Local Invariant Feature Detectors: A Survey

Local Invariant Feature Detectors: A Survey

Tinne Tuytelaars

Department of Electrical Engineering
Katholieke Universiteit Leuven
B-3001 Leuven, Belgium
Tinne.Tuytelaars@esat.kuleuven.be

Krystian Mikolajczyk

School of Electronics and Physical Sciences
University of Surrey
Guildford, UK
K.Mikolajczyk@surrey.ac.uk

now
the essence of knowledge

Boston – Delft

Foundations and Trends® in Computer Graphics and Vision

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is T. Tuytelaars and K. Mikolajczyk, Local Invariant Feature Detectors: A Survey, Foundations and Trends® in Computer Graphics and Vision, vol 3, no 3, pp 177–280, 2007

ISBN: 978-1-60198-138-7

Foundations and Trends® in Computer Graphics and Vision
Volume 3 Issue 3, 2007

Editorial Scope

Foundations and Trends® in Computer Graphics and Vision will publish survey and tutorial articles in the following topics:

- Rendering: Lighting models; Forward rendering; Inverse rendering; Image-based rendering; Non-photorealistic rendering; Graphics hardware; Visibility computation
- Shape: Surface reconstruction; Range imaging; Geometric modelling; Parameterization;
- Mesh simplification
- Animation: Motion capture and processing; Physics-based modelling; Character animation
- Sensors and sensing
- Image restoration and enhancement
- Segmentation and grouping
- Feature detection and selection
- Color processing
- Texture analysis and synthesis
- Illumination and reflectance modeling
- Shape Representation
- Tracking
- Calibration
- Structure from motion
- Motion estimation and registration
- Stereo matching and reconstruction
- 3D reconstruction and image-based modeling
- Learning and statistical methods
- Appearance-based matching
- Object and scene recognition
- Face detection and recognition
- Activity and gesture recognition
- Image and Video Retrieval
- Video analysis and event recognition
- Medical Image Analysis
- Robot Localization and Navigation

Information for Librarians

Foundations and Trends® in Computer Graphics and Vision, 2007, Volume 3, 4 issues. ISSN paper version 1572-2740. ISSN online version 1572-2759. Also available as a combined paper and online subscription.

Foundations and Trends® in
Computer Graphics and Vision
Vol. 3, No. 3 (2007) 177–280

DOI: 10.1561/0600000017

Local Invariant Feature Detectors: A Survey

Tinne Tuytelaars[1] and Krystian Mikolajczyk[2]

[1] *Department of Electrical Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium, Tinne.Tuytelaars@esat.kuleuven.be*

[2] *School of Electronics and Physical Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK, K.Mikolajczyk@surrey.ac.uk*

Abstract

In this survey, we give an overview of invariant interest point detectors, how they evolved over time, how they work, and what their respective strengths and weaknesses are. We begin with defining the properties of the ideal local feature detector. This is followed by an overview of the literature over the past four decades organized in different categories of feature extraction methods. We then provide a more detailed analysis of a selection of methods which had a particularly significant impact on the research field. We conclude with a summary and promising future research directions.

Preface: The Local Features Paradigm

Interest points have become increasingly popular over the last few years. Today, they are the preferred strategy for solving a wide variety of problems, from wide baseline matching and the recognition of specific objects to the recognition of object classes. Additionally, similar ideas have been applied to texture recognition, scene classification, robot navigation, visual data mining, and symmetry detection, to name just a few application domains.

Yet, in spite of their recent success and gain in popularity, local features can hardly be called *novel*. In fact, they have been around since the late 1970s — even though different terminology was used at the time and the level of invariance was less than what we typically work with today. Indeed, the term “interest points” has been introduced by Moravec back in 1979 [155], and there exists a long tradition in corner, blob, and edgel detectors — all of which fall under the general category of “local features.”

Interest points were popular in the past mainly because of their efficiency, information content, and invariance. However, the recent upraise of local feature based approaches is not so much due to the locality of the features nor to their increased levels of invariance. We claim it is rather caused by a shift in paradigm on how to use such features. Previously, the stress was on accurate (even subpixel) localization and search for correspondences, and on the gain in efficiency by considering only a carefully chosen subset of pixels. These arguments still hold today. Yet on top of that, the emphasis moved toward representing the image content in a robust and flexible way, with image understanding the primordial goal.

The introduction of powerful local descriptors by Lowe [126] had a significant impact on the popularity of local features. Interest points combined with local descriptors started to be used as a black box providing reliable and repeatable measurements from images for a wide range of applications. The vision community soon realized the local descriptors computed on the interest points can capture the essence of a scene without the need for a semantic-level segmentation. Separating the different foreground objects from the background is a very hard problem indeed — a problem that probably cannot be solved in a generic way using low-level features only. Assuming you have such a segmentation available prior to the actual image interpretation thus results in a chicken-and-egg problem. However, representing the image as a set of overlapping local regions, this problem can be circumvented, as it yields an implicit segmentation: since the features are local, some of them cover part of the foreground object(s) and can be considered as relevant, while others fall on the background or on object boundaries and can be considered as irrelevant. It is the task of the subsequent higher-level processing steps to filter out the relevant information or at least to be robust to the (sometimes high) percentage of outliers. This new way of looking at local features has opened up a whole new range of applications and has brought us a step closer to cognitive-level image understanding.

This survey focuses on the feature detectors only, with the emphasis on local features well suited for image understanding applications. Local descriptors will be discussed in another survey.

Contents

1

Introduction

In this section, we discuss the very nature of local (invariant) features. What do we mean with this term? What is the advantage of using local features? What can we do with them? What would the ideal local feature look like? These are some of the questions we attempt to answer.

1.1 What are Local Features?

A local feature is an image pattern which differs from its immediate neighborhood. It is usually associated with a change of an image property or several properties simultaneously, although it is not necessarily localized exactly on this change. The image properties commonly considered are intensity, color, and texture. Figure 1.1 shows some examples of local features in a contour image (left) as well as in a grayvalue image (right). Local features can be points, but also edgels or small image patches. Typically, some measurements are taken from a region centered on a local feature and converted into descriptors. The descriptors can then be used for various applications.

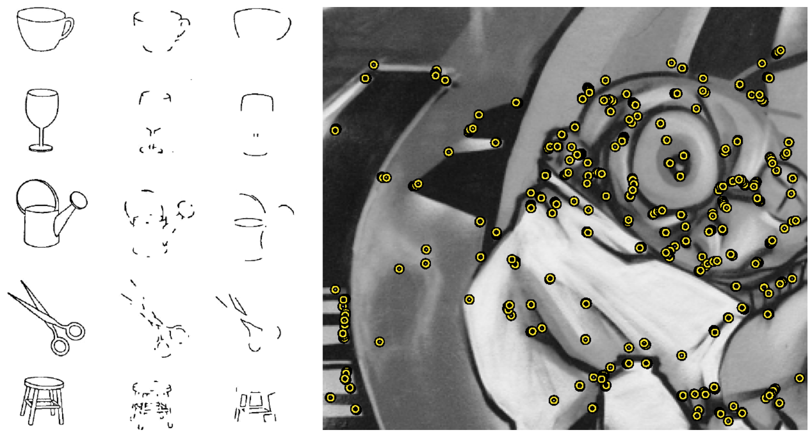

Fig. 1.1 Importance of corners and junctions in visual recognition [20] and an image example with interest points provided by a corner detector (cf. Section 3.2).

1.2 Why Local Features?

As discussed shortly in the preface, local (invariant) features are a powerful tool, that has been applied successfully in a wide range of systems and applications.

