

ICIP 2005 Tutorial

**Image Registration: A Survey and  
Recent Advances**

Presented by

**Barbara Zitová  
Jan Flusser  
Filip Šroubek**

Institute of Information Theory and Automation  
Academy of Sciences of the Czech Republic  
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic

E-mail: {zitova,flusser,sroubekf}@utia.cas.cz

This is complementary material for ICIP'05 tutorial "Image Registration: A Survey and Recent Advances". It aims to present a review of recent as well as classic image registration methods. Image registration is the process of overlaying two or more images of the same scene taken at different times, under different lighting conditions, from different viewpoints and/or by different sensors. The major registration purpose is to remove or suppress geometric distortions between the reference and sensed images. Image registration is a crucial step in all image analysis tasks in which the final information is gained by combining various data sources. Image registration is required, among others, in remote sensing (multispectral classification, image fusion, environmental monitoring, change detection, image mosaicing, weather forecasting, integrating information into GIS), in medicine (combining CT and NMR data to obtain more complete information about the patient, monitoring of tumor growth, comparison of the patient data with anatomical atlases), in cartography (map updating), and in military applications (target localization, missile navigation).

**Keywords:** Image registration, feature detection, feature matching, mapping function, re-sampling

# Part I: Survey

## 1 Introduction

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It geometrically aligns two images - the reference and sensed images. The present differences between images are introduced due to different imaging conditions. Image registration is a crucial step in all image analysis tasks in which the final information is gained from the combination of various data sources like in image fusion, change detection, and multichannel image restoration. Typically, registration is required in remote sensing (multispectral classification, environmental monitoring, change detection, image mosaicing, weather forecasting, creating super-resolution images, integrating information into geographic information systems - GIS), in medicine (combining CT and NMR data to obtain more complete information about the patient, monitoring tumor growth, treatment verification, comparison of the patient's data with anatomical atlases), in cartography (map updating), and in computer vision (target localization, automatic quality control), to name a few.

During the last decades, image acquisition devices have undergone rapid development and growing amount and diversity of obtained images invoked the research on automatic image registration. A comprehensive survey of image registration methods was published in 1992 by Brown [26]. This material is based on the survey [237] covering relevant approaches introduced later and in this way mapping the current development of registration techniques. According to the database of the Institute of Scientific Information (ISI), in the last ten years more than 1000 papers were published on the topic of image registration. Methods published before 1992 that became classic or introduced key ideas, which are still in use, are included as well to retain the continuity and to give complete view of image registration research. We do not contemplate to go into details of particular algorithms or describe results of comparative experiments, rather we want to summarize main approaches and point out interesting parts of the registration methods.

In Section 2 various aspects and problems of image registration will be discussed. Both area-based and feature-based approaches to feature selection are described in Section 3. Section 4 reviews the existing algorithms for feature matching. Methods for mapping function design are given in Section 5. Finally, Section 6 surveys main techniques for image transformation and resampling. Evaluation of the image registration accuracy is covered in Section 7. Section 8 concludes main trends in the research on registration methods and offers the outlook for the future.

## 2 Image registration methodology

Image registration, as it was mentioned above, is widely used in remote sensing, medical imaging, computer vision etc. In general, its applications can be divided into four main

groups according to the manner of the image acquisition:

**Different viewpoints (multiview analysis):**

Images of the same scene are acquired from different viewpoints. The aim is to gain larger a 2D view or a 3D representation of the scanned scene.

Examples of applications: *Remote sensing* – mosaicing of images of the surveyed area. *Computer vision* – shape recovery (shape from stereo).

**Different times (multitemporal analysis):**

Images of the same scene are acquired at different times, often on regular basis, and possibly under different conditions. The aim is to find and evaluate changes in the scene which appeared between the consecutive image acquisitions.

Examples of applications: *Remote sensing* – monitoring of global land usage, landscape planning. *Computer vision* – automatic change detection for security monitoring, motion tracking. *Medical imaging* - monitoring of the healing therapy, monitoring of the tumor evolution.

**Different sensors (multimodal analysis):**

Images of the same scene are acquired by different sensors. The aim is to integrate the information obtained from different source streams to gain more complex and detailed scene representation.

Examples of applications: *Remote sensing* – fusion of information from sensors with different characteristics like panchromatic images, offering better spatial resolution, color/multispectral images with better spectral resolution, or radar images independent of cloud cover and solar illumination. *Medical imaging* – combination of sensors recording the anatomical body structure like MRI (magnetic resonance image), ultrasound or CT (computer tomography) with sensors monitoring functional and metabolic body activities like PET (positron emission tomography), SPECT (single photon emission computed tomography) or MRS (magnetic resonance spectroscopy). Results can be applied, for instance, in radiotherapy and nuclear medicine.

**Scene to model registration:**

Images of a scene and a model of the scene are registered. The model can be a computer representation of the scene, for instance maps or digital elevation models (DEM) in GIS, another scene with similar content (another patient), "average" specimen, etc. The aim is to localize the acquired image in the scene/model and/or to compare them.

Examples of applications: *Remote sensing* – registration of aerial or satellite data into maps or other GIS layers. *Computer vision* – target template matching with real-time images, automatic quality inspection. *Medical imaging* – comparison of the patient's image with digital anatomical atlases, specimen classification.

Due to the diversity of images to be registered and due to various types of degradations it is impossible to design a universal method applicable to all registration tasks. Every method

should take into account not only the assumed type of geometric deformation between the images but also radiometric deformations and noise corruption, required registration accuracy and application-dependent data characteristics.

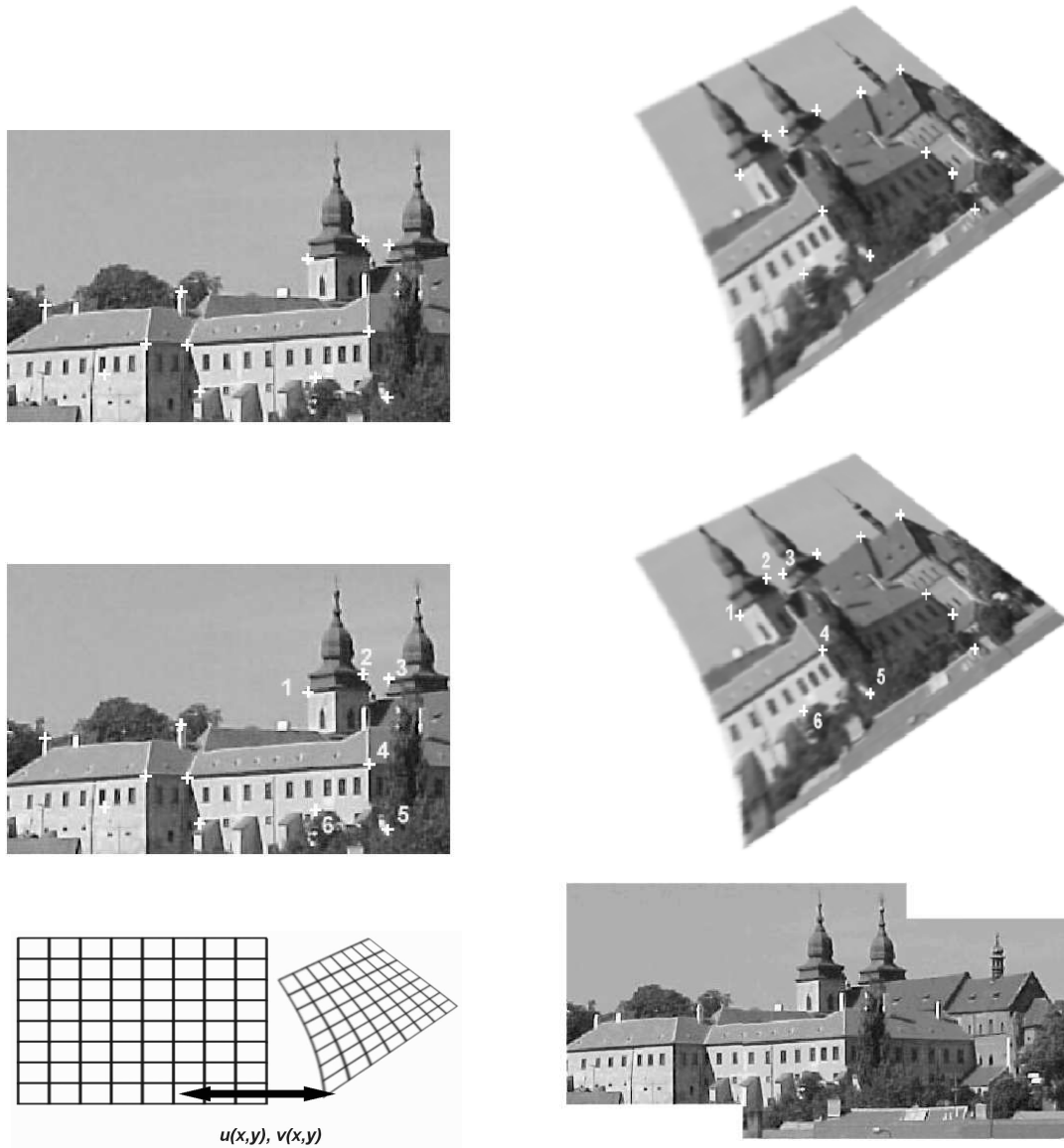


Figure 1: **Four steps of image registration:** Top row - feature detection (corners were used as the features in this case). Middle row - *feature matching* by invariant descriptors (the corresponding pairs are marked by numbers). Bottom left - *transform model estimation* exploiting the established correspondence. Bottom right - *image resampling and transformation* using appropriate interpolation technique.

Nevertheless, the majority of the registration methods consists of the following four steps (see Fig.1):

- *Feature detection.* Salient and distinctive objects (closed-boundary regions, edges, contours, line intersections, corners, etc.) are manually or, preferably, automatically

detected. For further processing, these features can be represented by their point representatives (centers of gravity, line endings, distinctive points), which are called *control points* (CPs) in the literature.

- *Feature matching.* In this step, the correspondence between the features detected in the sensed image and those detected in the reference image is established. Various feature descriptors and similarity measures along with spatial relationships among the features are used for that purpose.
- *Transform model estimation.* The type and parameters of the so-called *mapping functions*, aligning the sensed image with the reference image, are estimated. The parameters of the mapping functions are computed by means of the established feature correspondence.
- *Image resampling and transformation.* The sensed image is transformed by means of the mapping functions. Image values in non-integer coordinates are computed by the appropriate interpolation technique.

The implementation of each registration step has its typical problems. First, we have to decide what kind of features is appropriate for the given task. The features should be distinctive objects, which are frequently spread over the images and which are easily detectable. Usually, the physical interpretability of the features is demanded. The detected feature sets in the reference and sensed images must have enough common elements, even in situations when the images do not cover exactly the same scene or when there are object occlusions or other unexpected changes. The detection methods should have good localization accuracy and should not be sensitive to the assumed image degradation. In an ideal case, the algorithm should be able to detect the same features in all projections of the scene regardless of the particular image deformation.

In the feature matching step, problems caused by an incorrect feature detection or by image degradations can arise. Physically corresponding features can be dissimilar due to the different imaging conditions and/or due to the different spectral sensitivity of the sensors. The choice of the feature description and similarity measure has to consider these factors. The feature descriptors should be invariant to the assumed degradations. Simultaneously, they have to be discriminable enough to be able to distinguish among different features as well as sufficiently stable so as not to be influenced by slight unexpected feature variations and noise. The matching algorithm in the space of invariants should be robust and efficient. Single features without corresponding counterparts in the other image should not affect its performance.

The type of the mapping functions should be chosen according to the *a priori* known information about the acquisition process and expected image degradations. If no *a priori* information is available, the model should be flexible and general enough to handle all possible degradations which might appear. The accuracy of the feature detection method, the reliability of feature correspondence estimation, and the acceptable approximation error

need to be considered too. Moreover, the decision about which differences between images have to be removed by registration has to be done. It is desirable not to remove the differences we are searching for if the aim is a change detection. This issue is very important and extremely difficult.

Finally, the choice of the appropriate type of resampling technique depends on the trade-off between the demanded accuracy of the interpolation and the computational complexity. The nearest-neighbor or bilinear interpolation are sufficient in most cases; however, some applications require more precise methods.

Because of its importance in various application areas as well as because of its complicated nature, image registration has been the topic of much recent research. The historically first survey paper [66] covers mainly the methods based on image correlation. Probably the most exhaustive review of the general-purpose image registration methods is in [26]. Registration techniques applied particularly in medical imaging are summarized in [204], [128], [90], and [116]. In [9] the surface based registration methods in medical imaging are reviewed. Volume-based registration is reviewed in [40]. The registration methods applied mainly in remote sensing are described and evaluated in [61], [84] and [111]. Big evaluation project of different registration methods was run in Vanderbilt university [217].

Registration methods can be categorized with respect to various criteria. The ones usually used are the application area, dimensionality of data, type and complexity of assumed image deformations, computational cost, and the essential ideas of the registration algorithm. Here, the classification according to the essential ideas is chosen, considering the decomposition of the registration into the described four steps. The techniques exceeding this four-step framework are covered according to their major contribution.

## 3 Feature detection

Formerly, the features were objects manually selected by an expert. During an automation of this registration step, two main approaches to feature understanding have been formed.

### 3.1 Area-based methods

Area-based methods put emphasis rather on the feature matching step than on their detection. No features are detected in these approaches so the first step of image registration is omitted. The methods belonging to this class will be covered in sections corresponding to the other registration steps.

### 3.2 Feature-based methods

The second approach is based on the extraction of salient structures – features – in the images. Significant regions (forests, lakes, fields), lines (region boundaries, coastlines, roads, rivers) or points (region corners, line intersections, points on curves with high curvature) are understood as features here. They should be distinct, spread all over the image and efficiently detectable in both images. They are expected to be stable in time to stay at fixed positions during the whole experiment.

The comparability of feature sets in the sensed and reference images is assured by the invariance and accuracy of the feature detector and by the overlap criterion. In other words, the number of common elements of the detected sets of features should be sufficiently high, regardless of the change of image geometry, radiometric conditions, presence of additive noise, and of changes in the scanned scene. The "remarkableness" of the features is implied by their definition. In contrast to the area-based methods, the feature-based ones do not work directly with image intensity values. The features represent information on higher level. This property makes feature-based methods suitable for situations when illumination changes are expected or multisensor analysis is demanded.

**Region features.** The region-like features can be the projections of general high contrast closed-boundary regions of an appropriate size [75], [54], water reservoirs, and lakes [74], [92], buildings [96], forests [174], urban areas [170] or shadows [24]. The general criterion of closed-boundary regions is prevalent. The regions are often represented by their centers of gravity, which are invariant with respect to rotation, scaling, and skewing and stable under random noise and gray level variation.

Region features are detected by means of segmentation methods [144]. The accuracy of the segmentation can significantly influence the resulting registration. Goshtasby et al. [75] proposed a refinement of the segmentation process to improve the registration quality. The segmentation of the image was done iteratively together with the registration; in every iteration, the rough estimation of the object correspondence was used to tune the segmentation parameters. They claimed the subpixel accuracy of registration could be achieved.

Recently, selection of region features invariant with respect to change of scale caught attention. Alhichri and Kamel [2] proposed the idea of virtual circles, using distance transform. Affinely invariant neighborhoods were described in [203], based on Harris corner detector [141] and edges (curved or straight) going through detected corners. Different approach to this problem using Maximally Stable Extremal Regions based on homogeneity of image intensities was presented by Matas et al. [132].

**Line features.** The line features can be the representations of general line segments [216], [137], [96], object contours [117], [36], [77], coastal lines, [177], [129], roads [119] or elongated anatomic structures [213] in medical imaging. Line correspondence is usually expressed by pairs of line ends or middle points.

Standard edge detection methods, like Canny detector [28] or a detector based on the Laplacian of Gaussian [131], are employed for the line feature detection. The survey of existing edge detection method together with their evaluation can be found in [236]. Li et al. [117] proposed to exploit the already detected features in the reference image (optical data) for the detection of lines in the sensed images (SAR images with speckle noise, which is a typical degradation present in this type of data). They applied elastic contour extraction. The comparison of different operators for the feature edge detection and the ridge detection in multimodal medical images is presented by Maintz et al. [127], [126].

**Point features.** The point features group consists of methods working with line intersections [184], [207], road crossings [170], [82], centroids of water regions, oil and gas pads [199],



high variance points [45], local curvature discontinuities detected using the Gabor wavelets [233], [130], inflection points of curves [3], [11], local extrema of wavelet transform [60], [94], the most distinctive points with respect to a specified measure of similarity [120], and corners [20], [215], [96].

The core algorithms of feature detectors in most cases follow the definitions of the "point" as line intersection, centroid of closed-boundary region or local modulus maxima of the wavelet transform. Corners form specific class of features, because "to-be-a-corner" property is hard to define mathematically (intuitively, corners are understood as points of high curvature on the region boundaries). Much effort has been spent in developing precise, robust, and fast method for corner detection. A survey of corner detectors can be found in [234], [181], [164] and the most up-to-date and exhaustive in [165]. The latter also analyzes localization properties of the detectors. Corners are widely used as control points mainly because of their invariance to imaging geometry and because they are well perceived by a human observer.

Kitchen and Rosenfeld [105] proposed to exploit the second-order partial derivatives of the image function for corner detection. Dreschler and Nagel [43] searched for the local extrema of the Gaussian curvature. However, corner detectors based on the second-order derivatives of the image function are sensitive to noise. Thus Förstner [64] developed a more robust, although time consuming, corner detector, which is based on the first-order derivatives only. The reputable Harris detector (also referred to as the Plessey detector) [141] is in fact its inverse. The application of the Förstner detector is described in [112], where it is used for the registration of dental implants images.

More intuitive approach was chosen by Smith and Brady [182] in their robust SUSAN method. As the criterion they used the size of the area of the same color as that of the central pixel. Trajkovic and Hedley [201] designed their operator using the idea that the change of the image intensity at the corners should be high in all directions. Recently, Zitová et al. [239] proposed a parametric corner detector, which does not employ any derivatives and which was designed to handle blurred and noisy data. Rohr et al. designed corner detectors, even for 3-D data, allowing user interaction [167].

