Evaluation of Salient Point Techniques

N. Sebe[1], Q. Tian[2], E. Loupias[3], M. Lew[1], and T. Huang[2]

[1] LIACS, Leiden University, Leiden, The Netherlands,
{nicu, mlew}@liacs.nl
[2] Beckman Institute, University of Illinois at Urbana-Champaign, USA
{qitian, huang}@ifp.uiuc.edu
[3] Laboratoire Reconnaissance de Formes et Vision, INSA-Lyon, France
loupias@rfv.insa-lyon.fr

Abstract. In image retrieval, global features related to color or texture are commonly used to describe the image content. The problem with this approach is that these global features cannot capture all parts of the image having different characteristics. Therefore, local computation of image information is necessary. By using salient points to represent local information, more discriminative features can be computed. In this paper we compare a wavelet-based salient point extraction algorithm with two corner detectors using the criteria: repeatability rate and information content. We also show that extracting color and texture information in the locations given by our salient points provides significantly improved results in terms of retrieval accuracy, computational complexity, and storage space of feature vectors as compared to global feature approaches.

1 Introduction

Haralick and Shapiro [1] consider a point in an image *interesting* if it has two main properties: distinctiveness and invariance. This means that a point should be distinguishable from its immediate neighbors and the position as well as the selection of the interesting point should be invariant with respect to the expected geometric and radiometric distortions. Considering these properties, Schmid and Mohr [2] proposed the use of corners as interest points in image retrieval. Different corner detectors are evaluated and compared in [3] and the authors show that the best results are provided by the Harris corner detector [4].

Corner detectors, however, were designated for robotics and shape recognition and they have drawbacks when are applied to natural images. Visual focus points do not need to be corners: when looking at a picture, we are attracted by some parts of the image, which are the most meaningful for us. We cannot assume them to be located only in corner points, as is mathematically defined in most corner detectors. For instance, a smoothed edge may have visual focus points and they are usually not detected by a corner detector. Corners also gather in textured regions. The problem is that due to efficiency reasons only a preset number of points per image can be used in the indexing process. Since in this case most of the detected points will be in a small region, the other parts of the image may not be described in the index at all.

We aim for a set of interesting points called *salient points* that are related to any visual interesting part of the image whether it is smoothed or corner-like. Moreover, to describe different parts of the image, the set of salient points should not be clustered in few regions. We believe multi-resolution representation is interesting to detect salient points. Our wavelet-based salient points [5] are detected for smoothed edges and are not gathered in texture regions. Hence, they lead to a more complete image representation than corner detectors.

We also compare our wavelet-based salient point detector with the Harris corner detectors used by Schmid and Mohr [3]. In order to evaluate the "interestingness" of the points obtained with these detectors (as was introduced by Haralick and Shapiro [1]) we compute the repeatability rate and the information content. We are also interested in using the salient points in a retrieval scenario. Therefore, in a small neighborhood around the location of each point we extract local color and texture features and use only this information in retrieval. It is quite easy to understand that using a small amount of such points instead of all image pixels reduces the amount of data to be processed. Moreover, local information extracted in the neighborhood of these particular points is assumed to be more robust to classic transformations (additive noise, affine transformations including translation, rotation, and scale effects, partial visibility).

2 Wavelet-Based Salient Points

A wavelet-based salient point detector has been presented in our previous work [6]. Here we briefly present the outline of the algorithm and we show some examples of detected salient points.

