

Review Article

Building Local Features from Pattern-Based Approximations of Patches: Discussion on Moments and Hough Transform

Andrzej Sluzek

School of Computer Engineering, Nanyang Technological University, Blk N4, Nanyang Avenue, Singapore 639798

Correspondence should be addressed to Andrzej Sluzek, assluzek@ntu.edu.sg

Received 30 April 2008; Accepted 24 October 2008

Recommended by Simon Lucey

The paper overviews the concept of using circular patches as local features for image description, matching, and retrieval. The contents of scanning circular windows are approximated by predefined patterns. Characteristics of the approximations are used as feature descriptors. The main advantage of the approach is that the features are categorized at the detection level, and the subsequent matching or retrieval operations are, thus, tailored to the image contents and more efficient. Even though the method is not claimed to be scale invariant, it can handle (as explained in the paper) image rescaling within relatively wide ranges of scales. The paper summarizes and compares various aspects of results presented in previous publications. In particular, three issues are discussed in detail: visual accuracy, feature localization, and robustness against “visual intrusions.” The compared methods are based on relatively simple tools, that is, area moments and modified Hough transform, so that the computational complexity is rather low.

Copyright © 2009 Andrzej Sluzek. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

It has been well demonstrated in numerous reports on physiology of vision (e.g., [1, 2]) that, in general, humans perceive known objects as collections of local visual saliencies. Several theories differently explain details of the process (see the critical survey in [3]), but there is a common understanding that when a sufficient number of local features found in the observed image consistently match correspondingly similar features in a known object, the object would be recognized. Although optical illusions may happen in some cases, such a mechanism allows visual detection of known objects under various degrading conditions (occlusions, cluttered scenes, partial visibility due to poor illumination, etc.).

Even without the psychophysiological justification, low-level local features have been used in computer vision since the 1980s. Initially, they were primarily considered a mechanism for stereovision and motion tracking (e.g., [4, 5]) but later, the same approach was found useful for many other applications of machine vision (e.g., image matching, detection of partially hidden objects, visual information retrieval). Typical detectors of low-level local features are derived from differential properties of image intensities or

colors. The most popular detectors (e.g., Harris-Plessey [5] or SIFT [6]) are based on derivatives in spatial and/or scale domains and they do not retrieve any structural information from the image (even though, Harris-Plessey is often called a “corner detector”). However, there is a documented need for matching based on the local visual contents. For example, Mikolajczyk and Schmid in [7] presented cases of corresponding local features that are correctly detected but cannot be matched because of inadequate descriptors. Those features would be easily matched if the “visual similarity” between extracted patches can be quantified.

One of the most popular methods of image content matching is based on moment invariants which exist for various types of geometric and photometric distortions (e.g., [8, 9]). Several works employ them as descriptors of local features (e.g., [9, 10]) computed over circular windows (or windows of other regular shapes). Many alternative techniques for the local image content description exist as well. For example, local contrast measures have been reported (see [11]) as powerful descriptors in textured images. Another method, based on locally applied Radon filters, has been successfully used for description and recognition of human faces (see [12]). The above-mentioned approaches assume

that local features can be matched by extracting (and comparing) properties invariant under distortions present in the analyzed images. However, the actual concept of “visual similarity” goes beyond that.

According to Biederman (see [1]), humans recognize known objects by identifying certain classes of geometric patterns that are combinations of contour and region properties. Such patterns may have diversified shapes, but all instances of the same pattern have the same structural composition that can be parameterized (at least approximately) using several configuration and intensity/color parameters. The method discussed in our paper follows this idea (although we do not use *geons* proposed by Biederman). The main assumption is that visual saliencies (local features) of interest correspond to various local geometric patterns that may exist within analyzed images. Even if the image is noised or distorted, the patterns (if prominent enough) should remain visible, although their appearances may be corrupted.

As in the majority of local feature detectors, the proposed method employs a scanning window of a regular shape. For rotational invariance, circular windows are proposed, but the method can work using windows of other shapes as well (e.g., squares or hexagons). Generally, the windows are larger in other detectors (because more complex contents have to be identified within windows), but the actual size of scanning windows is of secondary importance (as explained in Section 4). The objective is to detect those locations of the scanning window, where the window content is “visually similar” to a pattern of interest and to find the best approximation of the window by this pattern, that is, to create an idealized local model of the image. Two simple examples are shown in Figure 1, where digital circular windows of 30-pixel radius are approximated by a corner and a *T-junction* (the patterns that can be clearly visible in the windows).

Such locally found approximations can be potentially very powerful features for identifying similar fragments in images, for detecting partially visible known objects, for visual information retrieval, and for other similar tasks.

This paper presents analysis, discussion, and exemplary results on how such approximation-based local features can be defined and detected in images. Although certain aspects of the presented method have been already published (e.g., [13, 14]), this is an attempt to summarize the results and to highlight the identified advantages and drawbacks. In particular, the following issues are explored:

- (1) building accurate pattern-based approximations in the presence of degrading effects (techniques based on area moments and on modified Hough transform are discussed in Section 2);
- (2) quantitative methods of estimating “visual similarity” between approximations and the approximated windows (both indirect approaches, moment similarities, similarities based on Hough transform, and direct methods, Radon transform and image correlation, are briefly overviewed in Section 3);
- (3) definition, accurate localization, and scale invariance of approximation-based features (based on results of 1 and 2) are discussed in Section 4.



FIGURE 1: Exemplary approximations of circular windows by patterns accurately corresponding to the actual visual contents of the windows.

In all sections, exemplary figures are used to illustrate the discussed effects and properties.

Preliminary concepts on how such approximation-based local features can be incorporated into image matching systems are briefly discussed only in Section 5 that concludes the paper.