The number of detected points can be very high, which increases the computational time necessary for the registration. Several authors proposed methods for an efficient selection of a subset of points (better than random) which does not degrade the quality of the resulting registration. Goshtasby [74] used only points belonging to a convex hull of the whole set. Lavine [109] proposed to use points forming the minimum spanning trees of sets. Ehlers [45] merged points into "clumps" – large dense clusters.

**Summary.** To summarize, the use of feature-based methods is recommended if the images contain enough distinctive and easily detectable objects. This is usually the case of applications in remote sensing and computer vision. The typical images contain a lot of details (towns, rivers, roads, forests, room facilities, etc). On the other hand, medical images are not so rich in such details and thus area-based methods are usually employed here. Sometimes, the lack of distinctive objects in medical images is solved by the interactive selection done by an expert or by introducing extrinsic features, rigidly positioned with respect to the patient (skin markers, screw markers, dental adapters etc.) [128]. The applicability of area-based and feature-based methods for images with various contrast and sharpness is

analyzed in [160]. Recently, registration methods using simultaneously both area-based and feature-based approaches have started to appear (for example see [89]).

## 4 Feature matching

The detected features in the reference and sensed images can be matched by means of the image intensity values in their close neighborhoods, the feature spatial distribution, or the feature symbolic description. Some methods, while looking for the feature correspondence, simultaneously estimate the parameters of mapping functions and thus merge the second and third registration steps.

In the following paragraphs, the two major categories (area-based and feature-based methods, respectively) are retained and further classified into subcategories according to the basic ideas of the matching methods.

### 4.1 Area-based methods

Area-based methods, sometimes called correlation-like methods or template matching [61] merge the feature detection step with the matching part. These methods deal with the images without attempting to detect salient objects. Windows of predefined size or even entire images are used for the correspondence estimation during the second registration step, see e.g. [153], [12], [4].

The limitations of the area-based methods originate in their basic idea. Firstly, the rectangular window, which is most often used, suits the registration of images which locally differ only by a translation. If images are deformed by more complex transformations, this type of the window is not able to cover the same parts of the scene in the reference and sensed images (the rectangle can be transformed to some other shape). Several authors proposed to use circular shape of the window for mutually rotated images. However, the comparability of such simple-shaped windows is violated too if more complicated geometric deformations (similarity, perspective transforms, etc.) are present between images.

Another disadvantage of the area-based methods refers to the "remarkableness" of the window content. There is high probability that a window containing a smooth area without any prominent details will be matched incorrectly with other smooth areas in the reference image due to its non-saliency. The features for registration should be preferably detected in distinctive parts of the image. Windows, whose selection is often not based on their content evaluation, may not have this property.

Classical area-based methods like cross-correlation exploit for matching directly image intensities, without any structural analysis. Consequently, they are sensitive to the intensity changes, introduced for instance by noise, varying illumination, and/or by using different sensor types.

**Correlation-like methods.** The classical representative of the area-based methods is the normalized cross-correlation (CC) and its modifications [154].

$$CC(i, j) = \frac{\sum_W (W - E(W))(I_{(i,j)} - E(I_{(i,j)}))}{\sqrt{\sum_W (W - E(W))^2} \sqrt{\sum_{I_{(i,j)}} (I_{(i,j)} - E(I_{(i,j)}))^2}}$$

This measure of similarity is computed for window pairs from the sensed and reference images and its maximum is searched. The window pairs for which the maximum is achieved are set as the corresponding ones (see Fig.2). If the subpixel accuracy of the registration is demanded, the interpolation of the CC measure values needs to be used. Although the cross-correlation based registration can exactly align mutually translated images only, it can also be successfully applied when slight rotation and scaling are present.

There are generalized versions of CC for geometrically more deformed images. They compute the cross-correlation for each assumed geometric transformation of the sensed image window [86] and are able to handle even more complicated geometric deformations than the translation – usually the similarity transform. Berthilsson [17] tried to register in this manner even affinely deformed images and Simper [179] proposed to use a divide and conquer system and the cross-correlation technique for registering images differing by perspective changes as well as changes due to the lens imperfections. The computational load, however, grows very fast with the increase of the transformation complexity. In case the images/objects to be registered are partially occluded the extended CC method based on increment sign correlation [102] can be applied [103].

Similar to the cross-correlation methods is the sequential similarity detection algorithm (SSDA) [12]. It uses the sequential search approach and a computationally simpler distance measure than the cross-correlation. It accumulates the sum of absolute differences of the image intensity values (matrix  $l_1$  norm) and applies the threshold criterion – if the accumulated sum exceeds the given threshold, the candidate pair of windows from the reference and sensed images is rejected and the next pair is tested. The method is likely to be less accurate than the cross-correlation but it is faster. Sum of squared differences similarity measure was used in [222] for iterative estimation of perspective deformation using piecewise affine estimates for image decomposed to small patches.

Recently big interest in the area of multimodal registration has been paid to the correlation ratio based methods. In opposite to classical cross-correlation, this similarity measure can handle intensity differences between images due to the usage of different sensors - multimodal images. It supposes that intensity dependence can be represented by some function. Comparison of this approach to several other algorithms developed for multimodal data can be found in [163]. In case of noisy images with certain characteristic (fixed-pattern noise), projection-based registration [27], working with accumulated image rows and columns, respectively, outperforms classical cross-correlation.

Huttenlocher et al. [99] proposed a method working with other type of similarity measure. They registered binary images (the output of an edge detector) transformed by translation or translation plus rotation, by means of the Hausdorff distance (HD). They compared the HD based algorithm with the cross-correlation. Especially on images with perturbed pixel locations, which are problematic for cross-correlation, HD outperforms the cross-correlation.

Two main drawbacks of the correlation-like methods are the flatness of the similarity

measure maxima (due to the self-similarity of the images) and high computational complexity. The maximum can be sharpened by preprocessing or by using the edge or vector correlation. Pratt [153] applied, prior to the registration, image filtering to improve the CC performance on noisy or highly correlated images. Van Wie [205] and Anuta [6] employed the edge-based correlation, which is computed on the edges extracted from the images rather than on the original images themselves. In this way, the method is less sensitive to intensity differences between the reference and sensed images, too. Extension of this approach, called vector-based correlation, computes the similarity measures using various representations of the window.

Despite the limitations mentioned above, the correlation-like registration methods are still often in use, particularly thanks to their easy hardware implementation, which makes them useful for real-time applications.

**Fourier methods.** If an acceleration of the computational speed is needed or if the images were acquired under varying conditions or they are corrupted by frequency-dependent noise, then Fourier methods are preferred rather than the correlation-like methods. They exploit the Fourier representation of the images in the frequency domain. The phase correlation method is based on the Fourier Shift Theorem [23] and was originally proposed for the registration of translated images. It computes the cross-power spectrum of the sensed and reference images and looks for the location of the peak in its inverse (see Fig.2).

$$\frac{\mathcal{F}(f)\mathcal{F}(g)^*}{|\mathcal{F}(f)\mathcal{F}(g)^*|} = e^{2\pi i(ux_0+vy_0)}$$

The method shows strong robustness against the correlated and frequency dependent noise and nonuniform, time varying illumination disturbances. The computational time savings are more significant if the images, which are to be registered, are large.

De Castro and Morandi [29] introduced an extension of the phase correlation for additional rotation transform. If the change of image scale is present too, the images can be registered using the combination of polar-log mapping of the spectral magnitude (which corresponds to Fourier-Mellin transform) and the phase correlation [159], [31] or cepstrum filter [112]. The applications of the extended algorithm in remote sensing (SPOT images) and medical imaging (MR images) are described in [31]. The testing of the accuracy of the method in simulated conditions (registration of deformed and noisy images of ocular fundus) was performed with satisfying results [34]. Affinely distorted images were registered by means of phase correlation and log-polar mapping in [221]. Application of phase correlation in 3D is described in [124]. Another application exploiting the Fourier transform is described in [6]. The authors proposed to compute the correlation in frequency domain. This method can handle multimodal images when applied to the edge representations instead of the original graylevel images. Extension of phase correlation to subpixel registration by means of the analytic expression of phase correlation on downsampled images was introduced by Foroosh et al. [63].

**Mutual information methods.** The mutual information (MI) methods are the last group of the area-based methods to be reviewed here. They have appeared recently and represent the leading technique in multimodal registration. Registration of multimodal images is the

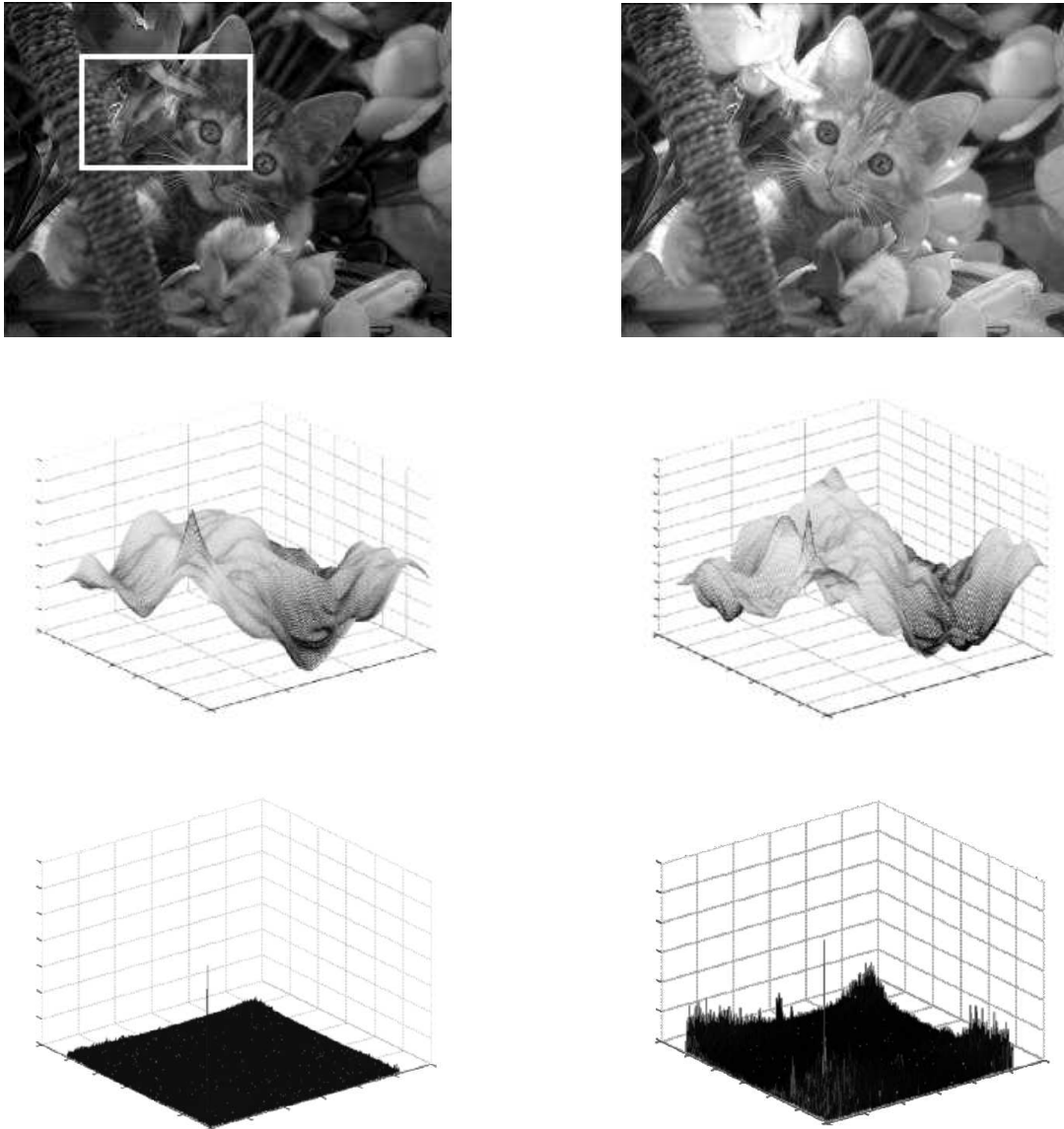


Figure 2: **Area-based matching methods:** registration of small template to the whole image using *normalized cross-correlation* (middle row) and *phase correlation* (bottom row). The maxima identify the matching positions. The template is of the same spectral band as the reference image (the graphs on the left depict red–red channel matching) and of different spectral band (the graphs on the right demonstrate red–blue channel matching). In a general case the normalized cross-correlation could fail in case of multimodal data.

difficult task, but often necessary to solve, especially in medical imaging. The comparison of anatomical and functional images of the patient’s body can lead to a diagnosis which would be impossible to gain otherwise. Remote sensing often makes use of the exploitation of more sensor types, too.

The mutual information, originating from the information theory, is a measure of statistical dependency between two data sets and it is particularly suitable for registration of

images from different modalities. Mutual information between two random variables  $X$  and  $Y$  is given by

$$MI(X, Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X, Y),$$

where  $H(X) = -E_X(\log(P(X)))$  represents entropy of random variable and  $P(X)$  is the probability distribution of  $X$ . The method is based on the maximization of MI (see Fig.3) Often the speed up of the registration is implemented, exploiting the coarse-to-fine resolution strategy (the pyramidal approach).

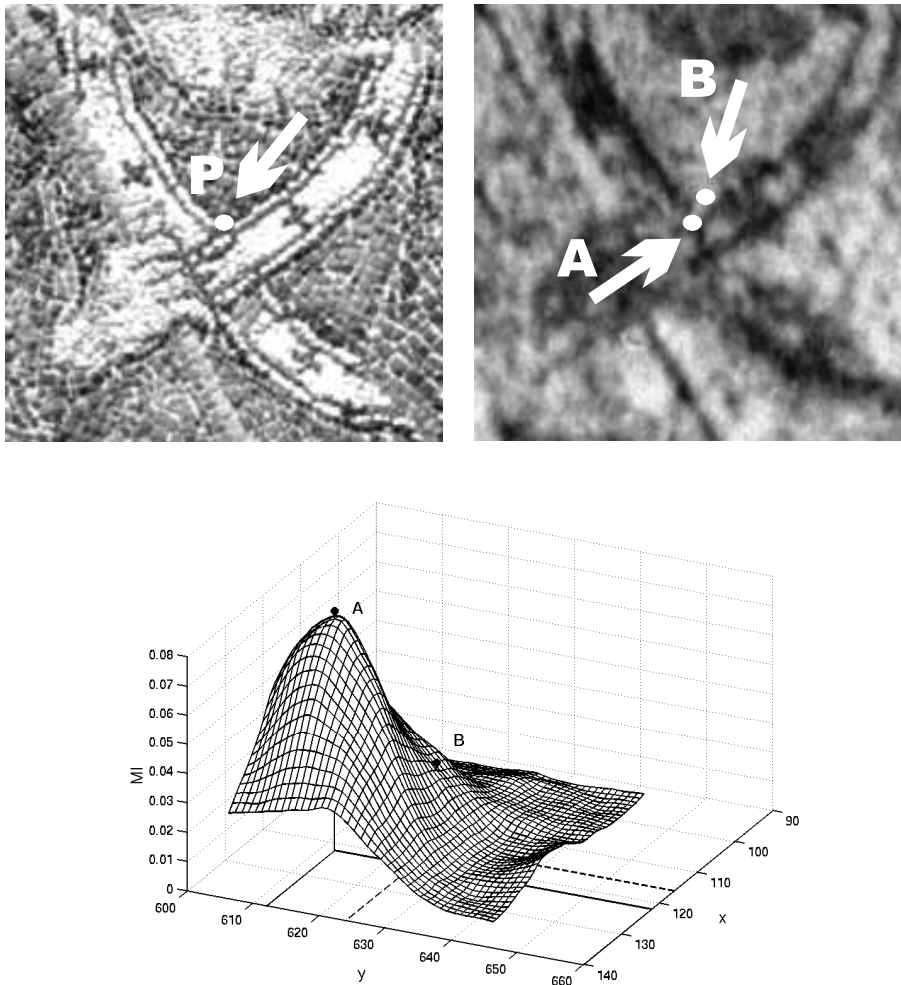


Figure 3: **Mutual information:** MI criterion (bottom row) computed in the neighborhood of point  $P$  between new and old photographs of the mosaic (top row). Maximum of MI shows the correct matching position (point  $A$ ). Point  $B$  indicates the false matching position selected previously by the human operator. The mistake was caused by poor image quality and by complex nature of the image degradations.

One of the first articles proposing this technique is Viola and Wells [210]. The authors described the application of MI for the registration by magnetic resonance images as well as

for the 3-D object model matching to the real scene. MI was maximized using the gradient descent optimization method. Thévenaz and Unser [196], [195], [197] tried to combine various approaches, solving individual steps of MI registration. They employed the Parzen window for the joint probability computation and the Jeeves method ([196]) or the Marquardt-Levenberg method ([195]) to maximize the MI. To speed up the computation, they used spline pyramids [197]. Ritter et al. [161] used hierarchical search strategy together with simulated annealing to find the maximum of the MI. Studholme et al. [186] compared three similar measures of information dependency – the joint entropy, the MI, and the normalized MI (a new idea, claimed to outperform the others). They applied discrete histogram estimates of the joint probability distribution instead of application of the Parzen window used in previous work. Maximization was achieved by using a multiresolution hill climbing algorithm. They registered MR-CT and MR-PET images of a human brain. Maes et al. [125] optimized the MI by means of the Brent’s method and the Powell’s multidimensional direction set method to register MR, CT, and PET images of a human brain that differ by similarity transform. MI applied to breast MR images is described in [171]. The authors proposed to model the global deformation present between the images by a combination of the affine transformations and the spline-based free form deformation. Likar and Pernus [121] studied the performance of different methods for the joint probability estimation in registration of muscle fibre images. The comparison of the basic MI registration with the version employing the coarse-to-fine speed up was done in [151]. The comparison of MI to six other registration methods including the normalized cross-correlation and the gradient correlation is described in [150]. The relation of MI to other area-based similarity (correlation coefficients, correlation ratio) measures is described in [162] using the formulation of maximum likelihood estimation problem.

The above mentioned MI methods work with the entire image data and directly with image intensities. Rangarajan et al. [157] applied MI on extracted features (points of the area borders), but this approach is still rare. Similar to MI, coming from the theory of information, is similarity measure based on cross-entropy [235].

