Wavelet-based Salient Points for Image Retrieval

E. Loupias[1*] N. Sebe[+] S. Bres[*] J.-M. Jolion[*]

[+]*Leiden Institute of Advanced Computer Science*
Leiden University, The Netherlands
nicu@wi.leidenuniv.nl

[*]*Laboratoire Reconnaissance de Formes*
et Vision, INSA Lyon
{loupias,jolion}@rfv.insa-lyon.fr

Abstract

The use of interest points in content-based image retrieval allows image index to represent local properties of the image. Classic corner detectors can be used for this purpose. However, they have drawbacks when applied to various natural images for image retrieval, because visual features need not be corners and corners may gather in small regions.

We present a salient point detector that extract points where variations occur in the image, whether they are corner-like or not. It is based on wavelet transform to detect global variations as well as local ones.

The wavelet-based salient points are evaluated for image retrieval with a retrieval system using texture features. Retrieval results are compared for different point detectors.

1 Introduction

We are interested in content-based image retrieval in general image databases. The query is an image (*iconic search*), and the retrieved images should be *similar* to the query. Since high-level concepts (objects, feelings, etc.) cannot be extracted automatically from the image without specific knowledge, we use an *image similarity* based on low-level features (such as color, texture and shapes).

To allow fast querying, an image is "summarized" by a set of features, the *image index*. *Global features* give information about the image as a whole [1,2]. The use of *local features* is an interesting alternative, since it leads to an index based on local properties of the image. This approach is also attractive for sub-image search.

The feature extraction is limited to a subset of the image pixels, the *interest points*, where the image information is supposed to be the most important [3,4,5,6]. This paper focuses on the selection of points that are significant to compute features for indexing.

Corner detectors are commonly used for indexing [3,5]. Corners are usually defined as points where gradient is high in multiple orientations. This definition leads to corner detectors based on local derivatives [7,8,9]. Another approach is based on local neighborhood properties [10,11].

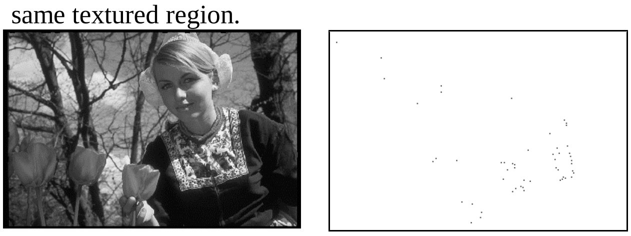
Corner detectors, which are designed for robotics or shape recognition, have drawbacks when applied to various natural images for image retrieval.

Visual focus points need not be corners: visual meaningful feature is not necessarily located in a corner point. For instance in Figure 1, the fox fur is too smoothed to be detected as a corner.

(a) Fox image (b) 100 corners (Harris)

Figure 1. Image with smoothed edges

Corners may gather in small regions: in various natural images, regions may well contain textures (trees, shirt patterns, etc.), where a lot of corners are detected (*cf.* Figure 2). To limit the indexing computation time, the number of points is preset. Most of the corners are in the same textured region.



(a) Dutch image (b) 100 corners (Harris)

Figure 2. Image with texture in the Dutch dress

With corner detectors, both examples lead to an incomplete representation, where some parts of the image are not described in the index.

[1] Guest period in Leiden University was supported by Région Rhône-Alpes (EURODOC grant)

For these reasons, corner points may not represent the most interesting subset of pixels for image indexing. Indexing points should be related to any visual "interesting" part of the image, whether it is smoothed or corner-like. To describe different parts of the image, the set of interesting point should not be clustered in few regions.

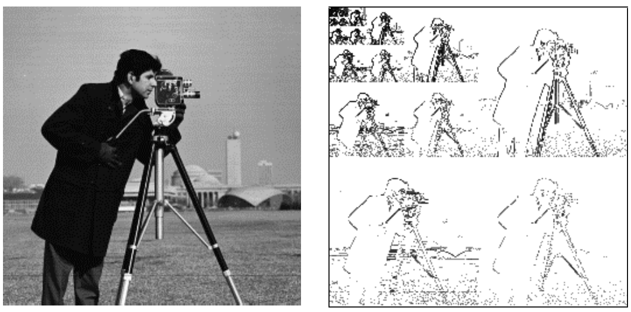
From now on, we will refer to these points as ***salient points***, which are not necessarily corners. We will avoid the term *interest points*, which is ambiguous, since it was previously used in the literature as *corner*. Wavelet representations, which express image variations at different resolutions, are attractive to extract salient points.