2. Pattern-Based Approximations of Circular Patches

We assume that patterns of interest are defined by circular patches containing certain geometric structures. Patches of other regular shapes (e.g., squares, hexagons, etc.) can be considered as well, but circular patches are more universal because of their rotational invariance. Several examples of patterns of interest are given in Figure 2.

As shown in the figure, patterns are defined over circles of an arbitrary radius R , and each instance of a pattern is represented (within the general characteristic of the pattern) by several configuration parameters (defining its geometry) and several intensities (or colors) describing the pattern's visual appearance. The number of parameters (i.e., the complexity of patterns) is not limited, but patterns with 2–4 configurations parameters (and similar numbers of intensities/colors) are the most realistic ones for scanning windows of a limited diameter. All examples shown in Figure 2 are such patterns. For example, a *T-junction* pattern is defined by three colors C_1 , C_2 , and C_3 , the angular width β_1 , and the orientation angle β_2 .

When an image is analyzed, we attempt to approximate contents of a scanning window by the available patterns. The pattern-based local features are found at locations where “the best approximations” exist. Parameters of those approximations would be used as descriptors of the features. In our researches, the radius of scanning windows ranges between 7 and 25 pixels. Smaller windows do not provide enough resolution for patterns with fine details, while larger windows unnecessarily increase computational costs.

Formally, the pattern-based approximation consists in computing the optimum configuration parameters and intensities/colors for a given content of the scanning circular window. Knowing the optimum parameters, we can synthesize the pictorial form of the approximation (as shown in Figure 1). The synthesized images are used mainly for visualization (to estimate how accurately, from the human perspective, the original image has been approximated) and, generally, are not needed for other purposes.

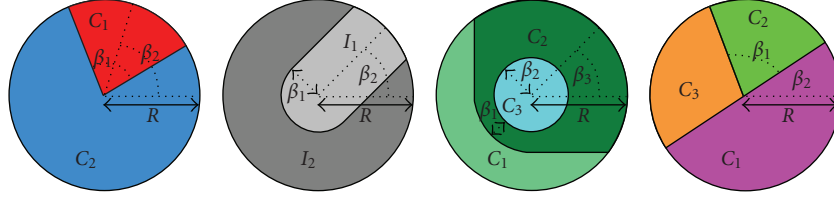


FIGURE 2: Exemplary types of patterns. Configuration parameters (β) and intensity/color parameters (I/C) are indicated for each pattern.

2.1. Moment-Based Approximations. Our previous papers (e.g., [13]) presented a moment-based technique for producing approximations for various patterns. It was based on the observation that configuration parameters and intensity/color parameters of patterns can be expressed as functions of low-order moments computed over the whole circle. For example, the angular width β_1 of a *corner* pattern (i.e., one of its configuration parameters, see Figure 2) is equal to

$$\beta_1 = 2 \arcsin \sqrt{1 - \frac{16[(m_{20} - m_{02})^2 + 4m_{11}^2]}{9R^2(m_{10}^2 + m_{01}^2)}}, \quad (1)$$

while the orientation angle β_2 for a *T-junction* pattern (see Figure 2) satisfies

$$m_{01} \cos \beta_2 - m_{10} \sin \beta_2 = \pm \frac{4}{3R} \sqrt{(m_{20} - m_{02})^2 + 4m_{11}^2}, \quad (2)$$

where m_{pq} are moments of $p + q$ order computed within the system of coordinates placed in the window center.

Intensities of the approximations can be also expressed using moments. For example, three intensities of a *T-junction* pattern (see Figure 2) satisfy the following system of linear equations:

$$\begin{aligned} \frac{2m_{00}}{R^2} &= I_1\pi + I_2\beta_1 + I_3(\pi - \beta_1), \\ \frac{3m_{10}}{R^3} &= -2I_1c_2 + I_2(c_2 - c_{2-1}) + I_3(c_2 + c_{2-1} - 2s_2), \\ \frac{3m_{01}}{R^3} &= -2I_1s_2 + I_2(s_2 - s_{2-1}) + I_3(s_2 + s_{2-1} + 2c_2), \end{aligned} \quad (3)$$

where c_x and s_x indicate $\cos \beta_x$ and $\sin \beta_x$, correspondingly.

Alternatively, when the configuration parameters are already known, we can estimate the intensities/colors of the approximations by averaging intensities/colors of the corresponding regions within the approximated patch.

Equations (1)–(3) (and their counterparts for other patterns) are basically the same for both grey-level and color images. The only difference is that for color images, moments are 3D vectors (moments computed for RGB components), so that the expressions should be modified accordingly (details are discussed in [15]).

The expressions derived for a certain pattern can be applied to a circular image of any content, and the obtained values (if the solutions exist, e.g., (1) or (2) may not have any solution) become parameters of the approximation of the given image by this pattern.



FIGURE 3: Exemplary moment-based approximations for a *corner* pattern.

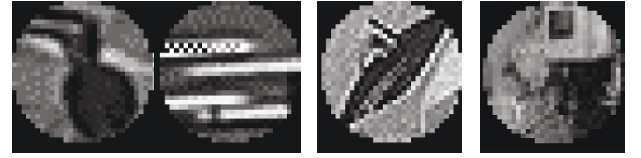


FIGURE 4: Circular images for which approximations do not exist for *corner* (2 examples), *T-junction*, and *pierced round corner* patterns, correspondingly.

This method has several advantages. First, it produces accurate approximations even for textured images (where other techniques, e.g., the corner approximations discussed in [16], fail) and for heavily blurred patterns where visual identification of a pattern is difficult even for a human eye (see examples in Figure 3 for *corner* patterns).

The method can also identify windows which cannot be approximated by the pattern of interest (the corresponding equations have no solutions). Exemplary circular images for which approximations cannot be found are given in Figure 4.

