}\mathbf{H}\mathbf{f} + \mathbf{G}\mathbf{w} \\ &= \mathbf{f} + (\mathbf{G}\mathbf{H} - \mathbf{I})\mathbf{f} + \mathbf{G}\mathbf{w} \\ &= \mathbf{f} + \mathbf{e}_r + \mathbf{e}_w. \end{aligned} \quad (14)$$

When regularization of \mathbf{G} is used to attenuate the noise \mathbf{e}_w , at the same time the restoration filter will deviate from the inverse of \mathbf{H} , and their product is not unity, resulting in

image dependent ringing \mathbf{e}_r . Adaptive restoration can be used to overcome the restrictions of the LSI restoration (Banham & Katsaggelos 1997).

In addition, also a boundary value problem causes error: the result of the convolution of the ideal image and the blur PSF always extends beyond the borders of the observed image captured by a finite size sensor. So, part of the information needed to restore the image is not available. Because the PSF of the restoration filter has a very large support, typically comparable to the image size, this error is not only seen at the borders but it causes oscillatory error in the whole image. In spatial domain restoration, the missing data outside the observed image can be estimated by extrapolating the image. In frequency domain restoration, the Fourier transform assumes periodicity of the image, which can be interpolated from top to bottom and from left to right to make it transit smoothly (Lagendijk & Biemond 2005). Reeves (2005) proposes a method that sums the results of two restorations to cancel the errors. Iterative restoration techniques can explicitly include also the unknown borders into the model, but they are computationally more demanding (Reeves 2005).

Total variation regularization is a regularization method which has achieved popularity in recent years (Karl 2005). The criterion to be minimized is

$$J(\hat{\mathbf{f}}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \alpha \|\mathbf{D}\mathbf{f}\|_1, \quad (15)$$

where \mathbf{D} represents discrete approximation of a gradient operator so that the elements $\mathbf{D}\mathbf{f}$ are brightness changes in the image, and the sum of these values is the total amount of change the signal goes through, i.e., total variation. Total variation regularization penalizes the total amount of gradient in the image. So, the result can contain localized steep gradients preserving image edges. When the local derivative is small, greater smoothness is imposed in these regions, while for regions with a large local derivative, larger gradients are allowed in the solution. The difficulty of total variation regularization is that it leads to a non-linear optimization problem, which has to be solved iteratively (Karl 2005).

In the discussion above, it was assumed that PSF of the blur $h(\mathbf{n})$ is known, which is not the case in general. This leads to a blind image deconvolution problem in which both the blur PSF and the original image are estimated. Typically the deblurring result is unsatisfactory if the estimate of the PSF support size deviates only a few percent from the correct size (Zamfir *et al.* 2007). It may be possible to deduce the PSF analytically, based on the parameters of the camera motion, defocus etc., but usually the PSF is

deduced from the observed image.

There are two main approaches for the blind deconvolution. First, identifying the PSF separately from the image, and using it with classical deblurring algorithms; and secondly, simultaneously estimating the PSF and the original image. The first class of the methods commonly assumes a parametric model of the blur, such as those described in Section 2.1, and estimates the parameters by investigating points or edges in the blurred image or the zero pattern in the Fourier domain (Kundur & Hatzinakos 1996). The latter class of the methods include approaches such as autoregressive moving average (ARMA) parameter estimation methods using maximum-likelihood estimation or generalized cross validation and nonparametric deterministic constrained methods which are iterative (Kundur & Hatzinakos 1996). One approach to the problem is multichannel blind deconvolution, which combines information from different images, blurred by different PSFs to overcome the problem of the frequency domain zeros at specific locations. Typically, the images need to be accurately registered for multichannel approaches, but also a multichannel blind deconvolution method which does not require accurate registration of the images has been proposed (Šroubek & Flusser 2005).

As can be noted from this short presentation, deblurring of the images presents lots of challenges and possible sources of new artifacts. Although the research on image deblurring has a very long tradition, and very sophisticated methods have been developed, the original image can never be exactly recovered due to the nature of the problem. This justifies a blur invariant analysis of the images.

3 Global blur and geometric invariants

As discussed in Chapter 2, images contain various degradations due to non-ideal imaging system and the prevailing imaging conditions. In Section 2.1, we presented the model of image blurring in the LSI system (2). Often images of the same scene differ also due to viewpoint and illumination change. In general, the relation between an observed image $g(\mathbf{n})$ and the ideal image $f(\mathbf{n})$ is given by $g(\mathbf{n}) = \mathcal{D}[f(\mathbf{n})]$, where \mathcal{D} is a degradation operator. In the case of an LSI system, \mathcal{D} can be given by

$$g[\tau(\mathbf{n})] = (f * h)(\mathbf{n}) + w(\mathbf{n}), \quad (16)$$

where $\mathbf{n}' = \tau(\mathbf{n})$ represents the transformation of the spatial coordinates, $h(\mathbf{n})$ is the PSF of the imaging system, and $w(\mathbf{n})$ is additive noise.

In this chapter, we discuss global image features that are invariant to blur and geometrical transformations, given by $h(\mathbf{n})$ and $\tau(\mathbf{n})$, respectively. The features are computed from the whole image, and are thus referred to as global: their value changes if the pattern is altered locally at any place. These features can be used to recognize and register images. In Section 3.1, we present the most essential global geometric invariants that are the foundation of the spatial and Fourier domain combined blur and geometric invariants presented in Sections 3.2 and 3.3, respectively.

3.1 Global geometric invariants

A pinhole camera is often used to approximate cameras with a real lens system. For a pinhole camera, the transformation of the image coordinates of planar objects under general motion is given by a non-linear projective transformation, which has eight parameters. The features that are used for pattern recognition are typically invariant to simpler linear transformations, such as, shift, rotation, and scaling, which together define a four parameter similarity transformation. Similarity transformation can be used for approximation of the projective transformation if the optical axis is perpendicular to a planar object. The features may also be invariant to a linear six parameter affine transformation which approximates well the projective transformation in the case of a planar object, when the distance between the object and the camera is large compared to the object dimensions.

3.1.1 Moment invariants

An important class of geometric invariants is based on the image moments. By definition, the central moments are invariant to image shift. For discrete image $f(\mathbf{n})$, the (p, q) 'th central moment is given by

$$\mu_{pq} = \sum_{\mathbf{n}} (n_1 - \bar{n}_1)^p (n_2 - \bar{n}_2)^q f(\mathbf{n}), \quad (17)$$

where (\bar{n}_1, \bar{n}_2) is the centroid of the image. For the regular moments, $(\bar{n}_1, \bar{n}_2) = (0, 0)$. Normalized central moments, given by

$$v_{pq} = \frac{\mu_{pq}}{\mu_{00}^{(p+q+2)/2}}, \quad (18)$$

are invariant to shift and scaling.

The first moment invariants that were used for visual pattern recognition were proposed by Hu (1961), who introduced seven moment invariants to rotation of 2-D objects consisting of second and third order moments. The four simplest of these invariants are given below:

$$\Phi_1 = \mu_{20} + \mu_{02}, \quad (19)$$

$$\Phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, \quad (20)$$

$$\Phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \quad (21)$$

$$\Phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2. \quad (22)$$

By substituting μ_{pq} in (19-22) with normalized central moments v_{pq} , invariants to similarity transformation, i.e., shift, rotation, and scaling, are obtained.

There are various applications and improvements of these invariants (Flusser 2006). The moment invariants have been applied for instance to character and object recognition (Belkasim *et al.* 1991) and image registration (Goshtasby 1985). Hupkens & de Clippeleir (1995) proposed normalization of the moment invariants for contrast invariance. Wong *et al.* (1995) presented a method of generating moment invariants of up to the fifth order. Wallin & Kiibler (1995) introduced a method for the generation of complete sets of rotation invariants using Zernike moments. Abu-Mostafa & Psaltis (1984) proposed the use of complex moments for rotation invariance. Complex moments are given by

$$c_{pq} = \sum_{\mathbf{n}} (n_1 + in_2)^p (n_1 - in_2)^q f(\mathbf{n}), \quad (23)$$

where i denotes the imaginary unit.

Flusser (2000) suggested forming the rotation invariants as follows: let $n \geq 1$ and k_i , p_i , and q_i be non-negative integers so that $\sum_{i=1}^n k_i(p_i - q_i) = 0$. Then the rotation invariants are obtained by

$$I = \prod_{i=1}^n c_{p_i, q_i}^{k_i}. \quad (24)$$

Flusser (2000) also introduced a method of forming a basis for the invariants, which is the smallest set of the invariants by which all others can be expressed.

Reiss (1991) and Suk & Flusser (1993) introduced independently moment invariants to affine transformation of images. Later Suk & Flusser (2004) proposed a graph based method of generating invariants systematically. Yang & Cohen (1999) proposed affine invariants based on cross-weighted moments. This approach results in more robust invariants than those based on regular moments, at the expense of computational cost.

3.1.2 Invariance in the Fourier domain

The global invariants presented in this subsection are based on 2-D Fourier transform of the images. 2-D discrete Fourier transform (DFT) of an N -by- N image is given by

$$F(\mathbf{u}) = \sum_{\mathbf{n}} f(\mathbf{n}) e^{-i2\pi \mathbf{u}^T \mathbf{n} / N}. \quad (25)$$

The Fourier domain invariants utilize the translation, rotation, and scaling properties of the DFT:

$$f(n_1 - t_1, n_2 - t_2) \Leftrightarrow F(u_1, u_2) e^{-i2\pi(u_1 t_1 + u_2 t_2) / N}, \quad (26)$$

$$f(s_1 n_1, s_2 n_2) \Leftrightarrow \frac{1}{|s_1 s_2|} F\left(\frac{u_1}{s_1}, \frac{u_2}{s_2}\right), \text{ and} \quad (27)$$

$$f(r, \theta + \theta_0) \Leftrightarrow F(\rho, \omega + \theta_0), \quad (28)$$

where $n_1 = r \cos \theta$, $n_2 = r \sin \theta$, $u_1 = \rho \cos \omega$, and $u_2 = \rho \sin \omega$.

Because translation of an image only has an effect on the phase, the magnitude $|F(\mathbf{u})|$ of the DFT is invariant to translation. Thus coefficients $|F(\mathbf{u})|$ are translation invariant features.

A basic approach for matching translated images is to compute cross-correlation in the Fourier domain using correlation theorem, i.e., by computing a cross power spectrum followed by the inverse transform (Gonzales & Woods 2002). Cross-correlation can be used for both recognition and registration.

Because the information pertaining to the image displacement is solely included into the phase of the cross power spectrum, it is possible to match images using only the normalized cross power spectrum. This method, called phase correlation, results in a much sharper correlation peak, and is invariant to uniform scaling or level shift of the image brightness function. Phase correlation is preferred to cross-correlation in the presence of band limited noise and an unknown signal, because all spectral phase terms are treated equally. On the other hand, for white noise, the cross-correlation function, which is dominated by the largest spectral components is optimal (Kuglin & Hines 1975).

Foroosh *et al.* (2002) introduced an extension of phase correlation to subpixel accuracy. Castro & Morandi (1987) proposed a method for matching shifted and rotated images, but without true rotation invariance. They computed the normalized cross power spectrum as a function of rotation angle difference. For the correct rotation the inverse transformed cross power spectrum is close to the impulse function, the location of which gives the shift between the images.

Rotation invariance can be obtained by mapping the image into polar coordinates so that the rotation transforms to a circular shift. Invariance to the circular shift is achieved by extracting the DFT magnitude. Scale invariance can be obtained by using the Mellin transform which transforms the scale change into shift, using logarithmic mapping. Invariance to the resulting shift is obtained by extracting the DFT magnitude.

The combined circular-Fourier radial-Mellin transform, called the Fourier-Mellin transform for short, is invariant to rotation and scaling (Wood 1996). In the Fourier-Mellin transform, the image $f(\mathbf{n})$ is log-polar mapped into $f(\ln(r), \theta)$, where $r = \sqrt{n_1^2 + n_2^2}$ and $\theta = \arctan(n_2/n_1)$. This transforms the rotation into circular shift along θ -axis, and scale change into shift along $\ln(r)$ -axis. Thus, the DFT magnitude of the log-polar mapped image is invariant to rotation and scaling of the original image (Sheng & Duvernoy 1986, de Bougrenet de la Tocnaye & Ghorbel 2004). Sheng & Duvernoy (1986) show that the moment invariants of Hu (1961) are a special case of the Fourier-Mellin transform.

There are couple of ways of combining shift invariance with rotation and scale invariance. Sheng & Duvernoy (1986) centralize the images to compensate for the

shift, and then apply the Fourier-Mellin transform. However, the centralization is very sensitive to clutter in the image. Casasent & Psaltis (1976) extracted first the DFT magnitude of the image. According to (26-28), the DFT magnitude is invariant to shift of the image but inverse scaled and rotated. To obtain invariance to remaining scaling and rotation, they applied the Fourier-Mellin transform. The magnitude of the resulting transform is invariant to shift, rotation, and scaling. Capodiferro *et al.* (1987) correlated the log-polar sampled amplitude spectra for matching of the images. This approach can be used to extract transformation parameters of the images for registration (sheng Chen *et al.* 1994, Reddy & Chatterji 1996): The location of the correlation peak gives the rotation and scaling parameters, while the translation parameters are solved by inverse rotating and scaling the other image, and then evaluating cross-correlation between the images.

Ben-Arie & Wang (1998) proposed a method for affine invariant object recognition, although the method does not involve complete affine invariant features. They use shift invariant DFT magnitude, which is log-log mapped to transform the uneven scaling in two directions into shifts along the axes. The rotations related to affine transformation are handled by correlating multiple rotated versions of these log-log sampled amplitude spectra. The method can be modified also for registration.