In the following, we distinguish three broad categories of feature detectors based on their possible usage. It is not exhaustive or the only way of categorizing the detectors but it emphasizes different properties required by the usage scenarios. *First*, one might be interested in a specific type of local features, as they may have a specific semantic interpretation in the limited context of a certain application. For instance, edges detected in aerial images often correspond to roads; blob detection can be used to identify impurities in some inspection task; etc. These were the first applications for which local feature detectors have been proposed. *Second*, one might be interested in local features since they provide a limited set of well localized and individually identifiable anchor points. What the features actually represent is not really relevant, as long as their location can be determined accurately and in a stable manner over time. This is for instance the situation in most matching or tracking applications, and especially for camera calibration or 3D reconstruction. Other application domains include pose estimation, image alignment or mosaicing. A typical example here are the features used in the KLT tracker [228]. *Finally*, a set of local features can be used as a robust image representation, that allows to recognize objects or scenes without the need for segmentation. Here again, it does not really matter what the features actually represent. They do not even have to be localized precisely, since the goal is not to match them on an individual basis, but rather to analyze their statistics. This way of exploiting local features was first reported in the seminal work of [213] and [210] and soon became very popular, especially in the context of object recognition (both for specific objects as well as for category-level recognition). Other application domains include scene classification, texture analysis, image retrieval, and video mining.

Clearly, each of the above three categories imposes its own constraints, and a good feature for one application may be useless in the context of a different problem. These categories can be considered when searching for suitable feature detectors for an application at hand. In this survey, we mainly focus on the second and especially the third application scenario.

Finally, it is worth noting that the importance of local features has also been demonstrated in the context of object recognition by the human visual system [20]. More precisely, experiments have shown that removing the corners from images impedes human recognition, while removing most of the straight edge information does not. This is illustrated in Figure 1.1.

1.3 A Few Notes on Terminology

Before we discuss feature detectors in more detail, let us explain some terminology commonly used in the literature.

1.3.1 Detector or Extractor?

Traditionally, the term *detector* has been used to refer to the tool that extracts the features from the image, e.g., a corner, blob or edge detector. However, this only makes sense if it is *a priori* clear what the corners, blobs or edges in the image are, so one can speak of "false detections" or "missed detections." This only holds in the first usage scenario mentioned earlier, not for the last two, where *extractor* would probably be semantically more correct. Still, the term *detector* is widely used. We therefore also stick to this terminology.

1.3.2 Invariant or Covariant?

A similar discussion holds for the use of “invariant” or “covariant.” A function is invariant under a certain family of transformations if its value does not change when a transformation from this family is applied to its argument. A function is covariant when it commutes with the transformation, i.e., applying the transformation to the argument of the function has the same effect as applying the transformation to the output of the function. A few examples may help to explain the difference. The area of a 2D surface is invariant under 2D rotations, since rotating a 2D surface does not make it any smaller or bigger. But the orientation of the major axis of inertia of the surface is covariant under the same family of transformations, since rotating a 2D surface will affect the orientation of its major axis in exactly the same way. Based on these definitions, it is clear that the so-called local scale and/or affine invariant features are in fact only covariant. The descriptors derived from them, on the other hand, are usually invariant, due to a normalization step. Since the term local invariant feature is so widely used, we nevertheless use “invariant” in this survey.

1.3.3 Rotation Invariant or Isotropic?

A function is isotropic at a particular point if it behaves the same in all directions. This is a term that applies to, e.g., textures, and should not be confused with rotational invariance.

1.3.4 Interest Point, Region or Local Feature?

In a way, the ideal local feature would be a point as defined in geometry: having a location in space but no spatial extent. In practice however, images are discrete with the smallest spatial unit being a pixel and discretization effects playing an important role. To localize features in images, a local neighborhood of pixels needs to be analyzed, giving all local features some implicit spatial extent. For some applications (e.g., camera calibration or 3D reconstruction) this spatial extent is completely ignored in further processing, and only the location derived from the feature extraction process is used (with the location sometimes determined up to sub-pixel accuracy). In those cases, one typically uses the term *interest point*.

However, in most applications those features also need to be described, such that they can be identified and matched, and this again calls for a local neighborhood of pixels. Often, this neighborhood is taken equal to the neighborhood used to localize the feature, but this need not be the case. In this context, one typically uses the term *region* instead of interest point. However, beware: when a local neighborhood of pixels is used to describe an interest point, the feature extraction process has to determine not only the location of the interest point, but also the size and possibly the shape of this local neighborhood. Especially in case of geometric deformations, this significantly complicates the process, as the size and shape have to be determined in an invariant (covariant) way.

In this survey, we prefer the use of the term *local feature*, which can be either points, regions or even edge segments.

1.4 Properties of the Ideal Local Feature

Local features typically have a spatial extent, i.e., the local neighborhood of pixels mentioned above. In contrast to classical segmentation, this can be any subset of an image. The region boundaries do not have to correspond to changes in image appearance such as color or texture. Also, multiple regions may overlap, and "uninteresting" parts of the image such as homogeneous areas can remain uncovered.

Ideally, one would like such local features to correspond to semantically meaningful object parts. In practice, however, this is unfeasible, as this would require high-level interpretation of the scene content, which is not available at this early stage. Instead, detectors select local features directly based on the underlying intensity patterns.

Good features should have the following properties:

- *Repeatability*: Given two images of the same object or scene, taken under different viewing conditions, a high percentage of the features detected on the scene part visible in both images should be found in both images.
- *Distinctiveness/informativeness*: The intensity patterns underlying the detected features should show a lot of variation, such that features can be distinguished and matched.
- *Locality*: The features should be local, so as to reduce the probability of occlusion and to allow simple model approximations of the geometric and photometric deformations between two images taken under different viewing conditions (e.g., based on a local planarity assumption).
- *Quantity*: The number of detected features should be sufficiently large, such that a reasonable number of features are detected even on small objects. However, the optimal number of features depends on the application. Ideally, the number of detected features should be adaptable over a large range by a simple and intuitive threshold. The density of features should reflect the information content of the image to provide a compact image representation.
- *Accuracy*: The detected features should be accurately localized, both in image location, as with respect to scale and possibly shape.
- *Efficiency*: Preferably, the detection of features in a new image should allow for time-critical applications.

Repeatability, arguably the most important property of all, can be achieved in two different ways: either by invariance or by robustness.

- *Invariance*: When large deformations are to be expected, the preferred approach is to model these mathematically if possible, and then develop methods for feature detection that are unaffected by these mathematical transformations.
- *Robustness*: In case of relatively small deformations, it often suffices to make feature detection methods less sensitive to

such deformations, i.e., the accuracy of the detection may decrease, but not drastically so. Typical deformations that are tackled using robustness are image noise, discretization effects, compression artifacts, blur, etc. Also geometric and photometric deviations from the mathematical model used to obtain invariance are often overcome by including more robustness.

1.4.1 Discussion

Clearly, the importance of these different properties depends on the actual application and settings, and compromises need to be made.

Repeatability is required in all application scenarios and it directly depends on the other properties like invariance, robustness, quantity etc. Depending on the application increasing or decreasing them may result in higher repeatability.

Distinctiveness and locality are competing properties and cannot be fulfilled simultaneously: the more local a feature, the less information is available in the underlying intensity pattern and the harder it becomes to match it correctly, especially in database applications where there are many candidate features to match to. On the other hand, in case of planar objects and/or purely rotating cameras (e.g., in image mosaicing applications), images are related by a global homography, and there are no problems with occlusions or depth discontinuities. Under these conditions, the size of the local features can be increased without problems, resulting in a higher distinctiveness.

Similarly, an increased level of *invariance* typically leads to a reduced *distinctiveness*, as some of the image measurements are used to lift the degrees of freedom of the transformation. A similar rule holds for *robustness versus distinctiveness*, as typically some information is disregarded (considered as noise) in order to achieve robustness. As a result, it is important to have a clear idea on the required level of invariance or robustness for a given application. It is hard to achieve high invariance and robustness at the same time and invariance, which is not adapted to the application, may have a negative impact on the results.

Accuracy is especially important in wide baseline matching, registration, and structure from motion applications, where precise correspondences are needed to, e.g., estimate the epipolar geometry or to calibrate the camera setup.