**Optimization methods.** Finding the minimum of dissimilarity measure (penalty function) or the maximum of similarity measure is a multidimensional optimization problem, where the number of dimensions corresponds to the degrees of freedom of the expected geometrical transformation. The only method yielding global extreme solution is an exhaustive search over the entire image. Although it is computationally demanding, it is often used if only translations are to be estimated.

In case of transformations with more degrees of freedom or in case of more complex similarity measures, sophisticated optimization algorithms are required, which help to localize the maxima or minima, respectively. The application of Gauss-Newton numerical minimization algorithm for minimizing the sum of squared differences is described in [175], where the projective geometric deformation was used. In [210] maxima of mutual information was found using the gradient descent optimization method. Levenberg-Marquardt optimization method was applied in [173] to minimize the variance in intensities of corresponding pixels. The images were registered by means of the projective transformation model plus the lens distortion model. The combination of the Levenberg-Marquardt method and the sum of the squared differences metric is described in [194]. Similarly, Wolberg and Zokai [222] used

this combination for registering of perspectively deformed images. The Powell’s multidimensional direction set method [100] is applied in Maes et al. [125]. Starink and Backer [183] tried to minimize a dissimilarity measure defined on point pairs by means of simulated annealing. Another optimization method, suited for multimodal data registration was introduced in [101] and its applicability was proved in combination with mutual information and correlation ratio. Again, the optimization methods can be speeded up by the pyramidal approach.

There should be noted one thing with respect to these optimization methods. Sometimes next to the dissimilarity measure term the formula to be minimized contains as well so-called regularization or penalty term, which interconnects the transformation and data to be transformed (see [85]). These two terms together form the cost function (energy) associated with the registration and the aim of the optimization methods is to minimize it. In literature such methods can be referred to as *energy minimization methods*. The regularization term is usually omitted in case of rigid body transforms, but in nonrigid transformations such as elastic or fluid registration methods, described more in detail in Section 5, is present.

## 4.2 Feature-based methods

We assume that two sets of features in the reference and sensed images represented by the control points (CPs) (points themselves, end points or centers of line features, centers of gravity of regions, etc.) have been detected. The aim is to find the pair-wise correspondence between them using their spatial relations or various descriptors of features.

**Methods using spatial relations.** Methods based primarily on the spatial relations among the features are usually applied if detected features are ambiguous or if their neighborhoods are locally distorted. The information about the distance between the CPs and about their spatial distribution is exploited.

Goshtasby in [74] described the registration based on the graph matching algorithm. He was evaluating the number of features in the sensed image that, after the particular transformation, fall within a given range next to the features in the reference image. The transformation parameters with the highest score were then set as a valid estimate.

Clustering technique, presented by Stockman et al. [184], tries to match points connected by abstract edges or line segments. The assumed geometrical model is the similarity transform. For every pair of CPs from both the reference and sensed images, the parameters of the transformation which maps the points on each other are computed and represented as a point in the space of transform parameters. The parameters of transformations that closely map the highest number of features tend to form a cluster, while mismatches fill the parameter space randomly. The cluster is detected and its centroid is assumed to represent the most probable vector of matching parameters. Mapping function parameters are thus found simultaneously with the feature correspondence. Local errors do not influence globally the registration process. The clustering technique was implemented, for example, in [75] and [30].

Barrow et al. [14] introduced the chamfer matching for image registration. Line features detected in images are matched by means of the minimalization of the generalized distance



between them. Borgefors [22] proposed an improved version, where better measure of correspondence – the sequential distance transform together with the root mean square average – was applied. The algorithm employs also the pyramidal speed-up.

Even that this overview does not intend to cover 3D registration methods, here the well-known Iterative Closest Point (ICP) algorithm, introduced by Besl and McKay [18] is mentioned, because it represents a key approach for registering 3D shapes (including free-form curves and surfaces).

**Methods using invariant descriptors.** As an alternative to the methods exploiting the spatial relations, the correspondence of features can be estimated using their description, preferably invariant to the expected image deformation (see Fig.4). The description should fulfill several conditions. The most important ones are *invariance* (the descriptions of the corresponding features from the reference and sensed image have to be the same), *uniqueness* (two different features should have different descriptions), *stability* (the description of a feature which is slightly deformed in an unknown manner should be close to the description of the original feature), and *independence* (if the feature description is a vector, its elements should be functionally independent). However, usually not all these conditions have to (or can) be satisfied simultaneously and it is necessary to find an appropriate trade-off.

Features from the sensed and reference images with the most similar invariant descriptions are paired as the corresponding ones. The choice of the type of the invariant description depends on the feature characteristics and the assumed geometric deformation of the images. While searching for the best matching feature pairs in the space of feature descriptors, the minimum distance rule with thresholding is usually applied. If a more robust algorithm is needed, the matching likelihood coefficients [51], which can better handle questionable situations, can be an appropriate solution. Guest et. al. proposed to select features according to the reliability of their possible matches [83].

The simplest feature description is the image intensity function itself, limited to the close neighborhood of the feature [1], [112]. To estimate the feature correspondence, authors computed the cross-correlation on these neighborhoods. Other types of similarity measures can be used, too. Zheng and Chellapa make use of the correlation coefficients [233]. They assumed the similarity geometric deformation. In their approach, firstly the rotation between images was compensated by the estimation of the illuminant direction and then the coarse-to-fine correlation based registration was performed. In [238] the mutual information was used for the improvement of the feature correspondence.

The following references are examples of the intuitive descriptions, which usually do not fulfill some of the mentioned criteria of invariant descriptors. Sester et. al. [174] proposed to describe forests, used as the region features, by elongation parameter, compactness, number of holes, and several characteristics of the minimum bounding rectangle. To register stars with a catalog, Murtagh [138] assigned to every point feature the description of the spatial distribution of other features lying around. Vujovic and Brzakovic in [213] represented every detected feature (elongated structure intersections) by its signature formed by the longest structure and angles between all other structures, participating in the intersection. Similarly, Zana [229] described each feature point by means of angles between relevant intersecting lines. Montesinos et. al. [136] proposed to use differential descriptors of the image function in the neighborhood of the detected CPs. Yang and Cohen [227] used border triangles generated

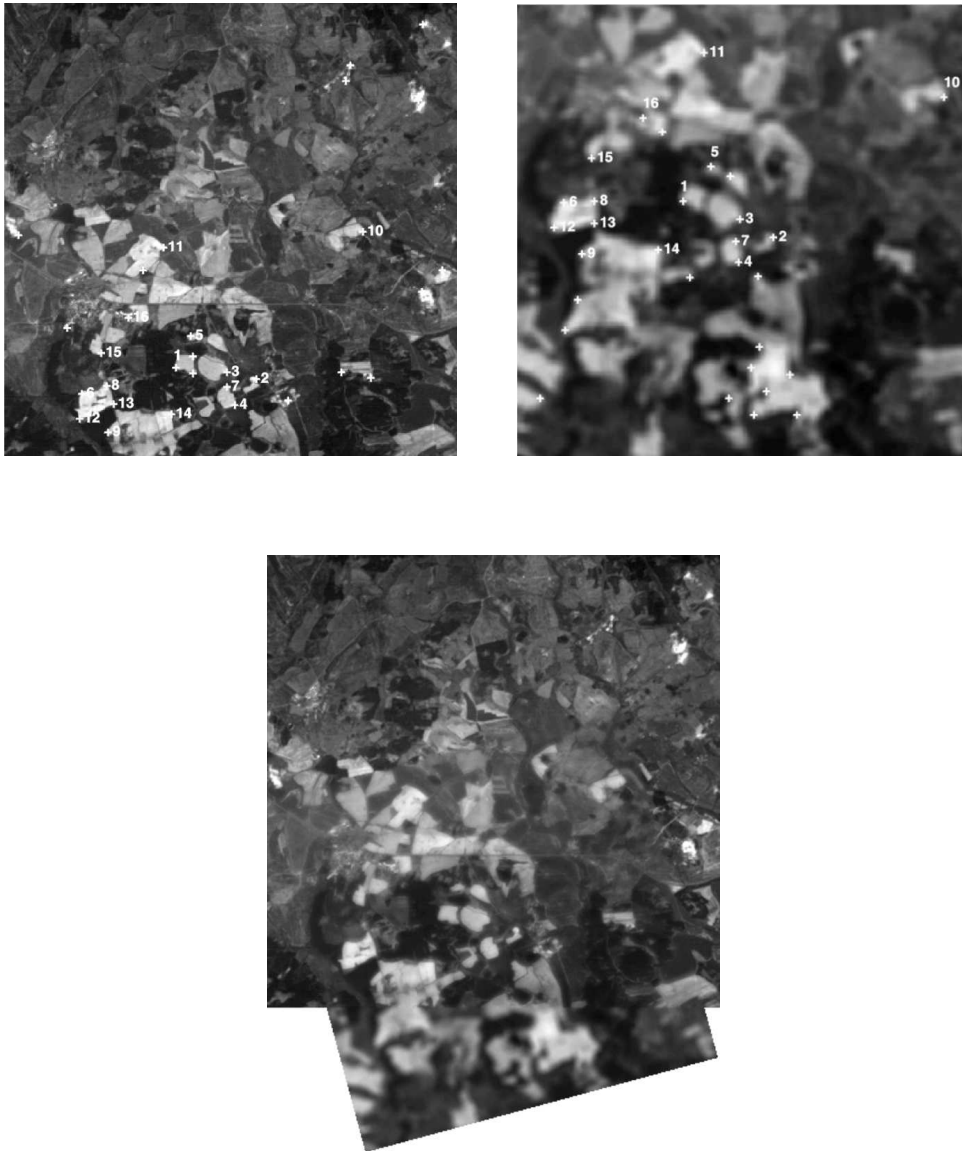


Figure 4: **Feature-based method using invariant descriptors:** in these two satellite images, control points (corners) were matched using invariants based on complex moments [57]. The numbers identify corresponding CP's. The bottom image shows the registration result.

by object convex hull and computed on them affine geometric invariants.

Many authors used closed-boundary regions as the features. In principle, any invariant and discriminative enough shape descriptor can be employed in region matching. Peli [149] proposed simple and fast description by radial shape vector but the usage of this method is limited to star-shape regions only. A generalized shape description in a form of a binary matrix was proposed in [68], [189]. In [75], the shape matrices were used for registration of rotated and scaled satellite images. In [117] a chain code representation of contours was

proposed as the invariant description and a chain code correlation-like measure was used for finding the correspondence. Skea et al. [180] represented non-collinear triplets of CPs by the sphericity. Suk [187] proposed the invariant shape description of the regions represented by polygons and further developed this approach in [188]

A large group of methods uses moment-based invariants for description of closed-boundary region features. Considering the most often assumed deformations, Hu [97] introduced moment invariants to the similarity transform.

Flusser and Suk derived the affine transform invariants [53] and used them successfully for registration of SPOT and Landsat images [54]. Holm [92] extracted closed boundary regions and proposed to represent them by their perimeter, area, compactness, moments, and moment invariants. Bhattacharya [20] suggested the application of complex moments. Brivio et al. [24] modeled shadow structures in mountain images by means of their inertia ellipses. The ellipses are here described by their area, inclination of the main axis and ellipticity. All these attributes are functions of moments. Li et al. [117] used first two Hu's moments as preselectors for matching of closed contours. The candidate matches were tested using the chain code representation of the contours. A similar method was described in [35], where the moment invariants are used together with the chain codes. Sato and Cipolla [172] computed directly, without correspondence estimation, the parameters of the present geometric deformations (an affine transform was expected) using the circular moments of distribution of the line features orientation. They combined moments and the scale-space representation of the images. Recently, Flusser and Suk [55] introduced a new class of moment invariants that are invariant to image blurring and demonstrated their performance in registering SPOT and AVHRR satellite images. Bentoutou [16] et al. registered mutually shifted and blurred digital subtraction angiography images using these invariants. Flusser et al. further developed this approach in [57] by introducing the combined blur-rotation invariants. In [52] they generalized the previous invariants to register 3-D images.

Invariant combination of the basic geometric properties of features can form geometrically oriented descriptors. Govindu et al. [77] represented the extracted contours from possibly rotated images by the slopes of tangents in the contour points. They did not look for contour correspondence, but only for the distributions of the proposed descriptors. By comparison of the corresponding distributions from the reference and sensed images the mutual image rotation can be estimated. They derived a similar type of descriptors for the affine transform, too. A detailed study of the practical aspects of the proposed method can be found in [76]. Wang and Chen [216] computed the histogram of line-length ratios and the histogram of angle differences of any two line segments in the reference and sensed images. They assumed the similarity transform. Griffin and Alexopoulos [80] used the ratio of the smallest enclosing circle radii, the difference of the locations of centroids of gravity, and the sorting of the neighbors lexicographically according to the angle with the  $x$  axis and the distance from the centroid. All these methods skip the step of finding the feature correspondence and establish directly the mapping function parameters.

Hsieh et al. [95] used the angle histogram computed on line feature points for the compensation of rotation difference. After the removal of the rotation difference, the feature point correspondence is found by means of cross-correlation. They compared their rotation compensation with the one described in [233].

Shekhar et al. [176] combined different types of features and their descriptors. They

decomposed the present geometric deformation into elementary steps and then estimated transform parameters using the feature consensus – every type of feature descriptor votes for the corresponding value of the parameter. The value of the parameter which maximizes the number of votes over all descriptor types is then chosen.

Ventura et al. [209] described image features by various descriptors (ellipticity, angle, thinness, etc.) and represented relations among them by a multi-value logical tree (MVLT). Then they compared the MVLTs of the reference and sensed images to find the feature correspondence. MVLTs are applied also in [24], together with moment invariants.

Invariant descriptors can be used as well in situations, when no precedent feature detection was done and the invariants are successively computed for the window sliding across the whole image [55]. For translated and rotated images, Goshtasby [69] proposed to calculate the moment invariants [97] from the circular-shaped windows and then to apply the cross-correlation criterion on the moment window representation. A similar idea was used earlier by Wong and Hall [224]. Along with the moment-based window description, they applied hierarchical search strategy to match radar and optical images.

**Relaxation methods.** A large group of the registration methods is based on the relaxation approach, as one of the solutions to the consistent labeling problem (CLP): to label each feature from the sensed image with the label of a feature from the reference image, so it is consistent with the labeling given to the other feature pairs [135]. The process of recalculating the pair figures of merit, considering the match quality of the feature pairs and of matching their neighbors, is iteratively repeated until a stable situation is reached. The reference work was done by Ranade and Rosenfeld [156]. Here, the displacement of the feature sets transformed by a certain geometric transformation defines the figures of merit of the feature pairs. This method can handle shifted images and it tolerates local image distortions.

Wang et al. [215] extended the classical relaxation by including the description of the corner features. They used corner sharpness, contrast, and slope. This algorithm allows to handle translation and rotation distortions in the images, but it is computationally demanding. Medioni and Nevatia [133] used line features and their descriptors (coordinates, orientation, average contrast). Cheng and Huang [33] proposed a star-based registration which considers individual feature points along with all links to their neighbors. Ton and Jain [199] speeded up the algorithm by integrating the MergeSort concept. Their method works with shifted and rotated images. Relaxation based method even for similarity transformed images was proposed, for example, by Cheng [32], Ogawa [142] and Li [118]. Different relaxation methods are compared in [155].

Another solution to the CLP problem and consequently to the image registration is backtracking, where consistent labeling is generated in recursive manner. A registration method based on backtracking is described in [135].

**Pyramids and wavelets.** We conclude the discussion about the feature matching by mentioning some works that try to reduce the computational cost due to the large image size by means of pyramidal approach.

First attempts were done back in 1977. Vanderbrug and Rosenfeld concentrated in their work on the amount of computation needed for the window pair testing. In [206], they used a subwindow first to find probable candidates of the corresponding window in the reference

image and then the full-size window was applied. They discussed the appropriate choice of the subwindow size to minimize the expected computational cost. In [169] they proposed to use first both the sensed and the reference images at a coarser resolution and then, on locations with small error measure, to match higher resolution images. Althof et al. [4] proposed to decrease the necessary computational load by taking just a sparse regular grid of windows for which the cross correlation matching is performed. These techniques are simple examples of pyramidal methods.

In general, this coarse-to-fine hierarchical strategy applies the usual registration methods, but it starts with the reference and sensed images on a coarse resolution (generated using Gaussian pyramids, simple averaging or wavelet transform coefficients, among others). Then they gradually improve the estimates of the correspondence or of the mapping function parameters while going up to the finer resolutions. At every level, they considerably decrease the search space and thus save the necessary computational time. Another important advantage resides in the fact that the registration with respect to large-scale features is achieved first and then small corrections are made for finer details. On the other hand, this strategy fails if a false match is identified on a coarser level. To overcome this, a backtracking or consistency check should be incorporated into the algorithms.

Combining cross-correlation with the pyramidal approach that exploits a summing pyramid (the pixel value at a coarser level corresponds to the summation of the pixel values on the previous level), a median pyramid, and an averaging pyramid was proposed in [219], [233] and [37] respectively. Wong and Hall [225] combined the SSDA method with pyramidal speed-up. Wang and Chen [216] extracted features (centroids of closed boundary regions) at every resolution level and found the parameters of geometric deformation from the histogram of angle differences and line-length ratios, as mentioned above. Thévenaz et al. applied a cubic spline based pyramid along with the minimization of the mean square intensity difference between the images [193] and the mutual information maximization [196], respectively. Sharma and Pavel [175] used the multiresolution Laplacian pyramid for the infrared and radar images registration. Kumar et al. [107] combined different types of pyramids (Laplacian, Gaussian) with different similarity measures (cross-correlation, sum of squared differences) to register aerial video sequences. Nonlinear min-max filters applied in a pyramidal scheme was used in [178].

Recently, wavelet decomposition of the images was recommended for the pyramidal approach due to its inherent multiresolution character. Methods can differ in the type of the applied wavelet and the set of wavelet coefficients used for finding the correspondence. Most frequently used methods decompose the image recursively into four sets of coefficients (LL, HL, LH, HH) by filtering the image successively with two filters, a low-pass filter L and a high-pass filter H, both working along the image rows and columns.