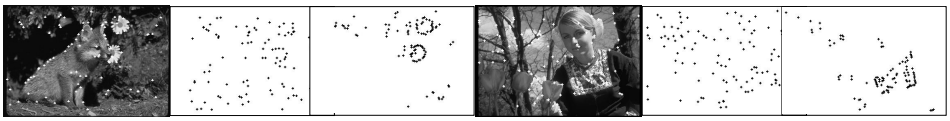
A wavelet is an oscillating and attenuated function with zero integral. We study the image f at the scales (or resolutions) 1/2, 1/4, ..., 2^j, $j \in Z$ and $j \leq -1$. The wavelet detail image $W_{2^j} f$ is obtained as the convolution of the image with the wavelet function dilated at different scales. We consider orthogonal wavelets with compact support. First, this assures that we have a complete and non-redundant representation of the image. Second, we know from which signal points each wavelet coefficient at the scale 2^j was computed. We can further study the wavelet coefficients for the same points at the finer scale 2^{j+1}. There is a set of coefficients at the scale 2^{j+1} computed with the same points as a coefficient $W_{2^j} f(n)$ at the scale 2^j. We call this set of coefficients the children $C(W_{2^j} f(n))$ of the coefficient $W_{2^j} f(n)$. The children set in one dimension is:

$$C(W_{2^j} f(n)) = \{W_{2^{j+1}} f(k), 2n \leq k \leq 2n + 2p - 1\} \tag{1}$$

where p is the wavelet regularity, $0 \leq n < 2^j N$, and N the length of the signal.

Each wavelet coefficient $W_{2^j} f(n)$ is computed with $2^{-j} p$ signal points. It represents their variation at the scale 2^j. Its children coefficients give the variations of some particular subsets of these points (with the number of subsets depending on the wavelet). The most salient subset is the one with the highest wavelet coefficient at the scale 2^{j+1}, that is the maximum in absolute value of $C(W_{2^j} f(n))$. In our salient point extraction algorithm, we consider this maximum and look at his highest child. Applying recursively this process, we select a coefficient $W_{2^{-1}} f(n)$ at the finer resolution 1/2. Hence, this coefficient represents $2p$ signal points. To select a salient point

Fig. 1. Salient points extraction: spatial support of tracked coefficients



Daubechies4 Haar Harris Daubechies4 Haar Harris

Fig. 2. Salient points examples. For Daubechies4 and Haar salient points are detected for smooth edges (fox image) and are not gathered in textured regions (girl image).

from this tracking, we choose among these $2p$ points the one with the highest gradient (Figure 1). We set its saliency value as the sum of the absolute value of the wavelet coefficients in the track:

$$saliency = \sum_{k=1}^{-j} |C^{(k)}(W_{2^j} f(n))|, -\log_2 N \leq j \leq -1 \tag{2}$$

The tracked point and its saliency value are computed for every wavelet coefficient. A point related to a global variation has a high saliency value, since the coarse wavelet coefficients contribute to it. A finer variation also leads to an extracted point, but with a lower saliency value. We then need to threshold the saliency value, in relation to the desired number of salient points. We first obtain the points related to global variations; local variations also appear if enough salient points are requested.

The salient points extracted by this process depend on the wavelet we use. Haar is the simplest wavelet function, so is the fastest for execution. The larger the spatial support of the wavelet, the more the number of computations. Nevertheless, some localization drawbacks can appear with Haar due to its non-overlapping wavelets at a given scale. This can be avoided with the simplest overlapping wavelet, Daubechies4. Examples of salient points extracted using Daubechies4, Haar, and Harris detectors are shown in Figure 2. Note that while for Harris the salient points lead to an incomplete image representation, for the other two detectors the salient points are detected for smooth edges (as can be seen in the fox image) and are not gathered in texture regions (as can be seen in the girl image). Hence, they lead to a more complete image representation.

3 Repeatability and Information Content

Repeatability is defined by the image geometry. Given a 3D point P and two projection matrices M_1 and M_2, the projections of P into two images I_1 and I_2 are $p_1 = M_1 P$ and $p_2 = M_2 P$. The point p_1 detected in image I_1 is repeated in image I_2 if the corresponding point p_2 is detected in image I_2. To measure the repeatability, a unique relation between p_1 and p_2 has to be established. In the case of a planar scene this relation is defined by an homography: $p_2 = H_{21} p_1$.