The above mentioned Fourier-Mellin transform of the DFT magnitude is the approach most often used for similarity transformation invariance. This approach has some disadvantages. First, it does not use full information of the image spectrum, as only amplitude spectrum is used. By discarding phase spectrum, basically half of the information is lost. Secondly, the amplitude spectrum is more sensitive to illumination changes compared to the phase spectrum. Instead of amplitude spectrum, it is possible to use other higher order spectra, given by

$$\Psi_n(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = F^*(\mathbf{s}) \prod_{i=1}^n F(\mathbf{u}_i), \quad (29)$$

where \mathbf{u}_i , with $i = 1, \dots, n$, are vectors in the 2-D frequency space, and $\mathbf{s} = \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_n$. These higher order spectra are shift invariant (Chandran *et al.* 1997). When $n = 1$ we get the power spectrum and further with the value $n = 2$ we get the bispectrum.

The bispectrum can be used instead of the amplitude spectrum in conjunction with the Fourier-Mellin transform in similarity transformation invariant image recognition and registration (Tsatsanis & Giannakis 1992). While the bispectrum is shift invariant, it retains both amplitude and phase information, and unlike amplitude spectrum, the im-

age can be reconstructed from it, except for the shift (Dianat & Rao 1990, Sadler 1992). The bispectrum is four dimensional, and in practice only a 2-D slice of the bispectrum is used. This does not lose any essential information, and the image reconstruction is still possible (Petropulu & Pozidis 1998). It is also possible to use a phase-only bispectrum which gives robustness to illumination variations (Heikkilä 2004).

3.1.3 Other invariants

Some very simple invariants exist, even for affine transformation of the images, such as the mean gray value or the normalized gray value histogram. The problem with these invariants is their lack of discrimination power: there are too few of them and also they lose the information on the spatial distribution of the image intensity. Rahtu *et al.* (2005a) solved this problem by introducing the multiscale autoconvolution method. They form multiple scaled versions of the images and combine them by multiplications (Rahtu *et al.* 2005b) or convolutions (Rahtu *et al.* 2005a) to form new images which retain the affine relationships. By varying the scales, they can form arbitrarily many new images. Subsequently, they extract simple affine invariant features, such as the mean gray value, from these images and can get arbitrarily many features. These features also retain spatial information on intensity distribution.

Kadyrov & Petrou (2001) proposed a method called the trace transform to generate a large number of global invariants based on a particular combination of three functionals applied to image function. With an appropriate choice of the functionals, one can make the result invariant to a chosen transformation. They applied the method for similarity transformed images, and later for affine transformed images (Kadyrov & Petrou 2004). The method can also be used for registration of images by choosing functionals that are sensitive to the transformation of the images.

3.2 Blur invariant features in the spatial domain

3.2.1 Blur invariants based on image moments

Despite the vast amount of research on invariant pattern recognition, there are not many approaches that are invariant to image blurring. The most well known class of blur invariant features, proposed by Flusser & Suk (1998), is based on geometric image moments. These invariants combine the geometric invariance of the moment invariants,

presented in Section 3.1.1, with blur invariance. More precisely, the features are invariant to blur which has centrally symmetric PSF. As shown in Section 2.1, motion, out of focus, and atmospheric turbulence blur have centrally symmetric PSFs. The moment invariants were proposed first for motion blur (Flusser & Suk 1994), and for circularly symmetric blur (Flusser *et al.* 1996) separately. The combined blur and geometric moment invariants have been used for both recognition and registration of images.

The basic form of moment invariants $C(p, q)$ is the following (Flusser & Suk 1998): If $(p + q)$ is even, then

$$C(p, q) = 0. \quad (30)$$

If $(p + q)$ is odd, then

$$C(p, q) = \mu_{pq} - \frac{1}{\mu_{00}} \sum_{n=0}^p \sum_{\substack{m=0 \\ 0 < n+m < p+q}}^q \binom{p}{n} \binom{q}{m} C(p-n, q-m) \cdot \mu_{nm}, \quad (31)$$

where $p + q$ is the order of the invariants and μ_{pq} is a central moment of the image (17).

When moment invariants are computed using central moments, they are also invariant to image shift. Invariance to contrast change is achieved using normalization $C(p, q)/\mu_{00}$ and invariance to image scaling by replacing central moments μ_{pq} with normalized central moments given by (18). Flusser & Suk (1998) proposed also invariants in the Fourier domain, but note that they are unstable and cannot be used in practice (see Section 3.3.1).

The moment based blur invariants have been used for registration of satellite images (Flusser & Suk 1998), object recognition (Flusser *et al.* 1996), and for registration of rotated X-ray images by computing the invariants for several rotation angles (Bentoutou *et al.* 2002). In Paper I, the moment invariants were used for detecting a blurred, but otherwise unchanged background between two images in a sequence. The idea was to replace the blurred background in the image using a sharp image in the sequence. In this way, the need for deblurring was limited only to the moving objects. The authors proposed also a method of weighting the invariants according their estimated signal-to-noise ratio.

Stern *et al.* (2002) suggested two moment based methods for motion blur invariant object recognition. The first method recovers the moments of the original image using the relationship of the moments of the blur PSF and moments of the blurred image. The PSF of the blur must be known. The second method evaluates the central moments

normal to the motion direction which results in blur invariance, but requires the motion direction to be known.

Zhang *et al.* (2000b) used the same relationship of the central moments of the original and blurred image in conjunction with blur variant moments for a focus measure. Wee & Paramesran (2007) proposed blur invariant features based on the orthogonal Legendre moments. The derivation of the invariants resembles the derivation of the invariants in (Flusser & Suk 1998).

Some of the odd order rotation invariants proposed by Hu (1961) and Wong *et al.* (1995) are also invariant to blur (Flusser & Suk 1998). These include, for example, invariants (21) and (22). Flusser & Zitová (1999) introduced blur and rotation invariants based on the complex moments. These invariants are defined as follows: let $n \geq 1$ and k_i , p_i , and q_i be non-negative integers so that $(p_i + q_i)$ is odd and that $\sum_{i=1}^n k_i(p_i - q_i) = 0$. Then the invariants to blur and rotation are given by

$$J = \prod_{i=1}^n K(p_i, q_i)^{k_i}, \quad (32)$$

where $K(p_i, q_i)$ is defined similarly to $C(p, q)$ in (31) but using complex moments c_{pq} , given by (23), instead of central moments μ_{pq} . The authors also presented an approach to forming a complete basis of the invariants. Invariance to translation is achieved using central coordinates in the definition of the complex moments. Invariance to contrast change and scaling can be obtained similarly to the basic moment invariants of (31).

The combined blur and rotation invariants have been applied to the registration of satellite images (Flusser & Boldys 1999), and the estimation of the planar motion of the camera (Zitová & Flusser 2002). In these applications, the invariants are computed in a circular neighborhood of the control points which are then matched. Bentoutou & Taleb (2005) have used the odd order invariants of Hu (1961) and Wong *et al.* (1995) for matching the control points in registration of blurred digital subtraction angiography X-ray images.

Zhang *et al.* (2000a) developed a method for blur and affine invariant registration using normalization of the image into a canonical form, in conjunction with the blur invariant moments. Later Zhang *et al.* (2002b) proposed blur and affine invariants based on similar normalization. Suk & Flusser (2003) combined the affine moment invariants (Reiss 1991, Suk & Flusser 1993) and blur invariants (Flusser & Zitová 1999) for blur and affine invariants by the following formulation: if $A(\mu_{00}, \dots, \mu_{PQ})$ is an affine moment invariant, then $A(C(0,0), \dots, C(P,Q))$, where $C(0,0) = \mu_{00}$ and other

blur invariants $C(p, q)$ are defined by (31), is a combined blur-affine invariant. The performance of the invariants was demonstrated in the classification of deformed images of digits. Flusser *et al.* (2003) proposed also invariants for n -dimensional images and applied them to 3-D medical image registration.

Because features $C(p, q)$ are invariant to centrally symmetric blur, they also cannot distinguish centrally symmetric images. Flusser & Zitová (2004) tackled this limitation by proposing invariants to convolution with circularly symmetric PSF and rotation. These invariants can recognize also centrally symmetric objects, which are not circularly symmetric. Another limitation of the invariants is their sensitivity to noise and image clutter, which becomes more severe as the order of the invariants increases (Pawlak 1992).

One more limitation is the error due to boundary effect: pixels laying near the boundary of the blurred image are affected also by the pixels laying outside of the image due to blur. Boundary effect becomes more noticeable for larger blur PSFs, and typically leads to erroneous recognition if the ratio between PSF size and the image size is larger than 0.15 (Flusser & Suk 1998). Candocia (2004) has examined the effect of the boundary error on the blur invariant moment features in detail, and found out that the boundary effect results in a larger error for the moment invariants of higher orders.

Quite recently, Metari & Deschenes (2008) proposed two new classes of blur and geometric invariant descriptors. These descriptors are invariant to blur which has a PSF symmetric with respect to the diagonals, i.e., $h(x_1, x_2) = h(x_2, x_1)$. This includes out of focus and Gaussian blur, but not motion blur. The first class of the invariants is built on simple ratios of central moments, or Mellin transforms of the images, and they are invariant to blur, uniform scaling, and contrast change. Descriptors based on central moments are also invariant to image shift. The second class of the invariants is build on ratios of the central or complex moments of the images. In addition to shift invariance, the central moment based features are invariant to blur and anisotropic scaling, and the complex moment based invariants to blur, rotation and uniform scaling. According to Metari & Deschenes (2008), their invariants provide better discrimination power with much less computing time compared to the earlier moment based descriptors with similar geometric invariance properties (Flusser & Suk 1998, Flusser & Zitová 1999, Suk & Flusser 2003).

3.2.2 Other blur invariants in the spatial domain

A color image descriptor that is robust to blur, and invariant to illuminant color changes was introduced by van de Weijer & Schmid (2006). They reaffirm the observation that the ratios of some image derivatives are robust to smoothing with a Gaussian, and they utilize this to obtain blur robustness for specific color derivative based color constant ratios. In theory, the robustness is obtained for Gaussian blur, but according to their results, the method can also moderately tolerate other kinds of blurs.

There are also various methods for registration of blurred images which are actually not invariant, but robust to blur. Most of them assume a Gaussian PSF which may be varying spatially. Images of scenes with objects at various depths typically contain this kind of blur. These registration methods are often used to precede image fusion, which means the integration of the information contained in multiple images. Kubota *et al.* (1999) and Kubota & Aizawa (2005) proposed a two stage registration method of shifted, rotated, scaled and differently focused images in which the degree of blur may vary locally. In the first step, they form a Gaussian pyramid of the images and register them globally, starting from the lowest resolution using a full search and then refining the result gradually. The use of the Gaussian pyramid is supposed to reduce the effect of Gaussian blur. In the second step, they divide the images by rectangular grid into blocks. The block corners are locally matched between the images by choosing the best match among various Gaussian blurred versions of the images. The quadrilateral blocks are then mapped between the images by the locations of the four corners.

Myles & da Vitoria Lobo (1998) propose a method which iteratively recovers the parameters of affine motion and defocus blur. The blur PSF shape is assumed to be Gaussian or a disk. Arora *et al.* (2008) propose a method for image registration by alignment of the local phase iteratively at various spatial frequencies and directions. The local phase is computed at each pixel location by filtering the image with Gabor wavelets. An overdetermined system of equations is then used to recover any type of transformation between the images. The images have to be roughly registered for the algorithm to converge. The authors argue that the method is robust to smoothly varying illumination, noise and blur with an even and real PSF.

3.3 Fourier domain blur invariants

The blur invariants based on image moments, discussed in Section 3.2.1, have been extensively studied and have many applications. Although Flusser & Suk (1998) presented invariants to centrally symmetric blur also in the Fourier domain, these have received much less attention due to problems in their practical implementation. The Fourier domain invariants from Flusser & Suk (1998) are presented in Section 3.3.1. In the rest of Section 3.3, we introduce new developments of the Fourier domain invariants, which are among the main contributions of this thesis. These methods are covered in more detail in the original Papers I, II, III, IV, V, VI, and IX.

3.3.1 Blur invariants based on phase tangent

If the magnitude and phase parts of the image blurring process (3) are separated, and the noise term is neglected, this results in

$$|G(\mathbf{u})| = |F(\mathbf{u})| \cdot |H(\mathbf{u})| \quad \text{and} \quad (33)$$

$$\phi_g(\mathbf{u}) = \phi_f(\mathbf{u}) + \phi_h(\mathbf{u}), \quad (34)$$

where $\phi(\mathbf{u})$ denotes the phase spectrum. In the following, we consider only phase (34).

If we assume that the blur PSF $h(\mathbf{n})$ is centrally symmetric, namely $h(\mathbf{n}) = h(-\mathbf{n})$, its Fourier transform is always real-valued, and as a consequence its phase is only a two-valued function, given by

$$\phi_h(\mathbf{u}) = \begin{cases} 0 & \text{if } H(\mathbf{u}) \geq 0 \\ \pi & \text{if } H(\mathbf{u}) < 0. \end{cases} \quad (35)$$

Because of the periodicity of the tangent

$$\tan[\phi_g(\mathbf{u})] = \tan[\phi_f(\mathbf{u}) + \phi_h(\mathbf{u})] = \tan[\phi_f(\mathbf{u})]. \quad (36)$$

Thus, $\tan[\phi_g(\mathbf{u})]$ is invariant to the convolution of the original image with any centrally symmetric PSF. This includes blur models introduced in Section 2.1. We call these invariant features phase-tangent invariants and denote them by

$$T(\mathbf{u}) = \tan[\phi_f(\mathbf{u})] = \frac{F^I(\mathbf{u})}{F^R(\mathbf{u})}, \quad (37)$$

where F^I and F^R denote the imaginary and real parts of $F(\mathbf{u})$, respectively. These invariants were proposed for object recognition by Flusser & Suk (1998), but it should be noted that Hayes *et al.* (1980) suggested already earlier the blur invariant tangent of the phase to be used in phase only reconstruction of images blurred by a centrally symmetric PSF.