There are also disadvantages of the moment-based approximation technique. First, in many cases, it produces an approximation even though the visual inspection clearly indicates that the window content is not similar to the given pattern. Several examples of such scenarios are given in Figure 5.

Secondly, the quality of approximations may be strongly affected by “visual intrusions,” that is, unwanted additions to the image content caused by other objects, illumination effects, or just by the natural nonuniformity of images. A relatively mild effect of “visual intrusion” is shown in Figure 6(a), where a dark stripe affects accuracy of the *corner* approximation produced by the method. A much worse situation can be seen in Figure 6(b), where an external object enters the circular window and completely distorts the approximation by a 90° *T-junction* pattern (even though, the shape of the actual junction within the image is not affected by the intrusion).

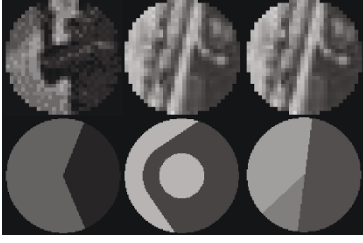


FIGURE 5: Visually incorrect approximations of circular images by corner, pierced round corner, and T-junction patterns.

Moment-based approximations are also difficult mathematically. Equations for calculating approximations parameters (similar to (1)–(3)) should be individually designed for each type of patterns. Even for relatively simple patterns, polynomial expressions of higher orders are needed. For example, approximations by *pierced round corner* pattern (see Figure 2) use 4th-order polynomial equations. Moreover, the limited number of low-level moments (higher-level moments are too sensitive to noise and digitization effects) naturally limits the number of parameters, that is, the complexity of patterns. It is very difficult to find a reasonably simple solution for patterns with more than 3 configuration parameters (and the corresponding number of intensities/colors).

2.2. Approximations Based on Hough Transform. Patterns considered in this work can be represented as unions of grey-level/color regions and contours defining boundaries between those regions (see Figure 2). Thus, an alternative method of building pattern approximations can be based on contour detection techniques. Several similar attempts have been reported previously (e.g., [17]), but our objective is to develop a tool suitable for patterns more complex than typically considered corners or junctions.

We propose to use a well-known Hough transform with modifications addressing needs and constraints of the problem. First of all, the calculations are performed within the limited area of scanning windows so that, in order to provide enough data, all images pixels are involved instead of contour pixels only. This technique (preliminarily proposed in [18]) exploits directional properties of image gradients.

Assume that Hough transform is built for the family of 2D curves specified by

$$f(x, y, a_1, \dots, a_n) = 0 \quad (4)$$

equations with n parameters a_1, \dots, a_n .

Each pixel (x_0, y_0) of $I(x, y)$ image contributes to (A_1, \dots, A_n) accumulator in the parameter space the dot product of the image gradient $\vec{\nabla}I$ and the unit vector normal to the hypothetical curve (both taken at (x_0, y_0) coordinates):

$$\begin{aligned} \text{Acc}(A_1, \dots, A_n) &= \text{Acc}(A_1, \dots, A_n) + \vec{\nabla}I(x_0, y_0) \\ &\quad \circ \text{norm}(f(x_0, y_0, A_1, \dots, A_n)). \end{aligned} \quad (5)$$

Thus, regardless the gradient magnitude, only the gradient components orthogonal to the expected curve are actually taken into account. For example, if concentric circles (or their arcs) are detected, only the radial components of the gradient are taken into account, while for detecting radial segments, only the components that are orthogonal to radials (see Figure 7).

We additionally increase the contribution of pixels proportionally to their distance from the circle's center because of poorer angular resolution in the central part of digital circles. A somehow similar problem has been handled in [19] by using polar coordinates.

After contours of a pattern-based approximation have been extracted, intensities/colors of the corresponding regions can be estimated using the methods described in Section 2.1.

There are several advantages of using Hough transform for building pattern-based approximations of circular images. In particular, the approximation results are generally much less sensitive to “visual intrusions.” Figure 8 shows examples where in spite of intrusions distorting the “idealized” contents of circular windows, the accuracy of approximations is very good, much better than by using moment expressions.

Moreover, approximations can be often obtained even if the pattern areas differ only in textures, while the average intensities/colors are identical. An illustrative example of such case (where corners can be hardly identified) is shown in Figure 9.

Another important advantage is that Hough transform-based approximations can be decomposed and built incrementally. In many cases, contours defining the pattern boundaries consist of fragments that can be detected (using Hough transform) separately. The configuration parameters of already found contour components can be used as default values for detection of subsequent fragments.

A pattern shown in Figure 10 (*sharp pierced corner*) has four configuration parameters (orientation β , angular width α , radius of the hole r , and distance d). Search in a 4D parameter space would be computationally expensive. However, the corner component of the boundary can be identified using only a 2D space (the orientation angle β and the angular width α). Given the orientation angle β , the hole parameters can be found in another 2D space (radius r and distance d).

Some weaknesses of this method also exist. In particular, approximations built using Hough transform may have random, incorrect configurations in heavily blurred images. An example is given in Figure 11.

It can be eventually concluded that techniques for building pattern-based approximations of patches can be based on both integral (moments) and gradient (Hough transform) properties of approximated images. However, gradient-based mechanisms should be considered the tool of primary importance.