The percentage of detected points which are repeated is the *repeatability rate*. A repeated point is not in general detected exactly at position p_2, but rather in some neighborhood of p_2. The size of this neighborhood is denoted by ε and repeatability within this neighborhood is called ε-*repeatability*. The set of point pairs (p_2, p_1) which correspond within an ε-neighborhood is $P(\varepsilon) = \{(p_2, p_1) | dist(p_2, H_{21} p_1) < \varepsilon\}$. Considering N, the total number of points detected, the repeatability rate is:

$$r(\varepsilon) = \frac{|D(\varepsilon)|}{N},\ 0 \leq r(\varepsilon) \leq 1 \tag{3}$$

We would also like to know how much average information content a salient point "has" as measured by its greylevel pattern. The more distinctive the greylevel patterns are, the larger the entropy is. In order to have rotation invariant descriptors

for the patterns, we chose to characterize salient points by local greyvalue rotation invariants which are combinations of derivatives. We computed the "local jet" [7] which is consisted of the set of derivatives up to N^{th} order. These derivatives describe the intensity function locally and are computed stably by convolution with Gaussian derivatives. The local jet of order N at a point $\mathbf{x} = (x, y)$ for an image I and a scale σ is defined by: $J^N[I](\mathbf{x}, \sigma) = \{L_{i_1 \dots i_n}(\mathbf{x}, \sigma) | (\mathbf{x}, \sigma) \in I \times R^+\}$, where $L_{i_1 \dots i_n}(\mathbf{x}, \sigma)$ is the convolution of image I with the Gaussian derivatives $G_{i_1 \dots i_n}(\mathbf{x}, \sigma)$, $i_k \in \{x, y\}$.

In order to obtain invariance under the group $SO(2)$ (2D image rotation), Koenderink and van Doorn [7] computed differential invariants from the local jet:

$$\boldsymbol{\nu}[0 \dots 3] = \begin{bmatrix} L_x L_x + L_y L_y \\ L_{xx} L_x L_x + 2 L_{xy} L_x L_y + L_{yy} L_y L_y \\ L_{xx} + L_{yy} \\ L_{xx} L_{xx} + 2 L_{xy} L_{xy} + 2 L_{yy} L_{yy} \end{bmatrix} \tag{4}$$

The computation of entropy requires a partitioning of the space of $\boldsymbol{\nu}$. Partitioning is dependent on the distance measure between descriptors and we consider the approach described by Schmid, et al. [3]. The distance we used is the Mahalanobis distance given by: $d_M(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) = \sqrt{(\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2)^T \Lambda^{-1} (\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2)}$, where $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ are two descriptors and Λ is the covariance of $\boldsymbol{\nu}$. The covariance matrix Λ is symmetric and positive definite. Its inverse can be decomposed into $\Lambda^{-1} = P^T D P$ where D is diagonal and P an orthogonal matrix. Furthermore, we can define the square root of Λ^{-1} as $\Lambda^{-1/2} = D^{1/2} P$ where $D^{1/2}$ is a diagonal matrix whose coefficients are the square roots of the coefficients of D. The Mahalanobis distance can then be rewritten as: $d_M(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) = ||D^{1/2} P (\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2)||$. The distance d_M is the norm of difference of the normalized vectors: $\boldsymbol{\nu}_{norm} = D^{1/2} P \boldsymbol{\nu}$. This normalization allows us to use equally sized cells in all dimensions. This is important since the entropy is directly dependent on the partition used. The probability of each cell of this partition is used to compute the entropy of a set of vectors $\boldsymbol{\nu}$.

In the experiments we used a set of 1000 images taken from the Corel database and we compared 4 salient point detectors. In Section 2 we introduced two salient point detectors using wavelets: Haar and Daubechies4. For benchmarking purposes we also considered the Harris corner detector [4] and a variant of it called PreciseHarris, introduced by Schmid, et al. [3].