Quantity is particularly useful in some class-level object or scene recognition methods, where it is vital to densely cover the object of interest. On the other hand, a high number of features has in most cases a negative impact on the computation time and it should be kept within limits. Also robustness is essential for object class recognition, as it is impossible to model the intra-class variations mathematically, so full invariance is impossible. For these applications, an accurate localization is less important. The effect of inaccurate localization of a feature detector can be countered, up to some point, by having an extra robust descriptor, which yields a feature vector that is not affected by small localization errors.

1.5 Global versus Local Features

Local invariant features not only allow to find correspondences in spite of large changes in viewing conditions, occlusions, and image clutter (wide baseline matching), but also yield an interesting description of the image content for image retrieval and object or scene recognition tasks (both for specific objects as well as categories). To put this into context, we briefly summarize some alternative strategies to compute image representations including global features, image segments, and exhaustive and random sampling of features.

1.5.1 Global Features

In the field of image retrieval, many global features have been proposed to describe the image content, with color histograms and variations thereof as a typical example [237]. This approach works surprisingly well, at least for images with distinctive colors, as long as it is the overall composition of the image as a whole that the user is interested in, rather than the foreground object. Indeed, global features cannot distinguish foreground from background, and mix information from both parts together.

Global features have also been used for object recognition, resulting in the first appearance-based approaches to tackle this challenging problem. Turk and Pentland [245] and later Murase and Nayar [160] proposed to compute a principal component analysis of a set of model images and to use the projections onto the first few principal components as descriptors. Compared to the purely geometry-based approaches tried before, the results of the novel appearance-based approach were striking. A whole new range of natural objects could suddenly be recognized. However, being based on a global description, image clutter and occlusions again form a major problem, limiting the usefulness of the system to cases with clean backgrounds or where the object can be segmented out, e.g., relying on motion information.

1.5.2 Image Segments

An approach to overcome the limitations of the global features is to segment the image in a limited number of regions or segments, with each such region corresponding to a single object or part thereof. The best known example of this approach is the blobworld system, proposed in [31], which segments the image based on color and texture, then searches a database for images with similar "image blobs." An example based on texture segmentation is the wide baseline matching work described in [208].

However, this raises a chicken-and-egg problem as image segmentation is a very challenging task in itself, which in general requires a high-level understanding of the image content. For generic objects, color and texture cues are insufficient to obtain meaningful segmentations.

1.5.3 Sampled Features

A way to deal with the problems encountered with global features or image segmentations, is to *exhaustively sample* different subparts of the image at each location and scale. For each such image subpart, global features can then be computed. This approach is also referred to as a *sliding window* based approach. It has been especially popular in the context of face detection, but has also been applied for the recognition of specific objects or particular object classes such as pedestrians or cars.

By focusing on subparts of the image, these methods are able to find similarities between the queries and the models in spite of changing backgrounds, and even if the object covers only a small percentage of the total image area. On the downside, they still do not manage to cope with partial occlusions, and the allowed shape variability is smaller than what is feasible with a local features based approach. However, by far the biggest drawback is the inefficiency of this approach. Each and every subpart of the image must be analyzed, resulting in thousands or even millions of features per image. This requires extremely efficient methods which significantly limits the scope of possible applications.

To overcome the complexity problems more sparse *fixed grid sampling* of image patches was used (e.g., [30, 62, 246, 257]). It is however difficult to achieve invariance to geometric deformations for such features. The approach can tolerate some deformations due to dense sampling over possible locations, scales, poses etc. 00, but the individual features are not invariant. An example of such approach are multi-scale interest points. As a result, they cannot be used when the goal is to find precise correspondences between images. However, for some applications such as scene classification or texture recognition, they may well be sufficient. In [62], better results are reported with a fixed grid of patches than with patches centered on interest points, in the context of scene classification work. This can be explained by the dense coverage, as well as the fact that homogeneous areas (e.g., sky) are also taken into account in the fixed grid approach which makes the representation more complete. This dense coverage is also exploited in [66], where a fixed grid of patches was used on top of a set of local invariant features in the context of specific object recognition, where the latter supply an initial set of correspondences, which then guide the construction of correspondences for the former.

In a similar vein, rather than using a fixed grid of patches, a *random sampling* of image patches can also be used (e.g., [97, 132, 169]). This gives a larger flexibility in the number of patches, the range of scales or shapes, and their spatial distribution. Good scene recognition results are shown in [132] based on random image patches. As in the case of fixed grid sampling, this can be explained by the dense coverage which ignores the localization properties of features. Random patches are in fact a subset of the dense patches, and are used mostly to reduce the complexity. Their repeatability is poor hence they work better as an addition to the regular features rather than as a stand alone method.

Finally, to overcome the complexity problems while still providing a large number of features with better than random localization [140, 146] proposed to sample features uniformly from edges. This proved useful for dealing with wiry objects well represented by edges and curves.

1.6 Overview of this Survey

This survey article consists of two parts. First, in Section 2, we review local invariant feature detectors in the literature, from the early days in computer vision up to the most recent evolutions. Next, we describe a few selected, representative methods in more detail. We have structured the methods in a relatively intuitive manner, based on the type of feature extracted in the image. Doing so, we distinguish between corner detectors (Section 3), blob detectors (Section 4), and region detectors (Section 5). Additionally, we added a section on various detectors that have been designed in a computationally efficient manner (Section 6). With this structure, we hope the reader can easily find the type of detector most useful for his/her application. We conclude the survey with a qualitative comparison of the different methods and a discussion of future work (Section 7).

To the novice reader, who is not very familiar with local invariant feature detectors yet, we advice to skip Section 2 at first. This section has been added mainly for the more advanced reader, to give further insight in how this field evolved and what were the most important trends and to add pointers to earlier work.

References

[1] A. Almansa and T. Lindeberg, "Fingerprint enhancement by shape adaptation of scale–space operators with automatic scale selection," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2027–2042, 2000.

[2] L. Alvarez and F. Morales, "Affine morphological multiscale analysis of corners and multiple junctions," *International Journal of Computer Vision*, vol. 2, no. 25, pp. 95–107, 1997.

[3] I. M. Anderson and J. C. Bezdek, "Curvature and tangential deflection of discrete arcs: A theory based on the commutator of scatter matrix pairs and its application to vertex detection in planar shape data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 27–40, 1984.

[4] N. Ansari and E. J. Delp, "On detecting dominant points," *Pattern Recognition*, vol. 24, no. 5, pp. 441–451, 1991.

[5] H. Asada and M. Brady, "The curvature primal sketch," *Pattern Analysis and Applications*, vol. 8, no. 1, pp. 2–14, 1986.

[6] F. Attneave, "Some informational aspects of visual perception," *Psychological Review*, vol. 61, pp. 183–193, 1954.

[7] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, "Uniqueness of the gaussian kernel for scale-space filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 1, pp. 26–33, 1986.

[8] S. C. Bae, I.-S. Kweon, and C. Don Yoo, "COP: A new corner detector," *Pattern Recognition Letters*, vol. 23, no. 11, pp. 1349–1360, 2002.

[9] R. Bajcsy, "Computer identification of visual surface," *Computer and Graphics Image Processing*, vol. 2, pp. 118–130, 1973.

[10] R. Bajcsy and D. A. Rosenthal, *Visual and Conceptual Focus of Attention Structured Computer Vision.* Academic Press, 1980.

[11] J. Bala, K. DeJong, J. Huang, H. Vafaie, and H. Wechsler, "Using learning to facilitate the evolution of features for recognizing visual concepts," *Evolutionary Computation*, vol. 4, no. 3, pp. 297–311, 1996.

[12] J. Bauer, H. Bischof, A. Klaus, and K. Karner, "Robust and fully automated image registration using invariant features," *International Society for Photogrammetry and Remote Sensing*, 2004.

[13] A. Baumberg, "Reliable feature matching across widely separated views," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 774–781, 2000.

[14] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *International Journal on Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[15] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proceedings of the European Conference on Computer Vision*, pp. 404–417, 2006.

[16] P. R. Beaudet, "Rotationally invariant image operators," in *Proceedings of the International Joint Conference on Pattern Recognition*, pp. 579–583, 1978.