In this paper, we discuss only relatively simple techniques with low computational complexity. Although more complex mathematical models have been proposed for the same or similar problems (e.g., [12, 17, 20], etc.), we believe that for

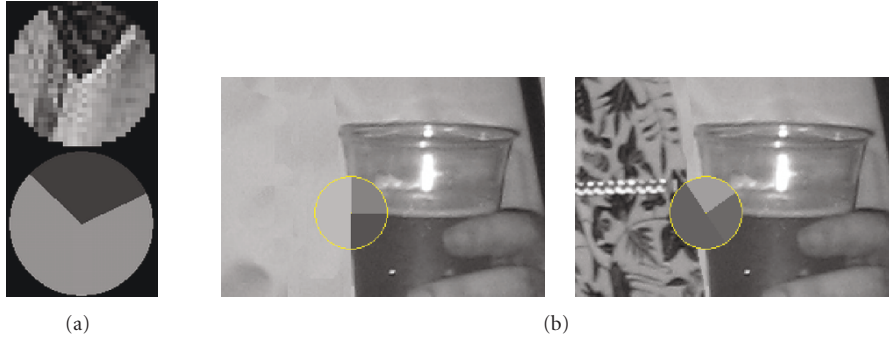


FIGURE 6: Examples of (a) a mild distortion and (b) strong distortion of the moment-based approximations caused by “visual intrusions.”

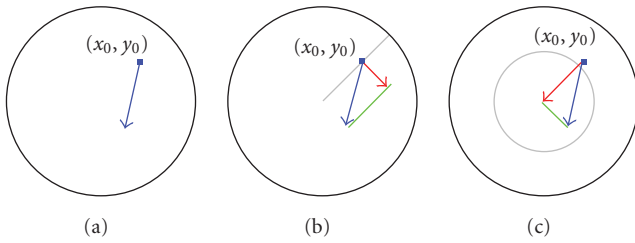


FIGURE 7: (a) Exemplary intensity gradient and (b) its contribution to the Hough accumulator when detecting radial lines and (c) detecting concentric circles.

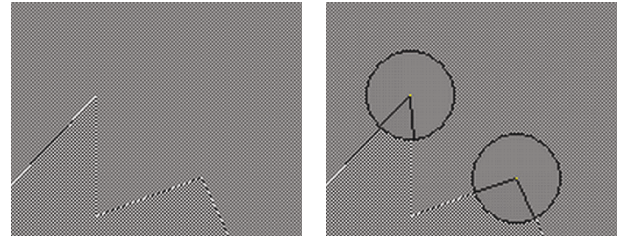


FIGURE 9: Examples of corners produced by texture differences only. The approximations have been accurately found based on Hough transform.

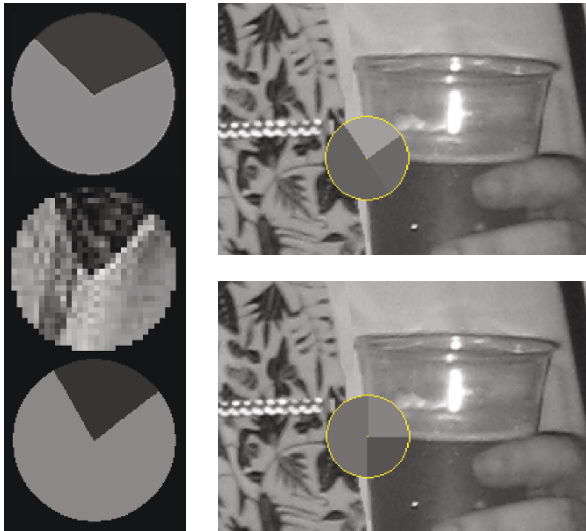


FIGURE 8: Comparison between moment-based approximations (top row) and approximation based on modified Hough transform (bottom row) in case of “visual intrusions.”

the majority of intended applications, the methods discussed in this paper provide at least satisfactory solutions.

3. Accuracy of Approximations

The main objective of building pattern-based approximations of patches is to obtain robust local features, that is, features that can be reliably detected in images that are distorted

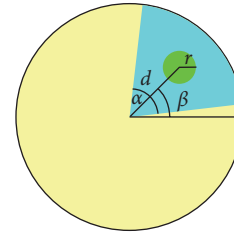


FIGURE 10: A pattern with four configurations parameters. A 4D parameter space used for Hough transform-based approximation building can be decomposed into two 2D problems.

and degraded by various effects. This assumption would be justified if the approximations are actually similar to the approximated fragments. However, as shown in Section 2 (e.g., Figures 5 and 6), visual appearances of approximations may strongly differ from the approximated images. Such approximations are obviously useless, because the visual structures of the original images are lost.

Therefore, there is a need to quantify similarity between approximations and approximated patches. Only those image locations where the highest similarity exists between window contents and their approximations would be used as the local features of interest. The similarity measures should obviously correspond to the “visual similarity” (i.e., the similarity subjectively estimated by a human observer) between images. Additionally, the measures should be simple



FIGURE 11: Corner-based approximations of a blurred image obtained using moments (left) and Hough transform (right).

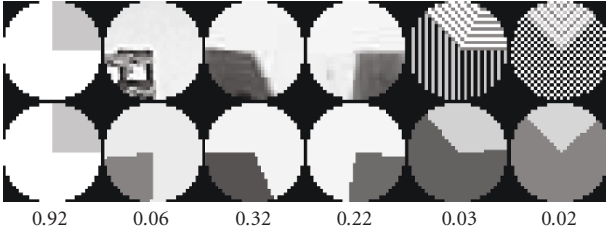


FIGURE 12: Examples of *corner* approximations of similar visual quality but different similarity measures (based on cross-correlation).

enough to be repetitively applied to the window scanning images.

The most straightforward similarity measure would be a cross-correlation which does not even need normalization because we expect roughly the same colors/intensities in circular patches and in their pattern-based approximations. However, as discussed in [13], neither the overall cross-correlation (i.e., computed over the whole patch) nor any combination of regional cross-correlations (i.e., computed separately for each region of the approximation) is a reliable measure. Figure 12 shows several circular patches and their *corner* approximations. Visually, all approximations are equally similar to the approximated patches, but the correlation-based similarities (given in Figure 12) are very different. Therefore, even though the features can be found as local maxima of the similarity values, the correspondence between the visual similarity and the similarity measure is very poor.