3.1 Results for Repeatability

Before we can measure the repeatability of a particular detector we first had to consider typical image alterations such as image rotation and image scaling. In both cases, for each image we extracted the salient points and then we computed the average repeatability rate over the database for each detector.

In the case of image rotation, the rotation angle varied between 0° and 180°. The repeatability rate in a ε=1 neighborhood for the rotation sequence is displayed in Figure 3.

The detectors using wavelet transform (Haar and Daubechies4) give better results compared with the other ones. Note that the results for all detectors are not very dependent on image rotation. The best results are provided by Daubechies4 detector.

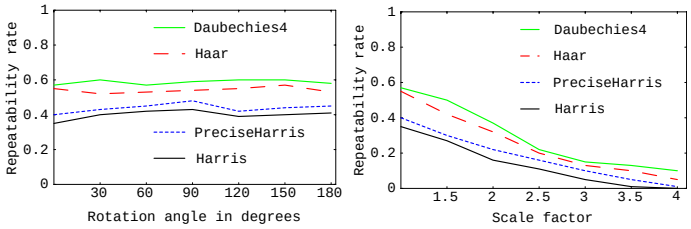


Fig. 3. Repeatability rate for image rotation (left) and scale change (right) (ε=1)

In the case of scale changes, for each image we considered a sequence of images obtained from the original image by reducing the image size so that the image was aspect-ratio preserved. The largest scale factor used was 4.

The repeatability rate for scale change is presented in Figure 3. All detectors are very sensitive to scale changes. The repeatability is low for a scale factor above 2 especially for Harris and PreciseHarris detectors. The detectors based on wavelet transform provide better results compared with the other ones.

3.2 Results for Information Content

For each detector we computed the salient points for the set of images and characterized each point by a vector of local greyvalue invariants (cf. Eq. (4)). The invariants were normalized and the entropy of the distribution was computed. The cell size in the partitioning was the same in all dimensions and it was set to 20. The σ used for computing the greylevel invariants was 3. We also considered random points in our comparison. For each image in the database we computed the mean number m of salient points extracted by different detectors and then we selected m random points using a uniform distribution.

The results are given in Table 1. The detector using the Daubechies4 wavelet transform has the highest entropy and thus the salient points obtained are the most distinctive. The results obtained for Haar wavelet transform are almost as good. The results obtained with PreciseHarris detector are better than the ones obtained with Harris but worse than the ones obtained using the wavelet transform. Moreover, the results obtained for all of the salient points detectors are significantly better than those obtained for random points. The difference between the results of Daubechies4 and random points is about a factor of two.

Detector	**Entropy**
Haar	6.0653
Daubechies4	6.1956
Harris	5.4337
PreciseHarris	5.6975
Random	3.124

Table 1. The information content for different detectors

In summary, the most "interesting" salient points were detected using the Daubechies4 detector. These points have the highest information content and proved to be the most robust to rotation and scale changes. Therefore, in our next

experiments we will consider this detector and as benchmark the PreciseHarris corner detector.

4 Content-Based Retrieval

Our next goal is to use the salient points in a content-based retrieval scenario. We consider a modular approach: the salient points are first detected for each image in the database and then feature vectors are extracted from a small neighborhood around each salient point. This approach assures the independence of the salient point extraction techniques and the feature extraction procedure and gives the user the liberty to use any features he wants for a specific application [8]. In our experiments in constructing the feature vectors we used color moments because they provide a compact characterization of color information and they are more robust and efficient in content-based retrieval than the well-known color histograms [9] and Gabor texture features because they are extensively used for texture characterization [10, 11]. Of course the wavelet coefficients used during the salient point detection can be also used in constructing the feature vectors.

The number of salient points extracted will clearly influence the retrieval results. We performed experiments (not presented here due to space limitation) in which the number of salient points varied from 10 to several hundreds and found out that when using more than 50 points, the improvement in accuracy we obtained did not justify the computational effort involved. Therefore, in the experiments, 50 salient points were extracted for each image.