The moment invariants and the phase-tangent invariants have a relationship which can be presented by expanding the tangent of the Fourier transform phase into power series (Flusser & Suk 1998):

$$\tan[\phi_f(\mathbf{u})] = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_{kj} u_1^k u_2^j, \quad (38)$$

where $F(\mathbf{u}) \neq 0$, $\phi_f(\mathbf{u}) \neq \pm\pi/2$, and

$$a_{kj} = \frac{(-1)^{(k+j-1)/2} \cdot (-2\pi)^{k+1} \cdot M(k, j)}{k! \cdot j! \cdot m_{00}}, \quad (39)$$

when $k + j$ is odd and $a_{kj} = 0$ otherwise. $M(k, j)$ is similar to invariants (31) but constructed using regular moments.

Despite the relationship between the moment invariants and the Fourier domain invariants, their suitability for practical use is different. Similarly to the moment invariants, also the Fourier domain invariants suffer from error caused by boundary effect. However, a major disadvantage is the noise sensitivity of the phase-tangent invariants (Flusser & Suk 1998). This is due to the unstable properties of the tangent function near its discontinuing points:

$$\lim_{\phi \rightarrow \frac{\pi}{2}^-} \tan(\phi) = \lim_{\phi \rightarrow \frac{3\pi}{2}^-} \tan(\phi) = \infty, \text{ and} \quad (40)$$

$$\lim_{\phi \rightarrow \frac{\pi}{2}^+} \tan(\phi) = \lim_{\phi \rightarrow \frac{3\pi}{2}^+} \tan(\phi) = -\infty, \quad (41)$$

As can be seen in (40) and (41), when the real part of $F(\mathbf{u})$ is close to zero, only a small change in its value due to noise alters the value of $\tan[\phi_f(\mathbf{u})]$ significantly, even from ∞ to $-\infty$. Because of this instability, the Fourier phase-tangent invariants do not have any reported applications as such.

3.3.2 *Weighted blur invariants*

To overcome the noise sensitivity of the phase-tangent invariants $T(\mathbf{u})$ (37), Paper I introduces a method of weighting them according to their estimated signal-to-noise ratio. The weighting is done by using the Mahalanobis distance instead of the Euclidean distance between the feature vectors in classification of the images. The Mahalanobis distance is given by

$$\text{distance} = \mathbf{d} \cdot \mathbf{C}_T^{-1} \cdot \mathbf{d}^T, \quad (42)$$

where vector \mathbf{d} contains the differences of the invariants of two images and \mathbf{C}_T is the covariance matrix of the invariants. The covariance matrix of the invariants cannot be computed directly because the invariants are a non-linear function of the image data, and thus the covariance matrix is computed using linearization

$$\mathbf{C}_T \approx \mathbf{J} \cdot \mathbf{C} \cdot \mathbf{J}^T, \quad (43)$$

where \mathbf{C} is a $2N_T$ -by- $2N_T$ covariance matrix of $F^I(\mathbf{u}_i)$ and $F^R(\mathbf{u}_i)$, $i = 0, \dots, N_T$, and \mathbf{J} is a block diagonal Jacobian matrix containing the partial derivatives of the invariants $T(\mathbf{u}_i)$ with respect to $F^I(\mathbf{u}_i)$ and $F^R(\mathbf{u}_i)$. N_T is the number of invariants used.

The results of the experiments in Paper I show that weighting improves the classification results of the frequency domain phase-tangent invariants significantly close to the performance of the moment invariants. The classification results of the moment invariants were also improved slightly with the use of weighting.

3.3.3 *Novel Fourier domain blur invariants*

Despite the instability of the phase-tangent invariants, the Fourier domain phase has good properties for pattern recognition. Most of the important features of a signal are preserved by phase without amplitude information, and under certain conditions phase only information may even be used to reconstruct an image to within a scale factor (Oppenheim & Lim 1981). Magnitude spectrum, on the other hand, is quite similar between different images of real scenes.

In Paper II, the Fourier domain invariants were derived without the discontinuous tangent-function. Instead, it was noticed that the blur invariants can be constructed by raising the normalized Fourier transform to an even power. If noise is neglected, this

results in equality

$$\begin{aligned}
[e^{-i\phi_g(\mathbf{u})}]^{2n} &= e^{-i2n\phi_g(\mathbf{u})} \\
&= e^{-i2n\phi_f(\mathbf{u})} e^{-i2n\phi_h(\mathbf{u})} \\
&= [e^{-i\phi_f(\mathbf{u})}]^{2n},
\end{aligned} \tag{44}$$

where n is an integer and the PSF of the blur $h(\mathbf{n})$ is assumed to be centrally symmetric so that (35) holds. In other words, any even multiple of the Fourier transform phase modulo 2π is invariant to blur. When $n = 1$ the invariants are

$$\mathcal{B}(\mathbf{u}) = 2\phi_f(\mathbf{u}) \bmod 2\pi. \tag{45}$$

According to Euler's formula, and following the notation in (37), (44) can be decomposed into imaginary and real parts ($n = 1$):

$$-\sin(2\phi_f(\mathbf{u})) = -\frac{2F^I(\mathbf{u})F^R(\mathbf{u})}{[F^I(\mathbf{u})]^2 + [F^R(\mathbf{u})]^2} \tag{46}$$

$$\cos(2\phi_f(\mathbf{u})) = \frac{2[F^R(\mathbf{u})]^2}{[F^I(\mathbf{u})]^2 + [F^R(\mathbf{u})]^2} - 1 \tag{47}$$

Equations (46) and (47) do not have similar stability problems to those seen in (37), (40), and (41). Hence invariants $\mathcal{B}(\mathbf{u})$ were superior in the experiments of Paper II compared to the phase-tangent invariants (37). $\mathcal{B}(\mathbf{u})$ also outperformed the moment based blur invariants (31) due to their robustness to noise.

It is possible to improve the classification performance of the proposed Fourier phase invariants by using the weighting scheme proposed in Paper IX. This scheme is similar to the one presented in Paper I for moment invariants and for the phase-tangent invariants (see Section 3.3.2), but Paper XI presents the scheme more rigorously. The experiments in Paper IX compare the moment invariants, phase-tangent invariants, and the novel Fourier invariants with and without weighting. The weighting improves the classification result of the proposed invariants by up to 10 %. They also outperform the weighted moment and phase-tangent invariants clearly.

3.3.4 *Blur invariant phase correlation*

A similar approach to that used to develop the blur invariants in Paper II was used in Paper III for image registration. Traditional phase correlation registration (Kuglin &

Hines 1975) is based on the inverse Fourier transform of the normalized cross power spectrum of the images. Our method, called blur invariant phase correlation (BIPC), proceeds in a similar way, but first the normalized cross power spectrum of images $f_1(\mathbf{n})$ and $f_2(\mathbf{n}) = f_1(\mathbf{n} - \mathbf{t})$ is raised to the second (or any even) power, which makes it invariant to blur, similar to (44):

$$\begin{aligned}
 P(\mathbf{u}) &= \left(\frac{F_2(\mathbf{u})F_1^*(\mathbf{u})}{|F_2(\mathbf{u})||F_1^*(\mathbf{u})|} \right)^2 \\
 &= e^{-i2\phi_{f_2}(\mathbf{u})} e^{i2\phi_{f_1}(\mathbf{u})} \\
 &= e^{-i4\pi\mathbf{u}^T\mathbf{t}/N}.
 \end{aligned} \tag{48}$$

The last line in (48) is reached by the Fourier shift theorem (26). The inverse Fourier transform of (48) is a delta function $\delta(\mathbf{n} - 2\mathbf{t})$ centered at $2\mathbf{t}$. The position of the peak of the delta function is easy to detect and corresponds to the spatial shift between images $f_1(\mathbf{n})$ and $f_2(\mathbf{n})$ multiplied by 2. As mentioned in Section 2.1, we assume that also motion blur PSF is centrally symmetric with respect to the origin. This results in registration of the mass centers of the objects in original and blurred images. To be precise, if we register images f_1 and f_2 , of which f_2 is motion blurred by PSF of length L in angle θ , the registered position of f_2 will deviate by shift $\frac{L}{2}$ in direction $\theta + \pi$ from the correct physical position.

In the experiments of Paper III, the BIPC method was superior in comparison to phase correlation for registration of images containing centrally symmetric blur. It also outperformed the template matching registration by moment invariants, and was much faster to compute: the moment invariants have to be evaluated for every possible translation of the template, while BIPC is computed only once for the whole images using the 2-D fast Fourier transform (FFT).

3.3.5 *Blur and shift invariants*

In this section, we introduce the blur-translation invariants presented in Paper II. The downside of constructing the blur invariants using the phase spectrum is its sensitivity to translation of the images in contrast to the amplitude spectrum. Because the amplitude spectrum cannot be made invariant to blur, we have to use other spectral representations to gain shift invariance. Higher order spectra are defined by (29). We used bispectrum which is attained by using the value $z = 2$ in (29), namely

$$\Psi_2(\mathbf{u}_1, \mathbf{u}_2) = F(\mathbf{u}_1)F(\mathbf{u}_2)F^*(\mathbf{u}_1 + \mathbf{u}_2). \quad (49)$$

The bispectrum is invariant to translation, but also retains the phase information. A disadvantage is that if $F(\mathbf{u})$ is an N -by- N DFT of an image, the bispectrum becomes a four-dimensional N -by- N -by- N -by- N matrix. However, it is possible to take only 2-D slices of the bispectrum, which contain basically the same information (Dianat & Rao 1990, Petropulu & Pozidis 1998). We use a slice with $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{u}$, which becomes

$$\Psi_2(\mathbf{u}) = F(\mathbf{u})^2 F^*(2\mathbf{u}), \quad (50)$$

where the samples at points $2\mathbf{u}$ can be extracted utilizing the periodicity of DFT.

It is possible to construct a blur invariant phase-only bispectrum slice where the exponentials containing the phase are raised to the second power, similar to (44), namely

$$S(\mathbf{u}) = e^{-i2[2\phi_f(\mathbf{u}) - \phi_f(2\mathbf{u})]}. \quad (51)$$

The blur-translation invariants used in Paper II are given by the doubled phase of the bispectrum slice modulo 2π , namely

$$\mathcal{I}(\mathbf{u}) = 2[2\phi_f(\mathbf{u}) - \phi_f(2\mathbf{u})] \bmod 2\pi. \quad (52)$$

Similar to blur invariants (45), if the size of the DFT is N -by- N , we have approximately $N^2/2$ non-redundant blur-translation invariants available at once.

The image classification performance of the blur-translation invariants can be improved by weighting similar to the blur invariants as presented in Paper IX. According to the experiment, the classification accuracy of the blur-translation invariants was increased by up to 20 % through the use of weighting.

3.3.6 *Blur and similarity transformation invariants*

The blur invariant bispectrum slices (51) have the same scaling and rotation properties, (27) and (28), as the Fourier transform (except the scale factor in (27) which is $|ab|^{-3}$ for a bispectrum slice). Because the blur invariant bispectrum slice is also translation invariant, it can be used instead of the amplitude spectrum to obtain a similarity transformation invariant image descriptor by using the Fourier-Mellin transform as proposed by Casasent & Psaltis (1976) and Sheng Chen *et al.* (1994).

In our approach, presented in Paper V, the Fourier-Mellin transform is applied to the blur invariant bispectrum slice. As a result, both scale and orientation changes of the image transform into shifts in the log-polar sampled blur invariant bispectrum slice. Invariance to these shifts is obtained by computing again a shift invariant bispectrum slice. When applied to recognition of images which contain centrally symmetric blur, this method outperforms the blur and similarity transformation invariant moment descriptors which are more sensitive to noise.

The same approach can be used also for image registration, as demonstrated in Paper IV. In this case, we extract the shift between the log-polar sampled blur invariant bispectrum slices and decode the rotation and scaling parameters. The other image is inverse rotated and scaled according to estimated parameters. After this the images differ by shift, which is easy to solve using BIPC, presented in Section 3.3.4 and in Paper III. The experiments showed that for images containing centrally symmetric blur, the proposed method clearly outperforms Fourier-Mellin registration (sheng Chen *et al.* 1994), which is not invariant to blur.

A downside of the proposed method is the computational load: because of the uneven sampling of the log-polar transform, the DFTs have to be large to maintain the accuracy in the logarithmically sampled scale axis. This problem is emphasized when the phase spectrum is used, because it is not as smooth as the amplitude spectrum.

3.3.7 Blur and affine invariant object recognition

The method for blur and affine invariant object recognition, presented in Paper VI, does not form complete invariant features as do the methods in Papers II and V. The method is based on the approach introduced by Ben-Arie & Wang (1998), and its modified version used by Petrou & Kadyrov (2004). Ben-Arie & Wang (1998) use log-log sampled magnitude spectra of images which are invariant to shift, and transform the horizontal and vertical scaling into shifts along the axes. Invariance to rotations related to affine transformation is obtained using the rotated sets of the target and test images. Our method replaces the amplitude spectrum with a blur invariant bispectrum slice, and also has some other differences.

When an image is affine transformed, also the Fourier transform, as well as the bispectrum slice of the image, undergo a similar transformation. The blur invariant bispectrum slice $S_a(u_1^a, u_2^a)$ corresponding to an affine transformed image is given by

$$S_a(u_1^a, u_2^a) = S(u_1, u_2), \quad (53)$$

where

$$\begin{pmatrix} u_1^a \\ u_2^a \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \quad (54)$$

Thus, the blur invariant bispectrum slices of the two images, related by an affine transformation, differ by rotation α , scaling by a diagonal matrix, and by second rotation β . By logarithmically sampling the blur-translation invariant bispectrum slice along both axes, the uneven scaling by the diagonal matrix is converted into a 2-D shift. Thus, for correct rotations α and β , the log-log sampled blur invariant bispectrum slices differ only by a shift. To compare the images with these rotations α and β , we computed a set of target and test images with rotations α and β , respectively, where $\alpha, \beta \in \{0, 0 + \delta, \dots, \pi - \delta\}$ and δ is a small angle. To find the matching images, we compared the log-log sampled bispectrum slices of the rotated target, and test images pairwise using phase correlation.