Moreover, to effectively use the cross-correlation as a similarity measure, the approximation images should be synthesized (with the resolution corresponding to the size of patches).

Thus, alternative similarity measures with lower computational complexity have been proposed and tested. Similarity of low-level moments and similarity of Radon transforms have been reported in [21, 22], correspondingly. They provide more uniform correspondence between “visual quality” of approximations and computed similarities (exemplary results showing a simultaneous deterioration of both “visual quality” and computed similarities are shown in Figure 13). These measures are not sensitive to (uniformly distributed) noise, so that their global maxima can be used to determine positions of the pattern-based local features.

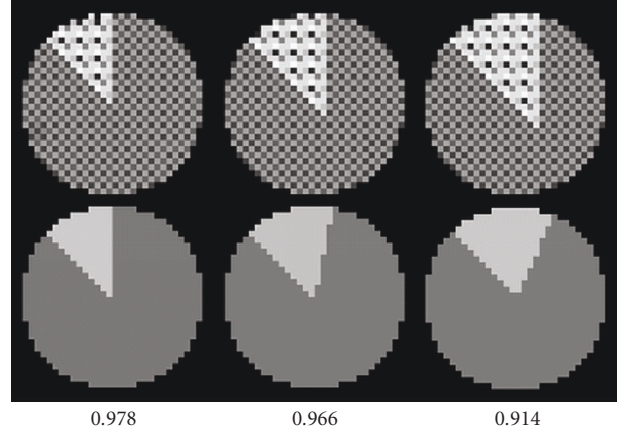


FIGURE 13: *Corner* approximations of gradually deteriorating both “visual quality” and computed similarity (similarity measure based on low-order moments).

It should be noticed, however, that in Figure 13, the similarity values change very slowly, much slower than the visual similarity that deteriorates rapidly. This is a significant disadvantage of such measures, as further discussed in Section 4.

Moreover, all abovementioned similarity measures are very sensitive to visual intrusions, so that even accurate approximations (e.g., built using Hough transform) may not be recognized as such.

An entirely different similarity measure can be proposed if Hough transform is used for building pattern-based approximations. For accurate approximations, the content of the winning bin in the parameter space is usually a prominent spike, while for less accurate approximations, the spike is less protruding. Thus, after testing several other approaches also based on Hough transform, we propose the similarity measure as *the ratio of the winning bin height over the sum of all bins' contents*. Exemplary results given in Figure 14 show how significantly this ratio changes when the scanning window moves away from the actual pattern location. In this example, 90° *T-junction* pattern has been deliberately selected because it needs only a 1D parameter space.

Currently, we consider this measure of similarity superior to other tested approaches, as far as the feature localization is concerned. However, this is not an absolute measure, that is, its values fluctuate significantly when the image is noised, even if the noise neither affects the “visual quality” of the pattern nor modifies the produced approximations. A self-explaining example (with the approximations superimposed over the original images) is shown in Figure 15. Thus, localization of the pattern-based features should be again based on detecting *local maxima* of similarity.

In the future applications, we plan a combination of this similarity measure with secondary area-based measures (Radon transform or moments). The primary measure would be used to localize the feature candidates. The secondary measure would provide the (absolute-value) estimate of whether the local maximum of the primary measure is

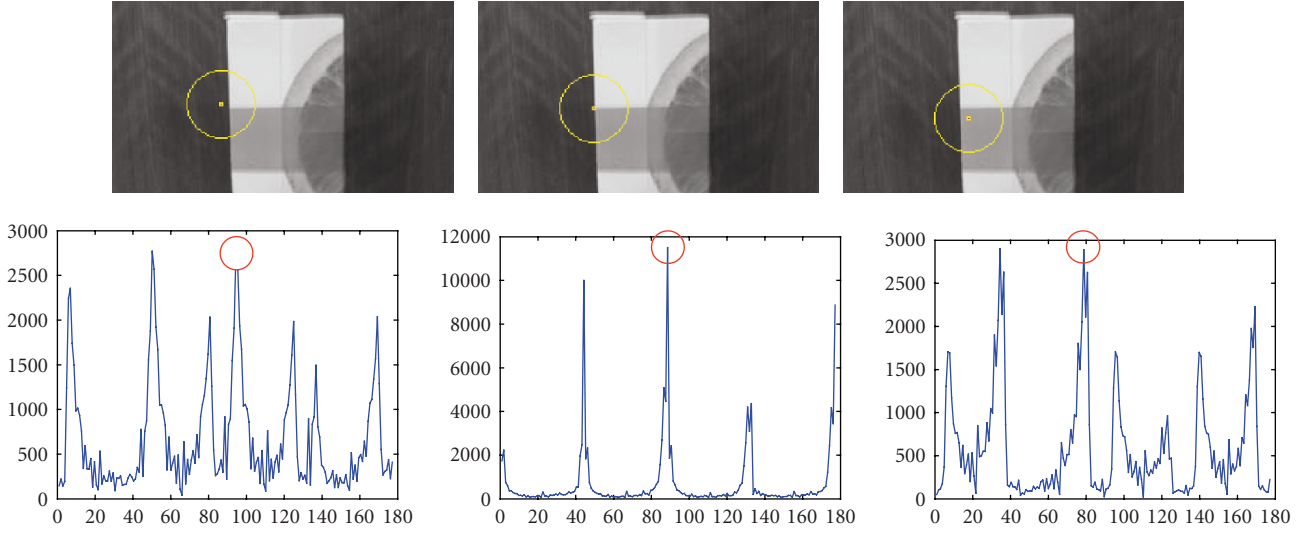


FIGURE 14: Three locations of the scanning window and the corresponding parameter space values (bin contents) for Hough transform of 90° *T-junction* pattern. The central column shows the window at the position matching the actual junction.