Turcajova and Kautsky [202] tested various orthogonal and biorthogonal wavelets (they used LL coefficients) together with cross-correlation on a regular grid of points to register affine transformed images. Spline biorthogonal wavelets and Haar wavelet outperformed others. Fonseca and Costa [60] detected the modulus maxima of LH and HL coefficients and looked for the maxima of the correlation coefficients, computed from LL coefficients in small surroundings of the detected maxima. Djamdji et al. [41] use just HH coefficients. Le Moigne [110] applied the Daubechies wavelet to register Landsat images and AVHRR data. They extracted LH and HL frequency coefficients and found the correspondence by means

of cross-correlation. Liu et al. [123] proposed the application of Gabor wavelet transform and Gaussian model of registration residua. You and Bhattacharya [228] use the maximum compact fuzzy sets of wavelet coefficients as features and Hausdorff distance as similarity measure. The robustness of the registration by means of the Daubechies and Haar wavelets was studied in [185].

**Summary.** Area-based methods are preferably applied when the images have not many prominent details and the distinctive information is provided by graylevels/colors rather than by local shapes and structure. Area-based methods have two principal limitations. Reference and sensed images must have somehow "similar" intensity functions, either identical (and then correlation-like methods can be used) or at least statistically dependent (this typically occurs in multimodal registration).

From the geometric point of view, only shift and small rotation between the images are allowed when using area-based methods (although the area-based methods can be generalized to full rotation and scaling, it is practically meaningless because of an extreme computational load). To speed up the searching, area-based methods often employ pyramidal image representations and sophisticated optimization algorithms to find the maximum of the similarity matrix.

Feature-based matching methods are typically applied when the local structural information is more significant than the information carried by the image intensities. They allow to register images of completely different nature (like aerial photograph and map) and can handle complex between-image distortions. The common drawback of the feature-based methods is that the respective features might be hard to detect and/or unstable in time. The crucial point of all feature-based matching methods is to have discriminative and robust feature descriptors that are invariant to all assumed differences between the images.

## 5 Transform model estimation

After the feature correspondence has been established the mapping function is constructed. It should transform the sensed image to overlay it over the reference one. The correspondence of the control points from the sensed and reference images together with the fact that the corresponding CP pairs should be as close as possible after the sensed image transformation are employed in the mapping function design.

The task to be solved consists of choosing the type of the mapping function (see Fig.5) and its parameter estimation. The type of the mapping function should correspond to the assumed geometric deformation of the sensed image, to the method of image acquisition (e.g. scanner dependent distortions and errors) and to the required accuracy of the registration (the analysis of error for rigid-body point-based registration was introduced in [48]).

In special situations when the geometric deformation is partially known, e.g. when there exists a model for the distortion caused by the acquisition device and/or the scene geometry, the pre-correction based on the inverse of the deformation can be performed (for example, in [177], [190], and [98] the authors model the Earth's shape and rotation, the satellite orbit and the scanning geometry of the sensor).

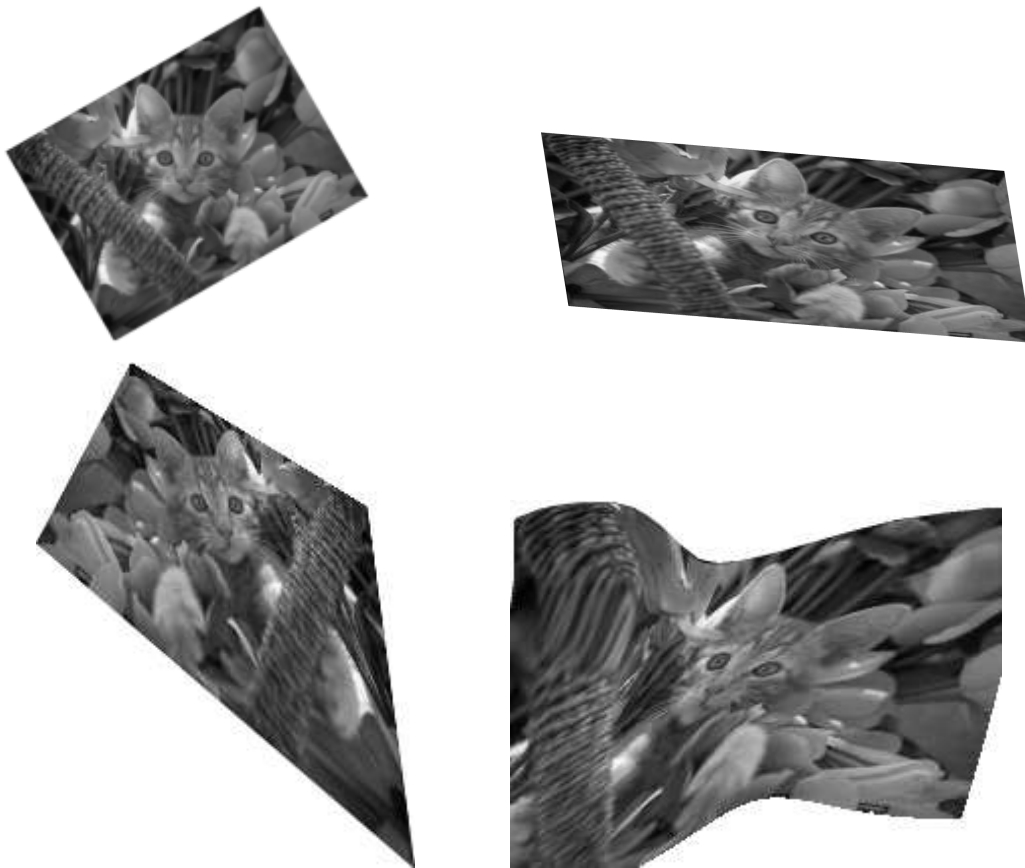


Figure 5: **Examples of various mapping functions:** – similarity transform (top left), affine transform (top right), perspective projection (bottom left), and elastic transform (bottom right).

Models of mapping functions can be divided into two broad categories according to the amount of image data they use as their support. Global models use all CPs for estimating one set of the mapping function parameters valid for the entire image. On the other hand, the local mapping functions treat the image as a composition of patches and the function parameters depend on the location of their support in the image. It leads to the tessellation of the image, usually a triangulation, and to the defining of parameters of the mapping function for each patch separately.

From another point of view, mapping functions can be categorized according to the accuracy of overlaying of the CPs used for computation of the parameters. Interpolating functions map the sensed image CPs on the reference image CPs exactly, whereas approximating functions try to find the best trade-off between the accuracy of the final mapping and other requirements imposed on the character of the mapping function. Since the CP coordinates are usually supposed not to be precise, the approximation model is more common.

**Global mapping models.** One of the most frequently used global models uses bivariate polynomials of low degrees. Similarity transform (Fig.5 - top left) is the simplest model – it consists of rotation, translation and scaling only

$$u = s(x \cos(\varphi) - y \sin(\varphi)) + t_x$$

$$v = s(x \sin(\varphi) + y \cos(\varphi)) + t_y$$

This model is often called "shape-preserving mapping" because it preserves angles and curvatures and is unambiguously determined by two control points.

Slightly more general but still linear model is an affine transform (Fig.5 - top right)

$$u = a_0 + a_1x + a_2y$$

$$v = b_0 + b_1x + b_2y,$$

which can map a parallelogram onto a square. This model is defined by three non-collinear control points, preserves straight lines and straight line parallelism. It can be used for multiview registration assuming the distance of the camera to the scene is large in comparison to the size of the scanned area, the camera is perfect (a pin-hole camera), the scene is flat, and the present geometric distortion has no local factors. If the condition on the distance of the camera from the scene is not satisfied the perspective projection model (Fig.5 - bottom left)

$$u = \frac{a_0 + a_1x + a_2y}{1 + c_1x + c_2y}$$

$$v = \frac{b_0 + b_1x + b_2y}{1 + c_1x + c_2y}$$

should be used. This model exactly describes a deformation of a flat scene photographed by a pin-hole camera the optical axis of which is not perpendicular to the scene. It can map a general quadrangle onto a square while preserving straight lines and is determined by four independent control points.

Slight violations of these assumptions may lead to the use of the second or the third-order polynomial models. Higher order polynomials usually are not used in practical applications because they may unnecessarily warp the sensed image in areas away from the control points when aligning with the reference image.

In general, the number of CPs is usually higher than the minimum number required for the determination of the mapping function. The parameters of the mapping functions are then computed by means of the least-square fit, so that the polynomials minimize the sum of squared errors at the CPs. Such mapping functions do not map the CPs onto their counterparts exactly. This approach was proved to be very effective and accurate for satellite images, for instance.



**Local mapping models.** However, a global polynomial mapping cannot properly handle images deformed locally. This happens, for instance, in medical imaging and in airborne imaging. The least square technique averages out the local geometric distortion equally over the entire image which is not desirable. Local areas of the image should be registered with the available information about the local geometric distortion in mind.

Several authors have shown the superiority of the local or at least locally sensitive registration methods above the global ones in such situations (Goshtasby [72], Ehlers and Fogel [46], Wiemker [220], and Flusser [50], among others). The weighted least square and weighted mean methods [72] gain the ability to register images locally by introducing the slight variation to the original least square method. The local methods called piecewise linear mapping [70] and piecewise cubic mapping [71], together with the Akima’s quintic approach [220], apply the combination of the CP-based image triangulation and of the collection of local mapping functions each valid within one triangle. These approaches belong to the group of the interpolating methods.

**Mapping by radial basis functions.** Radial basis functions are representatives of the group of global mapping methods but they are able to handle even locally varying geometric distortions. The resulting mapping function has a form of a linear combination of translated radially symmetric function plus a low-degree polynomial

$$u = a_0 + a_1x + a_2y + \sum_{i=1}^N c_i g(\mathbf{x}, \mathbf{x}_i)$$

and similarly for  $v$ .

Originally they were developed for the interpolation of irregular surfaces. Their name ”radial” reflects an important property of the function value at each point – it depends just on the distance of the point from the control points, not on its particular position. Multiquadrics, reciprocal multiquadrics, Gaussians, Wendland’s functions, and thin-plate splines are several examples of the radial basis functions used in image registration.

The application of the multiquadrics in the airborne remote sensing, together with the comparison to the third order polynomial method, is described in [46]. Its comparison to the Akima’s method is presented in [220]. The medical application of multiquadrics is shown in [122]. Wendland’s functions applied in medical image registration appear in [62]. These functions have very small global influence and even significant local deformations can be well registered by this approach. This property is advantageous for registering medical images, where changes occur mainly locally.

The most often used representatives of the radial basis functions are the thin-plate splines (TPS), where the radial terms have the form

$$g(\mathbf{x}, \mathbf{x}_i) = \|\mathbf{x} - \mathbf{x}_i\|^2 \ln(\|\mathbf{x} - \mathbf{x}_i\|)$$

(see Duchon [44] and Wahba [214] for the respective mathematical background). Although they had been used in mechanics and engineering for decades [87], they were introduced to image analysis community by Grimson [81] and Bookstein [21]. The TPS can be viewed as a very thin plate, which is fixed at the positions determined by the CPs in the reference image in the heights given by the  $x$  or  $y$  coordinates of the corresponding CPs in the sensed image

[73]. The TPS minimizes the quadratic variation functional of potential energy that reflects the amount of function variation and which should be small for a good mapping function. The type of registration can be chosen – exact interpolation [73], approximation [168] or generalized approximation taking into account that the anisotropic landmark errors [166] are possible. A comprehensive study focused on TPS-based registration of medical images can be found in [165].

The TPS registration gives good results but the computations can be very time consuming, namely if the number of CPs is high. Considerable attention has been paid to the methods decreasing the complexity of the TPS evaluation while preserving reasonable accuracy. Flusser [50] proposed an adaptive approximation of the TPS on square or triangular regions by simpler functions. Beatson and Newsam [15] adapted an earlier method by Greengard [78] for multipole expansion of radial functions. Powell [152] reduced the computational cost by the TPS tabulation. Barrodale et al. [13] paid attention to fast and robust calculation of the TPS coefficients.

Numerous papers have dealt with the comparison of the performance of thin-plate splines and other mapping functions. In [73], [220] and [59] they are compared with polynomials and multiquadrics when registering aerial images. The comparison of the TPS with the Gaussian radial basis functions and with multiquadrics considering their local properties was done in [8]. It was concluded that the TPS have favorable properties when used as mapping functions for image registration, while the other radial basis functions are more convenient for other applications such as image warping and morphing [8].

The TPS are not the only representatives of the spline family used for the mapping function design. A linear combination of translated cubic B-splines was used for the registration of the echo planar images [108]. Another type of spline-based function, the elastic body spline (EBS), was proposed in [39]. It evolved from the description of the equilibrium displacements of homogeneous isotropic elastic material subjected to a load. Body tissues depicted in the image data to be registered often have properties of elastic material. The EBS was used for the registration of 3-D MRI images of breasts. The authors claimed the EBS had outperformed the TPS in their experiments.

**Elastic registration.** Another approach to the registration of images with considerable complex and/or local distortions is *not* to use any parametric mapping functions, where the estimation of the geometric deformation is reduced to the search for the "best" parameters. This idea were introduced by Bajcsy et al. [10] and is often called *elastic registration* (Fig.5 - bottom right).

The images are viewed as pieces of a rubber sheet, on which external forces stretching the image and internal forces defined by stiffness or smoothness constraints are applied to bring them into alignment with the minimal amount of bending and stretching. The feature matching and mapping function design steps of the registration are done simultaneously. This is one of the advantages of elastic methods, because feature descriptors invariant to complicated deformations are not known and the feature correspondence is difficult to establish in the traditional way. The registration is achieved by locating the minimum energy state in an iterative fashion. A pyramidal approach is often applied. The external forces can be derived from the local optimization of the similarity function which is defined by the intensity values or by the correspondence of boundary structures [38], among others.

In [148], no external forces were used and the prescribed displacements, derived from the correspondence of boundary structures, were incorporated to the elastic image deformation.

Disadvantage of elastic registration is in situations when image deformations are very localized. This can be handled by means of *fluid registration*. Fluid registration methods make use of the viscous fluid model to control the image transformation. The reference image is here modelled as a thick fluid that flows out to match the sensed image under the control of the derivative of a Gaussian sensor model. This approach is mainly used in medical applications [25]. The weakness of this approach is blurring introduced during the registration process. Lester and Arridge [115] proposed to use fluid model just for finding the correspondence of control points and then process the very transformation by means of the thin plate splines. Comparison of three methods for fluid-based registration can be found in [223].

Another examples of nonrigid methods are *diffusion-based registration*, *level sets registration*, and *optical flow based registration*. The diffusion registration handles object contours and other features as membranes, setting the geometrical constraints. Three variations of this approach are described in [198]. Different solution was proposed by Andersen and Nielsen [5]. Vemuri et al. [208] introduced elastic registration method, based on evolution of level sets, moving along their respective normals. Finally, the optical flow approach was originally motivated by estimation of relative motion between images [19]. The class of optical flow registration covers very large number of methods and is beyond the scope of this survey.

## 6 Image resampling and transformation

The mapping functions constructed during the previous step are used to transform the sensed image and thus to register the images. The transformation can be realized in a forward or backward manner. Each pixel from the sensed image can be directly transformed using the estimated mapping functions. This approach, called a forward method, is complicated to implement, as it can produce holes and/or overlaps in the output image (due to the discretization and rounding). Hence, the backward approach is usually chosen. The registered image data from the sensed image are determined using the coordinates of the target pixel (the same coordinate system as of the reference image) and the inverse of the estimated mapping function. The image interpolation takes place in the sensed image on the regular grid. In this way neither holes nor overlaps can occur in the output image.

The interpolation itself is usually realized via convolution of the image with an interpolation kernel. An optimal interpolant – 2-D *sinc* function – is hard to implement in practice because of its infinite extent. Thus, many simpler interpolants of bounded support have been investigated in the literature. In order to reduce the computational cost, preferably separable interpolants have been considered. The separability enables to replace an  $m \times m$  2-D convolution by  $(m + 1)$  1-D convolutions which is much faster.

The nearest neighbor function, the bilinear and bicubic functions, quadratic splines [200], [42], cubic B-splines [93], higher-order B-splines [113], Catmull-Rom cardinal splines [104], [193], Gaussians [7], and truncated sinc functions [191] belong to the most commonly used interpolants. Meijering et al. [134] investigated higher-order polynomial kernels (quintic and

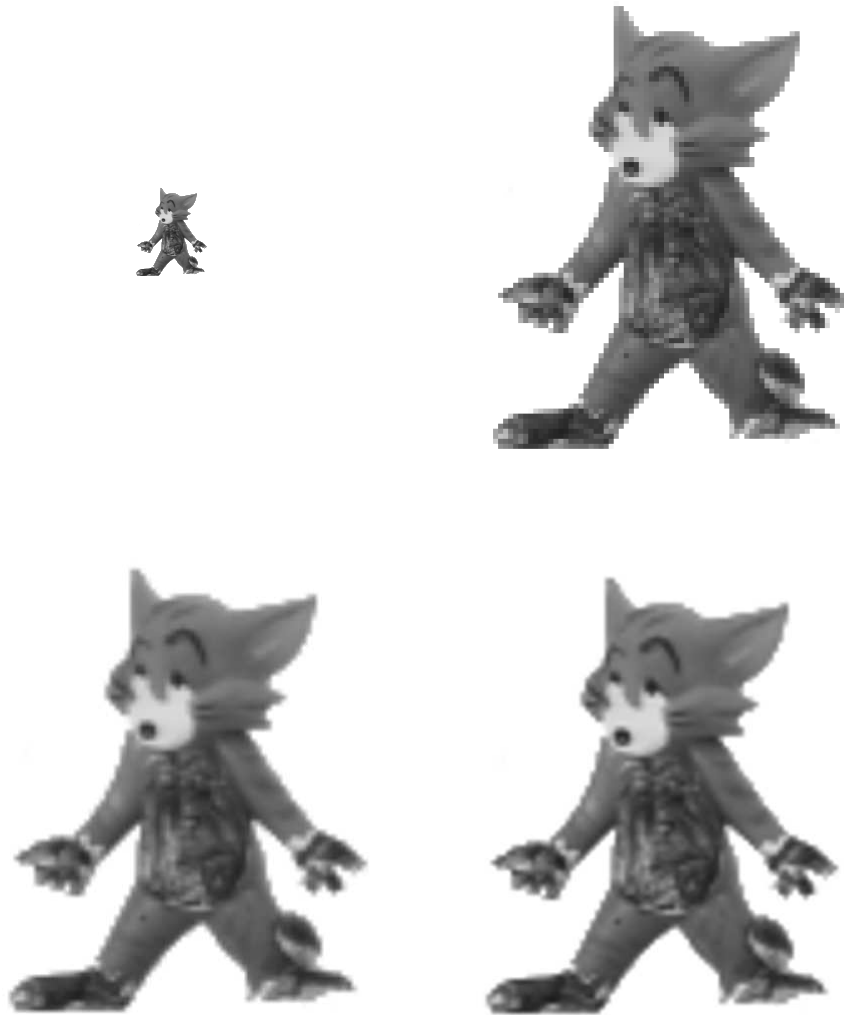


Figure 6: **Image interpolation methods:** the original image (top left) was enlarged five times using three different interpolation techniques - nearest neighbor (top right), bilinear (bottom left), and bicubic (bottom right).

septic). However, their experiments showed only marginal improvement in comparison with cubic interpolation at an highly increased computational cost.