For feature extraction, we considered the set of pixels in a small neighborhood around each salient point. In this neighborhood we computed the color moments (in a 3×3 neighborhood) and the Gabor moments (in a 9×9 neighborhood). For convenience, this approach is denoted as the Salient W (wavelet) approach when Daubechies4 detector is used and as the Salient C (corner) approach when the PreciseHarris corner detector is used. For benchmarking purposes we also considered the results obtained using the color moments and the wavelet moments [10] extracted over the entire image (denoted as Global CW approach) and the results obtained using the color moments and the Gabor moments [11] extracted over the entire image (denoted as Global CG approach).

The overall similarity distance D_j for the j^{th} image in the database is obtained by linearly combining the similarity distance of each individual feature:

$$D_j = \sum_i W_i S_j(f_i), \quad \text{with} \quad S_j(f_i) = (\mathbf{x}_i - \mathbf{q}_i)^T (\mathbf{x}_i - \mathbf{q}_i) \quad \text{and} \quad j = 1, \ldots, N \qquad (5)$$

where N is the total number of images in the database and $\mathbf{x}_i$ and $\mathbf{q}_i$ are the i^{th} feature (e.g. $i = 1$ for color and $i = 2$ for texture) vector of the j^{th} image in the database and the query, respectively. The low-level feature weights W_i for color and texture in Eq. (5) are set to be equal.

4.1 Results

In the first experiments we considered a database of 479 images (256×256 pixels in size) of color objects such as domestic objects, tools, toys, food cans, etc [12]. As ground truth we used 48 images of 8 objects taken from different camera viewpoints (6 images for a single object). Both color and texture information

were used. The Salient approaches, the Global CW approach, and the Global CG approach were compared. Color moments were extracted either globally (the Global CW and Global CG) or locally (the Salient approaches). For wavelet texture representation of the Global CW approach, each input image was first fed into a wavelet filter bank and was decomposed into three wavelet levels, thus 10 de-correlated subbands. For each subband, the mean and standard deviation of the wavelet coefficients were extracted. The total number of wavelet texture features was 20. For the Salient approaches, we extracted Gabor texture features from the 9×9 neighborhood of each salient point. The dimension of the Gabor filter was 7×7 and we used 2 scales and 6 orientations/scale. The first 12 features represented the averages over the filter outputs and the last 12 features were the corresponding variances. Note that these features were independent so that they had different ranges. Therefore, each feature was then Gaussian normalized over the entire image database. For the Global CG approach, the global Gabor texture features were extracted. The dimension of the global Gabor filter was 61×61. We extracted 36 Gabor features using 3 scales and 6 orientations/scale. The first 18 features were the averages over the filters outputs and the last 18 features were the corresponding variances.

In Figure 4 we show an example of a query image and the similar images from the database retrieved with various ranks. The Salient point approaches outperform both the Global CW approach and the Global CG approach. Even when the image was taken from a very different viewpoint, the salient points captured the object details enough so the similar image was retrieved with a good rank. The Salient W approach shows better retrieval performance than the Salient C approach. The Global CG approach provides better performance than the Global CW approach. This fact demonstrates that Gabor feature is a very good feature for texture characterization. Moreover, it should also be noted that: (1) the Salient point approaches only use the information from a very small part of the image, but still achieve a good representation of the image. For example, in our object database $9 \times 9 \times 50$ pixels were used to represent the image. Compared to the Global approaches (all 256×256 pixels were used), the Salient approaches only use less than 1/16 of the whole image pixels. (2) Compared to the Global CG approach, the Salient approaches have much less computational complexity.

Query

Salient W	1	2	4	15	21
Salient C	1	3	7	18	25
Global CW	2	7	12	25	33
Global CG	2	5	9	20	27

Fig. 4. Example of images of one object taken from different camera viewpoints and the corresponding ranks of each individual image using different approaches

Table 2 shows the retrieval accuracy for the object database. Each of the 6 images from the 8 classes was considered as query image and the average retrieval accuracy was calculated.