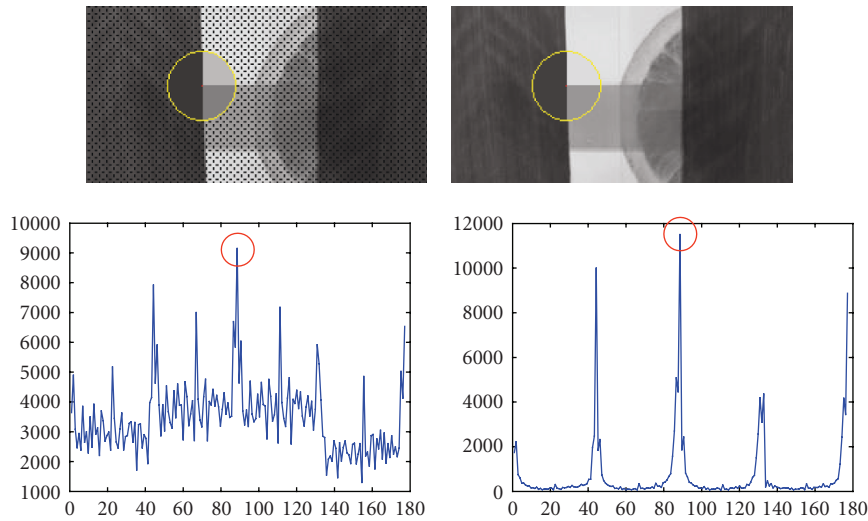


FIGURE 15: Changes of the bin contents for Hough transform of 90° *T-junction* pattern caused by a high-frequency noise added. The original results are shown for the reference.

actually a high-quality approximation or whether it should be ignored.

We can, thus, conclude that while accurate pattern-based approximations can be found relatively easily, it is more difficult to quantify the accuracy of approximations in a manner corresponding to a visual assessment by human observers. Measures are needed that (1) produce similarities proportional to the “visual quality” of approximations, as perceived by humans, (2) are insensitive to noises degrading the overall quality of images, (3) are robust against visual intrusions that do not affect the actual patterns of interest, and (4) produce sharp maxima for the actual locations of the patterns. The existing measures are not fully satisfactory yet, and we believe that a further development of similarity measures is an interesting topic of practical importance.

4. Approximation-Based Local Features

4.1. Detection and Localization. Based on the explanations given in Sections 2 and 3, the definition of approximation-based local features is straightforward.

A local feature (of radius R) defined by pattern P exists at a location (x, y) within the analyzed image I if:

- (1) the approximation by pattern P of the circular window of radius R located at (x, y) exists;
- (2) similarity between the approximation and the window content reaches a local maximum at (x, y) ;
- (3) (optional) the value of the absolute similarity measure (see Section 3) exceeds a predefined threshold.

Configuration and intensity/color parameters of the approximation are considered descriptors of the feature.

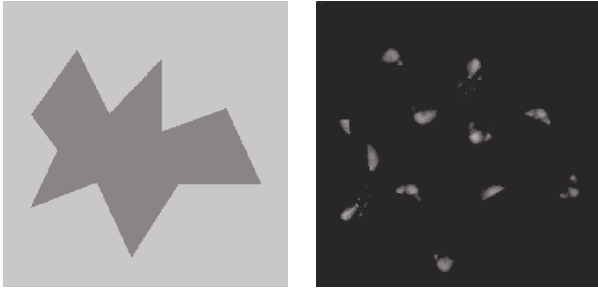


FIGURE 16: Localization problems for *corner* features using area-based similarity measures.

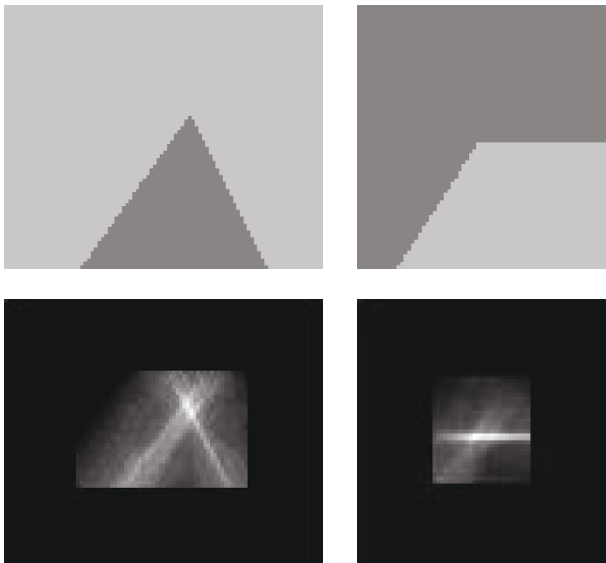


FIGURE 17: Localization of selected *corner* features using the similarity measure based on Hough transform.

In practice, implementation details of the above definition can vary. For example, it is well known that standard keypoint detectors (e.g., Harris-Plessey or SIFT) produce significant numbers of keypoints in typical images of natural quality. It is, therefore, possible to select the keypoints produced by such detectors as preliminary candidates for approximation-based local features and apply the method only to these locations. The advantage of such an approach is that the only task is to build the approximations and to estimate their accuracy (the localization of feature candidates is performed by the keypoint detector). Another recommended option is to scan images using larger position increments and to conduct a pixel-by-pixel search only around locations where approximations are found with a reasonable accuracy.

It should be noted, nevertheless, that both in the original method and in its improved variants, the same location can produce several pattern-based features. This happens if the window content can be approximated with a comparatively similar accuracy by several patterns.