Several survey papers on resampling techniques have been published in the last years. A detailed investigation and comparison of methods was carried out in [146] for 2-D images and in [79] for 3-D data. Thevenaz et al. [191] paid attention to the elimination of undesired interpolation artifacts. Lehman et al. [114] published a survey article covering main interpolation methods (various versions of *sinc* functions, nearest neighbor, linear, quadratic, cubic, cubic B-spline, Lagrange and Gaussian kernels) with the emphasis on medical imaging applications. They compared them using the spatial and Fourier analysis and tested the computational complexity as well as interpolation errors. Most recently, Thevenaz et

al. [192] have proposed a different approach to image resampling. Unlike the other methods, their resampling functions do not necessarily interpolate the image graylevels. They rather interpolate values calculated as certain functions of the graylevels. The authors have demonstrated this approach outperforms traditional interpolation techniques.

Even though the bilinear interpolation is outperformed by higher-order methods in terms of accuracy and visual appearance of the transformed image, it offers probably the best trade-off between accuracy and computational complexity and thus it is the most commonly used approach. Cubic interpolation is recommended when the geometric transformation involves a significant enlargement of the sensed image. Nearest neighbor interpolation should be avoided in most cases because of artifacts in the resampled image. The only exclusion is when the image to be transformed contains low number of intensities and we do not want to introduce "synthetic" graylevels/colors by higher order interpolation.

## 7 Evaluation of the image registration accuracy

Regardless of the particular images, the used registration method, and the application area, it is highly desirable to provide the user with an estimate how accurate the registration actually is. The accuracy evaluation is a non-trivial problem, partially because the errors can be dragged into the registration process in each of its stages and partially because it is hard to distinguish between registration inaccuracies and actual physical differences in the image contents. In this Section, we review basic error classes and methods for measuring the registration accuracy.

**Localization error.** Displacement of the CP coordinates due to their inaccurate detection is called *localization error*. Being an intrinsic error of the detection method, the localization error cannot be measured directly on the given image. However, the mean precision of most CP detection methods is known for various image types from computer simulation studies and ground truth comparisons. This can be used for estimation of the expected localization error in the particular case. Localization error can be reduced by selecting an "optimal" feature detection algorithm for the given data but usually there is a trade-off between the number of detected CP candidates and the mean localization error. Sometimes we prefer to have more CP with higher localization error rather than only few of them, yet detected more precisely.

**Matching error.** *Matching error* is measured by the number of false matches when establishing the correspondence between CP candidates. It is a serious mistake which usually leads to failure of the registration process and should be avoided. Fortunately, in most cases it can be ensured by robust matching algorithms. False match can be identified by consistency check, where two different matching methods are applied to the same set of the CP candidates. Only those pairs found by the both methods are considered as valid CP pairs, the other candidate points are excluded from the further processing. In case there is no other reliable matching method, false CP pairs can be identified by cross-validation. In each step, we exclude one pair from the set of CP's and calculate the mapping parameters (translation vector and rotation angle for instance). Then we check how well the excluded points are

mapped one to the other by this model. If their displacement is below a given threshold, they are accepted as a valid CP pair.

**Alignment error.** By the term *alignment error* we denote the difference between the mapping model used for the registration and the actual between-image geometric distortion. Alignment error is always present in practice because of two different reasons. The type of the chosen mapping model may not correspond to the actual distortion and/or the parameters of the model were not calculated precisely. The former case is caused by lack of a priori information about the geometric distortion while the latter originates from the insufficient number of CP's and/or their localization errors.

Alignment error can be evaluated in several ways. The simplest measure is a mean square error at the CP's (CPE). Although commonly used, it is not good alignment error measure. In fact, it only quantifies how well the CP coordinates can be fitted by the chosen mapping model. For any CP set, zero CPE can be reached just by selection of a mapping model with enough degrees of freedom (this well-known phenomenon is in numerical analysis called "overfitting"). On the other hand, large CPE can be caused by CP localization errors and does not necessarily reflect poor registration accuracy.

Very similar to the CPE is so called Test Point error (TPE). Test points are control points that were deliberately excluded from the calculation of the mapping parameters. TPE cannot be set to zero by overfitting which makes it more meaningful than CPE. However, the localization error of the test points may negatively affect this measure. This method can be used only if a sufficient number of the CP's is available. Otherwise, the exclusion of several CP's may result in inaccurate estimation of mapping parameters. In medical applications, CP's can be far from the region of interest. Thus, Fitzpatrick et al. [47], [49] proposed to detect anatomical points within the region of interest and to use them as test points (the called them "target points"). The concept of TPE can be extended such that the distance between corresponding "test" lines or surfaces is measured [140], [147].

Another approach to estimation of alignment accuracy is consistency check using multiple cues. Here, the image registered by the method under investigation is compared (by a proper metric in the image space) with the same image registered by another comparative method. As the comparative method we use preferably "gold standard method", which is a method commonly believed to be the best in the particular application area or for the given image type (gold standard method then plays a role similar to ground truth). This approach is often used in medical imaging [47], [218]. In application areas where any "gold standard" does not exist, like in remote sensing, computer vision, and industrial inspection, we take as the comparative method any method of different nature. Small difference between the registration results then indicates (although does not guarantee) good registration accuracy.

Different consistency check can be employed when a set of at least two sensed images is registered to the same reference [34], [65], [226], [91]. The sensed images can be also registered among themselves using the same set of CP's, which provides another set of mapping parameters. Using transitivity of mappings, we obtain for each sensed image two sets of mapping parameters, i.e. two registered images, which should be theoretically the same. The displacement of the test points can serve as a quality measure.

Finally, the oldest method of registration accuracy estimation – visual assessment by a domain expert – should be mentioned. It is still in use at least as a complement of the above

mentioned objective error measures.

Estimation of accuracy of registration algorithms is an substantial part of registration process. Without quantitative evaluation, no registration method can be accepted for practical utilization. A lot of work has been done on validation of rigid-body registration (the most comprehensive case-study is probably the Vanderbilt University project [218]) while validation of nonlinear, local and elastic registration methods is still at the beginning.

## 8 Current trends and outlook for the future

Image registration is one of the most important tasks when integrating and analyzing information from various sources. It is a key stage in image fusion, change detection, super-resolution imaging, and in building image information systems, among others. This paper gives a survey of the classical and up-to-date registration methods, classifying them according to their nature as well as according to the four major registration steps. Although a lot of work has been done, automatic image registration still remains an open problem. Registration of images with complex nonlinear and local distortions, multimodal registration, and registration of  $N$ -D images (where  $N > 2$ ) belong to the most challenging tasks at this moment.

When registering images with nonlinear, locally dependent geometric distortions, we are faced with two basic problems – how to match the control points and what mapping functions to use for registration. While the second one can be solved at least on theoretical level by using appropriate radial basis functions, the first problem is generally unsolvable due to its nature. Since the between-image deformations can be arbitrary, we cannot use any automatic matching method. Another conceptual question here is how can we distinguish between image deformations and real changes of the scene.

In multimodal registration, mutual information technique has become a standard reference, mainly in medical imaging. However, being an area-based technique, the mutual information has principal limitations. To overcome them, some authors combined the mutual information with other, preferably feature-based, methods to gain higher robustness and reliability. To speed up the computation, they often employed pyramidal image representation along with fast optimization algorithms. Unfortunately, when the images have significant rotation and/or scaling differences, these methods either fail or become extremely time expensive. The future development on this field could pay more attention to the feature-based methods, where appropriate invariant and modality-insensitive features can provide good platform for the registration. Besides, we trust that new application-specific methods utilizing particular sensor characteristics appear soon in remote sensing.

The major difficulty of  $N$ -D image registration resides in its computational complexity. Although the speed of computers has been growing, the need to decrease the computational time of methods persists. The complexity of methods as well as the size of data still grows (the higher resolution, higher dimensionality, larger size of scanned areas). Moreover, the demand for higher robustness and accuracy of the registration usually enforces solutions utilizing the iterations or backtracking, which also produces increase of computational complexity of the method.

# Part II: Recent advances

A very challenging area is registration and fusion of blurred images. Since imaging sensors and other devices have their physical limits and observational conditions use to be imperfect, the acquired image represents only a degraded version of the original scene. Most degradations are caused by such factors as wrong focus, motion of the scene, media turbulence, and limited spatial and spectral resolution of the sensor. Such degradations usually result in a blurring of the image and can be modelled as a convolution of the ideal image with an unknown point-spread function (PSF) of the imaging system. Blurred images require special registration techniques, which must be insensitive to image blurring.

## 1 Registration of blurred images

The methods for registration of blurred images can be, as well as the general-purpose registration methods, divided in two groups – area-based and feature-based ones. Regardless of the particular technique, all feature extraction methods, similarity measures, and matching algorithms used in the registration process should be invariant to image blurring.

Area-based methods do not search for particular landmarks in the images. They try to estimate directly the between-image translation and rotation. Myles and Lobo [139] proposed an iterative method working well if a good initial estimate of the transformation parameters is available. Zhang et al. [230], [231] proposed to estimate the registration parameters by bringing the images into canonical form. Since blur-invariant moments were used to define the normalization constraints, neither the type nor the level of the blur influences the parameter estimation. Kubota et al. [106] proposed a two-stage registration method based on hierarchical matching, where the amount of blur is considered as another parameter of the search space. Zhang and Blum [232] proposed an iterative multiscale registration based on optical flow estimation in each scale, claiming that optical flow estimation is robust to image blurring. All area-based methods require considerable (or even complete) spatial overlap of the images to yield reliable results, which is their major drawback.

Feature-based blur-invariant registration methods have appeared very recently, just after the first paper on the moment-based blur-invariant features [56]. Originally, these features could only be used for registration of mutually shifted images [55], [16]. The proposal of their rotational-invariant version [57] in combination with a robust detector of salient points [239] led to the registration methods that are able to handle blurred, shifted and rotated images [58], [52].

We present here an example how registration of blurred images can help to estimate the motion parameters and the current position of the camera in the real indoor scene. The camera, which takes one frame every second, moves parallel to the wall and rotates. The system monitors the scene and looks for alien objects which may appear in the view of the camera. When such an object appears in front of the camera, the camera is automatically focused on it. Thus, the rest of the scene becomes out-of-focused (see Fig. 7). Position



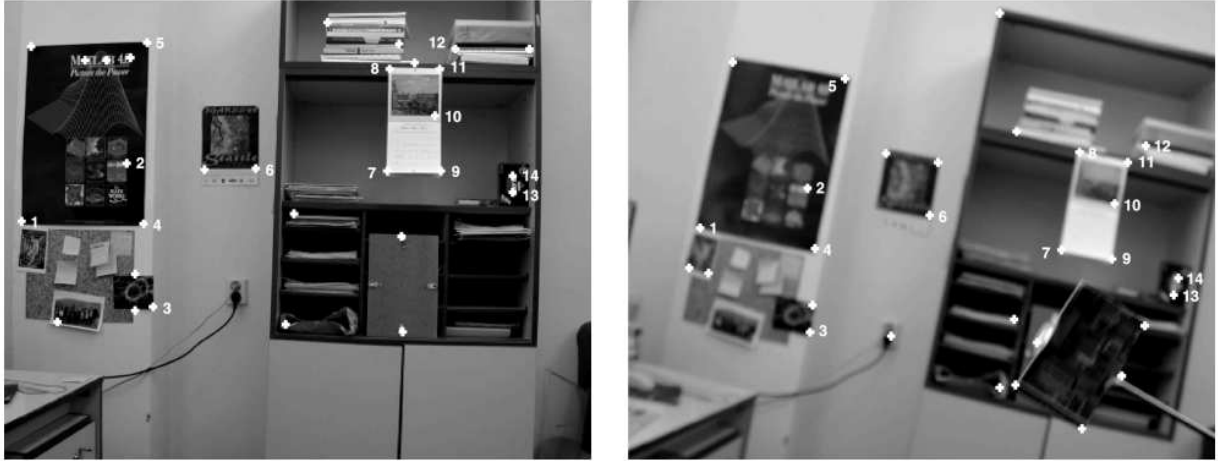


Figure 7: Images of an indoor scene. Camera at the initial position (left); camera at an unknown position, the scene is out-of-focused due to the inserted object (right). 30 CPC's are marked by crosses, those which form the corresponding CP pairs are numbered.

estimation can be resolved via registration of the initial (reference) frame and the image acquired at the moment when the object has appeared. The registration parameters (mutual rotation and shift of the reference and current frames) then unambiguously determine the actual camera position.

The motion estimation proceeds as follows. First, control point candidates (CPC's) are detected both in the initial and the current frames. Significant corners and other corner-like dominant points are considered as the candidates. To detect them, a method developed particularly for blurred images [239] was employed. In this experiment we detected 30 CPC's in each frame.

Secondly, the correspondence between the CPC's sets must be established and the candidates having no counterparts should be rejected. To do that, a vector of invariants is computed for each CPC over its circular neighborhood (here, moment invariants described in [57] were employed) and then the CPC's are matched in the space of the invariants. Once the control point correspondence has been established, their positions are refined by local search in their neighborhoods.

Finally, as soon as the control point correspondence is established and their coordinates refined, we can find an "optimal" rotational-translational mapping whose parameters are calculated via least-square fit. Knowing these parameters and the initial camera position, the current position can be easily estimated.

## 2 Fusion of blurred images

Registration of blurred images is very often just a preprocessing step for image fusion. The term *fusion* means in general an approach to extraction of information spontaneously adopted in several domains. The goal of image fusion is to integrate information into one new image containing information the quality of which cannot be achieved otherwise.

We assume that several differently blurred images of the same scene are available. In that case, fusion means obtaining one "sharp" image. This can be done by *Multichannel Blind Deconvolution* (MBD) technique. However, almost all MBD methods require a "perfect" registration of input images, which is far from being realistic in practice. Although the above-cited registration methods are very sophisticated and can be applied almost to all types of images, the result tends to be rarely perfect. The registration error usually varies from subpixel values to few pixels. Here we present a fusion method which can handle between-image misregistrations up to few pixels or, in other words, which performs accurate registration and deconvolution at the same time.

Let us define the model in a more rigorous form. Assume that the  $k$ -th acquired image  $z_k$  can be modeled by blurring the "ideal" image  $u$  and shifting the result by an unknown vector  $(a_k, b_k) = t_k$ , i.e.,

$$z_k(x + a_k, y + b_k) = (u * h_k)(x, y) + n_k(x, y), \quad (1)$$

where  $h_k$  is an unknown PSF having a character of a low-pass filter, and  $n_k$  denotes noise.

Images degraded according to this model cannot be fused by standard multichannel blind deconvolution (MBD) methods proposed in the literature; see [88, 67, 143, 158, 145, 211] for a survey and other references. If they were applied, the channel misregistrations would lead to strong artifacts in the fused image. On the other hand, the misregistrations considered in this model are too small to be fully removed by image registration techniques (in case of blurred images, registration methods usually can suppress large spatial misalignments but seldom reach subpixel accuracy).

Fusion of images degraded according to this model requires special blind deconvolution methods, which can – in addition to the deconvolution itself – identify and compensate the between-channel misregistration. A successful method based on a stochastic approach was introduced in [212] and is described below.

### Maximum *a posteriori* restoration

Eq. (1) can be expressed in a equivalent form

$$z_i(x, y) = u(x, y) * h_i(x - a_i, y - b_i) + n_i(x, y),$$

which can be further rewritten as

$$z_i(x, y) = (u * g_i)(x, y) + n_i(x, y), \quad (2)$$

where  $g_i$  is a shifted version of the original PSF  $h_i$ :  $g_i(x, y) = h_i(x - a_i, y - b_i)$ . We can therefore work only with  $g_i$ . However, in this case the estimate of the blur size has to include also the maximum shift between the channels. Since this is difficult to determine, standard MBD techniques, which require an accurate blur-size estimate, cannot be applied in their present form. One can deduce that the problem of channel misalignment is solved if the MBD method is robust to blur size overestimation.

To overcome difficulties connected with the blur size, we adopt a stochastic approach and formulate restoration as a maximum *a posteriori* (MAP) estimation. We assume that

$u, g = \{g_1, \dots, g_N\}$  and  $z = \{z_1, \dots, z_N\}$  are random vector fields with given probability density functions (pdf)  $p(u)$ ,  $p(g)$  and  $p(z)$ , respectively, and we look for such realizations of  $u$  and  $g$  which maximize the *a posteriori* probability  $p(u, g|z)$ . It is safe to assume that  $u$  and  $g$  are statistically independent and  $n_i$  is additive white (i.e. uncorrelated) Gaussian noise (AWGN). Using the Bayes rule, the relation between *a priori* densities  $p(u)$ ,  $p(g)$  and the *a posteriori* density becomes  $p(u, g|z) \propto p(z|u, g)p(u)p(g)$ . The conditional pdf  $p(z|u, g)$  follows from our model (2) and from our assumption of AWGN. The MAP estimation is then equivalent to minimizing  $L(u, g) = -\log(p(u, g|z))$  and we can write

$$L(u, g) = \sum_{i=1}^N \frac{1}{2\sigma_i^2} \|g_i * u - z_i\|^2 - \log(p(u)) - \log(p(g)), \quad (3)$$

This equation can be solved by means of genetic algorithms that are robust but also very complex. Instead, we have adopted an approach of alternating minimizations over  $u$  and  $g$ . The advantage of this scheme lies in its simplicity. If the priors  $p(u)$  and  $p(g)$  are properly selected, each term in (3) becomes convex and the derivatives w.r.t.  $u$  and  $g$  are easy to calculate. One can then use the method of conjugate gradients to find a solution in every minimization step.