Top	6	10	20
Salient W	61.2	75.2	85.7
Salient C	58.9	73.8	83.2
Global CW	47.3	62.4	71.7
Global CG	58.3	73.4	82.8

Table 2. Retrieval accuracy (%) using 48 images from 8 classes for object database

Results in Table 2 show that using the salient point information the retrieval results are significantly improved (>10%) compared to the Global CW approach. When compared to the Global CG approach, the retrieval accuracy of the Salient W approach is 2.9%, 2.8%, and 2.9% higher in the top 6, 10, and 20 images, respectively. The Salient C approach has approximatively 2.5% lower retrieval accuracy comparing with the Salient W approach. Additionally, the Salient approaches have much lower computational complexity and 33.3% less storage space of feature vectors than the Global CG approach. Although the global wavelet texture features are fast to compute, their retrieval performance is much worse than the other methods. Therefore, in terms of overall retrieval accuracy, computational complexity, and storage space of feature vectors, the Salient W approach is best among all the approaches.

In our second experiments we considered a database consisted of 4013 various images covering a wide range of natural scenes such as animals, buildings, paintings, mountains, lakes, and roads. In order to perform quantitative analysis, we randomly chose 15 images from a few categories, e.g., building, flower, tiger, lion, road, forest, mountain, sunset and use each of them as queries. For each category, we measured how many hits, i.e. how many similar images to the query were returned in the top 20 retrieved images.

Figure 5 shows the average number of hits for each category using the Global CW approach, the Global CG approach, and the Salient W approach. Clearly the Salient approach has similar performance comparing with the Global CG approach and outperforms the Global CW approach when the first five categories are considered. For the last three categories, which are forest, mountain, and sunset, the global approaches (both Global CW and Global CG) perform better than the Salient approach because now the images exhibit more global characteristics and therefore, the global approaches can capture better the image content.

As noted before, the Salient approach uses only a very small part of the image to extract the features. Therefore, comparing with the global approaches the Salient approach has much less computational complexity. Regarding the storage space of feature vectors, the number of Gabor texture features used in the Salient approach and the Global approach were 24 and 36, respectively. This does not have a big effect for small database. However, for very large image databases, the storage space used for these texture features will surely make big

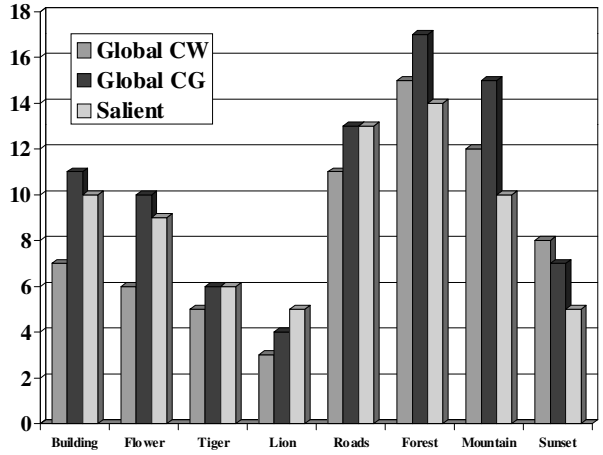


Fig. 5. The average number of hits for each category using the global color and wavelet moments (Global CW), the global color and Gabor moments (Global CG) and the Salient W approach (Salient)

difference. As to the color features, both approaches have the same number of features.

5 Discussion

In this paper we compared a wavelet-based salient point extraction algorithm with two corner detectors using the criteria: repeatability rate and information content. Our points have more information content and better repeatability rate than the Harris corner detector. Moreover, the detectors have significantly more information content than randomly selected points.

We also show that extracting color and texture information in the locations given by our salient points provides significantly improved results in terms of retrieval accuracy, computational complexity, and storage space of feature vectors as compared to global feature approaches. Our salient points are interesting for image retrieval because they are located in visual focus points and therefore, they capture the local image information.