Previous point detectors make use of multiresolution representation. Chen *et al.* consider two different resolutions to extract corners [12]. In image retrieval context, contrast-based points are extracted in [4]. However, a lot of points are also extracted in textured regions because these regions are contrasted. Points are extracted with a specific wavelet in [6]. But since only a given scale is used, different resolutions features cannot be detected.

2 From wavelet transform to salient points

The wavelet transform is a multiresolution representation that expresses image variations at different scales. For wavelet theory, see [13]. For wavelet description and algorithms, see [14].

A wavelet is an *oscillating* and *attenuated* function (its integral is equal to zero). We study the image f at the scales (or resolutions) ½, ¼, … 2^j, $j \in \mathbb{Z}$ and $j \leq -1$. The wavelet *detail image* $W_{2^j} f$ is the convolution of the image with the wavelet function dilated at different scales.



(a) Cameraman image (b) Haar transform

Figure 3. Haar transform

Here we consider *orthogonal wavelets*, which lead to a complete and non-redundant representation of the image. A wavelet can also have a *compact support*: its value is zero outside a bounded interval. The simplest orthogonal compactly supported wavelet is the Haar wavelet, which is the discontinuous crenel function. Daubechies discovered wavelets, with any regularity $p > 1$, that are also orthogonal and compactly supported [15].

Salient point extraction

The wavelet representation gives information about the variations in the signal at different scales. In our retrieval context, we would like to extract salient points from any part of the image where "something" happens in the signal at any resolution. A high wavelet coefficient (in absolute value) at a coarse resolution corresponds to a region with high global variations. The idea is to find a relevant point to represent this global variation by looking at wavelet coefficients at finer resolutions.

Since we use wavelets with a compact support, we know from which signal points each wavelet coefficient at the scale 2^j was computed, and we can study the wavelet coefficients for the same points at the finer scale 2^{j+1}. Indeed 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. This occurs because of the compact support of the wavelet (see [16] for details). We call this set of coefficients the *children* $C\left(W_{2^j} f(n)\right)$ of the coefficient $W_{2^j} f(n)$. The children set in one dimension[2] is:

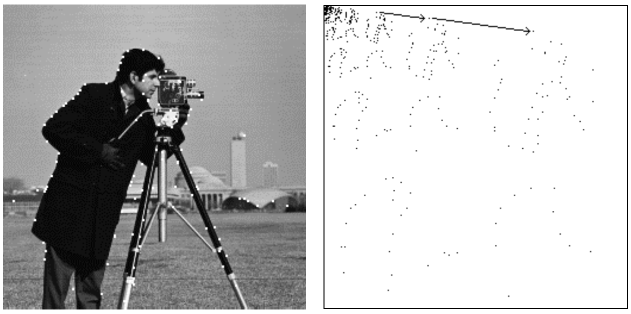
$$C\left(W_{2^j} f(n)\right) = \left\{W_{2^{j+1}} f(k),\ 2n \leq k \leq 2n + 2p - 1\right\},$$
$0 \leq n < 2^j N$ (N is the length of the signal).

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 higher wavelet coefficient at the scale 2^{j+1}, that is the maximum in absolute value of $C\left(W_{2^j} f(n)\right)$. In our salient point extraction algorithm, we consider this maximum, and look at its highest child. Applying recursively this process, we select a coefficient $W_{2^{-1}} f(n)$ at the finer resolution ½ (*cf.* Figure 4). Hence, this coefficient only represents $2p$ signal points. To select a salient point from this tracking, we choose among these $2p$ points the one with the highest gradient. We set its *saliency value* as the sum of the absolute value of the wavelet coefficients in the track:

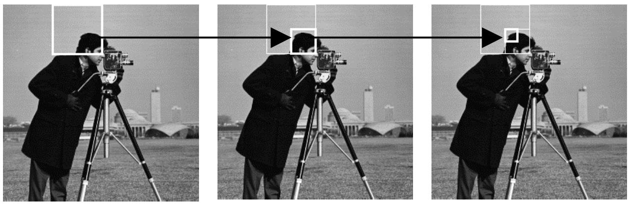
$$saliency = \sum_{k=1}^{-j} \left| C^{(k)}\left(W_{2^j} f(n)\right) \right|,\ 0 \leq n < 2^j N,\ -J_{\max} \leq j \leq -1 .$$

[2] For clarity we use one-dimensional signals. Extension to two dimensions and signals with length not restricted to a power of 2, in addition to algorithm complexity, are discussed in [16].

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 just need to threshold the saliency value, in relation to the desired number of salient. We first obtain the points related to global variations; local variations also appear if enough salient points are requested.