Unless feature candidates are prelocated by an external keypoint detector, the similarity values are used to localize the approximation-based features. Unfortunately, as

indicated in Section 3, the area-based similarity measures (i.e., cross-correlation, moments, and Radon transforms) do not perform well in this problem. Even in high-quality images, there is a tendency to detect clusters of pixels with comparable similarity values instead of producing sharp maxima. The actual location of the feature would be somewhere within a cluster, but the similarity variations are so small (see the example in Figure 13) that a minor noise, a small distortion, or even digitization effects may shift the maximum of similarity to a distant part of a cluster. Figure 16 shows clusters produced by *corner* approximations for an exemplary image of perfect (digitally) quality. Note highly uniform similarities (represented by intensities) within the clusters.

However, similarity measures based on Hough transform localize features with pixel accuracy (we do not consider subpixel accuracy although certain possibilities are discussed in [8]). Exemplary results for two corners from Figure 16 are given in Figure 17 (similarities are again represented by intensity levels). Figure 18 shows an exemplary 256×256 image and several pattern-based features detected within this image.

4.2. Are Approximation-Based Features Scale-Invariant? Approximations discussed in this paper are built over circular images of radius R . Therefore, in principle, the method is not scale invariant. Any change of radius (or image rescaling) may result in different sets, different descriptors, and/or different localization of approximation-based features.

However, from the practical perspective, the proposed features should be considered scale invariant within a certain range of scales. Figure 19 shows an exemplary image with several approximations obtained for windows of two significantly different diameters. The results given in Figure 19 illustrate a more general property of approximation-based features. As long as the scanning windows are large enough to include the approximating patterns but small enough so that the patterns are not visually suppressed by prominent features from the neighboring areas, the size of scanning window is actually not important for detecting pattern-based features. Of course there are certain limits but we conclude from the preliminary experiments that for typical images, the radius of scanning windows can vary within approximately 50–200% range without significant changes in the results. Most of detected approximation-based features are the same, and their characteristics (parameters of approximations) also remain unaffected.

Because the numbers of approximation-based features extracted from a single image are rather large (depending primarily on the image complexity and the number of available patterns), many of the features become effectively scale invariant in the sense explained above.

5. Summary

Currently, the most prospective area of application for approximation-based feature is visual information retrieval

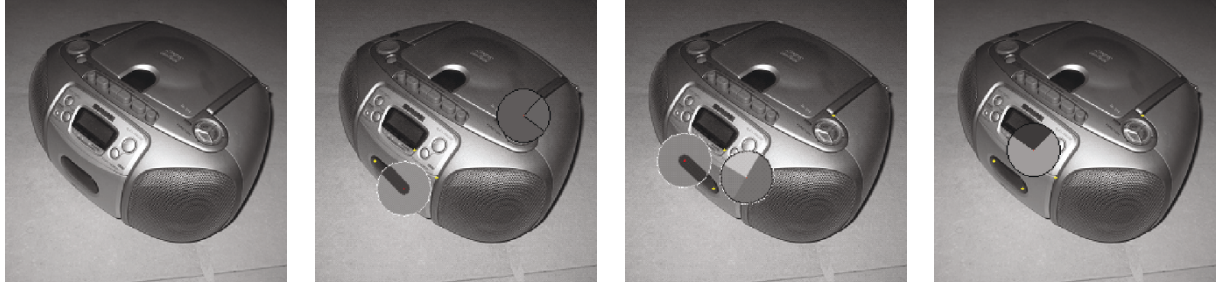


FIGURE 18: A 256×256 image and several approximation-based local features detected (shown in three images for better visibility).

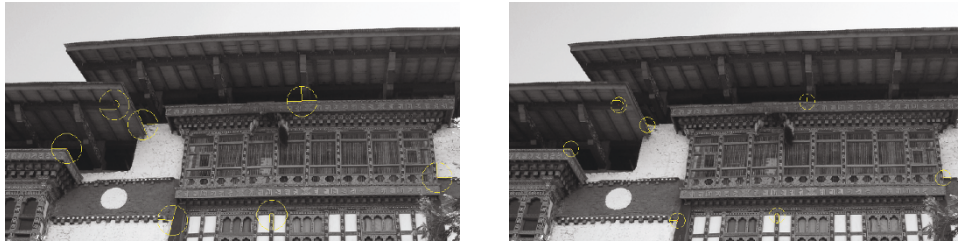


FIGURE 19: Exemplary pattern-based local features obtained by using scanning windows of significantly different sizes.

(VIR). Although computations used in the proposed algorithms are simple, the amount of data to be processed (moments and/or Hough transforms calculated over scanning windows of significant sizes applied to large images, determining similarities between approximations, and windows, etc.) is prohibitively large for typical real-time tasks (e.g., for vision-based search operations in exploratory robotics). Thus, the advantages of approximation-based features reflect primarily our VIR experiences and goals.

We envisage that database images will be preprocessed, that is, approximation-based features are pre-detected and memorized in the database together with the images. Such feature detection and memorization for all database images can be done offline whenever computational resources are available. The additional memory requirements are insignificantly small compared to the memory needed to store the images themselves. New types of approximation-based features can be incrementally added to the databases when approximation builders for new patterns become available.

The proposed features are a natural candidate for matching images since they provide local visual semantics of the analyzed images. Whenever a query image is submitted, it would be processed in the same way. Subsequently, local feature extracted from the query image would be matched against the database features. If enough evidence is found that the local semantics of the query image and of a database image are similar (e.g., approximations by the same patterns are extracted at correspondingly matching locations and descriptors of the approximations are correspondingly consistent), the images may contain visually similar fragments. Because the configuration descriptors of the features are considered more significant than colors/intensities, images containing visually similar fragments can be matched even

if they are seen in completely different visual conditions (nonuniform changes of illuminations, different coloring, etc.). Nevertheless, variations of the matching algorithm are possible (depending on the applications), so that colors/intensities can be considered important descriptors as well.