[17] S. Belongie, J. Malik, and J. Puzicha,, "Shape context: A new descriptor for shape matching and object recognition," in *Proceedings of the Neural Information Processing Systems*, pp. 831–837, 2000.

[18] H. L. Beus and S. S. H. Tiu, "An improved corner detection algorithm based on chain-coded plane curves," *Pattern Recognition*, vol. 20, no. 3, pp. 291–296, 1987.

[19] D. J. Beymer, "Finding junctions using the image gradient," in *International Conference Computer Vision and Pattern Recognition*, pp. 720–721, 1991.

[20] I. Biederman, "Recognition-by-components: A theory of human image understanding," *Psychological Review*, vol. 2, no. 94, pp. 115–147, 1987.

[21] Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N–D images," in *Proceedings of the International Conference on Computer Vision,* vol. 1, pp. 105–112, 2001.

[22] P. Brand and R. Mohr, "Accuracy in image measure," *SPIE Conference on Videometrics III*, vol. 2350, pp. 218–228, 1994.

[23] V. Brecher, R. Bonner, and C. Read, "A model of human preattentive visual detection of edge orientation anomalies," in *Proceedings of the SPIE Conference of Visual Information Processing: From Neurons to Chips*, vol. 1473, pp. 39–51, 1991.

[24] L. Bretzner and T. Lindeberg, "Feature tracking with automatic selection of spatial scales," *Computer Vision and Image Understanding*, vol. 71, no. 3, pp. 385–392, 1998.

[25] C. R. Brice and C. L. Fennema, "Scene analysis using regions," *Artificial Intelligence*, vol. 1, pp. 205–226, 1970.

[26] M. Brown and D. Lowe, "Recognising panoramas," in *Proceedings of the International Conference on Computer Vision*, pp. 1218–1227, 2003.

[27] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 9, no. 4, pp. 532–540, 1983.

[28] C. Cabani and W. J. MacLean, "Implementation of an affine-covariant feature detector in field-programmable gate arrays," in *Proceedings of the International Conference on Computer Vision Systems*, 2007.

[29] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[30] P. Carbonetto, N. de Freitas, and K. Barnard, "A statistical model for general contextual object recognition," in *Proceedings of the European Conference on Computer Vision, part I*, pp. 350–362, 2004.

[31] C. Carson, S. Belongie, S. Greenspan, and J. Malik, "Blobworld: Image segmentation using expectation-maximization and its application to image querying," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1026–1038, 2002.

[32] K. R. Cave and J. M. Wolfe, "Modeling the role of parallel processing in visual search," *Cognitive Psychology*, vol. 22, pp. 225–271, 1990.

[33] S. P. Chang and J. H. Horng, "Corner point detection using nest moving average," *Pattern Recognition*, vol. 27, no. 11, pp. 1533–1537, 1994.

[34] D. Chapman, "Vision, instruction and action," Technical Report AI-TR-1204, AI Laboratory, MIT, 1990.

[35] C.-H. Chen, J.-S. Lee, and Y.-N. Sun, "Wavelet transformation for gray-level corner detection," *Pattern Recognition*, vol. 28, no. 6, pp. 853–861, 1995.

[36] M. Chen and P. Yan, "A multiscaling approach based on morphological filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, no. 7, pp. 694–700, 1989.

[37] W.-C. Chen and P. Rockett, "Bayesian labelling of corners using a grey–level corner image mode," in *Proceedings of the International Conference on Image Processing*, pp. 687–690, 1997.

[38] O. Chomat, V. Colin deVerdière, D. Hall, and J. Crowley, "Local scale selection for gaussian based description techniques," in *Proceedings of the European Conference on Computer Vision, Dublin, Ireland*, pp. 117–133, 2000.

[39] J. J. Clark and N. J. Ferrier, "Modal control of attentive vision system," in *Proceedings of the International Conference on Computer Vision*, pp. 514–523, 1988.

[40] C. Coelho, A. Heller, J. L. Mundy, D. A. Forsyth, and A. Zisserman, *An Experimental Evaluation of Projective Invariants*. Cambridge, MA: MIT Press, 1992.

[41] J. Cooper, S. Venkatesh, and L. Kitchen, "The dissimilarity corner detectors," *International Conference on Advanced Robotics*, pp. 1377–1382, 1991.

[42] J. Cooper, S. Venkatesh, and L. Kitchen, "Early jump-out corner detectors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 8, pp. 823–828, 1993.

[43] T. F. Cootes and C. Taylor, "Performance evaluation of corner detection algorithms under affine and similarity transforms," in *Proceedings of the British Machine Vision Conference*, 2001.

[44] J. J. Corso and G. D. Hager, "Coherent regions for concise and stable image description," in *Proceedings of the Conference on Computer Vision and Pattern Recognition,* vol. 2, pp. 184–190, 2005.

[45] J. C. Cottier,, "Extraction et appariements robustes des points d'intérêt de deux images non etalonnées," Technical Report, LIFIA-IMAG-INRIA, Rhone-Alpes, 1994.

[46] T. Cour and J. Shi, "Recognizing objects by piecing together the segmentation puzzle," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2007.

[47] J. L. Crowley and A. C. Parker, "A representation for shape based on peaks and ridges in the difference of low pass transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 2, pp. 156–170, 1984.

[48] J. L. Crowley and A. C. Sanderson, "Multiple resolution representation and probabilistic matching of 2D gray–scale shape," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 113–121, 1987.

[49] S. M. Culhane and J. Tsotsos, "An attentional prototype for early vision," in *Proceedings of the European Conference on Computer Vision*, pp. 551–560, 1992.

[50] D. Cyganski and J. A. Or, "Application of tensor theory to object recognition and orientation determination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, pp. 662–673, 1985.

[51] E. R. Davies, "Application of the generalised hough transform to corner detection," *IEE Proceedings*, vol. 135, no. 1, pp. 49–54, 1988.

[52] R. Deriche and T. Blaszka, "Recovering and characterizing image features using an efficient model-based approach," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 530–535, 1993.

[53] R. Deriche and G. Giraudon, "Accurate corner detection: An analytical study," in *Proceedings International Conference on Computer Vision*, pp. 66–70, 1990.

[54] R. Deriche and G. Giraudon, "A computational approach for corner and vertex detection," *International Journal of Computer Vision*, vol. 10, no. 2, pp. 101–124, 1993.

[55] P. Dias, A. Kassim, and V. Srinivasan, "A neural network based corner detection method," *in IEEE International Conference on Neural Networks*, vol. 4, pp. 2116–2120, 1995.

[56] M. Donoser and H. Bischof, "Efficient maximally stable extremal region (MSER) tracking," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 553–560, 2006.

[57] L. Dreschler and H. Nagel, "Volumetric model and 3D-trajectory of a moving car derived from monocular TV-frame sequence of a street scene," *In Computer Graphics and Image Processing*, vol. 20, pp. 199–228, 1982.

[58] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. Wiley–Interscience, 1973.

[59] Y. Dufournaud, C. Schmid, and R. Horaud, "Matching images with different resolutions," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 612–618, 2000.

[60] J. Dunham, "Optimum uniform piecewise linear approximation of planar curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 1, pp. 67–75, 1986.

[61] J. Q. Fang and T. S. Huang, "A corner finding algorithm for image analysis and registration," in *Proceedings of AAAI Conference*, pp. 46–49, 1982.

[62] L. FeiFei and P. Perona, "A Bayesian hierarchical model for learning natural scene categories," in *Proceedings of the Conference on Computer Vision and Pattern Recognition,* vol. 2, pp. 524–531, 2005.

[63] F. Y. Feng and T. Pavlidis, "Finding 'vertices' in a picture," *Computer Graphics and Image Processing*, vol. 2, pp. 103–117, 1973.

[64] F. Y. Feng and T. Pavlidis, "Decomposition of polygons into simpler components: Feature generation for syntactic pattern recognition," *IEEE Transactions on Computers*, vol. C-24, 1975.