The same method can be used in principle for registration of blurred images. In the object recognition experiments with blurred images, the proposed method outperformed the method introduced by Ben-Arie & Wang (1998) which is not invariant to blur. The proposed method was also shown to be more robust to noise compared to blur-affine moment invariants (Suk & Flusser 2003), and resulted in better classification accuracy for real images containing a small amount of perspective distortion. The method is unfortunately quite slow, because the size of the DFTs must be large to maintain accuracy in the case of non-uniform log-log sampling.

We present here an experiment that is not included in Paper VI. In this experiment, we systematically compared our novel method and blur-affine moment invariants for classification of images containing non-linear projective transformation. The methods were implemented as proposed in Paper VI and in the paper by Suk & Flusser (2003), and the nearest neighbor classification was used. The target images were 94 500-by-500 fish pictures on a black background (Kadyrov & Petrou 2001). The test images were degraded versions of the target images. The degradation included computer simulated motion blur of length 30 pixels in a random direction and projective transformation:

$$(n_1^p, n_2^p) = \left(\frac{p_{11}n_1 + p_{12}n_2 + p_{13}}{p_{31}n_1 + p_{32}n_2 + p_{33}}, \frac{p_{21}n_1 + p_{22}n_2 + p_{23}}{p_{31}n_1 + p_{32}n_2 + p_{33}} \right), \quad (55)$$

where $p_{31}n_1 + p_{32}n_2 + p_{33} \neq 0$. The parameters were $p_{11} = p_{22} = p_{33} = 1$, $p_{12} = p_{13} = p_{21} = p_{23} = 0$, and $p_{31} = p_{32} = c$. Non-zero p_{31} and p_{32} cause the transformation to be non-linear. In the experiments, the value of c was in the range $[0, 2 \cdot 10^{-3}]$. Examples of the target and test images are shown in Figure 6. Figure 7 shows the classification accuracy of the two methods varying the value of c . As can be seen, the proposed method can tolerate much more projective transformation. This result is in line with the last experiment in Paper VI which is done for real images containing also in some part non-linear transformation.

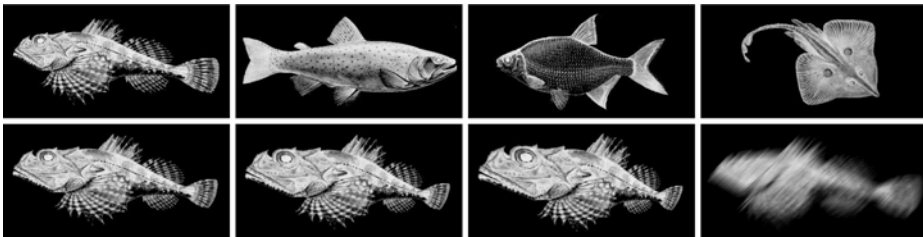


Fig 6. Top row: four examples of the 94 fish images. Bottom row: three projective transformed versions ($c = \{1, 1.5, 2\} \cdot 10^{-3}$) and one motion blurred (blur length is 30 pixels) version of the first image. (In the experiments, the test images were both projective transformed and blurred.)

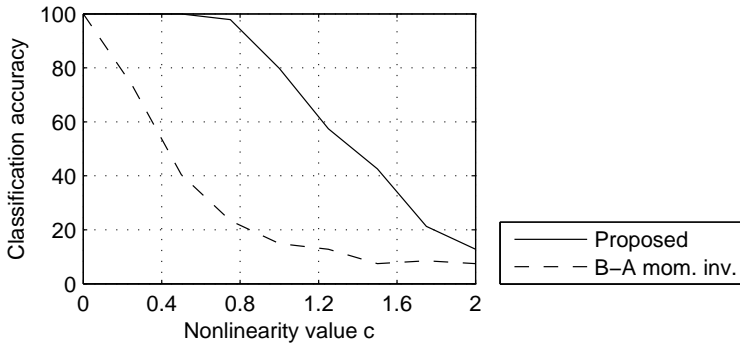


Fig 7. The classification accuracy of nearest neighbor classification of motion blurred and projective transformed images using the proposed Fourier domain blur-affine invariant method and blur-affine moment invariants.

3.4 Discussion

As already mentioned, there are not many approaches for blur invariant pattern recognition. The prevailing methods are based either on geometric image moments or Fourier domain phase. In this thesis, new Fourier domain methods have been developed. According to our findings, the Fourier phase based methods are more robust to image noise. In the case of blur and blur-translation invariants, the Fourier phase based methods are faster to compute using FFT in comparison to moment based methods. Especially image registration is faster using blur invariant phase correlation than using template matching with the moment invariants. On the other hand, object recognition using the similarity and affine transformation invariant Fourier phase methods is very slow due to the large DFT sizes that have to be used to maintain the accuracy in the case of logarithmic sampling of the spectrum.

Both the moment and the Fourier phase features are invariant to blurring with a centrally symmetric PSF. Although in the ideal case, motion, out of focus, and atmospheric turbulence result in centrally symmetric blur, the PSF of the motion blur is not centered on the origin in reality. As mentioned in Section 3.3.4, this has an effect on the registration result obtained by BIPC: the registered position of the blurred image deviates from the correct physical position by a shift of half blur length in a direction opposite to the motion. To find out the correct translation, the length and angle of the motion blur PSF should be known. The methods for estimating the parameters of the motion blur PSF were discussed shortly in Section 2.2. On the other hand, BIPC registers the mass centers of the objects in images, which is what also all non-blur invariant registration methods do implicitly in the case of blurred images. This may be a more desirable registration result in some applications. The requirement of central symmetric blur may have also other limitations in practice. Especially for long exposures, camera motion may be non-linear, resulting in a blur PSF which is not symmetric with respect to its center. This results in deviation of the feature values, and may lead to errors in recognition. Also, the methods, which are invariant to centrally symmetric blur, cannot separate centrally symmetric objects or images by definition.

Another source of error, for both moment and Fourier phase invariants, is the finite size of the observed blurred images: the support of the blur PSF always extends beyond the borders of the observed image, and thus some information needed for exact blur invariance is not available. In object recognition, the objects need to be segmented from the background accurately, which is assumed to have been already done. Segmentation

similarly loses part of the information. This source of error is only avoided if the objects lay on a uniform background. This could be the case, for example, for an orb in space or some object in the sky.

Blur invariants in general have had quite a limited number of practical applications. This may be due to limitations of the present invariants or the fact that in practice it is often possible to obtain sharp images and there is no need for blur invariants. Still there are some situations in which blurring of the images is inevitably present. One is the wide spreading use of low quality cameras, for example, in mobile devices. In these devices, the exposure time is often relatively long resulting to motion blur, and also the focusing may be inaccurate. Also an inevitable cause for out of focus blur is a scene with large depth variations: a lens can be focused only for a fixed distance. In Chapter 5, we present potential applications of the global blur invariants for multichannel deconvolution preprocessing and detection of image frauds.

4 Local blur insensitive descriptors

The image descriptors discussed in this chapter are extracted from the local neighborhoods of pixels, in contrast to the global features that are computed from the whole image. Local descriptors have been successfully applied for image registration, stereo matching, object recognition and categorization, and texture recognition (Mikolajczyk & Schmid 2005). Local descriptors can be computed for every pixel, which is typical for texture analysis (Ojala *et al.* 2002b), or only around interest points, which is usually the case in registration and recognition applications (Belongie *et al.* 2002). These two approaches can be referred to as dense and sparse computation of the local features (Schmid *et al.* 2005). The features used for these two approaches are typically slightly different. Local features have many desired properties over global features. Local features are more robust to non-global geometric transformations of images, and especially methods based on sparse features can be made robust to partial occlusions of images, parallax, background clutter, and they may not require segmentation. In Section 4.1, we will shortly introduce some of the methods for dense and sparse approaches, concentrating on their invariance properties. Then in Section 4.2, we present our blur insensitive local descriptors.

4.1 Local descriptors

In texture analysis, local descriptors are typically computed for every image pixel. These types of local descriptors are usually based on analysis of the distributions of local image characteristics, local frequency analysis, or analysis of local derivatives (Mikolajczyk & Schmid 2005). Histograms are often used to describe distributions of local image characteristics (Hadjidemetriou *et al.* 2004). The simplest approach is to use the histogram of the pixel values or the mean, and variance of this histogram (Mikolajczyk & Schmid 2005). One group of the traditional distribution based texture analysis methods is based on co-occurrence matrices, which describe local intensity value relations (Haralick *et al.* 1973). Frequency analysis of the images can be performed either in the spatial domain using masks, or using the local Fourier analysis (Randen & Husøy 1999).

In practice, a texture can occur in various orientations, scales, seen from a different

viewpoint, and with varying illumination. Optimally the local features should be invariant to these distortions. Various methods with different invariance properties have been developed during the past decades. In the following, some examples are given. For a review, see Turtinen (2007), Zhang & Tan (2002), Ojala *et al.* (2002b), and Tuceryan & Jain (1998).

Global illumination changes are often handled by normalization of the input image Turtinen (2007). Hadjidemetriou *et al.* (2004) used histogram equalization for illumination invariance before feature extraction. Ojala *et al.* (1996) proposed a method called the local binary pattern (LBP) which produces a code based on binary quantization of the pixels in a circle around the center pixel the value of which is used as a threshold. The resulting LBP codes, representing the local neighborhood shape, are computed around every pixel, and are presented as a histogram. The LBP is invariant to monotonic gray-scale changes.

One of the early rotation invariant methods for texture analysis, proposed by Davis (1981), used a polarogram which is a function of a co-occurrence matrix with a fixed length displacement vector and variable orientation. Statistics of the responses of orientation and scale tunable Gabor filters (Manjunathi & Ma 1996), which are basically multiscale edge and line detectors, are often used for texture analysis. There are extensions of this method for rotation invariant analysis which use 2-D polar Gabor filters (Haley & Manjunath 1999) or circularly shift the feature vector so that the features corresponding to the dominant orientation are placed first in the vector (Zhang *et al.* 2002a). Greenspan *et al.* (1994) used a steerable filter pyramid (Freeman & Adelson 1991) for rotation invariant texture classification. They extract the outputs of the filters as a function of orientation. In this representation, rotation of the image transforms to a circular shift. The invariance to this shift is obtained by using the DFT magnitude of this representation. Fountain & Tan (1998) compute a gradient image and construct a histogram of gradient directions. Rotation of the image causes the histogram to shift circularly, and thus the DFT magnitude of the histogram is invariant to rotation. Ojala *et al.* (2002b) incorporated rotation invariance to the LBP by rotating the binary codes representing the circular neighborhood to a minimum value. They also introduced multiresolution analysis using neighborhoods of different sizes.

Most of the existing statistical texture analysis methods assume that the texture images are acquired from the same viewpoint (Zhang & Tan 2002). Varma & Zisserman (2005) investigate 3-D texture classification under a changing viewpoint and illumination using a method which improves the earlier ideas of Cula & Dana (2004). In this

method, textures are modeled by joint distribution of filter responses. By clustering, exemplar filter responses are selected as textons, and are used to label every filter response and corresponding pixel in the image. Histograms of texton frequencies are then used as models of the training images. The same histogram is computed for test images, and the histograms are compared using nearest neighbor classification.

Instead of computing the local descriptors around every image pixel, it is also possible to select interest points from the images and compute descriptors only sparsely around them. Sparse descriptors are typically used for image registration and object recognition, but recently also for object category and texture classification (Zhang *et al.* 2007).

When local descriptors are computed around interest points, the interest points have to be detected first from each image, preferably invariant to the viewpoint. One well known rotation invariant interest point detector is the Harris detector (Harris & Stephens 1988) which detects points where signal change is significant in orthogonal directions, i.e., corners and junctions. Lindeberg (1998) proposed a scale invariant blob-like region detector. Lindeberg & Gårding (1997) showed how to obtain affine covariant elliptical regions which are further mapped to affine invariant circular regions. Mikolajczyk & Schmid (2004) extended the Harris detector for scale and affine invariant normalized regions around interest points. Matas *et al.* (2002) proposed another affine invariant detector to detect so called maximally stable extremal regions which are connected components of appropriately thresholded image. Also intensity extrema, edges, and entropy have been used for affine invariant region detection (Tuytelaars & Gool 2004, Kadir *et al.* 2004). The interest regions can also be rotated according to the dominant gradient orientation to obtain rotation invariance (Lowe 2004). Also photometric invariance to affine transformation of pixel intensities $f(\mathbf{n})$, i.e., $af(\mathbf{n}) + b$, can be obtained by normalization (Mikolajczyk & Schmid 2005). For this reason, the region descriptors do not necessarily have to be invariant to these transformations. For a review of interest point detectors, see Schmid & Roger Mohr (2000) and Mikolajczyk *et al.* (2005).

For the selection of the interest points, one also needs a method that rejects the outlier interest points, which do not fit to the global transformation between the images. This can be done, for example, by using the well known RANSAC algorithm (Forsyth & Ponce 2003). This makes the registration and recognition methods based on sparse features robust to local perturbation of some of the interest points.

The results for the dense descriptors for texture classification cannot be directly transferred to sparse region description. The regions often contain a single structure

in contrast to texture, which contains repeated patterns with statistical dependencies (Mikolajczyk & Schmid 2005). Many of the interest point descriptors are distribution based, and use histograms to represent different characteristics of the interest point neighborhood. For computational reasons, the histograms need to be shorter when they are computed for every interest point and not just once for the whole image, as is the case with dense descriptors.

Probably the most well known local descriptor is the scale invariant feature transform (SIFT) proposed by Lowe (2004). The method combines a scale invariant region detector and a descriptor, which is based on 3-D histogram of gradient locations and orientations in a grid spanning the detected region. The contribution to the histogram is weighted by the magnitude of the gradient. Mikolajczyk & Schmid (2005) proposed a gradient location-orientation histogram (GLOH), which is very similar to the SIFT. It uses a log-polar location grid instead of a Cartesian one, and applies principal component analysis (PCA) to reduce the descriptor size. PCA-SIFT is another improvement of the SIFT proposed by Ke & Sukthankar (2004). Heikkilä *et al.* (2009) have proposed a center symmetric local binary pattern (CS-LBP) method, which is a modification of the LBP method for interest points, and results in much shorter histograms. The histograms are computed for every cell in a location grid, and the final descriptor is a 3-D histogram of the CS-LBP feature locations and values. The shape context proposed by Belongie *et al.* (2002) is based on the 3-D histogram of the edge point locations and orientations in a log-polar grid.