For content-based retrieval, a fixed number of salient points (50 points in this paper) were extracted for each image. Color moments and Gabor moments were extracted from the 3×3 and the 9×9 neighborhood of the salient points, respectively. For benchmark purpose, the Salient point approaches were compared to the global color and wavelet moment (Global CW) approach and the global color and Gabor moments (Global CG) approach.

Two experiments were conducted and the results show that: (1) the Salient approaches have better performance than the Global CW approach. The Salient approaches proved to be robust to the viewpoint change because the salient points were located around the object boundaries and captured the details inside the objects, neglecting the background influence; (2) The Salient approaches have similar performance compared to the Global CG approach in terms of the retrieval accuracy. However, the Salient approaches achieve the best performance in the overall considerations of retrieval accuracy, computational complexity, and

storage space of feature vectors. The last two factors will have very important influence for very large image databases; (3) Better retrieval results are obtained when Daubechies4 salient points are used compared with Harris corners. This shows that our wavelet-based points can capture better the image content.

Our experimental results also show that the global Gabor features perform much better than the global wavelet features. This fact is consistent with the results of the other researchers in the field proving that Gabor features are very good candidates for texture characterization.

In conclusion, the novel contribution of this paper is in showing that a wavelet-based salient point technique beats the current leading method which uses the PreciseHarris corner detector [3] with respect to the area of content-based retrieval. In addition, we show that the wavelet-based salient point technique outperforms global feature methods, because the salient points are able to capture the local feature information and therefore, they provide a better characterization for the scene content. Moreover, the salient points are more "interesting" (as defined by Haralick and Shapiro [1]) than the Harris corner points since they are more distinctive and invariant.

In our future work, we plan to explore salient point extraction techniques which mimic the way the humans extract information in an image. This will hopefully lead to more semantically meaningful results. Moreover, we plan to extract shape information in the location of the salient points making the retrieval more accurate. We also intend to automatically determine the optimal number of the salient points needed to be extracted for each image.

References

1. Haralick, R., Shapiro, L.: Computer and Robot Vision II. Addison-Wesley (1993)
2. Schmid, C., Mohr, R.: Local grayvalue invariants for image retrieval. IEEE Trans on Patt Anal and Mach Intell **19** (1997) 530–535
3. Schmid, C., Mohr, R., Bauckhage, C.: Evaluation of interst point detectors. I J Comp Vis **37** (2000) 151–172
4. Harris, C., Stephens, M.: A combined corner and edge detector. Alvey Vis Conf (1993) 147–151
5. Loupias, E., Sebe, N.: Wavelet-based salient points: Applications to image retrieval using color and texture features. In: Visual'00. (2000) 223–232
6. Tian, Q., Sebe, N., Lew, M., Loupias, E., Huang, T.: Image retrieval using wavelet-based salient points. Journal of Electronic Imaging **10** (2001) 835–849
7. Koenderink, J., van Doorn, A.: Representation of local geometry in the visual system. Biological Cybernetics **55** (1987) 367–375
8. Sebe, N., Tian, Q., Loupias, E., Lew, M., Huang, T.: Color indexing using wavelet-based salient points. In: IEEE Workshop on Content-based Access of Image and Video Libraries. (2000) 15–19
9. Stricker, A., Orengo, M.: Similarity of color images. SPIE - Storage and Retrieval for Image and Video Databases III **2420** (1995) 381–392
10. Smith, J., Chang, S.F.: Transform features for texture classification and discrimination in large image databases. Int Conf on Imag Process **3** (1994) 407–411
11. Ma, W., Manjunath, B.: A comparison of wavelet transform features for texture image annotation. Int Conf on Imag Process **2** (1995) 256–259
12. Gevers, T., Smeulders, A.: PicToSeek: Combining color and shape invariant features for image retrieval. IEEE Trans Imag Process **20** (2000) 102–119