[65] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 264–271, 2003.

[66] V. Ferrari, T. Tuytelaars, and L. Van Gool, "Simultaneous object recognition and segmentation by image exploration," in *Proceedings of the European Conference on Computer Vision*, pp. 40–54, 2004.

[67] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, "Scale and the differential structure of images," *Image and Vision Computing*, vol. 10, pp. 376–388, 1992.

[68] P.-E. Forssen and D. Lowe, "Shape descriptors for maximally stable extremal regions," in *Proceedings of the International Conference on Computer Vision*, pp. 59–73, 2007.

[69] W. Förstner, "A framework for low level feature extraction," in *Proceedings of the European Conference on Computer Vision*, pp. 383–394, 1994.

[70] W. Förstner and E. Gülch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *Intercommission Conference on Fast Processing of Photogrammetric Data*, pp. 281–305, 1987.

[71] F. Fraundorfer and H. Bischof, "Evaluation of local detectors on non-planar scenes," in *Proceedings of the Austrian Association for Pattern Recognition Workshop*, pp. 125–132, 2004.

[72] H. Freeman, "A review of relevant problems in the processing of line-drawing data," in *Automatic Interpretation and Classification of Images*, (A. Graselli, ed.), pp. 155–174, Academic Press, 1969.

[73] H. Freeman, "Computer processing of line drawing images," *Surveys*, vol. 6, no. 1, pp. 57–97, 1974.

[74] H. Freeman and L. S. Davis, "A corner-finding algorithm for chain-coded curves," *IEEE Transactions on Computers*, vol. 26, pp. 297–303, 1977.

[75] M. Galun, E. Sharon, R. Basri, and A. Brandt, "Texture segmentation by multiscale aggregation of filter responses and shape elements," in *Proceedings of the International Conference on Computer Vision,* vol. 2, pp. 716–725, 2003.

[76] P. Gaussier and J. P. Cocquerez, "Neural networks for complex scene recognition: Simulation of a visual system with several cortical areas," in *Proceedings of the International Joint Conference on Neural Networks,* vol. 3, pp. 233–259, 1992.

[77] S. Ghosal and R. Mehrotra, "Zernike moment–based feature detectors," in *International Conference on Image Processing*, pp. 934–938, 1994.

[78] G. J. Giefing, H. Janssen, and H. Mallot, "Saccadic object recognition with an active vision system," in *Proceedings of the European Conference on Artificial Intelligence*, pp. 803–805, 1992.
[79] S. Gilles, *Robust Description and Matching of Image*. PhD thesis, University of Oxford, 1998.
[80] V. Gouet, P. Montesinos, R. Deriche, and D. Pelé, "Evaluation de détecteurs de points d'intérêt pour la couleur," in *12ème Congrès Francophone AFRIF–AFIA de Reconnaissance des Formes et Intelligence Artificielle*, pp. 257–266, 2000.
[81] S. Grossberg, E. Mingolla, and D. Todorovic, "A neural network architecture for preattentive vision," *IEEE Transactions on Biomedical Engineering*, vol. 36, pp. 65–84, 1989.
[82] A. Guidicci, "Corner characterization by differential geometry techniques," *Pattern Recognition Letters*, vol. 8, no. 5, pp. 311–318, 1988.
[83] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Addison-Wesley, pp. 453–507, 1993.
[84] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, pp. 147–151, 1988.
[85] T. I. Hashimoto, S. Tsujimoto, and S. Arimoto, "Spline approximation of line images by modified dynamic programming," *Transactions of IECE of Japan*, vol. J68, no. 2, pp. 169–176, 1985.
[86] X. C. He and N. H. C. Yung, "Curvature scale space corner detector with adaptive threshold and dynamic region of support," in *International Conference on Pattern Recognition*, pp. 791–794, 2004.
[87] F. Heitger, L. Rosenthaler, R. von der Heydt, E. Peterhans, and O. Kubler, "Simulation of neural contour mechanisms: From simple to end-stopped cells," *Vision Research*, vol. 32, no. 5, pp. 963–981, 1992.
[88] A. Heyden and K. Rohr, "Evaluation of corner extraction schemes using invariance method," in *Proceedings of the International Conference on Pattern Recognition*, pp. 895–899, 1996.
[89] S. Heymann, K. Maller, A. Smolic, B. Froehlich, and T. Wiegand, "SIFT implementation and optimization for general-purpose GPU," in *Proceedings of the International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, 2007.
[90] J. Hong and X. Tan, "A new approach to point pattern matching," in *Proceedings of the International Conference on Pattern Recognition*, vol. 1, pp. 82–84, 1988.
[91] R. Horaud, T. Skordas, and F. Veillon, "Finding geometric and relational structures in an image," in *Proceedings of the European Conference on Computer Vision*, pp. 374–384, 1990.
[92] D. Hubel and T. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *Journal of Physiology*, vol. 160, pp. 106–154, 1962.
[93] L. Itti and C. Koch, "Computational modeling of visual attention," *Nature Reviews Neuroscience*, vol. 2, no. 3, pp. 194–203, 2001.

[94] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 1254–1259, 1998.

[95] Q. Ji and R. M. Haralick, “Corner detection with covariance propagation,” in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 362–367, 1997.

[96] B. Julesz, “Textons, the elements of texture perception, and their interactions,” *Nature*, vol. 290, pp. 91–97, 1981.

[97] F. Jurie and B. Triggs, “Creating efficient codebooks for visual recognition,” in *Proceedings of the International Conference on Computer Vision*, pp. 604–610, 2005.

[98] T. Kadir and M. Brady, “Scale, saliency and image description,” *International Journal of Computer Vision*, vol. 45, no. 2, pp. 83–105, 2001.

[99] T. Kadir, M. Brady, and A. Zisserman, “An affine invariant method for selecting salient regions in images,” in *Proceedings of the European Conference on Computer Vision*, pp. 345–457, 2004.

[100] Y. Ke and R. Sukthankar, “PCA–SIFT: A more distinctive representation for local image descriptors,” in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 511–517, 2004.

[101] C. Kenney, B. Manjunath, M. Zuliani, G. Hewer, and A. Van Nevel, “A condition number for point matching with application to registration and post-registration error estimation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1437–1454, 2003.

[102] C. S. Kenney, M. Zuliani, and B. S. Manjunath, “An axiomatic approach to corner detection,” *International Conference on Computer Vision and Pattern Recognition*, pp. 191–197, 2005.

[103] W. Kienzle, F. A. Wichmann, B. Scholkopf, and M. O. Franz, “Learning an interest operator from human eye movements,” in *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop*, pp. 1–8, 2005.

[104] L. Kitchen and A. Rosenfeld, “Gray-level corner detection,” *Pattern Recognition Letters*, vol. 1, pp. 95–102, 1982.

[105] C. Koch and S. Ullman, “Shifts in selective visual attention: Towards the underlying neural circuitry,” *Human Neurobiology*, vol. 4, no. 4, pp. 219–227, 1985.

[106] J. J. Koenderink, “The structure of images,” *Biological Cybernetics*, vol. 50, pp. 363–396, 1984.

[107] R. Laganiere, “A morphological operator for corner detection,” *Pattern Recognition*, vol. 31, no. 11, pp. 1643–1652, 1998.

[108] D. J. Langridge, “Curve encoding and detection of discontinuities,” *Computer Graphics Image Processing*, vol. 20, pp. 58–71, 1982.

[109] I. Laptev and T. Lindeberg, “Tracking of multi-state hand models using particle filtering and a hierarchy of multi-scale image features,” in *Proceedings of Scale-Space and Morphology Workshop*, pp. 63–74, Lecture Notes in Computer Science, 2001.

[110] S. Lazebnik, C. Schmid, and J. Ponce, "Sparse texture representation using affine invariant neighborhoods," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 319–324, 2003.

[111] J. S. Lee, Y. N. Sun, C. H. Chen, and C. T. Tsai, "Wavelet based corner detection," *Pattern Recognition*, vol. 26, pp. 853–865, 1993.