Comparing to matching techniques based on other local features, the complexity of matching using the approximation-based features can be significantly reduced. Approximation-based features are categorized by approximating pattern so that only features approximated by the same patterns are the potential matches. Thus, the estimated number of attempted matches is reduced exponentially. Additionally, the method allows “targeted” image matching by using only a subset of available patterns (those representing the visual contents considered important in a given problem).

The issues of effective image matching using the approximation-based local features are not discussed in this paper. Generally, the techniques are similar to already known algorithms, for example, geometric hashing (see [23]) or methods used in [6, 7, 10].

The paper has presented only the principles of the proposed methods and approaches. Thus, no conclusive statistics on the method’s performances can be presented yet. Currently, the methods are integrated into a working platform that can be used for selected applications. One of the important issues is expansion of the list of available patterns so that complex images can be described by large numbers of more diversified features. It is our hope that the proposed approach can be developed into useful tools for visual data storage and retrieval systems (including internet browsers for visual contents). Further results of currently conducting researches will be addressed in future papers.

Acknowledgments

The results presented in the paper are done under A*STAR Science and Engineering Research Council Grant no. 072 134 0052. The financial support of SERC is gratefully acknowledged.

References

- [1] I. Biederman, "Recognition-by-components: a theory of human image understanding," *Psychological Review*, vol. 94, no. 2, pp. 115–147, 1987.
- [2] M. J. Tarr, H. H. Bülthoff, M. Zabinski, and V. Blanz, "To what extent do unique parts influence recognition across changes in viewpoint?" *Psychological Science*, vol. 8, no. 4, pp. 282–289, 1997.
- [3] S. Edelman, "Computational theories of object recognition," *Trends in Cognitive Sciences*, vol. 1, no. 8, pp. 296–304, 1997.
- [4] H. Moravec, "Rover visual obstacle avoidance," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, pp. 785–790, Vancouver, Canada, August 1981.
- [5] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the 4th Alvey Vision Conference (AVC '88)*, pp. 147–151, Manchester, UK, September 1988.
- [6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [7] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.
- [8] A. Sluzek, "Identification of planar objects in 3-D space from perspective projections," *Pattern Recognition Letters*, vol. 7, no. 1, pp. 59–63, 1988.
- [9] F. Mindru, T. Tuytelaars, L. van Gool, and T. Moons, "Moment invariants for recognition under changing viewpoint and illumination," *Computer Vision and Image Understanding*, vol. 94, no. 1–3, pp. 3–27, 2004.
- [10] Md. Saiful Islam and A. Sluzek, "Relative scale method to locate an object in cluttered environment," *Image and Vision Computing*, vol. 26, no. 2, pp. 259–274, 2008.
- [11] T. Maenpaa and M. Pietikainen, "Texture analysis with local binary patterns," in *Handbook of Pattern Recognition and Computer Vision*, C. H. Chen and P. S. P. Wang, Eds., pp. 197–216, World Scientific, Teaneck, NJ, USA, 3rd edition, 2005.
- [12] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Chen, "Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 1, pp. 786–791, Beijing, China, October 2005.
- [13] A. Sluzek, "On moment-based local operators for detecting image patterns," *Image and Vision Computing*, vol. 23, no. 3, pp. 287–298, 2005.
- [14] A. Sluzek, "A new local-feature framework for scale-invariant detection of partially occluded objects," in *Proceedings of the 1st Pacific Rim Symposium on Advances in Image and Video Technology (PSIVT '06)*, L.-W. Chang, W.-N. Lie, and R. Chiang, Eds., vol. 4319 of *Lecture Notes in Computer Science*, pp. 248–257, Springer, Hsinchu, Taiwan, December 2006.
- [15] A. Sluzek, "Approximation-based keypoints in colour images—a tool for building and searching visual databases," in *Proceedings of the 9th International Conference on Advances in Visual Information Systems (VISUAL '07)*, G. Qiu, C. Leung, X.-Y. Xue, and R. Laurini, Eds., vol. 4781 of *Lecture Notes in Computer Science*, pp. 5–16, Springer, Shanghai, China, June 2007.
- [16] S.-T. Liu and W.-H. Tsai, "Moment-preserving corner detection," *Pattern Recognition*, vol. 23, no. 5, pp. 441–460, 1990.
- [17] L. Parida, D. Geiger, and R. Hummel, "Junctions: detection, classification, and reconstruction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, pp. 687–698, 1998.
- [18] F. O'Gorman and M. B. Clowes, "Finding picture edges through collinearity of feature points," *IEEE Transactions on Computers*, vol. C-25, no. 4, pp. 449–456, 1976.
- [19] K. Murakami, Y. Maekawa, M. Izumida, and K. Kinoshita, "Fast line detection by the local polar coordinates using a window," *Systems and Computers in Japan*, vol. 38, no. 6, pp. 43–52, 2007.
- [20] M. A. Ruzon and C. Tomasi, "Edge, junction, and corner detection using color distributions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1281–1295, 2001.
- [21] A. Sluzek and Md. Saiful Islam, "New types of keypoints for detecting known objects in visual search tasks," in *Vision Systems: Application*, G. Obinata and A. Dutta, Eds., pp. 423–442, I-Tech, Vienna, Austria, 2007.
- [22] A. Sluzek, "Keypatches: a new type of local features for image matching and retrieval," in *Proceedings of the 16th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG '08)*, pp. 231–238, Plzen, Czech Republic, February 2008.
- [23] H. J. Wolfson and I. Rigoutsos, "Geometric hashing: an overview," *IEEE Computational Science & Engineering*, vol. 4, no. 4, pp. 10–21, 1997.