

Salient Points for Content Based Retrieval

N.Sebe¹

Q.Tian²

E.Loupias³

M.S.Lew¹

T.S.Huang²

¹Leiden Institute of Advanced
Computer Science,
Leiden, The Netherlands
{nicu, mlew}@liacs.nl

²Beckman Institute,
University of Illinois
at Urbana-Champaign, USA
{qitian, huang}@ifp.uiuc.edu

³Laboratoire 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. We show that extracting color and texture information in the locations given by our salient points provides significantly improved results as compared to global feature approaches. Moreover, our salient points lead to a more complete image representation than using corners.

1 Introduction

In a typical content-based image database retrieval application, the user has an image he or she is interested in and wants to find similar images from the entire database. First, for each image in the database, a feature vector characterizing some image properties is computed and stored in a feature database. Second, given a query image, its feature vector is computed, compared to the stored feature vectors, and images most similar to the query are returned to the user.

In general, the features are computed from the entire image. However, these global features cannot handle all parts of the image having different characteristics. Therefore, local computation of image information is necessary. Local features can be computed at different image scales to obtain an image index based on local properties of the image and they need to be sufficiently discriminative to "summarize" the local image information. These features are too time-consuming to be computed for each pixel in the image and therefore, the feature extraction should be limited to a subset of the image pixels, the interest points [2, 4], where the image information is supposed to be the most important. Besides saving time in the indexing process, these points may lead to a more discriminative index because they are related to the visually most important parts of the image.

Schmid and Mohr [2] proposed the use of corners as interest points in image retrieval. They compared different corner detectors [3] and showed that the best results were provided by the Harris corner detector [1]. Corner detectors, however, were designated for robotics and shape recognition and 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. Therefore, we aim for a set of interesting points called *salient points* [5] that are related to any visual interesting part of the image whether it is smoothed or corner-like.

We believe multi-resolution representation is interesting to detect salient points. We present a salient point extraction algorithm that uses the wavelet transform, which expresses image variations at different resolutions. Our wavelet-based salient points lead to a more complete image representation than corner detectors.

2 Wavelet-based Salient Points

The wavelet representation gives information about the variations in the image at different scales. We would like to extract salient points from any part of the image where something happens 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.

Our salient point extraction scheme is as follows:

- From each wavelet coefficient at a coarse level, trace the corresponding coefficients in the finer level.
- Extract the maximum coefficient and trace it recursively in the finer levels (see Figure 1).
- For each image pixel, compute the saliency value as the sum of the coefficients along the trace(s) ended in this point.
- Threshold to extract the most prominent salient points.

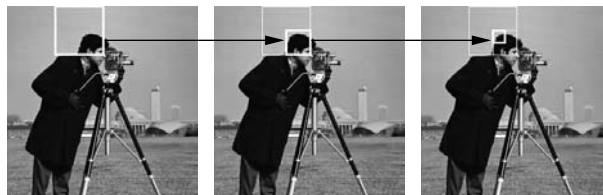


Figure 1: Salient points extraction: spatial support of tracked coefficients

The salient points extracted by this process depend on the wavelet we use. Haar is the simplest wavelet function,

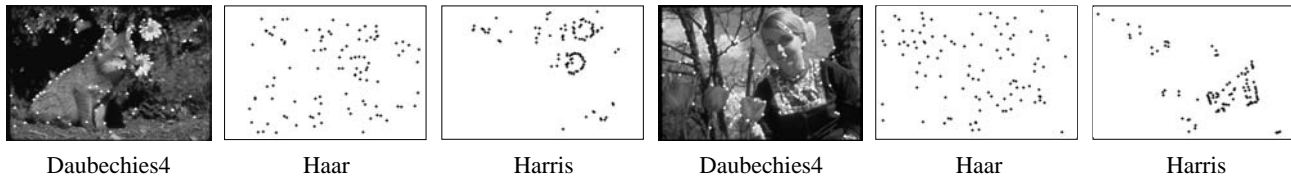


Figure 2: Saliency points examples. For Daubechies4 and Haar saliency points are detected for smooth edges (fox image) and are not gathered in textured regions (girl image).

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 saliency points extracted using Daubechies4, Haar, and Harris detectors are shown in Figure 2. Note that while for Harris the corners lead to an incomplete image representation, for the other two detectors the saliency points are detected for smooth edges and are not gathered in texture regions. Hence, they lead to a more complete image representation. An interactive demo is running at URL: <http://www.ifp.uiuc.edu/~nicu/demo.html>.

3 Content-based retrieval

We are interested in using the saliency points in a retrieval scenario. We consider a modular approach: the saliency points are first detected for each image in the database and then feature vectors are extracted from a small neighborhood around each saliency point. This approach assures the independence of the saliency point extraction techniques and the feature extraction procedure and gives the user the liberty to use any features he wants for a specific application.

In our implementation the user can choose the desired saliency point extraction technique (Haar, Daubechies4, Harris), the features to be used in retrieval (color moments, Gabor, and wavelet moments), and the way the features should be extracted (around saliency points or over the entire image). Moreover, relevance feedback is also supported. An example query is shown in Figure 3.

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, $S_j(f_i)$:

$$D_j = \sum_i W_i S_j(f_i) \quad j = 1, \dots, N \quad (1)$$

where N is the total number of images in the database and $S_j(f_i)$ is defined as:

$$S_j(f_i) = (\mathbf{x}_i - \mathbf{q}_i)^T (\mathbf{x}_i - \mathbf{q}_i) \quad (2)$$

where \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. (1) are set to be equal.

4 Summary

In this paper we show that extracting color and texture information in the locations given by our saliency points provides significantly improved results as compared to global



Figure 3: User Interface: An experimental result using the color moments extracted from the 3×3 neighborhood of the Daubechies4 saliency points (rank from the left to right and from top to bottom, the top left is the query image)

feature approaches. Our saliency points are interesting for image retrieval because they are located in visual focus points and therefore, they capture the local image information better than the corner extractors.

References

- [1] C. Harris and M. Stephens. A combined corner and edge detector. *Alvey Vis Conf*, pages 147–151, 1999.
- [2] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *IEEE Trans on Patt Anal and Mach Intell*, 19(5):530–535, 1997.
- [3] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *IJ Comp Vis*, 37(2):151–172, 2000.
- [4] N. Sebe, Q. Tian, E. Louprias, M.S. Lew, and T.S. Huang. Color indexing using wavelet-based saliency points. In *IEEE Workshop on Content-based Access of Image and Video Libraries*, pages 15–19, 2000.
- [5] Q. Tian, N. Sebe, M. S. Lew, E. Louprias, and T. S. Huang. Image retrieval using wavelet-based saliency points. *to appear in Journal of Electronic Imaging*, October 2001.