(a) 100 Haar salient points (b) Tracked coefficients



(c) Spatial support of tracked coefficients

Figure 4. Salient points extraction

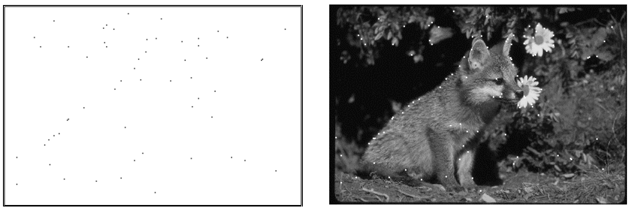
How to choose the wavelet for salient point extraction? Haar is the simplest wavelet function, so the fastest for execution. Some localization drawbacks can appear with Haar due to its non-overlapping wavelets at a given scale. This drawback can be avoided with the simplest overlapping wavelet, Daubechies 4. However, this kind of drawback is not likely in natural images.

3 Examples

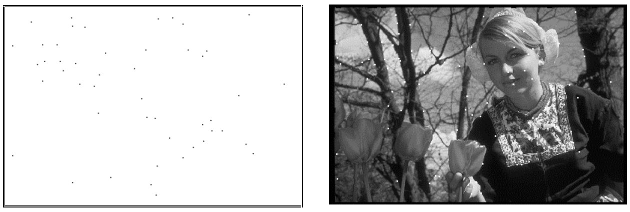
The salient points detected with the Haar transform are presented for the images used in Figure 1 and Figure 2 (*cf.* Figure 5). For each image the detected points are superimposed on the original image to evaluate salient points location.

Salient points are detected for smoothed edges (*cf.* Figure 5.a) and are not gathered in textured regions (*cf.* Figure 5.b). Hence they should lead to a more complete image representation than corner detectors. Similar behavior can be observed with Daubechies 4 wavelets.

Repeatability of the detection under typical alterations is a common evaluation criterion for corner detectors. Repeatability of our detector is comparable to other detectors. However this criterion may not be relevant in our context, because features stability is more important than geometric stability for image retrieval.



(a) 100 Haar salient points for the Fox image



(b) 100 Haar salient points for the Dutch image

Figure 5. Haar salient points examples

4 Evaluation for image retrieval

The best way to evaluate points detectors for image retrieval is to compare retrieval results obtained with each detector. The retrieval system is constituted by the *indexing* (points extraction and computation of local features to build image indexes) and the *querying* (based on a similarity measure between indexes).

Different retrieval systems and image databases are used in [16] to compare points detectors. Here we present results with an image retrieval system based on texture features [17]. Gabor features are computed for regions around the extracted points for different scales and orientations. Maximum amplitudes are used to build a set of histograms, which is the image index.

We use a database of 1505 various natural images. Each image belongs to an instinctive category (animals, flowers, landscapes, buildings, cities…). Very heterogeneous categories and images too different from the rest of the category are removed from the testset. Finally, we have a testset of 577 images in 9 classes.

We present the recall-precision graph, computed from different numbers of return images n. The system retrieves r images that belong to the same class C as the query ($r \leq n$). There are N_C images in the class C of the query. Then $P = r / n$ is the precision and $R = r / N_C$ the recall for this query. We use each testset image as a query, and use the average recall and precision for the graph (*cf.* Figure 6).

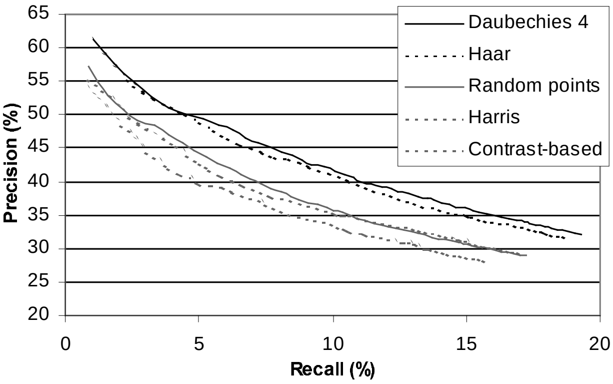


Figure 6. Retrieval results

We observe wavelet-based salient points perform better than other detectors for these features and this database. Random points are also used in the experiment: we randomly select points, and compute the Gabor features around these points. Their good result can be explained by their spreading in the image. For that reason they lead to a more complete representation of the image than some detectors. Obviously, the random points are very unlikely to be located in corners or edges point, but they are spread enough to represent these variations in the index. Good result of random points for indexing was observed with other databases and other local features [16]. These experiments show that the points spreading can be as important as the points location for image indexing (depending on the features). However, wavelet-based salient points, which are simultaneously spread and located, perform better than random points.