We continue our discussion with a fundamental question of selecting the proper priors  $p(u)$  and  $p(g)$ .

In most of the cases, we do not have any prior knowledge of the original image and therefore we must do with general image priors that only constrain local spatial behavior of the image. Commonly used general image priors are based on Markov random fields (MRF) and they are characterized by Gibbs distributions in the form  $p(u) \propto \exp(-F(u)/\lambda)$ , where  $\lambda$  is a constant and  $F$  is called the *energy function*. One can find various forms of the energy function in the literature: first derivative operator  $\int \|\nabla u\|^2$  (Tikhonov regularization), second derivative operator  $\int |\Delta u|^2$ , Huber MRFs, etc. However, the most promising results have been achieved for variational integrals of the form

$$F(u) = \int \phi(|\nabla u|),$$

where  $\phi$  is strictly convex, nondecreasing function that grows at most linearly. Examples of  $\phi(s)$  are  $s$  (total variation),  $\sqrt{1+s^2} - 1$  (hypersurface minimal function) or  $\log(\cosh(s))$ . Nonconvex functions, such as  $\log(1+s^2)$ ,  $s^2/(1+s^2)$  or  $\arctan(s^2)$  (Mumford-Shah functional) are also often used since they provide better results for segmentation problems.

If we adopted the same general approach as in the case of image priors, eq. (3) would be too “generous” (flat) and obtained results would be very poor. Unlike the image priors, the blur priors can take the advantage of the multichannel framework and derive  $p(g)$  from the fundamental multichannel constraint stated in [88, 67]. If we consider again Gibbs distribution, the energy function for the blur prior takes the form of

$$F(g) = \sum_{1 \leq i < j \leq N} \|z_i * g_j - z_j * g_i\|^2.$$

To further constrain this distribution and improve the stability in the case of overestimated blur sizes, it suffices to assume only blurs with positive values and thus set  $p(g) = 0$  for all  $g(x) < 0$ .

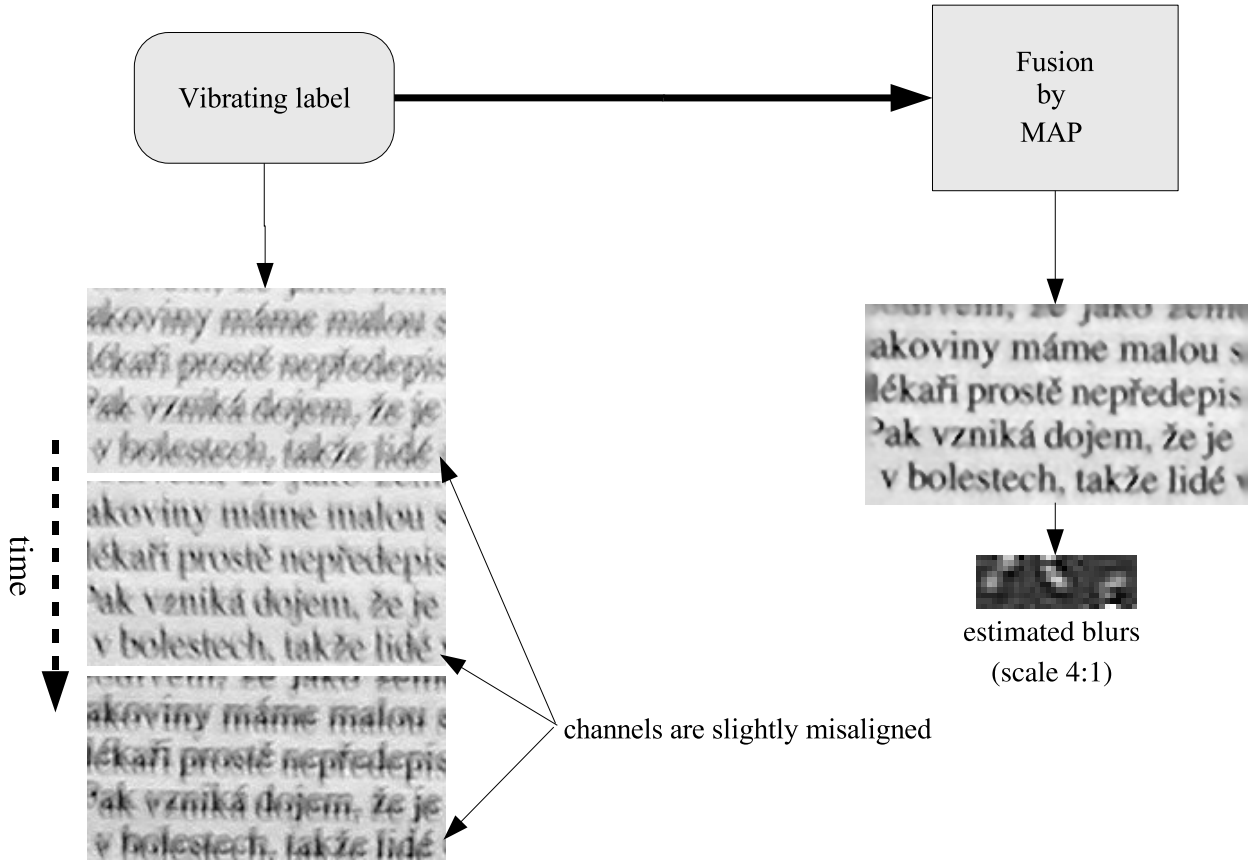


Figure 8: Slightly misaligned images: Three images of a text label attached to a vibrating machine were acquired with a standard digital camera under low-light conditions. One can observe motion blur of various kind and small spatial misalignments. The reconstructed text label is perfectly legible and the estimated shifted PSFs compensate for the misalignment. The results were obtained by the proposed MAP fusion method.

### Experiments with real data

We demonstrate the performance and broad applicability of the proposed MAP fusion method on three distinct examples: slightly misaligned images, severely misregistered images and low-resolution images. It is important to note that only real data without any preprocessing were used in all the examples. The first example in Fig. 8 reveals the capability of the proposed fusion method to perform deconvolution and simple registration in parallel. The second example in Fig. 9 illustrates that a registration step is often necessary in order to remove more complex geometric deformations. However, registration of blurred images is far from being perfect and a slight misalignment often remains. It is the MAP fusion method that is robust enough to cope with such cases.

Up to this point, we have concerned solely the deconvolution and misregistration problems. One can also apply the proposed method to low-resolution images and perform simultaneously resolution enhancement (*superresolution*). This can be done by introducing a

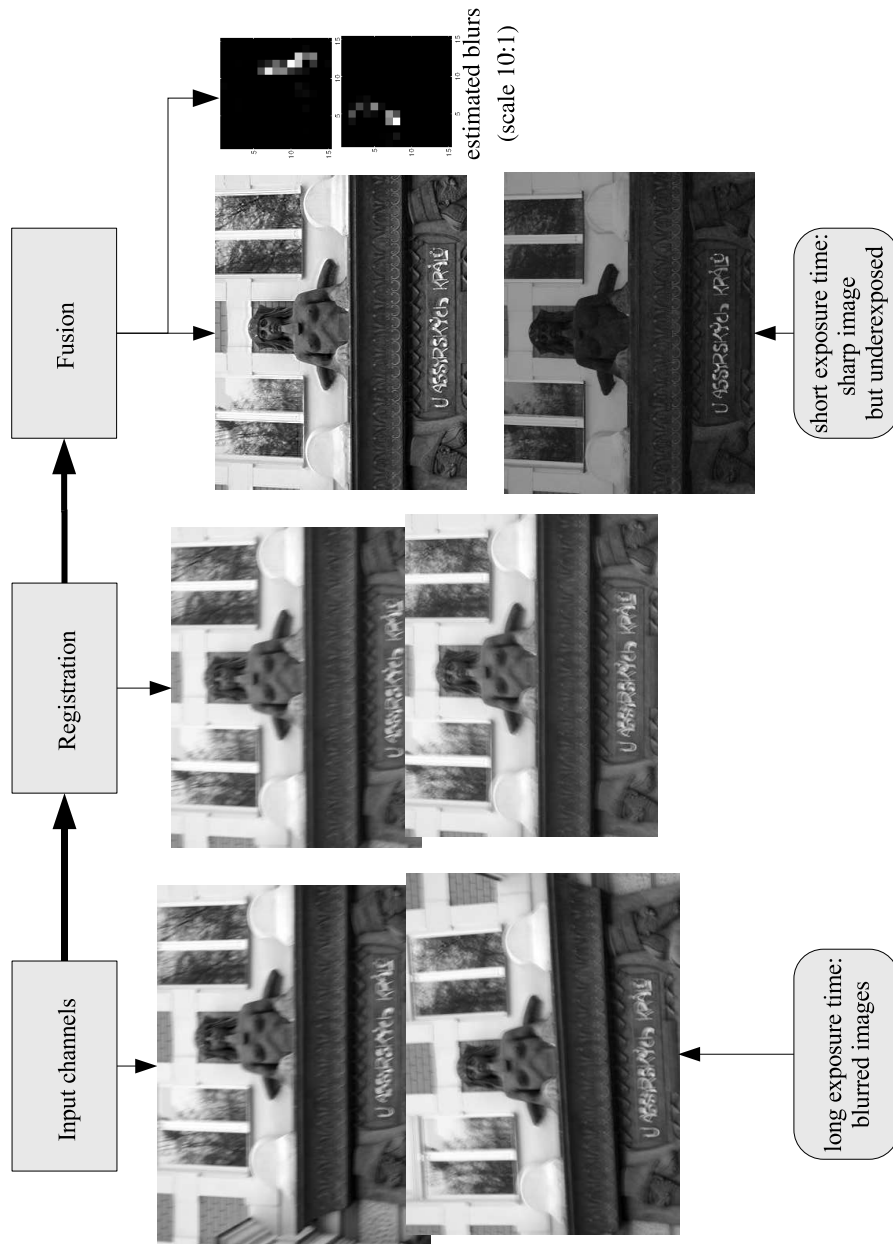


Figure 9: Severely misregistered images: Two long-exposure photos were taken from hand with a standard digital camera. Severe blur due to long exposure degrades the images. Complex misregistration due to the change of camera position is apparent and must be removed in a preprocessing registration step. The fused image is very sharp and clear and the estimated PSFs remove any misalignment left by the registration step. Compare the result with the short-exposure photo (right bottom) that is sharp but underexposed due to low-light conditions. The results were obtained by the proposed MAP fusion method.

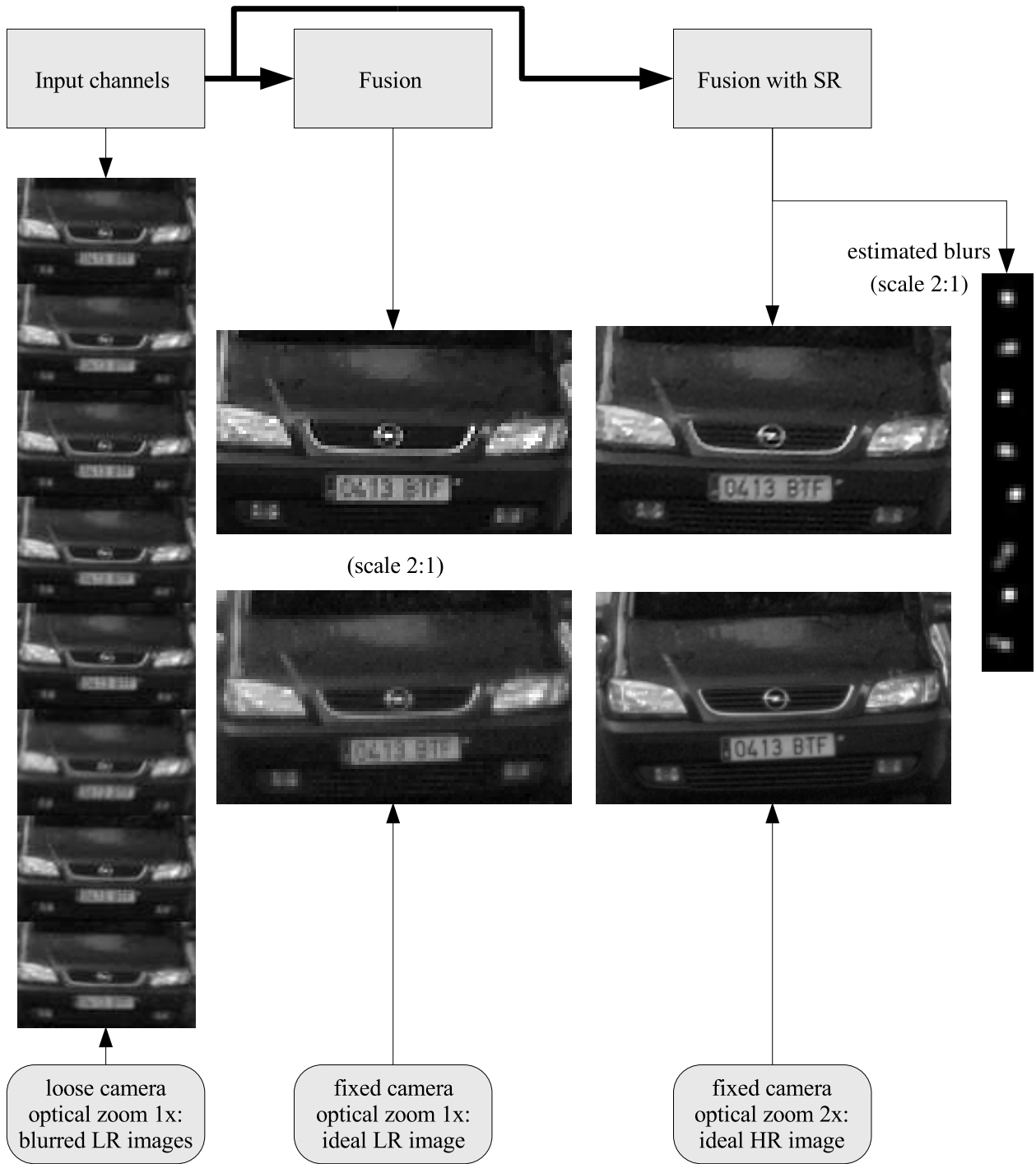


Figure 10: Low-resolution blurred images: Several photos (left column) of a car were taken from hand with a digital camera and optical zoom 1x. Photos are slightly blurred and misaligned due to the camera motion and they are of low resolution nature because of the small optical zoom. The middle column shows the result of MAP fusion (deconvolution and registration) and the “ideal” low-resolution photo acquired with the camera on a tripod and the same optical zoom 1x. The right column shows the results of enhanced MAP fusion (with resolution enhancement 2x) and the “ideal” photo acquired with the camera on a tripod and the optical zoom 2x. One can see that the reconstructed high-resolution image is comparable to the zoom shot.

*decimation* operator  $D$  into our acquisition model (2) and we obtain

$$z_i(x, y) = D(u * g_i)(x, y) + n_i(x, y),$$

Operator  $D$  models the low-resolution acquisition of digital sensors by performing convolution with a Gaussian mask (sensor blur) and downsampling. In this case, the log-function we minimize is similar to (3) except that  $D$  appears at several places. The third example depicted in Fig. 10 demonstrates the superresolution feature.

## References

- [1] S. Abdelsayed, D. Ionescu, and D. Goodenough. Matching and registration method for remote sensing images. *Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS'95*, pages 1029–1031, Florence, Italy, 1995.
- [2] H. S. Alhichri and M. Kamel. Virtual circles: A new set of features for fast image registration. *Pattern Recognition Letters*, 24:1181–1190, 2003.
- [3] W. S. Ibrahim Ali and F. S. Cohen. Registering coronal histological 2-D sections of a rat brain with coronal sections of a 3-D brain atlas using geometric curve invariants and B-spline representation. *IEEE Trans. Medical Imaging*, 17:957–966, 1998.
- [4] R. J. Althof, M. G. J. Wind, and J. T. Dobbins. A rapid and automatic image registration algorithm with subpixel accuracy. *IEEE Trans. Medical Imaging*, 16:308–316, 1997.
- [5] P. R. Andersen and M. Nielsen. Non-rigid registration by geometry-constrained diffusion. *Medical Image Analysis*, 5:81–88, 2001.
- [6] P. E. Anuta. Spatial registration of multispectral and multitemporal digital imagery using Fast Fourier Transform. *IEEE Trans. Geoscience Electronics*, 8:353–368, 1970.
- [7] C. R. Appledorn. A new approach to the interpolation of sampled data. *IEEE Trans. Medical Imaging*, 15:369–376, 1996.
- [8] N. Arad, N. Dyn, D. Reissfeld, and Y. Yeshurun. Image warping by radial basis functions: Application to facial expressions. *CVGIP: Graphical Models and Image Processing*, 56:161–172, 1994.
- [9] M. A. Audette, F. P. Ferrie, and T. M. Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical image analysis*, 4:201–217, 2000.
- [10] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.
- [11] S. Banerjee, D. P. Mukherjee, and D. D. Majumdar. Point landmarks for registration of CT and NMR images. *Pattern Recognition Letters*, 16:1033–1042, 1995.
- [12] D. I. Barnea and H. F. Silverman. A class of algorithms for fast digital image registration. *IEEE Trans. Computing*, 21:179–186, 1972.
- [13] I. Barrodale, D. Skea, M. Berkley, R. Kuwahara, and R. Poockert. Warping digital images using thin plate splines. *Pattern Recognition*, 26:375–376, 1993.
- [14] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf. Parametric correspondence and chamfer matching: Two new techniques for image matching. *Proc. 5th Int'l. Joint Conf. Art. Intell.*, pages 659–663, Cambridge, Massachusetts, 1977.
- [15] R. K. Beatson and G. N. Newsam. Fast evaluation of radial basis functions. *Computers Math. Appl.*, 24:7–19, 1992.
- [16] Y. Bentoutou, N. Taleb, M. Chikr El Mezouar, M. Taleb, and J. Jetto. An invariant approach for image registration in digital subtraction angiography. *Pattern Recognition*, 35:2853–2865, 2002.