[112] V. Lepetit and P. Fua, "Keypoint recognition using randomized trees," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1465–1479, 2006.

[113] L. Li and W. Chen, "Corner detection and interpretation on planar curves using fuzzy reasoning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 11, pp. 1204–1210, 1999.

[114] R.-S. Lin, C.-H. Chu, and Y.-C. Hsueh, "A modified morphological corner detector," *Pattern Recognition Letters*, vol. 19, no. 3, pp. 279–286, 1998.

[115] T. Lindeberg, "Scale-space for discrete signals," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 234–254, 1990.

[116] T. Lindeberg, "Detecting salient blob-like image structures and their scales with a scale-space primal sketch – a method for focus-of-attention," *International Journal of Computer Vision*, vol. 11, no. 3, pp. 283–318, 1993.

[117] T. Lindeberg, *Scale-Space Theory in Computer Vision.* Kluwer Academic Publishers, 1994.

[118] T. Lindeberg, "Direct estimation of affine image deformation using visual front-end operations with automatic scale selection," in *Proceedings of the International Conference on Computer Vision*, pp. 134–141, 1995.

[119] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.

[120] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.

[121] T. Lindeberg and J. Garding, "Shape-adapted smoothing in estimation of 3-D shape cues from affine deformations of local 2-D brightness structure," *Image and Vision Computing*, vol. 15, no. 6, pp. 415–434, 1997.

[122] H. Ling and D. Jacobs, "Deformation invariant image matching," in *Proceedings of the International Conference on Computer Vision,* vol. 2, pp. 1466–1473, 2005.

[123] S.-T. Liu and W.-H. Tsai, "Moment preserving corner detection," *Pattern Recognition*, vol. 23, no. 5, pp. 441–460, 1990.

[124] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 2, no. 60, pp. 91–110, 2004.

[125] D. G. Lowe, "Organization of smooth image curves at multiple scales," *International Conference on Computer Vision*, pp. 558–567, 1988.

[126] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the International Conference on Computer Vision*, pp. 1150–1157, 1999.

[127] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 959–973, 2003.

[128] B. Luo, A. D. J. Cross, and E. R. Hancock, "Corner detection via topographic analysis of vector potential," *Pattern Recognition Letters*, vol. 20, no. 6, pp. 635–650, 1998.

[129] J. Malik, S. Belongie, J. Shi, and T. Leung, "Textons, contours and regions: Cue integration in image segmentation," in *Proceedings of the International Conference on Computer Vision*, pp. 918–925, 1999.

[130] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanism," *Journal of Optical Society of America A*, vol. 7, no. 5, pp. 923–932, 1990.

[131] B. S. Manjunath, C. Shekhar, and R. Chellappa, "A new approach to image feature detection with applications," *Pattern Recognition*, vol. 29, no. 4, pp. 627–640, 1996.

[132] R. Maree, P. Geurts, J. Piater, and L. Wehenkel, "Random subwindows for robust image classification," in *Proceedings of the Conference on Computer Vision and Pattern Recognition,* vol. 1, pp. 34–40, 2005.

[133] D. Marr, *Vision.* USA, San Francisco, CA: W.H. Freeman and Company, 1982.

[134] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," in *Proceedings of the British Machine Vision Conference*, pp. 384–393, 2002.

[135] J. Matas, S. Obdrzalek, and O. Chum, "Local affine frames for wide-baseline stereo," in *Proceedings of 16th International Conference Pattern Recognition,* vol. 4, pp. 363–366, 2002.

[136] G. Medioni and Y. Yasumoto, "Corner detection and curve representation using cubic B-spline," *Computer Vision, Graphics and Image Processing*, vol. 39, no. 1, pp. 267–278, 1987.

[137] R. Mehrotra, S. Nichani, and N. Ranganathan, "Corner detection," *Pattern Recognition*, vol. 23, no. 11, pp. 1223–1233, 1990.

[138] K. Mikolajczyk, *Scale and Affine Invariant Interest Point Detectors.* PhD thesis, 2002. INRIA Grenoble.

[139] K. Mikolajczyk, B. Leibe, and B. Schiele, "Local features for object class recognition," in *Proceedings of the International Conference on Computer Vision*, pp. 525–531, 2005.

[140] K. Mikolajczyk, B. Leibe, and B. Schiele, "Multiple object class detection with a generative model," in *Proceedings the Conference on Computer Vision and Pattern Recognition*, pp. 26–36, 2006.

[141] K. Mikolajczyk and C. Schmid, "Indexing based on scale-invariant interest points," in *Proceedings of the International Conference on Computer Vision*, pp. 525–531, Vancouver, Canada, 2001.

[142] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector," in *Proceedings of the European Conference on Computer Vision*, pp. 128–142, 2002.

[143] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 1, no. 60, pp. 63–86, 2004.

[144] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[145] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1/2, pp. 43–72, 2005.

[146] K. Mikolajczyk, A. Zisserman, and C. Schmid, "Shape recognition with edge based features," in *Proceedings of the British Machine Vision Conference*, pp. 779–788, 2003.

[147] R. Milanese, *Detecting Salient Regions in an Image: From Biological Evidence to Computer Implementation.* PhD thesis, University of Geneva, 1993.

[148] R. Milanese, J.-M. Bost, and T. Pun, "A bottom-up attention system for active vision," in *Proceedings of the European Conference on Artificial Intelligence*, pp. 808–810, 1992.

[149] D. Milgram, "Computer methods for creating photomosaics," *IEEE Transactions on Computers*, vol. 23, pp. 1113–1119, 1975.

[150] F. Mohanna and F. Mokhtarian, "Performance evaluation of corner detection algorithms under affine and similarity transforms," in *Proceedings of the British Machine Vision Conference*, 2001.

[151] F. Mokhtarian and A. Mackworth, "Scale-based description of plannar curves and two-dimensional shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 1, pp. 34–43, 1986.

[152] F. Mokhtarian and R. Suomela, "Robust image corner detection through curvature scale-space," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1376–1381, 1998.

[153] P. Montesinos, V. Gouet, and R. Deriche, "Differential invariants for color images," in *Proceedings of the International Conference on Pattern Recognition*, pp. 838–840, 1998.

[154] H. Moravec, "Towards automatic visual obstacle avoidance," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 584–590, 1977.

[155] H. Moravec, "Visual mapping by a robot rover," in *Proceedings of the International Joint Conference on Artificial Intellingence*, pp. 598–600, 1979.

[156] H. Moravec, "Rover visual obstacle avoidance," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 785–790, 1981.

[157] P. Moreels and P. Perona, "Evaluation of features detectors and descriptors based on 3D objects," in *Proceedings of the International Conference on Computer Vision*, pp. 800–807, 2005.

[158] P. Moreels and P. Perona, "Evaluation of features detectors and descriptors based on 3D objects," *International Journal of Computer Vision*, vol. 73, no. 3, pp. 800–807, 2007.

[159] G. Mori, X. Ren, A. Efros, and J. Malik, "Recovering human body configurations: Combining segmentation and recognition," in *Proceedings of the Conference on Computer Vision and Pattern Recognition,* vol. 2, pp. 326–333, 2004.

[160] H. Murase and S. Nayar, "Visual learning and recognition of 3D objects from appearance," *International Journal on Computer Vision*, vol. 14, no. 1, pp. 5–24, 1995.

[161] E. Murphy-Chutorian and M. Trivedi, "N-tree disjoint-set forests for maximally stable extremal regions," in *Proceedings of the British Machine Vision Conference*, 2006.

[162] J. Mutch and D. G. Lowe, "Multiclass object recognition with sparse, localized features," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 11–18, 2006.

[163] H. H. Nagel, "Displacement vectors derived from second-order intensity variations in image sequences," *Computer Vision Graphics, and Image Processing*, vol. 21, pp. 85–117, 1983.