Other sparse local descriptors include steerable filters (Freeman & Adelson 1991) evaluated for local image patches around interest points (Mikolajczyk & Schmid 2005) and complex filters (Schaffalitzky & Zisserman 2002). Lazebnik *et al.* (2005) proposed a sparse method for texture analysis which uses the affine invariant Harris detector (Mikolajczyk & Schmid 2004) and the Laplacian blob detector (Gårding & Lindeberg 1996), and descriptors based on spin image and rotation invariant feature transform (RIFT). Spin image is a 2-D histogram of the distance from the center and the intensity value. RIFT is a generalization of the SIFT which uses rings around the center instead of a grid, and the gradient orientation is given relative to the direction pointing outward from the center. Halawani *et al.* (2006) use integral invariants based on relational kernel functions for object recognition, and show how they can be computed in a scale invariant circular region proposed by Lowe (2004). The invariants are in a way an extension of the LBP operator producing continuous values and are also robust to illumination changes.

To the author’s best knowledge, there are no local descriptors or other texture analysis methods that are claimed to be insensitive to blurring. In principle, the global moment invariants or the global Fourier domain invariants presented earlier could be used for texture analysis, but they are mainly targeted for global object recognition, not local analysis of images.

4.2 Blur insensitive descriptors

Next, we introduce the blur insensitive local phase quantization (LPQ) descriptor and its rotation invariant extension in Sections 4.2.1 and 4.2.2, respectively. These methods are discussed in detail in Papers VII and VIII. In our experiments, we have used the LPQ descriptor as a dense local feature for texture classification.

4.2.1 Local phase quantization

The local phase quantization (LPQ) descriptor is based on the quantized phase of the discrete Fourier transform (DFT) computed in local image windows, i.e., short term Fourier transform (STFT). Because the computation of the LPQ is presented thoroughly in Paper VII, we describe the steps of the method, using the illustrative chart shown in Figure 8.

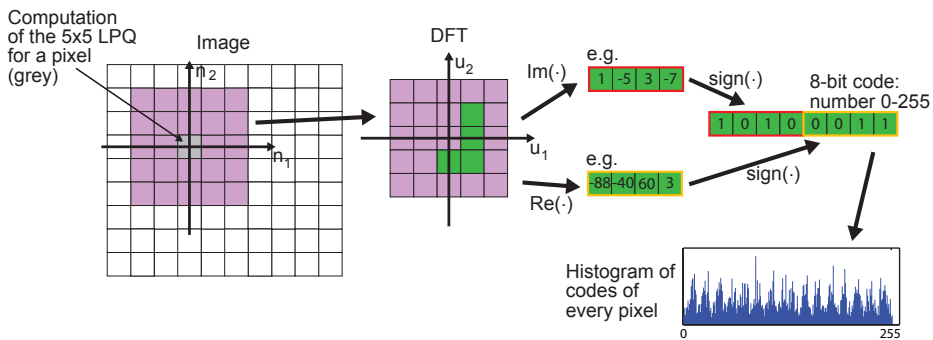


Fig 8. The LPQ method: The image illustrates the computation of the LPQ code for the gray pixel using 5-by-5 neighborhood. The codes for every pixel are added to a histogram.

For texture analysis, the LPQ descriptor is formed locally for each pixel. In Figure 8, it is computed for the pixel which is colored gray. The STFT coefficients are computed in the M -by- M neighborhood of the pixel for the lowest horizontal, vertical, and diagonal frequencies (1,0); (0,1); (1,1); and (1,-1) (in Figure 8 $M = 5$). Imaginary and real parts of these four frequency coefficients are binary quantized, based on their sign, resulting in an 8-bit binary number which is a decimal number between 0-255. Effectively, this codes only the quadrant of the phase of each coefficient, discarding amplitude information. The quantization can be made more efficient by decorrelating the frequency coefficients before quantization, as presented in Paper VII, because information is maximally preserved in scalar quantization, if the samples to be quantized are statistically independent. For decorrelation, it is assumed that the PSF of possible image blur is isotropic. In practical implementation, the frequency coefficients can be computed very efficiently with 1-D convolutions for the rows and columns of the image successively.

When LPQ is used for texture classification, the 8-bit codes for every image pixel are added up to a 256-bin histogram which describes the texture. Texture classification is then carried out based on the distance between the histograms, for example, by using the Chi square distance measure with the nearest neighbor classifier.

The generation of the codes and their histograms resembles the LBP method (Ojala *et al.* 2002b) which operates in the spatial domain. Similarly to the LBP, LPQ can be thought to combine statistical and structural texture analysis: the histogram is a statistical representation of the local patterns which correspond to the 256 codes. The phases of the horizontally, vertically, and two diagonally directed Fourier bases, shown in Figure 9 ($M = 32$), contribute to the pattern of each code. Figure 10 presents the spatial filters corresponding to codes 1, 2, 3, 4, 10, 37, 96, and 241 when $M = 32$. Note that in practice, M is much smaller and the filters coarser.

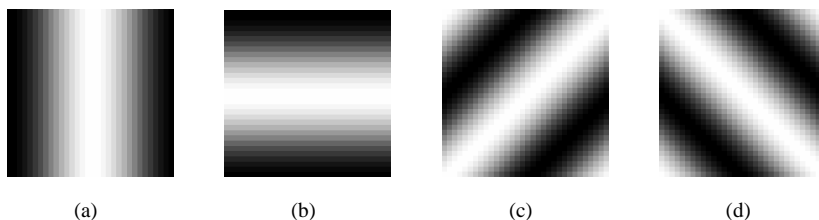


Fig 9. Fourier bases the phases of which contribute to the LPQ filters corresponding to the 256-codes.

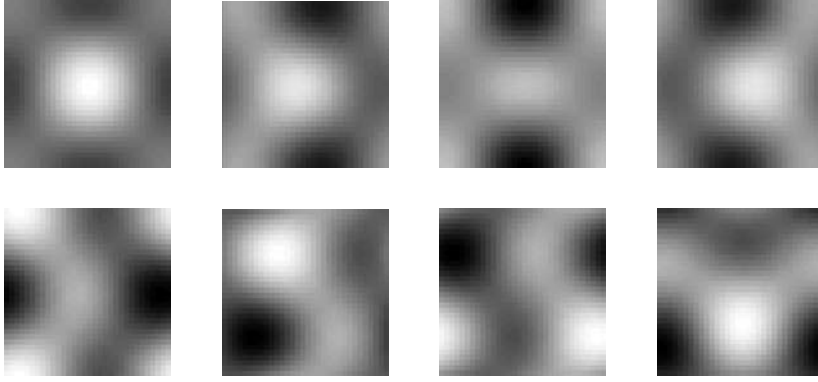


Fig 10. Filters corresponding to LPQ codes 1, 2, 3, 4, 10, 37, 96, and 241.

The codes produced by the LPQ operator are relatively insensitive to centrally symmetric blur. The blur insensitivity is obtained in a slightly different way compared to the global frequency domain blur invariants. According to (34) and (35) the phase of the observed image, given by $\phi_g(\mathbf{u})$, is invariant to centrally symmetric blur, and equals the phase of the ideal image, if the real valued frequency response $H(\mathbf{u})$ of the blur is non-negative:

$$\phi_g(\mathbf{u}) = \phi_f(\mathbf{u}) \quad \text{for all } H(\mathbf{u}) \geq 0. \quad (56)$$

In the case of ideal motion and out of focus blur, PSFs $h(\mathbf{n})$ have a rectangular and disk shape, respectively, as shown in Figure 2(a) and 2(b). This results in frequency responses $H(\mathbf{u})$, which oscillate between positive and negative values. Figure 3(a) and 3(b) illustrate the corresponding magnitude $|H(\mathbf{u})|$. The values of $H(\mathbf{u})$ are always positive at low frequencies before the first zero crossings, and thus satisfy (56). This is illustrated in the 1-D case for a cross section of $h(\mathbf{n})$ denoted by $h(n)$ in Figure 11 (top row). In 1-D, $H(u)$ is a sinc function with first zero crossing at frequency = (sampling frequency)/(blur length). In the case of a Gaussian PSF, which models atmospheric turbulence blur, $H(\mathbf{u})$ is also a Gaussian that always satisfies condition (56). 1-D Gaussian PSF $h(n)$ and $H(u)$ are illustrated in Figure 11 (bottom row).

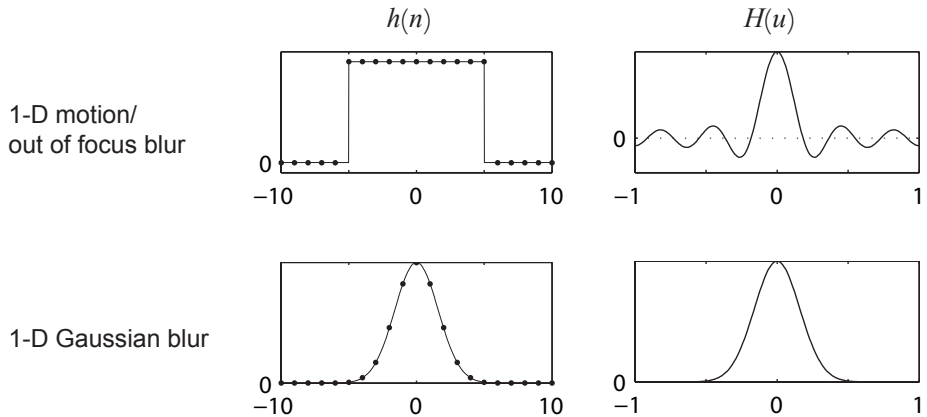


Fig 11. 1-D PSF $h(n)$ for motion and out of focus blur (top left), and for Gaussian blur (bottom left), and the corresponding frequency responses $H(u)$ on the right.

Because LPQ uses only the lowest frequency coefficients of the STFT, it is insensitive to centrally symmetric blur due to the arguments above. As M becomes larger LPQ becomes more insensitive to blur, but at the same time higher frequency information is lost. This is illustrated in Figure 13, where the image in Figure 12 is reconstructed from full phase spectrum 13(a), from LPQ codes 13(b), and from decorrelated LPQ codes 13(c) with $M = 3$ and in 13(d-f) similarly with $M = 7$. In 13(a), no phase information is lost. In Figure 13(b) information is lost due to quantization of the phase. As can be seen, decorrelation in Figure 13(c) improves the reconstruction. At the bottom row of Figure 13, where $M = 7$, the reconstructed images miss some high frequency information compared to the images of top row. In the experiments of Paper VII, we have used value $M = 7$, which seemed to be a good trade-off between discriminability and blur insensitivity.



Fig 12. The original *cameraman* image.

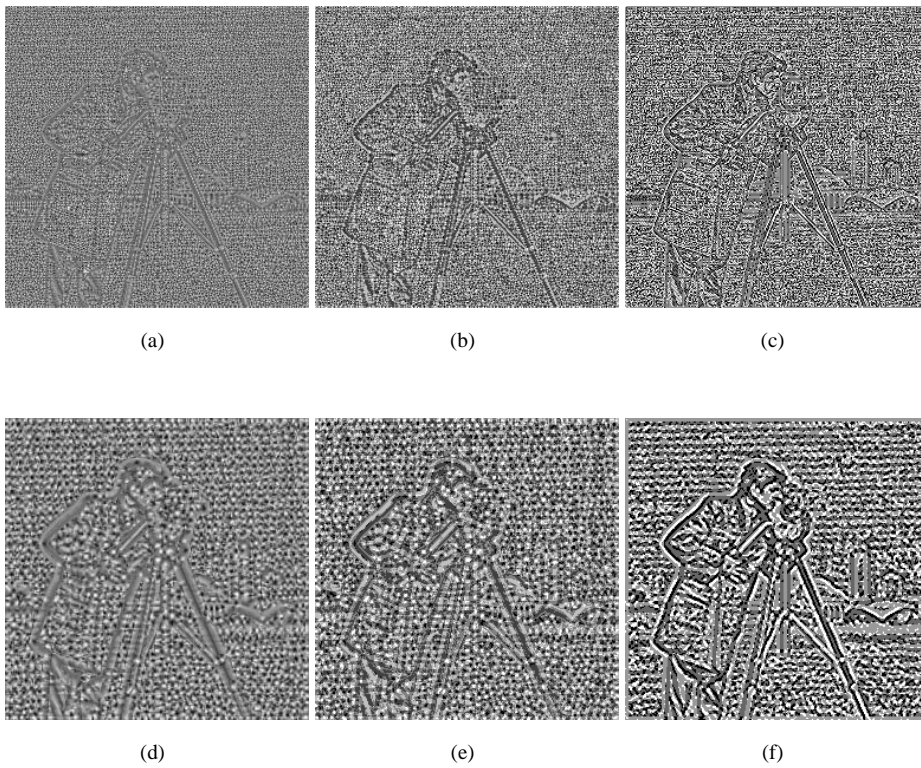


Fig 13. The cameraman (Figure 12) reconstructed from a full phase spectrum (a), from LPQ codes (b), and from decorrelated LPQ codes (c) when $M = 3$. (d), (e), and (f) contain similar reconstructions when $M = 7$.

The proposed method was compared with two well-known texture analysis operators, the LBP and a Gabor filter bank based method (Manjunathi & Ma 1996). The results of the texture classification experiments on the Outex texture database (Ojala *et al.* 2002a) showed that the LPQ method tolerates significantly more blurring than the other methods. In addition to that, LPQ also gave slightly better results for sharp texture images. Figure 14 shows an example texture image and two blurred versions of it used in the experiments.