- [17] R. Berthilsson. Affine correlation. *Proc. Int'l. Conf. on Pattern Recognition ICPR'98*, pages 1458–1461, Brisbane, Australia, 1998.
- [18] P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14:239–254, 1992.
- [19] S. S. Beuchemin and J. L. Barron. The computation of optical flow. *ACM Computing Surveys*, 27:433–467, 1995.
- [20] D. Bhattacharya and S. Sinha. Invariance of stereo images via theory of complex moments. *Pattern Recognition*, 30:1373–1386, 1997.
- [21] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11:567–585, 1989.
- [22] G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10:849–865, 1988.
- [23] R. N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, 1965.
- [24] P. A. Brivio, A. D. Ventura, A. Rampini, and R. Schettini. Automatic selection of control points from shadow structures. *Int'l. Journal of Remote Sensing*, 13:1853–1860, 1992.
- [25] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In *Proc. Visualization in Biomedical Computing (VBC'96)*, volume 1131, pages 267–276. Springer Lecture Notes in Computer Science, Hamburg, Germany, 1996.
- [26] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24:326–376, 1992.
- [27] S. C. Cain, M. M. Hayat, and E. E. Armstrong. Projection-based image registration in the presence of fixed-pattern noise. *IEEE Transactions on Image Processing*, 10:1860–1872, 2001.
- [28] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8:679–698, 1986.
- [29] E. D. Castro and C. Morandi. Registration of translated and rotated images using finite Fourier transform. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9:700–703, 1987.
- [30] S. H. Chang, F. H. Cheng, W. H. Hsu, and G. Z. Wu. Fast algorithm for point pattern matching: Invariant to translations, rotations and scale changes. *Pattern Recognition*, 30:311–320, 1997.
- [31] Q. Chen, M. Defrise, and F. Deconinck. Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 16:1156–1168, 1994.
- [32] F. H. Cheng. Point pattern matching algorithm invariant to geometrical transformation and distortion. *Pattern Recognition Letters*, 17:1429–1435, 1996.
- [33] J. K. Cheng and T. S. Huang. Image registration by matching relational structures. *Pattern Recognition*, 17:149–159, 1984.

- [34] A. V. Cideciyan. Registration of ocular fundus images. *IEEE Engineering in Medicine and Biology*, 14:52–58, 1995.
- [35] X. Dai and S. Khorram. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Trans. Geoscience and Remote Sensing*, 37:2351–2362, 1999.
- [36] X. Dai and S. Khorram. Development of a feature-based approach to automated image registration for multitemporal and multisensor remotely sensed imagery. *Int'l. Geoscience and Remote Sensing Symposium IGARSS'97*, pages 243–245, Singapore, 1997.
- [37] P. Dani and S. Chaudhuri. Automated assembling of images: Image montage preparation. *Pattern Recognition*, 28:431–445, 1995.
- [38] C. Davatzikos, J. L. Prince, and R. N. Bryan. Image registration based on boundary mapping. *IEEE Trans. Medical Imaging*, 15:112–115, 1996.
- [39] M. H. Davis, A. Khotanzad, D. P. Flaming, and S. E. Harms. A physics-based coordinate transformation for 3D image matching. *IEEE Trans. Medical Imaging*, 16:317–328, 1997.
- [40] L. Ding, A. Goshtasby, and M. Satter. Volume image registration by template matching. *Image and Vision Computing*, 19:821–832, 2001.
- [41] J. P. Djamdji, A. Bajaoui, and R. Maniere. Geometrical registration of images: The multiresolution approach. *Photogrammetric Engineering and Remote Sensing*, 53:645–653, 1993.
- [42] N. A. Dodgson. Quadratic interpolation for image resampling. *IEEE Trans. Image Processing*, 6:1322–1326, 1997.
- [43] L. Dreschler and H. Nagel. Volumetric model and 3-D trajectory of a moving car derived from monocular TV-frame sequence of a street scene. *Proc. Int'l. Joint Conf. Artificial Intelligence*, pages 692–697, Vancouver, Canada, 1981.
- [44] J. Duchon. Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces. *RAIRO Anal. Numér.*, 10:5–12, 1976.
- [45] M. Ehlers. Region-based matching for image registration in remote sensing databases. *Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS'91*, pages 2231–2234, Espoo, Finland, 1991.
- [46] M. Ehlers and D. N. Fogel. High-precision geometric correction of airborne remote sensing revisited: The multiquadric interpolation. *Proc. SPIE: Image and Signal Processing for Remote Sensing*, 2315:814–824, 1994.
- [47] J. M. Fitzpatrick. Detection failure, assessing success. *in: Medical Image Registration; Ed. by J. V. Hajnal and D. L. G. Hill, and D. J. Hawkes*, CRC Press, Baton Rouge, Florida:117–139, 2001.
- [48] J. M. Fitzpatrick and J. B. West. The distribution of target registration error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging*, 20:917–927, 2001.
- [49] J. M. Fitzpatrick, J. B. West, and C. R. Maurer Jr. Predicting error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging*, 17:694–702, 1998.

- [50] J. Flusser. An adaptive method for image registration. *Pattern Recognition*, 25:45–54, 1992.
- [51] J. Flusser. Object matching by means of matching likelihood coefficients. *Pattern Recognition Letters*, 16:893–900, 1995.
- [52] J. Flusser, J. Boldys, and B. Zitova. Moment forms invariant to rotation and blur in arbitrary number of dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:234–246, 2003.
- [53] J. Flusser and T. Suk. Pattern recognition by affine moment invariants. *Pattern Recognition*, 26:167–174, 1993.
- [54] J. Flusser and T. Suk. A moment-based approach to registration of images with affine geometric distortion. *IEEE Trans. Geoscience and Remote Sensing*, 32:382–387, 1994.
- [55] J. Flusser and T. Suk. Degraded image analysis: An invariant approach. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20:590–603, 1998.
- [56] J. Flusser, T. Suk, and S. Saic. Recognition of blurred images by the method of moments. *IEEE Transactions on Image Processing*, 5:533–538, 1996.
- [57] J. Flusser and B. Zitová. Combined invariants to linear filtering and rotation. *Int'l. Journal of Pattern Recognition and Artificial Intelligence*, 13:1123–1136, 1999.
- [58] J. Flusser, B. Zitová, and T. Suk. Invariant-based registration of rotated and blurred images. In I. S. Tammy, editor, *IEEE 1999 International Geoscience and Remote Sensing Symposium. Proceedings*, pages 1262–1264, Los Alamitos, June 1999. IEEE Computer Society.
- [59] D. N. Fogel. Image rectification with radial basis functions: Application to RS/GIS data integration. *Proc. 3rd Int'l. Conf. on Integrating GIS and Environmental Modelling*, CD-ROM, 19 pp., Santa Fe, New Mexico, 1996.
- [60] L. M. G. Fonseca and M. H. M. Costa. Automatic registration of satellite images. *Proc. Brazilian Symposium on Computer Graphic and Image Processing*, pages 219–226, Brazil, 1997.
- [61] L. M. G. Fonseca and B. S. Manjunath. Registration techniques for multisensor remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*, 62:1049–1056, 1996.
- [62] M. Fornefett, K. Rohr, and H. S. Stiehl. Radial basis functions with compact support for elastic registration of medical images. *Image and Vision Computing*, 19:87–96, 2001.
- [63] H. Foroosh, J. B. Zerubia, and M. Berthod. Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing*, 11:188–200, 2002.
- [64] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centers of circular features. *Proc. ISPRS Workshop on Fast Processing of Photogrammetric Data*, pages 281–305, Interlaken, Switzerland, 1986.
- [65] P. A. Freeborough, R. P. Woods, and N. C. Fox. Accurate registration of serial 3d mr brain images and its application to visualizing change in neurodegenerative disorders. *J. Comput. Assist. Tomogr.*, 20:1012–1022, 1996.

- [66] B. K. Ghaffary and A. A. Sawchuk. A survey of new techniques for image registration and mapping. *Proc. SPIE: Applications of Digital Image Processing*, 432:222–239, 1983.
- [67] G.B. Giannakis and R.W. Heath. Blind identification of multichannel FIR blurs and perfect image restoration. *IEEE Trans. Image Processing*, 9(11):1877–1896, November 2000.
- [68] A. Goshtasby. Description and discrimination of planar shapes using shape matrices. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 7:738–743, 1985.
- [69] A. Goshtasby. Template matching in rotated images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 7:338–344, 1985.
- [70] A. Goshtasby. Piecewise linear mapping functions for image registration. *Pattern Recognition*, 19:459–466, 1986.
- [71] A. Goshtasby. Piecewise cubic mapping functions for image registration. *Pattern Recognition*, 20:525–533, 1987.
- [72] A. Goshtasby. Image registration by local approximation methods. *Image and Vision Computing*, 6:255–261, 1988.
- [73] A. Goshtasby. Registration of images with geometric distortions. *IEEE Trans. Geoscience and Remote Sensing*, 26:60–64, 1988.
- [74] A. Goshtasby and G. C. Stockman. Point pattern matching using convex hull edges. *IEEE Trans. Systems, Man and Cybernetics*, 15:631–637, 1985.
- [75] A. Goshtasby, G. C. Stockman, and C. V. Page. A region-based approach to digital image registration with subpixel accuracy. *IEEE Trans. Geoscience and Remote Sensing*, 24:390–399, 1986.
- [76] V. Govindu and C. Shekhar. Alignment using distributions of local geometric properties. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21:1031–1043, 1999.
- [77] V. Govindu, C. Shekhar, and R. Chellapa. Using geometric properties for correspondence-less image alignment. *Proc. Int’l. Conf. on Pattern Recognition ICPR’98*, pages 37–41, Brisbane, Australia, 1998.
- [78] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73:325–348, 1987.
- [79] G. J. Grevera and J. K. Udupa. An objective comparison of 3D image interpolation methods. *IEEE Transactions on Medical Imaging*, 17:642–652, 1998.
- [80] P. M. Griffin and C. Alexopoulos. Point pattern matching using centroid bounding. *IEEE Trans. Systems, Man and Cybernetics*, 19:1274–1276, 1989.
- [81] W. E. L. Grimson. A computational theory of visual surface interpolation. *Phil. Trans. R. Soc. London*, B 298:395–427, 1982.
- [82] S. Grove and R. Tonjes. A knowledge based approach to automatic image registration. *Proc. IEEE Int’l. Conf. on Image Processing ICIP’97*, pages 228–231, Santa Barbara, California, 1997.

- [83] E. Guest, E. Berry, R. A. Baldock, M. Fidrich, and M. A. Smith. Robust point correspondence applied to two- and three-dimensional image registration. *IEEE Transaction on pattern analysis and machine intelligence*, 23:165–179, 2001.
- [84] E. Gülch. Results of test on image matching of ISPRS WG. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46:1–18, 1991.
- [85] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes. Medical image registration. *CRC Press, Baton Rouge, Florida*, ISBN 0-8493-0064-9, 2001.
- [86] H. Hanaizumi and S. Fujimura. An automated method for registration of satellite remote sensing images. *Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS'93*, pages 1348–1350, Tokyo, Japan, 1993.
- [87] R. L. Harder and R. N. Desmarais. Interpolation using surface splines. *Journal of Aircraft*, 9:189–191, 1972.
- [88] G. Harikumar and Y. Bresler. Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms. *IEEE Trans. Image Processing*, 8(2):202–219, February 1999.
- [89] P. Hellier and C. Barillot. Coupling dense and landmark-based approaches for non rigid registration. *IRISA research report*, PI 1368:30, France, 2000.
- [90] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. Medical image registration. *Physics in medicine and biology*, 46:R1–R45, 2001.
- [91] M. Holden, D. L. G. Hill, E. R. E. Denton, J. M. Jarosz, T. C. S. Cox, T. Rohlfing, J. Goodey, and D. J. Hawkes. Voxel similarity measures for 3d serial mr brain image registration. *IEEE Transactions on Medical Imaging*, 19:94–102, 2000.
- [92] M. Holm. Towards automatic rectification of satellite images using feature based matching. *Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS'91*, pages 2439–2442, Espoo, Finland, 1991.
- [93] H. S. Hou and H. C. Andrews. Cubic splines for image interpolation and digital filtering. *IEEE Trans. Acoustic, Speech and Signal Processing*, 26:508–517, 1978.
- [94] J. W. Hsieh, H. Y. M. Liao, K. C. Fan, and M. T. Ko. A fast algorithm for image registration without predetermining correspondence. *Proc. Int'l. Conf. on Pattern Recognition ICPR'96*, pages 765–769, Vienna, Austria, 1996.
- [95] J. W. Hsieh, H. Y. M. Liao, K. C. Fan, M. T. Ko, and Y. P. Hung. Image registration using a new edge-based approach. *Computer Vision and Image Understanding*, 67:112–130, 1997.
- [96] Y. C. Hsieh, D. M. McKeown, and F. P. Perlant. Performance evaluation of scene registration and stereo matching for cartographic feature extraction. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14:214–237, 1992.
- [97] M. K. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Information Theory*, 8:179–187, 1962.
- [98] R. B. Huseby, O. M. Halck, and R. Solberg. A model-based approach for geometrical correction of optical satellite images. *Proc. Int'l. Geosci. Rem. Sensing Symp. IGARSS'99*, pages 330–332, Hamburg, Germany, 1999.

- [99] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15:850–863, 1993.
- [100] Numerical Recipes in C. The art of scientific computing. <http://www.nr.com>.
- [101] M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5:143–156, 2001.
- [102] S. Kaneko, I. Murase, and S. Igarashi. Robust image registration by increment sign correlation. *Pattern Recognition*, 35:2223–2234, 2002.
- [103] S. Kaneko, Y. Satoh, and S. Igarashi. Using selective correlation coefficient for robust image registration. *Pattern Recognition*, 36:1165–1173, 2003.
- [104] R. G. Keys. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoustics, Speech and Signal Processing*, 29:1153–1160, 1981.
- [105] L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, 1:95–102, 1982.
- [106] A. Kubota, K. Kodama, and K. Aizawa. Registration and blur estimation methods for multiple differently focused images. *Proc. 1999 Int. Conf. on Image Proc.*, II:447–451, 1999.
- [107] R. Kumar, H. S. Sawhney, J. C. Asmuth, A. Pope, and S. Hsu. Registration of video to geo-referenced imagery. *Proc. Int'l. Conf. on Pattern Recognition ICPR'98*, pages 1393–1399, Brisbane, Australia, 1998.
- [108] J. Kybic, P. Thévenaz, and M. Unser. Multiresolution spline warping for EPI registration. *Proc. SPIE: Mathematical Imaging – Wavelet Applications in Signal and Image Processing*, pages 571–579, Denver, Colorado, 1999.
- [109] D. Lavine, B. Lambird, and L. Kanal. Recognition of spatial point patterns. *Pattern Recognition*, 16:289–295, 1983.
- [110] J. le Moigne. Parallel registration of multi-sensor remotely sensed imagery using wavelet coefficients. *Proc. SPIE: Wavelet Applications*, 2242:432–443, Orlando, Florida, 1994.
- [111] J. le Moigne. First evaluation of automatic image registration methods. *Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS'98*, pages 315–317, Seattle, Washington, 1998.
- [112] T. M. Lehmann. A two stage algorithm for model-based registration of medical images. *Proc. Int'l. Conf. on Pattern Recognition ICPR'98*, pages 344–352, Brisbane, Australia, 1998.
- [113] T. M. Lehmann, C. Göner, and K. Spitzer. Addendum: B-spline interpolation in medical image processing. *IEEE Transaction on medical imaging*, 20:660–665, 2001.
- [114] T. M. Lehmann, C. Görnner, and K. Spitzer. Survey: Interpolation methods in medical image processing. *IEEE Trans. Medical Imaging*, 18:1049–1075, 1999.
- [115] H. Lester and S. R. Arridge. Summarising fluid registration by thin-plate spline warps with many landmarks. In *Proceedings of Medical Image Understanding and Analysis (MIUA97)*. Oxford, 1997.

- [116] H. Lester and S. R. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, 32:129–149, 1999.
- [117] H. Li, B. S. Manjunath, and S. K. Mitra. A contour-based approach to multisensor image registration. *IEEE Trans. Image Processing*, 4:320–334, 1995.
- [118] S. Z. Li. Matching: Invariant to translations, rotations and scale changes. *Pattern Recognition*, 25:583–594, 1992.
- [119] S. Z. Li, J. Kittler, and M. Petrou. Matching and recognition of road networks from aerial images. *Proc. 2nd European Conf. on Computer Vision ECCV'92*, pages 857–861, St. Margherita, Italy, 1992.
- [120] B. Likar and F. Pernus. Automatic extraction of corresponding points for the registration of medical images. *Medical Physics*, 26:1678–1686, 1999.
- [121] B. Likar and F. Pernus. A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing*, 19:33–44, 2001.
- [122] J. A. Little, D. L. G. Hill, and D. J. Hawkes. Deformations incorporating rigid structures. *Computer Vision and Image Understanding*, 66:223–232, 1997.
- [123] J. Liu, B. C. Vemuri, and J. L. Marroquin. Local frequency representations for robust multimodal image registration. *IEEE Transactions on medical imaging*, 21:462–469, 2002.
- [124] L. Lucchese, G. Doretto, and G. M. Cortelazzo. A frequency domain technique for range data registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:1468–1484, 2002.
- [125] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Medical Imaging*, 16:187–198, 1997.
- [126] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever. Comparison of edge-based and ridge-based registration of CT and MR brain images. *Medical Image Analysis*, 1:151–161, 1996.
- [127] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever. Evaluation on ridge seeking operators for multimodality medical image matching. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:353–365, 1996.
- [128] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2:1–36, 1998.
- [129] H. Maitre and Y. Wu. Improving dynamic programming to solve image registration. *Pattern Recognition*, 20:443–462, 1987.
- [130] B. S. Manjunath, C. Shekhar, and R. Chellapa. A new approach to image feature detection with applications. *Pattern Recognition*, 29:627–640, 1996.
- [131] D. Marr and E. Hildreth. Theory of edge detection. *Proc. R. Soc. London*, B 207:187–217, 1980.