5 Discussion

We presented a salient point detector based on wavelets. The wavelet-based salient points are interesting for image retrieval, because they are located in many visual features (whether they are corner-like or not), without gathering in textured regions. We presented a retrieval experiment with Gabor features.

We used the Haar transform for point extraction, which is simple but may lead to bad localization. Daubechies wavelets avoid this drawback, but are not symmetric. Since orthogonality is not required in our approach, we could extend it to other wavelets that are compactly supported and symmetric.

Since points performance for indexing depends on the image database, detector choice for a specific database should be investigated, as well as random points relevance for local features extraction.

Wavelets are also attractive to extract image features for indexing. These local features would be more related to our salient points.

References

[1] B. Funt and G. Finlayson, "Color Constant Color Indexing", *IEEE Trans. on PAMI*, May 1995, Vol. 17, No. 5, pp. 522-529.

[2] F. Liu and R. Picard, "Periodicity, Directionality, and Randomness: Wold Features for Image Modeling and Retrieval", *IEEE Trans. on PAMI*, July 1996, Vol. 18, No. 7, pp. 722-733.

[3] C. Schmid and R. Mohr, "Local Grayvalue Invariants for Image Retrieval", *IEEE Trans. on PAMI*, May 1997, Vol. 19, No. 5, pp. 530-535.

[4] S. Bres and J.-M. Jolion, "Detection of Interest Points for Image Indexation", *3rd Int. Conf. on Visual Information Systems, Visual99*, Amsterdam, The Netherlands, June 2-4 1999, pp. 427-434.

[5] T. Tuytelaars and L. Van Gool, "Content-based Image Retrieval Based on Local Affinely Invariant Regions", *3rd Int. Conf. on Visual Information Systems, Visual99*, Amsterdam, The Netherlands, 2-4 June 1999, pp. 493-500.

[6] S. Bhattacharjee and T. Ebrahimi, "Image Retrieval Based on Structural Content", *Workshop on Image Analysis for Multimedia Interactive Services*, Heinrich-Hertz-Institut (HHI) Berlin, Germany, May 31 - June 1 1999.

[7] L. Kitchen and A. Rosenfeld, "Gray-Level Corner Detection", *Pattern Recognition Letters*, December 1982, Vol. 1, No. 2, pp. 95-102.

[8] C. Harris and M. Stephens, "A Combined Corner and Edge Detector", *Proc. of 4th Alvey Vision Conference*, 1988, pp. 147-151.

[9] Z. Zheng, H. Wang and E. Teoh, "Analysis of Gray Level Corner Detection", *Pattern Recognition Letters*, 1999, Vol. 20, pp. 149-162.

[10] S. Smith, J. Brady, "SUSAN - A New Approach to Low Level Image Processing", *International Journal of Computer Vision*, May 1997, Vol. 23, No. 1, pp. 45-78.

[11] B. Zitova, J. Kautsky, G. Peters and J. Flusser, "Robust Detection of Significant Points in Multiframe Images", *Pattern Recognition Letters*, 1999, Vol. 20, pp. 199-206.

[12] C.-H. Chen, J.-S. Lee and Y.-N. Sun, "Wavelet Transformation for Gray-level Corner Detection", *Pattern Recognition*, 1995, Vol. 28, No. 6, pp. 853-861.

[13] S. Mallat, "A Theory for Multiresolution Signal Decomposition : The Wavelet Representation", *IEEE Trans. on PAMI*, July 1989, Vol. 11, No. 7, pp. 674-693.

[14] E. Stollnitz, T. DeRose and D. Salesin, "Wavelets for Computer Graphics: A Primer, part 1", *IEEE Computer Graphics and Applications*, May 1995, Vol. 15, No. 3, pp. 76-84.

[15] I. Daubechies, "Orthonormal bases of compactly supported wavelets", *Communications on Pure and Applied Mathematics*, 1988, Vol. 41, pp. 909-996.

[16] E. Loupias and N. Sebe, "Wavelet-based Salient Points for Image Retrieval", *RR*, Laboratoire Reconnaissance de Formes et Vision, INSA Lyon, 1999. On-line http://rfv.insa-lyon.fr/~loupias/points/

[17] C. Wolf, "Content based Image retrieval using Interest Points and Texture Features", *RR 99.09*, Laboratoire Reconnaissance de Formes et Vision, INSA Lyon, 1999.