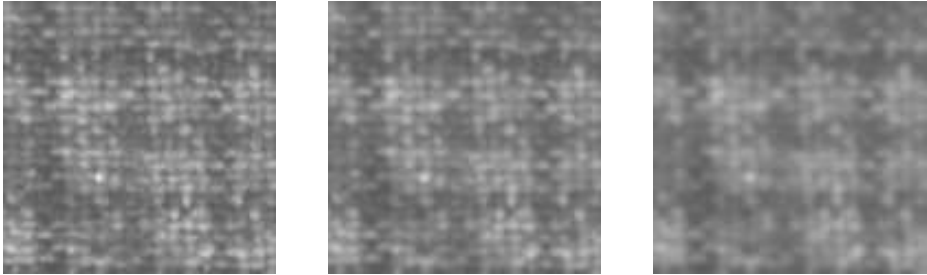


Fig 14. An example of the texture images used in the experiments (128-by-128) (left). Out of focus blurred versions of the same image with blur radii of one pixel (middle) and two pixels (right).

4.2.2 *Rotation invariant local phase quantization*

The rotation invariant LPQ method first estimates the characteristic orientation of each pixel for which the descriptor is computed. In the case of texture analysis, the characteristic orientation is estimated for every pixel in the image. The LPQ descriptor is then computed for each pixel in a direction corresponding to the characteristic orientation.

The characteristic orientation of a pixel is estimated using the quantized phase of the STFT coefficients which are computed in circular locations around the pixel. The phase of the coefficients is binary quantized based on the sign of the imaginary part. The characteristic orientation is then taken to be the angle corresponding to the circularly computed center of gravity of the vector containing the quantized phase of the frequency coefficients. In Paper VIII, also a faster approximation scheme for orientation estimation is presented. Both methods for estimating the characteristic orientation use only the low frequency phase and are thus insensitive to centrally symmetric blur similar to the standard LPQ. With the approximation scheme, the computational load

of the rotation invariant LPQ is close to standard LPQ. According to experiments, the performance of the rotation invariant LPQ is close to the standard LPQ for non-rotated blurred texture images. For rotated and blurred images, rotation invariant LPQ is clearly better compared to the rotation invariant LBP, and it is slightly better even without blur.

4.3 Discussion

Local image features have been used extensively for texture analysis, image registration, and object recognition in the computer vision field. Many authors have developed also features invariant to illumination and geometrical transformations, but the LPQ descriptors proposed in this thesis are the first features that are designed to be insensitive to blurring, more precisely centrally symmetric blurring.

In the case of blurred texture classification, the LPQ descriptor not only outperformed the LBP and the Gabor filter bank based comparison methods, but it surprised us by producing slightly better results also for sharp images. The same was true for the rotation invariant versions of the LPQ and LBP descriptors. The LPQ descriptor has similarities with the spatial domain LBP descriptor, and both are very fast to compute, computation of the LBP descriptor still being little faster. Also the method for blur insensitive estimation of the local orientation for rotation invariant LPQ can be implemented very fast using the proposed approximation method.

The LPQ descriptors are disturbed by similar factors compared to global blur invariants. First, the blurring should be centrally symmetric, which is not always the case, resulting in errors. Second, for the LPQ descriptor the boundary effect of blurring convolution is large due to the small window size which loses information and disturbs blur invariance. A larger window size results in greater blur insensitivity, but at the same time loses the high frequency information. For these reasons, the window size should be larger for blur PSFs of a greater extent. The advantage of the local LPQ descriptors is that the extent of the blur PSF can in principle be varying spatially. This may be advantageous for images containing different depths with different amounts of defocus blur.

In our texture classification experiments, we have computed the LPQ descriptors densely for every pixel of the image, and collected the resulting codes in a histogram which described the texture. The LPQ descriptor might be usable also as a sparse descriptor of interest points of blurred images. In Section 5, we present applications of the LPQ descriptor for face image recognition and for matching of stereo images.

5 Applications

Paper I introduced an application of blur invariant features for concealment of blurred background in an image sequence. In this application, the background of a blurred frame was replaced using a sharp frame containing the same scene. In this way the need for deblurring was limited only to the moving objects. In the following sections, we introduce some other possible applications of the blur invariants presented in this thesis.

5.1 Preprocessing for multichannel deconvolution

In Section 2.2, we introduced the problems related to reconstruction of blurred images by deconvolution. One approach to solving these problems is to use multichannel deconvolution, which uses multiple differently blurred images for restoration. In this approach, the lack of information on one channel, due to frequency domain zeros of the blur, can be supplemented by using the other channels blurred by different PSF. Most of the existing multichannel deconvolution methods are iterative or recursive, and they also assume spatial alignment of the channels and knowledge of the blur PSF (Šroubek & Flusser 2005). Šroubek & Flusser (2005) proposed a method for multichannel blind deconvolution of misaligned images. Šroubek *et al.* (2007) coupled the method with resolution enhancement, also called super-resolution. This method produces a good restoration result for multiple differently blurred images which may be displaced by a shift.

A shortcoming of the multichannel deconvolution method of Šroubek & Flusser (2005) is that it cannot use rotated or scaled input images. To tackle this limitation, we preceded the multichannel deconvolution algorithm with our blur invariant registration method presented in Section 3.3.6 and in Paper IV. This also enables the restoration of rotated and scaled images. Figure 15 shows an example of the application of the combined method. 200-by-200 images (a-c) are blurred by motion blur of five pixels in different directions, and in addition, (b) is rotated (4.13°) and shifted ($[1.23, 1.89]$ pixels), and (c) is scaled (0.90) and shifted ($[-1.24, 1.92]$ pixels). (d) presents the image reconstructed from images (a-c) using blur invariant registration (parameters: (b)

4.0°,1,[1.23,1.84]; (c) 0°,0.90,[-1.25,1.81]) followed by the blind multichannel deconvolution algorithm ¹.



(a)



(b)



(c)



(d)

Fig 15. (a-c) Degraded images: different motion blurs and similarity transformations. (d) Restored image: restored from images (a-c) using first the proposed blur invariant registration method, and then a multichannel deconvolution algorithm.

¹<http://zoi.utia.cas.cz/bsr-toolbox>

5.2 Image fraud detection

It is easy to modify images or produce frauds even by using some common image processing software. It may be in someone's interest, for example, to hide some parts of the image containing a specific object or person, or duplicate parts of the image such as crowd or smoke to make images more dramatic. Publically available image processing software contains such functionalities, for example, to conceal pimples in face images. Nowadays also images in the news are often modified or at least enhanced somehow, and it may be difficult to say what is an acceptable level of modification and what is altering the truth (Brugioni 1999).

The software usually performs the concealment of an object by extrapolating the surrounding image pixels, or replacing the patch using surrounding texture pointed to by the user. In the case of replacing an image patch, the borders of the patch have to be connected smoothly to the surrounding image to make it unnoticeable. Blurring of the image patch slightly is one way of achieving this. Blurring also alters the pixel values so that with a simple comparison, the patch cannot be said to be duplicated from another place (Mahdian & Saic 2006).

Blur invariant phase correlation can be applied to recognize a duplicated and blurred image patch if the blur PSF has centrally symmetric support, such as a 2-D Gaussian function. For this task, the image is divided into overlapping rectangular blocks which are matched between each other to find duplicates of some blocks. The block size can be varied. Figure 16 illustrates an example of this application when a block size of 70-by-70 and an overlap of 10 pixels is used: (a) shows an original image, (b) is a fraud containing a duplicated and blurred (Paint Shop Pro: Gaussian blur, radius 1) patch hiding a building, and (c) shows a block (white) which is detected to be duplicated (black) in another location.

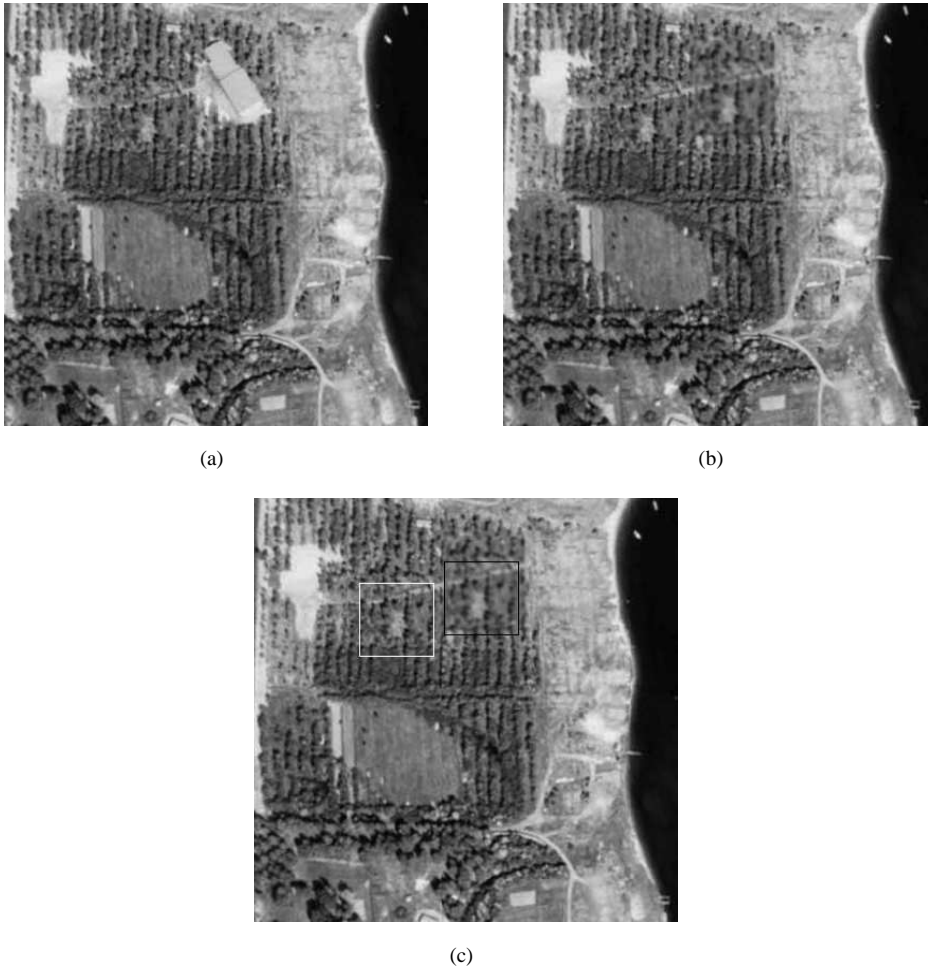


Fig 16. (a) Original image, (b) a fraud, and (c) duplicated texture detected: the original patch and its blurred duplicate are shown with white and black, respectively.

5.3 Face recognition

In recent years, face image analysis has attained wide interest in the field of computer vision research. The classical goals related to face image analysis are face detection, face recognition, and facial expression analysis, all of which have applications, for example, in surveillance and human computer interaction (Li & Jain 2005).

The most critical issue in face image analysis is to find efficient descriptors of the face appearance. The face images appear to be quite similar, and so the descriptors should be discriminative enough. On the other hand, the descriptors should be robust to typical image degradations, and also to variations in pose, illumination, and the person's age. There has been a lot of research on robust face image analysis, but the effect of image blurring has been mostly overlooked. No attempts to explicitly develop blur invariant descriptors for face image analysis have been made in the past (Ahonen *et al.* 2008).

Ahonen *et al.* (2008) proposed a method for recognition of blurred faces using the LPQ descriptor, introduced in Section 4.2.1 and in Paper VII. The approach is similar to that used by Ahonen *et al.* (2006) with the LBP descriptor. In this approach, the LPQ descriptor is first computed for every pixel of the face image. Then the image is divided into non-overlapping rectangular regions of size 8-by-8 pixels, as shown in Figure 17, and a histogram of the LPQ codes is computed independently within each of the regions. These histograms are then concatenated to form a global description of the face. As a preprocessing step, images are converted to gray-scale and registered using eye positions. The classification of the face images is done using Chi-square distance of the concatenated LPQ descriptor histograms and a nearest neighbor classifier.

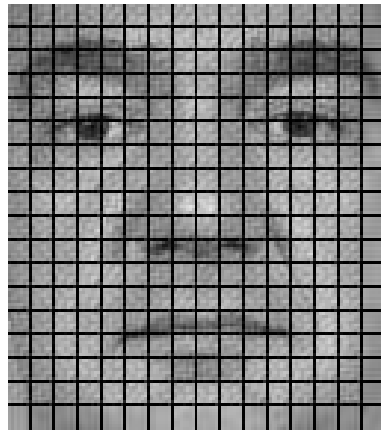


Fig 17. Histograms of LPQ descriptors are formed independently for rectangular regions of registered images.

The LPQ based method was compared with the LBP based method (Ahonen *et al.* 2006) using the CMU PIE (Pose, Illumination, and Expression) (Sim *et al.* 2003) and the Face Recognition Grand Challenge (FRGC) (experiment 1.0.4) (Phillips *et al.* 2005) datasets. The datasets contained various degradations including different facial expressions. In addition, the test images of the CMU PIE dataset were artificially Gaussian blurred. The FRGC dataset test images contain real out of focus blur, among other degradations. According to the experiment with the CMU PIE images, the LPQ based method outperformed the LBP based method clearly for blurred images, and was also slightly better without blurring. For the FRGC dataset, the LPQ based method was better than the LBP, and also outperformed the Local Ternary Pattern (LTP) (Tan & Triggs 2007), which is a modification of the LBP.

5.4 Stereo matching

An LPQ based approach has been applied also to blur and contrast invariant stereo matching (Pedone & Heikkilä 2008). Stereo matching is used in stereo vision to match corresponding points in two images representing the same scene from different viewpoints. Using these correspondences, it is possible to derive information about the 3-D structure of the scene.

In the most popular configuration, the stereo vision system contains two cameras with parallel imaging planes separated by a horizontal distance (Wang *et al.* 2002). The horizontal displacement of an image point between the images of the two cameras is called disparity, and is related to the 3-D depth of the corresponding scene point. Traditionally disparity is estimated for each image pixel by matching local horizontally shifted windows centered on the pixel of interest between the two images. The similarity measure used in matching can be, for example, the sum of absolute differences (SAD) (Scharstein & Szeliski 2002).