[164] M. Nakajima, T. Agui, and K. Sakamoto, "Pseudo-coding method for digital line figures," *Transactions of the IECE*, vol. J68–D, no. 4, pp. 623–630, 1985.

[165] U. Neisser, "Visual search," *Scientific American*, vol. 210, no. 6, pp. 94–102, 1964.

[166] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 2161–2168, 2006.

[167] J. A. Noble, "Finding corners," *Image and Vision Computing*, vol. 6, no. 2, pp. 121–128, 1988.

[168] J. A. Noble, *Descriptions of Image Surfaces*. PhD thesis, Department of Engineering Science, Oxford University, 1989.

[169] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in *Proceedings of the European Conference on Computer Vision*, pp. 490–503, 2006.

[170] H. Ogawa, "Corner detection on digital curves based on local symmetry of the shape," *Pattern Recognition*, vol. 22, no. 4, pp. 351–357, 1989.

[171] C. M. Orange and F. C. A. Groen, "Model-based corner detection," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 690–691, 1993.

[172] M. Pabst, H. J. Reitboeck, and R. Eckhorn, "A model of preattentive Region definition based on texture analysis," in *Models of Brain Function*, pp. 137–150, Cambridge, England: Cambridge University Press, 1989.

[173] K. Paler, J. Foglein, J. Illingworth, and J. Kittler, "Local ordered grey levels as an aid to corner detection," *Pattern Recognition*, vol. 17, no. 5, pp. 535–543, 1984.

[174] T. Pavlidis, *Structural Pattern Recognition*. Berlin, Heidelberg, NY: Springer-Verlag, 1977.

[175] T. Pavlidis, *Algorithms for Graphics and Image Processing*. Computer Science Press, 1982.

[176] T. Pavlidis, "Problems in recognition of drawings," *Syntactic and Structural Pattern Recognition*, vol. 45, pp. 103–113, 1988.

[177] M. Perdoch, J. Matas, and S. Obdrzalek, "Stable affine frames on isophotes," in *Proceedings of the International Conference on Computer Vision*, 2007.

[178] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, pp. 81–106, 1986.

[179] P. Rajan and J. Davidson, "Evaluation of corner detection algorithms," in *21th Southeastern Symposium on System Theory*, pp. 29–33, 1989.

[180] K. Rangarajan, M. Shah, and D. V. Brackle, "Optimal corner detection," *Computer Vision Graphics Image Processing*, vol. 48, pp. 230–245, 1989.

[181] A. Rattarangsi and R. T. Chin, "Scale-based detection of corners of planar curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 430–449, 1992.

[182] D. Reisfeld, H. Wolfson, and Y. Yeshurun, "Context free attentional operators: The generalized symmetry transform," *International Journal of Computer Vision*, vol. 14, no. 2, pp. 119–130, 1995.

[183] X. Ren, C. Fowlkes, and J. Malik, "Scale–invariant contour completion using conditional random fields," in *Proceedings of the International Conference on Computer Vision,* vol. 2, pp. 1214–1221, 2005.

[184] X. Ren and J. Malik, "Learning a classification model for segmentation," in *Proceedings of the International Conference on Computer Vision,* vol. 1, pp. 10–17, 2003.

[185] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nature Neuroscience*, vol. 2, pp. 1019–1025, 1999.

[186] B. Robbins and R. A. Owens, "2D feature detection via local energy," *Image Vision Comput*, vol. 15, no. 5, pp. 353–368, 1997.

[187] V. Roberto and R. Milanese, "Matching hierarchical structures in a machine vision system," *Intelligent Autonomous Systems*, pp. 845–852, 1989.

[188] K. Rohr, "Recognizing corners by fitting parametric models," *International Journal of Computer Vision*, vol. 9, no. 3, pp. 213–230, 1992.

[189] K. Rohr, "Localization properties of direct corner detectors," *Journal of Mathematical Imaging and Vision*, vol. 4, no. 2, pp. 139–150, 1994.

[190] K. Rohr, "On the precision in estimating the location of edges and corners," *Journal of Mathematical Imaging and Vision*, vol. 7, no. 1, pp. 7–22, 1997.

[191] A. Rosenfeld, "Picture processing by computer," *ACM Computing Surveys*, vol. 1, no. 3, pp. 147–176, 1969.

[192] A. Rosenfeld, "Digital image processing and recognition," *Digital Image Processing*, pp. 1–11, 1977.

[193] A. Rosenfeld and E. Johnston, "Angle detection on digital curves," *IEEE Transactions on Computers*, vol. C-22, pp. 875–878, 1973.

[194] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. Academic Press, Second Edition, 1982.

[195] A. Rosenfeld and M. Thurston, "Edge and curve detection for digital scene analysis," *IEEE Transactions on Computers*, vol. C-20, pp. 562–569, 1971.

[196] A. Rosenfeld, M. Thurston, and Y. H. Lee, "Edge and curve detection: Further experiments," *IEEE Transactions on Computers*, vol. C-21, pp. 677–715, 1972.

[197] A. Rosenfeld and J. S. Weszka, "An improved method of angle detection on digital curves," *IEEE Transactions on Computers*, vol. 24, no. 9, pp. 940–941, 1975.

[198] L. Rosenthaler, F. Heitger, O. Kubler, and R. von der Heydt, "Detection of general edges and keypoints," in *Proceedings of the European Conference on Computer Vision*, pp. 78–86, 1992.

[199] P. L. Rosin, "Representing curves at their natural scales," *Pattern Recognition*, vol. 25, pp. 1315–1325, 1992.

[200] P. L. Rosin, "Determining local natural scales of curves," *International Conference on Pattern Recognition*, 1994.

[201] P. L. Rosin, "Measuring corner properties," *Computer Vision and Image Understanding*, vol. 73, no. 2, pp. 292–307, 1999.

[202] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Proceedings of the International Conference on Computer Vision*, pp. 1508–1511, 2005.

[203] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Proceedings of the European Conference on Computer Vision*, pp. 430–443, 2006.

[204] C. A. Rothwell, A. Zisserman, D. A. Forsyth, and J. L. Mundy, "Planar object recognition using projective shape representation," *International Journal on Computer Vision*, vol. 16, pp. 57–99, 1995.

[205] W. S. Rutkowski and A. Rosenfeld, "A comparison of corner detection techniques for chain coded curves," Technical Report 623, Maryland University, 1978.

[206] P. A. Sandon, "Simulating visual attention," *Journal of Cognitive Neuroscience*, vol. 2, no. 3, pp. 213–231, 1990.

[207] P. V. Sankar and C. V. Sharma, "A parallel procedure for the detection of dominant points on digital curves," *Computer Graphics and Image Processing*, vol. 7, pp. 403–412, 1978.

[208] F. Schaffalitzky and A. Zisserman, "Viewpoint invariant texture matching and wide baseline stereo," in *Proceedings of the International Conference on Computer Vision*, pp. 636–643, 2001.

[209] F. Schaffalitzky and A. Zisserman, "Multi-view matching for unordered image sets," in *Proceedings of the European Conference on Computer Vision*, pp. 414–431, 2002.

[210] B. Schiele and J. L. Crowley, "Probabilistic object recognition using multidimensional receptive field histograms," in *Proceedings of the International Conference on Pattern Recognition,* vol. 2, pp. 50–54, 1996.

[211] B. Schiele and J. L. Crowley, "Recognition without correspondence using multi-dimensional receptive field histograms," *International Journal of Computer Vision*, vol. 36, no. 1, pp. 31–50, 2000.

[212] C. Schmid and R. Mohr, "Combining gray-value invariants with local constraints for object recognition," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 872–877, 1996.

[213] C. Schmid and R. Mohr, "Local gray-value invariants for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–534, 1997.

[214] C. Schmid, R. Mohr, and C. Bauckhage, "Comparing and evaluating interest points," in *Proceedings of the International Conference on Computer Vision*, pp. 230–235, 1998.

[215] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 151–172, 2000.