- [132] J. Matas, Š. Obdržálek, and O. Chum. Local affine frames for wide-baseline stereo. *In R. Kasturi, D. Laurendeau, and Suen C., editors :16th International Conference on Pattern Recognition ICPR 2002*, 4:363–366, 2002.
- [133] G. Medioni and R. Nevatia. Matching images using linear features. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:675–685, 1984.
- [134] E. H. W. Meijering, K. J. Zuiderveld, and M. A. Viergever. Image reconstruction by convolution with symmetrical piecewise  $n$ th-order polynomial kernels. *IEEE Trans. Image Processing*, 8:192–201, 1999.
- [135] R. S. Mitra and N. N. Murthy. Elastic maximal matching. *Pattern Recognition*, 24:747–753, 1991.
- [136] P. Montesinos, V. Gouet, R. Deriche, and D. Pelé. Matching color uncalibrated images using differential invariants. *Image and vision computing*, 18:659–671, 2000.
- [137] S. Moss and E. R. Hancock. Multiple line-template matching with EM algorithm. *Pattern Recognition Letters*, 18:1283–1292, 1997.
- [138] F. Murtagh. A feature based  $O(N^2)$  approach to point pattern matching. *Proc. Int'l. Conf. on Pattern Recognition ICPR'92*, pages 174–177, Hague, The Netherlands, 1992.
- [139] Z. Myles and N. V. Lobo. Recovering affine motion and defocus blur simultaneously. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(6):652–658, June 1998.
- [140] S. J. Nelson, M. R. Day, P. Buffone, L. L. Wald, T. F. Budinger, R. Hawkins, W. Dillon, S. Huhn, M. Prados, S. Chang, and D. B. Vigneron. Alignment of volume mri and high resolution f-18 flurodeoxyglucose pet images for evaluation of patients with brain tumors. *J Comp. Asst. Tomogr.*, 21:183–191, 1997.
- [141] A. Noble. Finding corners. *Image and Vision Computing*, 6:121–128, 1988.
- [142] H. Ogawa. Labeled point pattern matching by fuzzy relaxation. *Pattern Recognition*, 17:569–573, 1984.
- [143] Hung-Ta Pai and A.C. Bovik. On eigenstructure-based direct multichannel blind image restoration. *IEEE Trans. Image Processing*, 10(10):1434–1446, October 2001.
- [144] N. R. Pal and S. K. Pal. A review on image segmentation techniques. *Pattern Recognition*, 26:1277–1294, 1993.
- [145] G. Panci, P. Campisi, S. Colonnese, and G. Scarano. Multichannel blind image deconvolution using the bussgang algorithm: Spatial and multiresolution approaches. *IEEE Trans. Image Processing*, 12(11):1324–1337, November 2003.
- [146] J. A. Parker, R. V. Kenyon, and D. E. Troxel. Comparison of interpolating methods for image resampling. *IEEE Trans. Medical Imaging*, 2:31–39, 1983.
- [147] E. I. Parsai, K. M. Ayyangar, R. R. Dobelbower, and J. A. Siegel. Clinical fusion of three-dimensional images using bremsstrahlung spect and ct. *Journal of Nuclear Medicine*, 38:319–324, 1997.



- [148] W. Peckar, C. Schnorr, K. Rohr, and H. S. Stiehl. Two step parameter-free elastic image registration with prescribed point displacements. *Journal of Mathematical Imaging and Vision*, 10:143–162, 1999.
- [149] T. Peli. An algorithm for recognition and localization of rotated and scaled objects. *Proceedings of the IEEE*, 69:483–485, 1981.
- [150] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes. A comparison of similarity measures for use in 2D–3D medical image registration. *IEEE Trans. Medical Imaging*, 17:586–595, 1998.
- [151] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual information matching in multiresolution contexts. *Image and Vision Computing*, 19:45–52, 2001.
- [152] M. J. D. Powell. Tabulation of thin plate splines on a very fine two-dimensional grid. *Numerical Analysis Report of University of Cambridge, DAMTP/1992/NA2*, Cambridge, U.K., 1992.
- [153] W. K. Pratt. Correlation techniques of image registration. *IEEE Trans. Aerospace and Electronic Systems*, 10:353–358, 1974.
- [154] W. K. Pratt. *Digital Image Processing*. John Wiley, New York, 2nd ed., 1991.
- [155] K. E. Price. Relaxation matching techniques – a comparison. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 7:617–623, 1985.
- [156] S. Ranade and A. Rosenfeld. Point pattern matching by relaxation. *Pattern Recognition*, 12:269–275, 1980.
- [157] A. Rangarajan, H. Chui, and J. S. Duncan. Rigid point feature registration using mutual information. *Medical Image Analysis*, 4:1–17, 1999.
- [158] A. Rav-Acha and S. Peleg. Restoration of multiple images with motion blur in different directions. In *IEEE Workshop on Applications of Computer Vision (WACV)*, pages 22–27, 2000.
- [159] B. S. Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation and scale-invariant image registration. *IEEE Trans. Image Processing*, 5:1266–1271, 1996.
- [160] B. Rezaie and M. D. Srinath. Algorithms for fast image registration. *IEEE Trans. Aerospace and Electronic Systems*, 20:716–728, 1984.
- [161] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. van Saarloos. Registration of stereo and temporal images of the retina. *IEEE Trans. Medical Imaging*, 18:404–418, 1999.
- [162] A. Roche, G. Malandain, and N. Ayache. Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology*, 11:71–80, 2000.
- [163] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. *Proc. of First Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI’98), Lecture Notes in Computer Science, Cambridge, USA*, 1496:1115–1124, 1998.

- [164] K. Rohr. Localization properties of direct corner detectors. *Journal of Mathematical Imaging and Vision*, 4:139–150, 1994.
- [165] K. Rohr. *Landmark-Based Image Analysis: Using Geometric and Intensity Models*, volume 21 of *Computational Imaging and Vision Series*. Kluwer Academic Publishers, Dordrecht Boston London, 2001.
- [166] K. Rohr, H. S. Stiehl, T. M. Buzug, J. Weese, and M. H. Kuhn. Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transaction on medical imaging*, 20:526–534, 2001.
- [167] K. Rohr, H. S. Stiehl, R. Sprengel, W. Beil, T. M. Buzug, J. Wees, and M. H. Kuhn. Point based elastic registration of medical image data using approximating thin-plate splines. *Proc. Visualization in Biomedical Computing VBC'96*, pages 297–306, Hamburg, Germany, 1996.
- [168] K. Rohr, H. S. Stiehl, R. Sprengel, W. Beil, T. M. Buzug, J. Weese, and M. H. Kuhn. Nonrigid registration of medical images based on anatomical point landmarks and approximating thin-plate splines. *Proc. Aachen Workshop: Bildverarbeitung für die Medizin*, pages 41–46, Aachen, Germany, 1996.
- [169] A. Rosenfeld and G. J. Vanderbrug. Coarse-fine template matching. *IEEE Trans. Systems, Man and Cybernetics*, 7:104–107, 1977.
- [170] M. Roux. Automatic registration of SPOT images and digitized maps. *Proc. IEEE Int'l. Conf. on Image Processing ICIP'96*, pages 625–628, Lausanne, Switzerland, 1996.
- [171] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. J. Hawkes. Non-rigid registration of breast MR images using mutual information. *Proc. Medical Image Computing and Computer-Assisted Intervention MICCAI'98*, pages 1144–1152, Cambridge, Massachusetts, 1998.
- [172] J. Sato and R. Cipolla. Image registration using multi-scale texture moments. *Image and Vision Computing*, 13:341–353, 1995.
- [173] H. S. Sawhney and R. Kumar. True multi-image alignment and its applications to mosaicing and lens distortion correction. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21:235–243, 1999.
- [174] M. Sester, H. Hild, and D. Fritsch. Definition of ground control features for image registration using GIS data. *Proc. Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels*, CD-ROM, 7 pp., Columbus, Ohio, 1998.
- [175] R. K. Sharma and M. Pavel. Multisensor image registration. *Proc. Society for Information Display*, XXVIII:951–954, 1997.
- [176] C. Shekhar, V. Govindu, and R. Chellapa. Multisensor image registration by feature consensus. *Pattern Recognition*, 32:39–52, 1999.
- [177] D. Shin, J. K. Pollard, and J. P. Muller. Accurate geometric correction of ATSR images. *IEEE Trans. Geoscience and Remote Sensing*, 35:997–1006, 1997.
- [178] Y. Shinagawa and T. L. Kunii. Unconstrained automatic image matching using multiresolutional critical-point filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:994–1010, 1998.

- [179] A. Simper. Correcting general band-to-band misregistrations. *Proc. IEEE Int'l. Conf. on Image Processing ICIP'96*, 2:597–600, Lausanne, Switzerland, 1996.
- [180] D. Skea, I. Barrodale, R. Kuwahara, and R. Poeckert. A control point matching algorithm. *Pattern Recognition*, 26:269–276, 1993.
- [181] S. M. Smith. SUSAN low level image processing. <http://www.fmrib.ox.ac.uk/~steve/susan>.
- [182] S. M. Smith and J. M. Brady. SUSAN – A new approach to low level image processing. *Int'l. Journal of Computer Vision*, 23:45–78, 1997.
- [183] J. P. P. Starink and E. Baker. Finding point correspondence using simulated annealing. *Pattern Recognition*, 28:231 – 240, 1995.
- [184] G. Stockman, S. Kopstein, and S. Benett. Matching images to models for registration and object detection via clustering. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 4:229–241, 1982.
- [185] H. S. Stone, J. le Moigne, and M. McGuire. The translation sensitivity of wavelet-based registration. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21:1074–1081, 1999.
- [186] C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.
- [187] Suk T. and Flusser J. Vertex-based features for recognition of projectively deformed polygons. *Pattern Recognition*, 29:361–367, 1996.
- [188] Suk T. and Flusser J. Point-based projective invariants. *Pattern Recognition*, 33:251–261, 2000.
- [189] A. Taza and C. Y. Suen. Description of planar shapes using shape matrices. *IEEE Trans. Systems, Man, and Cybernetics*, 19:1281–1289, 1989.
- [190] O. Thepaut, K. Kpalma, and J. Ronsin. Automatic registration of ERS and SPOT multisensor images in a data fusion context. *Forest Ecology and Management*, 128:93–100, 2000.
- [191] P. Thévenaz, T. Blu, and M. Unser. Image interpolation and resampling. *Handbook of Medical Image Processing, Academic Press, in print*.
- [192] P. Thévenaz, T. Blu, and M. Unser. Interpolation revisited. *IEEE Trans. on Medical Imaging*, 19:739–758, 2000.
- [193] P. Thévenaz, U. E. Ruttimann, and M. Unser. A pyramidal approach to subpixel registration based on intensity. *IEEE Trans. Image Processing*, 7:27–41, 1998.
- [194] P. Thévenaz, U. E. Ruttimann, and M. Unser. Iterative multiscale registration without landmarks. *Proc. IEEE Int'l. Conf. on Image Processing ICIP'95*, pages 228–231, Washington D.C., 1995.
- [195] P. Thévenaz and M. Unser. An efficient mutual information optimizer for multiresolution image registration. *Proc. IEEE Int'l. Conf. on Image Processing ICIP'98*, pages 833–837, Chicago, Illinois, 1998.

- [196] P. Thévenaz and M. Unser. A pyramid approach to sub-pixel image fusion based on mutual information. *Proc. IEEE Int'l. Conf. on Image Processing ICIP'96*, pages 265–268, Lausanne, Switzerland, 1996.
- [197] P. Thévenaz and M. Unser. Spline pyramids for inter-modal image registration using mutual information. *Proc. SPIE: Wavelet Applications in Signal and Image Processing*, pages 236–247, San Diego, California, 1997.
- [198] J.-P. Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2:243–260, 1998.
- [199] J. Ton and A. K.Jain. Registering Landsat images by point matching. *IEEE Trans. Geoscience and Remote Sensing*, 27:642–651, 1989.
- [200] K. Toraichi, S. Yang, and R. Mori. Two-dimensional spline interpolation for image reconstruction. *Pattern Recognition*, 21:275–284, 1988.
- [201] M. Trajkovic and M. Hedley. Fast corner detection. *Image and Vision Computing*, 16:75–87, 1998.
- [202] R. Turcajová and J. Kautsky. A hierarchical multiresolution technique for image registration. *Proc. SPIE Mathematical Imaging: Wavelet Applications in Signal and Image Processing, Colorado*, 1996.
- [203] T. Tuytelaars and L. V. Gool. Matching widely separated views based on affinely invariant neighbourhoods. *International Journal of Computer Vision*, submitted, 2001.
- [204] P. A. van den Elsen, E.-J. D. Pol, and M. A. Viergever. Medical image matching – a review with classification. *IEEE Engineering in Medicine and Biology*, 12:26–39, 1993.
- [205] P. van Wie and M. Stein. A Landsat digital image rectification system. *IEEE Trans. Geoscience Electronics*, 15:130–136, 1977.
- [206] G. J. Vanderbrug and A. Rosenfeld. Two stage template matching. *IEEE Trans. Computers*, 26:384–393, 1977.
- [207] A. S. Vasileisky, B. Zhukov, and M. Berger. Automated image co-registration based on linear feature recognition. *Proc. 2nd Conf. Fusion of Earth Data*, pages 59–66, Sophia Antipolis, France, 1998.
- [208] B. C. Vemuri, J. Ye, Y. Chen, and C. M. Leonard. Image registration via level-set motion: Applications to atlas-based segmentation. *Medical Image Analysis*, 7:1–20, 2003.
- [209] A. D. Ventura, A. Rampini, and R. Schettini. Image registration by recognition of corresponding structures. *IEEE Trans. Geoscience and Remote Sensing*, 28:305–314, 1990.
- [210] P. Viola and W. M. Wells. Alignment by maximization of mutual information. *Int'l. Journal of Computer Vision*, 24:137–154, 1997.
- [211] F. Šroubek and J. Flusser. Multichannel blind iterative image restoration. *IEEE Trans. Image Processing*, 12(9):1094–1106, September 2003.
- [212] F. Šroubek and J. Flusser. Multichannel blind deconvolution of spatially misaligned images. *IEEE Trans. Image Processing*, 14(7):874–883, July 2005.

- [213] N. Vujovic and D. Brzakovic. Establishing the correspondence between control points in pairs of mammographic images. *IEEE Trans. Image Processing*, 6:1388–1399, 1997.
- [214] G. Wahba. *Spline Models for Observational Data*. Philadelphia; SIAM, 1990.
- [215] C. Y. Wang, H. Sun, S. Yadas, and A. Rosenfeld. Some experiments in relaxation image matching using corner features. *Pattern Recognition*, 16:167–182, 1983.
- [216] W. H. Wang and Y. C. Chen. Image registration by control points pairing using the invariant properties of line segments. *Pattern Recognition Letters*, 18:269–281, 1997.
- [217] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, Jr. C. R. Maurer, R. M. Kessler, and R. J. Maciunas. Retrospective intermodality registration techniques for images fo the head: Surface-based versus volume-based. *IEEE Transactions on Medical Imaging*, 18:144–150, 1999.
- [218] J. West et al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *J. Comput. Assist. Tomogr.*, 21:554–566, 1997.
- [219] A. P. Whichello and H. Yan. Document image mosaicing. *Proc. Int’l. Conf. on Pattern Recognition ICPR’98*, pages 1081–1084, Brisbane, Australia, 1998.
- [220] R. Wiemker, K. Rohr, L. Binder, R. Sprengel, and H. S. Stiehl. Application of elastic registration to imagery from airborne scanners. *Int’l. Archives for Photogrammetry and Remote Sensing*, XXXI–B4:949–954, 1996.
- [221] G. Wolberg and S. Zokai. Robust image registration using log-polar transform. *Proc. IEEE Intl. Conference on Image Processing*, page MA12.09, Canada, September 2000.
- [222] G. Wolberg and S. Zokai. Image registration for perspective deformation recovery. *SPIE Conf. on Automatic Target Recognition X*, page 12, Orlando, Florida, USA, April 2000.
- [223] G. Wollny and F. Kruggel. Computational cost of nonrigid registration algorithms based on fluid dynamics. *IEEE Transactions on medical imaging*, 21:946–952, 2002.
- [224] R. Y. Wong and E. L. Hall. Scene matching with invariant moments. *Computer Graphics and Image Processing*, 8:16–24, 1978.
- [225] R. Y. Wong and E. L. Hall. Sequential hierarchical scene matching. *IEEE Trans. Computers*, 27:359–366, 1978.
- [226] R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta. Automated image registration: I. general methods and intrasubject, intramodality validation. *J. Comput. Assist. Tomogr.*, 22:141–154, 1998.
- [227] Z. Yang and F. S. Cohen. Image registration and object recognition using affine invariants and convex hulls. *IEEE Transactions on Image Processing*, 8:934–946, 1999.
- [228] Y. You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *IEEE Transactions on Image Processing*, 5:416–428, 1996.
- [229] F. Zana and J. C. Klein. A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform. *IEEE Trans. Medical Imaging*, 18:419–428, 1999.

- [230] Y. Zhang, C. Wen, and Y. Zhang. Estimation of motion parameters from blurred images. *Pattern Recognition Letters*, 21:425–433, 2000.
- [231] Y. Zhang, C.Y. Wen, Y. Zhang, and Y.C. Soh. Determination of blur and affine combined invariants by normalization. *Pattern Recognition*, 35:211–221, 2002.
- [232] Z. Zhang and R.S. Blum. A hybrid image registration technique for a digital camera image fusion application. *Information Fusion*, 2:135–149, 2001.
- [233] Q. Zheng and R. Chellapa. A computational vision approach to image registration. *IEEE Trans. Image Processing*, 2:311–325, 1993.
- [234] Z. Zheng, H. Wang, and E. K. Teoh. Analysis of gray level corner detection. *Pattern Recognition Letters*, 20:149–162, 1999.
- [235] Y. M. Zhu. Volume image registration by cross-entropy optimization. *IEEE Transactions on Medical Imaging*, 21:174–180, 2002.
- [236] D. Ziou and S. Tabbone. Edge detection techniques - an overview. <http://citeseer.nj.nec.com/ziou97edge.html>, 1997.
- [237] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.
- [238] B. Zitova, J. Flusser, and F. Sroubek. Application of image processing for the conservation of the medieval mosaic. *Proceedings of the IEEE International Conference on Image Processing ICIP'02*, pages 993–996, Rochester 2002.
- [239] B. Zitová, J. Kautsky, G. Peters, and J. Flusser. Robust detection of significant points in multiframe images. *Pattern Recognition Letters*, 20:199–206, 1999.

## **Acknowledgment**

This work has been supported by the Ministry of Education of the Czech Republic under the project No. 1M6798555601 (Research Center DAR) and by the Grant Agency of the Czech Republic under the grants No. 102/04/0155 and No. 202/05/0242.