In the LPQ based method, LPQ codes are used as descriptors of the image windows to be matched. Unlike in the texture analysis approach introduced in Section 4.2.1, LPQ is computed for more than just the lowest frequencies of the DFTs of the image windows. The number of frequency samples used in the experiments was seven, corresponding to the shaded area in Figure 18. The phase of these coefficients was again quantized by two bits, resulting in a 14-bit code describing the window. For the matching of the windows, it is enough to compute the Hamming distance between the descriptor codes. The method has many parameters such as the window size and the

number of frequency coefficients, the choice of which is a trade-off between discriminability and blur insensitivity. In the experiments, the LPQ based method produced more accurate depth map compared to the traditional SAD based method in the case of motion, Gaussian, or defocus blurred images.

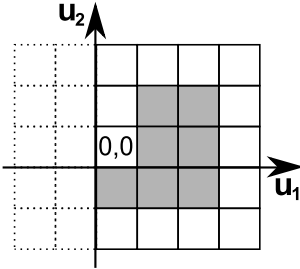


Fig 18. In stereo matching, the LPQ codes are computed around each pixel for the DFT coefficients in the locations colored gray.

6 Conclusions

Computer-aided pattern recognition and registration have many applications in different fields of science. In computer vision the patterns are image intensity functions, which are often degraded, making their recognition more difficult. Typical degradations include noise, occlusions, blur, and changes of the viewpoint. Ideally, the pattern recognition system should be invariant to these error sources. This thesis concentrated especially on blur invariance which is combined with invariance to the changing viewpoint. There are also various approaches to pattern recognition from which this thesis focused on feature based pattern recognition. In this approach, the patterns are described by a set of features which are extracted from the image.

Although there has been a lot of research on invariant pattern recognition, blur invariant pattern recognition has received less attention. Basically the only existing blur invariant features are based on geometric image moments, which are quite sensitive to noise and clutter in the images. An alternative to blur invariant pattern recognition would be deblurring of the images, followed by conventional non-blur invariant pattern recognition. However, deblurring is a difficult problem and introduces new artifacts to images.

In this thesis, we proposed novel features for blur invariant recognition and registration of images. The features are based on the Fourier transform phase of the images, and are invariant to centrally symmetric blur, similar to the image moment based blur invariants. We introduced features that are computed globally from the whole image, and features that are computed in local regions. For the global features, the blur invariance is obtained using the second power of the phase-only Fourier spectrum or bispectrum of the image. The frequency coefficients modified in this way are used as features. Local features are based on the short term Fourier transform computed in a local image windows. In this case, we used the phases of the lowest horizontal, vertical, and diagonal frequency coefficients, which are insensitive to blur. The phases of each of these four frequency coefficients are quantized into four levels, and are used to form an 8-bit descriptor for the local region.

In the original papers, the global features have been used for recognition and registration of blurred images. The global features are also combined with geometric invariances up to an affine transformation: shift invariance is obtained using a bispectrum,

rotation-scale invariance using log-polar mapping of the bispectrum slices, and affine invariance using rotated sets of the log-log mapped bispectrum slices. In the original papers, we performed various experiments which reinforced the invariance properties of the novel invariants, and showed that they are more robust to noise when compared with image moment based invariants. The proposed blur invariant phase correlation is a much faster approach to registration compared with moment invariants. Also recognition of shifted objects happens faster with the novel invariants.

In the original papers, the local features are used for classification of blurred image textures. In this application, the features are computed around every pixel in the images and are added up to a histogram that describes the texture. The local features were also extended for rotation invariance which was achieved by estimating dominant orientation of the local regions followed by computation of the oriented descriptors. In the experiments with blurred texture images, the novel features clearly outperformed some state of the art texture descriptors and were also slightly better for sharp textures.

We also proposed some applications of the invariants: pre-processing for multichannel blind deconvolution, fraud detection, face recognition, and stereo matching. We expect that many more applications for the proposed invariants exist. Some applications could be related to the low quality cameras of mobile devices with long shutter times and inaccurate focusing. Applications could be found also in other fields of pattern recognition and signal processing, possibly related to something other than 2-D patterns.

References

- Abu-Mostafa YS & Psaltis D (1984) Recognitive aspects of moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6(6): 698–706.
- Ahonen T, Hadid A & Pietikainen M (2006) Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(12): 2037–2041.
- Ahonen T, Rahtu E, Ojansivu V & Heikkilä J (2008) Rotation invariant local phase quantization for blur insensitive texture analysis. *Proc. International Conference on Pattern Recognition*, Tampa, Florida, 1–4.
- Arora H, Namboodiri AM & Jawahar C (2008) Robust image registration with illumination, blur and noise variations for super-resolution. *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, Nevada, 1301–1304.
- Banham MR & Katsaggelos AK (1997) Digital image restoration. *IEEE Signal Processing Magazine* 14(2): 24–41.
- Belkasim SO, Shridhar M & Ahmadi M (1991) Pattern recognition with moment invariants: a comparative study and new results. *Pattern Recognition* 24(12): 1117–1138.
- Belongie S, Malik J & Puzicha J (2002) Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(4): 509–522.
- Ben-Arie J & Wang Z (1998) Pictorial recognition of objects employing affine invariance in the frequency domain. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(6): 604–618.
- Bentoutou Y & Taleb N (2005) Automatic extraction of control points for digital subtraction angiography image enhancement. *IEEE Transactions on Nuclear Science* 52(1): 238–246.
- Bentoutou Y, Taleb N, Mezouar MCE, Taleb M & Jetto L (2002) An invariant approach for image registration in digital subtraction angiography. *Pattern Recognition* 35(12): 2853–2865.
- Brugioni DA (1999) *Photo Fakery: The History and Techniques of Photographic Deception and Manipulation*. Brassey's Inc, Dulles, VA.
- Candocia FM (2004) Moment relations and blur invariant conditions for finite-extent signals in one, two and N-dimensions. *Pattern Recognition Letters* 25(4): 437–447.
- Capodiferro L, Cusani R, Jacovitti G & Vascotto M (1987) A correlation based technique for shift, scale, and rotation independent object identification. *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, Texas, 12: 221–224.
- Casasent D & Psaltis D (1976) Position, rotation, and scale invariant optical correlation. *Applied Optics* 15(7): 1795–1799.
- Castro ED & Morandi C (1987) Registration of translated and rotated images using finite Fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9(5): 700–703.
- Chandran V, Carswell B, Boashash B & Elgar S (1997) Pattern recognition using invariants defined from higher order spectra: 2-D image inputs. *IEEE Transactions on Image Processing* 6(5): 703–712.
- Cula OG & Dana KJ (2004) 3D texture recognition using bidirectional feature histograms. *International Journal of Computer Vision* 59(1): 33–60.
- Davis LS (1981) Polarograms: A new tool for image texture analysis. *Pattern Recognition* 13(3): 219–223.

- de Bougrenet de la Tocnaye JL & Ghorbel F (2004) Scale-rotation invariant pattern recognition applied to image data compression. *Pattern Recognition Letters* 8(1): 55–58.
- Dianat SA & Rao RM (1990) Fast algorithms for phase and magnitude reconstruction from bispectra. *Optical Engineering* 29(5): 504–512.
- Flusser J (2000) On the independence of rotation moment invariants. *Pattern Recognition* 33(9): 1405–1410.
- Flusser J (2006) Moment invariants in image analysis. *Proc. World Academy of Science, Engineering and Technology*, 11: 196–201.
- Flusser J & Boldys J (1999) Registration of N-D images by blur invariants. *Proc. IEEE International Geoscience and Remote Sensing Symposium, Hamburg, Germany*, 1262–1264.
- Flusser J, Boldys J & Zitová B (2003) Moment forms invariant to rotation and blur in arbitrary number of dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(2): 234–246.
- Flusser J & Suk T (1994) Recognition of images degraded by linear motion blur without restoration. *Proc. International Workshop on Theoretical Foundations of Computer Vision, Dagstuhl, Germany*, 37–51.
- Flusser J & Suk T (1998) Degraded image analysis: An invariant approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(6): 590–603.
- Flusser J, Suk T & Saic S (1996) Recognition of blurred images by the method of moments. *IEEE Transactions on Image Processing* 5(3): 533–538.
- Flusser J & Zitová B (1999) Combined invariants to linear filtering and rotation. *International Journal of Pattern Recognition and Artificial Intelligence* 13(8): 1123–1136.
- Flusser J & Zitová B (2004) Invariants to convolution with circularly symmetric PSF. *Proc. International Conference on Pattern Recognition, Cambridge, UK*, 11–14.
- Foroosh H, Zerubia J & Berthod M (2002) Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing* 11(3): 188–200.
- Forsyth DA & Ponce J (2003) *Computer Vision: A Modern Approach*. Prentice Hall.
- Fountain SR & Tan TN (1998) Efficient rotation invariant texture features for content-based image retrieval. *Pattern Recognition* 31(11): 1725–1732.
- Freeman WT & Adelson EH (1991) The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13: 891–906.
- Gårding J & Lindeberg T (1996) Template matching in rotated images. *International Journal of Computer Vision* 17(2): 163–191.
- Gonzales RC & Woods R (1992) *Digital Image Processing*. Addison Wesley, Reading, MA.
- Gonzales RC & Woods R (2002) *Digital Image Processing*. Prentice-Hall, Upper Saddle River, NJ.
- Goshtasby A (1985) Template matching in rotated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 7(3): 338–344.
- Greenspan H, Belongie S, Goodman R & Perona P (1994) Rotation invariant texture recognition using a steerable pyramid. *Proc. International Conference on Pattern Recognition, Jerusalem, Israel*, 162–167.
- Hadjidemetriou E, Grossberg MD & Nayar SK (2004) Multiresolution histograms and their use for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(7): 831–847.
- Halawani A, Tamimi H, Burkhardt H & Zell A (2006) Using local integral invariants for object recognition in complex scenes. *Proc. International Conference on Image Analysis and*

- Recognition, Póvoa de Varzim, Portugal, II: 1–12.
- Haley GM & Manjunath BS (1999) Rotation-invariant texture classification using a complete space-frequency model. *IEEE Transactions on Image Processing* 8(2): 255–269.
- Haralick RM, Shanmugan KS & Dunstein I (1973) Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 3(6): 610–621.
- Harris CG & Stephens M (1988) A combined corner and edge detection. *Proc. Alvey Vision Conference, Manchester, UK*, 147–151.
- Hayes MH, Lim JS & Oppenheim AV (1980) Signal reconstruction from phase or magnitude. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28: 672–680.
- Heikkilä J (2004) Image scale and rotation from the phase-only bispectrum. *Proc. IEEE International Conference on Image Processing, Singapore*, 1783–1786.
- Heikkilä M, Pietikäinen M & Schmid C (2009) Description of interest regions with local binary patterns. *Pattern Recognition* 42(3): 425–436.
- Hu MK (1961) Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* 8(2): 179–187.
- Hupkens TM & de Clippeleir J (1995) Pattern recognition with moment invariants: a comparative study and new results. *Pattern Recognition* 16(4): 371–376.
- Jain AK, Duin RP & Mao J (2000) Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(1): 4–37.
- Kadir T, Zisserman A & Brady M (2004) An affine invariant salient region detector. *Proc. European Conference on Computer Vision, Prague, Czech Republic, I*: 228–241.
- Kadyrov A & Petrou M (2001) The trace transform and its applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(8): 811–828.
- Kadyrov A & Petrou M (2004) Affine invariant features from the trace transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(1): 30–44.
- Karl W (2005) Regularization in image restoration and reconstruction. In: Bovik A (ed) *Handbook of Image and Video Processing*, 183–202. Academic Press.
- Ke Y & Sukthankar R (2004) PCA-SIFT: a more distinctive representation for local image descriptors. *Proc. IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC*, 506–513.
- Kubota A & Aizawa K (2005) Reconstructing arbitrarily focused images from two differently focused images using linear filters. *IEEE Transactions on Image Processing* 14(11): 1848–1859.
- Kubota A, Kodama K & Aizawa K (1999) Registration and blur estimation methods for multiple differently focused images. *Proc. IEEE International Conference on Image Processing, Kobe, Japan*, 447–451.
- Kuglin CD & Hines DC (1975) The phase correlation image alignment method. *Proc. IEEE International Conference on Cybernetics and Society*, 163–165.
- Kundur D & Hatzinakos D (1996) Blind image deconvolution. *IEEE Signal Processing Magazine* 13(3): 43–64.
- Legendijk RL & Biemond J (2005) Basic methods for image restoration and identification. In: Bovik A (ed) *Handbook of Image and Video Processing*, 167–182. Academic Press.
- Lazebnik S, Schmid C & Ponce J (2005) A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(8): 1265–1278.
- Li SZ & Jain AK (2005) *Handbook of Face Recognition*. Springer, Berlin, Germany.
- Lindeberg T (1998) Feature detection with automatic scale selection. *International Journal of*