[216] S. Se, T. Barfoot, and P. Jasiobedzki, "Visual motion estimation and terrain modeling for planetary rovers," in *Proceedings of the International Symposium on Artificial Intelligence for Robotics and Automation in Space*, 2005.

[217] N. Sebe, T. Gevers, J. van de Weijer, and S. Dijkstra, "Corner detectors for affine invariant salient regions: Is color important," in *Proceedings of International Conference on Image and Video Retrieval*, pp. 61–71, 2006.

[218] N. Sebe, Q. Tian, E. Loupias, M. Lew, and T. Huang, "Evaluation of salient point techniques," *Image and Vision Computing*, vol. 21, no. 13–14, pp. 1087–1095, 2003.

[219] R. Sedgewick, *Algorithms*. Addison–Wesley, Second Edition, 1988.

[220] E. Seemann, B. Leibe, K. Mikolajczyk, and B. Schiele, "An evaluation of local shape-based features for pedestrian detection," in *Proceedings of the British Machine Vision Conference*, pp. 11–20, 2005.

[221] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 411–426, 2007.

[222] T. Serre, L. Wolf, and T. Poggio, "Object recognition with features inspired by visual cortex," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2005.

[223] A. Sha'ashua and S. Ullman, "Structural saliency: The detection of globally salient structures using a locally connected network," in *Proceedings of the International Conference on Computer Vision*, pp. 321–327, 1988.

[224] M. A. Shah and R. Jain, "Detecting time-varying corners," *Computer Vision, Graphics and Image Processing*, vol. 28, pp. 345–355, 1984.

[225] E. Sharon, A. Brandt, and R. Basri, "Segmentation and boundary detection using multiscale intensity measurements," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 469–476, 2001.

[226] F. Shen and H. Wang, "Corner detection based on modified Hough transform," *Pattern Recognition Letters*, vol. 23, no. 8, pp. 1039–1049, 2002.

[227] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[228] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 593–600, 1994.

[229] A. Singh and M. Shneier, "Gray-level corner detection a generalization and a robust real time implementation," in *Proceedings of the Computer Vision Graphics Image Processing*, vol. 51, pp. 54–69, 1990.

[230] S. N. Sinha, J. M. Frahm, M. Pollefeys, and Y. Genc, "GPU-based video feature tracking and matching," in *EDGE, Workshop on Edge Computing Using New Commodity Architectures*, 2006.

[231] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," in *Proceedings of the International Conference on Computer Vision*, pp. 1470–1478, 2003.

[232] S. M. Smith and J. M. Brady, "SUSAN — A new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 34, pp. 45–78, 1997.

[233] K. Sohn, J. H. Kim, and W. E. Alexander, "A mean field annealing approach to robust corner detection," *IEEE Transactions on Systems Man Cybernetics Part B*, vol. 28, pp. 82–90, 1998.

[234] J. Sporring, M. Nielsen, L. Florack, and P. Johansen, *Gaussian Scale-Space Theory*. Springer-Verlag, 1997.

[235] M. Stark and B. Schiele, "How good are local features for classes of geometric objects," in *Proceedings of the International Conference on Computer Vision*, 2007.

[236] B. Super and W. Klarquist, "Patch-based stereo in a general binocular viewing geometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 3, pp. 247–252, 1997.

[237] M. Swain and D. Ballard, "Color indexing," *International Journal in Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[238] C.-H. Teh and R. T. Chin, "On the detection of dominant points on digital curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 859–872, 1989.

[239] P. Tissainayagam and D. Suter, "Assessing the performance of corner detectors for point feature tracking applications," *Image and Vision Computing*, vol. 22, no. 8, pp. 663–679, 2004.

[240] M. Trajkovic and M. Hedley, "Fast corner detection," *Image and Vision Computing*, vol. 16, no. 2, pp. 75–87, 1998.

[241] A. M. Treisman, "Features and objects: The fourteenth Berlett memorial lecture," *Journal of Experimantal Psychology*, vol. 40A, pp. 201–237, 1988.

[242] W. Triggs, "Detecting keypoints with stable position, orientation and scale under illumination changes," in *Proceedings of the European Conference on Computer Vision,* vol. 4, pp. 100–113, 2004.

[243] L. Trujillo and G. Olague, "Synthesis of interest point detectors through genetic programming," *Genetic and Evolutionary Computation*, pp. 887–894, 2006.

[244] D. M. Tsai, "Boundary-based corner detection using neural networks," *Pattern Recognition*, vol. 30, pp. 85–97, 1997.

[245] M. A. Turk and A. P. Pentland, "Eigenfaces for face recognition," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 586–591, 1991.

[246] T. Tuytelaars and C. Schmid, "Vector quantizing feature space with a regular lattice," in *Proceedings of the International Conference on Computer Vision*, 2007.

[247] T. Tuytelaars and L. Van Gool, "Content-based image retrieval based on local affinely invariant regions," in *International Conference on Visual Information Systems*, pp. 493–500, 1999.

[248] T. Tuytelaars and L. Van Gool, "Wide baseline stereo matching based on local, affinely invariant regions," in *Proceedings of the British Machine Vision Conference*, pp. 412–425, 2000.
[249] T. Tuytelaars and L. Van Gool, "Matching widely separated views based on affine invariant regions," *International Journal of Computer Vision*, vol. 1, no. 59, pp. 61–85, 2004.
[250] J. van de Weijer, T. Gevers, and A. D. Bagdanov, "Boosting color saliency in image feature detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 150–156, 2006.
[251] A. Vedaldi and S. Soatto, "Features for recognition: Viewpoint invariance for non-planar scenes," in *Proceedings of the International Conference on Computer Vision*, pp. 1474–1481, 2005.
[252] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the Conference on Computer Vision and Pattern Recognition,* vol. 1, pp. 511–518, 2001.
[253] K. N. Walker, T. F. Cootes, and C. J. Taylor, "Locating salient object features," in *Proceedings of the British Machine Vision Conference*, 1998.
[254] H. Wang and M. Brady, "Real-time corner detection algorithm for motion estimation," *Image and Vision Computing*, vol. 13, no. 9, pp. 695–703, 1995.
[255] R. J. Watt, "Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus," *Journal of Optical Society of America*, vol. 4, no. 10, pp. 2006–2021, 1987.
[256] S. Winder and M. Brown, "Learning local image descriptors," *International Conference on Computer Vision and Pattern Recognition*, 2007.
[257] J. Winn, A. Criminisi, and T. Minka, "Object categorization by learned universal visual dictionary," in *Proceedings of the International Conference on Computer Vision,* vol. 2, pp. 1800–1807, 2005.
[258] A. P. Witkin, "Scale-space filtering," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1019–1023, 1983.
[259] M. Worring and A. W. M. Smeulders, "Digital curvature estimation," *CVGIP: Image Understanding*, vol. 58, no. 3, pp. 366–382, 1993.
[260] R. P. Wurtz and T. Lourens, "Corner detection in color images through a multiscale combination of end-stopped cortical cells," *Image and Vision Computing*, vol. 18, no. 6–7, pp. 531–541, 2000.
[261] X. Zhang, R. Haralick, and V. Ramesh, "Corner detection using the map technique," in *Proceedings of the International Conference on Pattern Recognition*, vol. 1, pp. 549–552, 1994.
[262] X. Zhang and D. Zhao, "A morphological algorithm for detecting dominant points on digital curves," *SPIE Proceedings, Nonlinear Image Processing 2424*, pp. 372–383, 1995.
[263] Z. Zheng, H. Wang, and E. Teoh, "Analysis of gray level corner detection," *Pattern Recognition Letters*, vol. 20, pp. 149–162, 1999.
[264] P. Zhu and P. M. Chirlian, "On critical point detection of digital shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 8, pp. 737–748, 1995.

[265] M. Zuliani, C. Kenney, and B. S. Manjunath, "A mathematical comparison of point detectors," in *Proceedings of the Computer Vision and Pattern Recognition Workshop*, p. 172, 2004.

[266] O. A. Zuniga and R. Haralick, "Corner detection using the facet model," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 30–37, 1983.