- Computer Vision 59(1): 61–85.
- Lindeberg T & Gårding J (1997) Shape-adapted smoothing in estimation of 3-D shape cues from affine deformations of local 2-D brightness structure. *Image and Vision Computing* 15(6): 415–434.
- Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 30(2): 79–116.
- Lu J & Yoshida Y (1999) Blurred image recognition based on phase invariants. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E82A(8): 1450–1455.
- Mahdian B & Saic S (2006) Detection of copy-move forgery using a method based on blur moment invariants. *Forensic Science International* 171(2-3): 180–189.
- Manjunathi BS & Ma WY (1996) Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(8): 837–842.
- Matas J, Chum O, Urbana M & Pajdla T (2002) Robust wide baseline stereo from maximally stable extremal regions. *Proc. British Machine Vision Conference, Cardiff, UK, 1*: 384–393.
- Metari S & Deschenes F (2008) New classes of radiometric and combined radiometric-geometric invariant descriptors. *IEEE Transactions on Image Processing* 17(6): 991–1006.
- Mikolajczyk K & Schmid C (2004) Scale and affine invariant interest point detectors. *International Journal of Computer Vision* 60(1): 63–86.
- Mikolajczyk K & Schmid C (2005) A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(10): 1615–1630.
- Mikolajczyk K, Tuytelaars T, Schmid C, Zisserman A, Matas J, Schaffalitzky F, Kadir T & Gool LV (2005) A comparison of affine region detectors. *International Journal of Computer Vision* 65(1-2): 43–72.
- Myles Z & da Vitoria Lobo N (1998) Recovering affine motion and defocus blur simultaneously. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(6): 652–658.
- Ojala T, Mäenpää T, Pietikäinen M, Viertola J, Kyllönen J & Huovinen S (2002a) Outex - new framework for empirical evaluation of texture analysis algorithms. *Proc. International Conference on Pattern Recognition, Quebec, Canada*, 701–706.
- Ojala T, Pietikäinen M & Harwood D (1996) A comparative study of texture measures with classification based on featured distribution. *Pattern Recognition* 29(1): 51–59.
- Ojala T, Pietikäinen M & Mäenpää T (2002b) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7): 971–987.
- Oppenheim AV & Lim JS (1981) The importance of phase in signals. *Proceedings of the IEEE* 69(5): 529–541.
- Pawlak M (1992) On the reconstruction aspects of moment descriptors. *IEEE Transactions on Information Theory* 38(6): 1698–1708.
- Pedone M & Heikkilä J (2008) Blur and contrast invariant fast stereo matching. *Proc. Advanced Concepts for Intelligent Vision Systems, Juan-les-Pins, France*, 883–890.
- Petropulu AP & Pozidis H (1998) Phase reconstruction from bispectrum slices. *IEEE Transactions on Signal Processing* 46(2): 527–530.
- Petrou M & Kadyrov A (2004) Affine invariant features from the trace transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(1): 30–44.
- Phillips PJ, Flynn PJ, Scruggs T, Bowyer KW, Chang J, Hoffman K, Marques J, Min J & Worek W (2005) Overview of the face recognition grand challenge. *Proc. IEEE Conference on*

- Computer Vision and Pattern Recognition, San Diego, CA, I: 947–954.
- Rahtu E, Salo M & Heikkilä J (2005a) Affine invariant pattern recognition using multi-scale autoconvolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(6): 908–918.
- Rahtu E, Salo M & Heikkilä J (2005b) A new efficient method for producing global affine invariants. *Proc. International Conference on Image Analysis and Processing*, Cagliari, Italy, 407–414.
- Randen T & Husøy JH (1999) Filtering for texture classification: A comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(4): 291–310.
- Reddy BS & Chatterji BN (1996) An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing* 5(8): 1266–1271.
- Reeves SJ (2005) Fast image restoration without boundary artifacts. *IEEE Transactions on Image Processing* 14(10): 1448–1453.
- Reiss TH (1991) The revised fundamental theorem of moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(8): 830–834.
- Sadler BM (1992) Shift and rotation invariant object reconstruction using the bispectrum. *Proc. Workshop on Higher-Order Spectral Analysis*, Vail, Colorado, 106–111.
- Schaffalitzky F & Zisserman A (2002) Multi-view matching for unordered image sets, or 'how do I organize my holiday snaps?'. *Proc. European Conference on Computer Vision*, Copenhagen, Denmark, 414–431.
- Scharstein D & Szeliski RS (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47(1-3): 7–42.
- Schmid C, Dorko G, Lazebnik S, Mikolajczyk K & Ponce J (2005) Pattern recognition with local invariant features. In: Chen C & Wang P (eds) *Handbook of Pattern Recognition and Computer Vision*, 3rd edition, 71–92. World Scientific Publishing Co.
- Schmid C & Roger Mohr CB (2000) Evaluation of interest point detectors. *International Journal of Computer Vision* 37(2): 151–172.
- Sheng Y & Duvernoy J (1986) Circular-Fourier-radial-Mellin transform descriptors for pattern recognition. *Journal of the Optical Society of America* 3(6): 885–888.
- sheng Chen Q, Defrise M & Deconinck F (1994) Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16(12): 1156–1168.
- Sim T, Baker S & Bsat M (2003) The CMU pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(12): 1615–1618.
- Stern A, Kruchakov I, Yoavi E, & Kopeika NS (2002) Recognition of motion-blurred images by use of the method of moments. *Applied Optics* 41(11): 2164–2171.
- Suk T & Flusser J (1993) Pattern recognition by affine moment invariants. *Pattern Recognition* 26(1): 167–174.
- Suk T & Flusser J (2003) Combined blur and affine moment invariants and their use in pattern recognition. *Pattern Recognition* 36(12): 2895–2907.
- Suk T & Flusser J (2004) Graph method for generating affine moment invariants. *Proc. International Conference on Pattern Recognition*, Cambridge, England, II: 192–195.
- Tan XY & Triggs B (2007) Enhanced local texture feature sets for face recognition under difficult lighting conditions. *Proc. International Workshop on Analysis and Modeling of Faces and Gestures*, 168–182.
- Tsatsanis MK & Giannakis GB (1992) Object and texture classification using higher order statis-

- tics. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(7): 733–750.
- Tuceryan M & Jain AK (1998) Texture analysis. In: Chen CH, Pau LF & Wang PSP (eds) *The Handbook of Pattern Recognition and Computer Vision*, 207–248. World Scientific Publishing Co.
- Turtinen M (2007) Learning and recognizing texture characteristics using local binary patterns. Dissertation. *Acta Universitatis Ouluensis C* 278.
- Tuytelaars T & Gool LJV (2004) Matching widely separated views based on affine invariant regions. *International Journal of Computer Vision* 59(1): 61–85.
- van de Weijer J & Schmid C (2006) Blur robust and color constant image decription. *Proc. IEEE International Conference on Image Processing*, Atlanta, Georgia, 993–996.
- Varma M & Zisserman A (2005) A statistical approach to texture classification from single images. *International Journal of Computer Vision* 62(1-2): 61–81.
- Šroubek F & Flusser J (2005) Multichannel blind deconvolution of spatially misaligned images. *IEEE Transactions on Image Processing* 14(7): 874–883.
- Šroubek F, Flusser J & Cristóbal G (2007) Multiframe blind deconvolution coupled with frame registration and resolution enhancement. In: Campisi P & Egiazarian K (eds) *Blind Image Deconvolution: Theory and Applications.*, 317–346. CRC Press.
- Wallin A & Kiibler O (1995) Complete sets of complex Zernike moment invariants and the role of the pseudoinvariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(11): 830–834.
- Wang Y, Ostermann J & Zhang YQ (2002) *Video processing and communications*. Prentice Hall, Upper Saddle River, New Jersey.
- Wee CY & Paramesran R (2007) Derivation of blur-invariant features using orthogonal Legendre moments. *IET Computer Vision* 1(2): 66–77.
- Wong WH, Siu WC & Lam KM (1995) Generation of moment invariants and their uses for character recognition. *Pattern Recognition Letters* 16(2): 115–123.
- Wood J (1996) Invariant pattern recognition: A review. *Pattern Recognition* 29(1): 1–17.
- Yang Z & Cohen FS (1999) Cross-weighted moments and affine invariants for image registration and matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(8): 804–814.
- Zamfir M, Zamfir A, Buzoloiu V & Drimborean A (2007) A study of the influence of the PSF accuracy on the quality of image deblurring. *Proc. International Symposium on Signals Circuits and Systems*, Iasi, Romania, 1–4.
- Zhang J, Marszałek M, Lazebnik S & Schmid C (2007) Local features and kernels for classification of texture and object categories: A comprehensive study. *International Journal of Computer Vision* 73(2): 213–238.
- Zhang J, Tan T & Ma L (2002a) Invariant texture segmentation via circular gabor filters. *Proc. International Conference on Pattern Recognition*, Quebec, Canada, 901–904.
- Zhang JG & Tan TN (2002) Brief review of invariant texture analysis methods. *Pattern Recognition* 35(3): 735–747.
- Zhang Y, Wen C & Zhang Y (2000a) Estimation of motion parameters from blurred images. *Pattern Recognition Letters* 21: 425–433.
- Zhang Y, Wen C, Zhang Y & Soh YC (2002b) Determination of blur and affine combined invariants by normalization. *Pattern Recognition* 35(1): 211–221.
- Zhang Y, Zhang Y & Wen C (2000b) A new focus measure method using moments. *Image and Vision Computing* 18: 959–965.

- Zitová B & Flusser J (2002) Estimation of camera planar motion from blurred images. Proc. IEEE International Conference on Image Processing, Rochester, New York, 329–332.
- Zitová B & Flusser J (2003) Image registration methods: a survey. Image and Vision Computing 21(11): 977–1000.

Original articles

- I Ojansivu V & Heikkilä J (2006) Motion blur concealment of digital video using invariant features. Proc. 8th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 2006), Antwerp, Belgium. Lecture Notes in Computer Science 4179: 35-45.
- II Ojansivu V & Heikkilä J (2007) Object recognition using frequency domain blur invariant features. Proc. 15th Scandinavian Conference on Image Analysis (SCIA 2007), Aalborg, Denmark. Lecture Notes in Computer Science 4522: 243-252.
- III Ojansivu V & Heikkilä J (2007) Image registration using blur-invariant phase correlation. IEEE Signal Processing Letters 14(7): 449-452.
- IV Ojansivu V & Heikkilä J (2007) Blur invariant registration of rotated, scaled and shifted images. Proc. 15th European Signal Processing Conference (EUSIPCO 2007), Poznań, Poland: 1755-1759.
- V Ojansivu V & Heikkilä J (2007) A method for blur and similarity transform invariant object recognition. Proc. 14th International Conference on Image Analysis and Processing (ICIAP 2007), Modena, Italy: 583-588.
- VI Ojansivu V & Heikkilä J (2008) A method for blur and affine invariant object recognition using phase-only bispectrum. Proc. 3rd International Conference on Image Analysis and Recognition (ICIAR 2008), Póvoa de Varzim, Portugal. Lecture Notes in Computer Science 5112: 527-536.
- VII Ojansivu V & Heikkilä J (2008) Blur insensitive texture classification using local phase quantization. Proc. 3rd International Conference on Image and Signal Processing (ICISP 2008), Cherbourg-Octeville, France. Lecture Notes in Computer Science 5099: 236-243.
- VIII Ojansivu V & Rahtu E & Heikkilä J (2008) Rotation Invariant Local Phase Quantization for Blur Insensitive Texture Analysis. Proc. 19th International Conference on Pattern Recognition (ICPR 2008), Tampa Bay, Florida, USA: 1-4.
- IX Ojansivu V & Heikkilä J (2009) Weighted DFT based blur invariants for pattern recognition. Proc. 16th Scandinavian Conference on Image Analysis (SCIA 2009), Oslo, Norway. Lecture Notes in Computer Science 5575: 71-80.

Reprinted with permission from Springer (I, II, VI, VII, IX), IEEE (III, V, VIII), and EURASIP (IV).

The original publications are not included in the electronic version of the dissertation.

ACTA UNIVERSITATIS OULUENSIS
SERIES C TECHNICA

322. Liedes, Toni (2009) Improving the performance of the semi-active tuned mass damper
323. Marina Tyunina & Orest Vendik (Eds.) (2009) Proceedings of the 16th International Student Seminar "Microwave and optical applications of novel phenomena and technologies", June 8–9, Oulu, Finland
324. Belt, Pekka (2009) Improving verification and validation activities in ICT companies—product development management approach
325. Harri Haapasalo & Hanna Kropsu-Vehkaperä (Eds.) (2009) The 3rd Nordic Innovation Research Conference - NIR 2008—IEM Arctic Workshop
326. Selek, István (2009) Novel evolutionary methods in engineering optimization—towards robustness and efficiency
327. Härkönen, Janne (2009) Improving product development process through verification and validation
328. Peiponen, Kai-Erik (2009) Optical spectra analysis of turbid liquids
329. Kettunen, Juha (2009) Essays on strategic management and quality assurance
330. Ahonen, Timo (2009) Face and texture image analysis with quantized filter response statistics
331. Uusipaavaliemi, Sari (2009) Framework for analysing and developing information integration. A study on steel industry maintenance service supply chain
332. Risikko, Tanja (2009) Safety, health and productivity of cold work. A management model, implementation and effects
333. Virtanen, Markku (2009) Mathematical modelling of flow and transport as link to impacts in multidiscipline environments
334. Liimatainen, Henrikki (2009) Interactions between fibres, fines and fillers in papermaking. Influence on dewatering and retention of pulp suspensions
335. Ghaboosi, Kaveh (2009) Intelligent medium access control for the future wireless networks
336. Möttönen, Matti (2009) Requirements engineering. Linking design and manufacturing in ICT companies
337. Leinonen, Jouko (2009) Analysis of OFDMA resource allocation with limited feedback

Book orders:
OULU UNIVERSITY PRESS
P.O. Box 8200, FI-90014
University of Oulu, Finland

Distributed by
OULU UNIVERSITY LIBRARY
P.O. Box 7500, FI-90014
University of Oulu, Finland

S E R I E S E D I T O R S

A
SCIENTIAE RERUM NATURALIUM

Professor Mikko Siponen

B
HUMANIORA

University Lecturer Elise Kärkkäinen

C
TECHNICA

Professor Hannu Heusala

D
MEDICA

Professor Helvi Kyngäs

E
SCIENTIAE RERUM SOCIALIUM

Senior Researcher Eila Estola

F
SCRIPTA ACADEMICA

Information officer Tiina Pistokoski

G
OECONOMICA

University Lecturer Seppo Eriksson

EDITOR IN CHIEF

University Lecturer Seppo Eriksson

PUBLICATIONS EDITOR

Publications Editor Kirsti Nurkkala

ISBN 978-951-42-9254-5 (Paperback)

ISBN 978-951-42-9255-2 (PDF)

ISSN 0355-3213 (Print)

ISSN 1796-2226 (Online)

