Wavelet-Based Queryng in Images Database

Mitko Kostov, Mile Petkovski and Ilija Jolevski[1]

***Abstract* – In this paper we present ways of organizing images into a database in order to be able to make fast querying. Multiresolution technique is used and the most important wavelet detail coefficients are selected to compose images pseudo-hashes. When searching for an image-query in the database, the pseudo-hashes of the images are considered instead of the images themselves.**

***Keywords* – Wavelets, multiresolution, images database, pseudo-hash, query.**

I. Introduction

Wavelet transforms have received significant attention recently from mathematicians, signal analysts and engineers as a new tool for feature extraction, signal and image compression, edge detection and denoising. Unlike the traditional Fourier techniques, wavelets are localized both in time and frequency domain. This feature makes them suitable for the analysis of nonstationary signals.

This paper considers a practical implementation of the wavelet transform for a fast searching in an images database. Images are decomposed in a few levels. Pseudo-hash information calculated from the images themselves are stored in a database. When searching for an image-query, its pseudo-hash is compared to pseudo-hashes from the database by using a simple sql statement. Selected images-candidates from the database match some defined criteria. This image processing system can be a good solution for small robots with restricted resources. Processing of images rather than comparing complete images is useful for robots incorporating visual servoing as an approach to guide robots using visual information.

The paper is organized as follows. The wavelet theory is summarized in Section 2. Section 3 presents algorithms for fast querying in an images database. The experimental results are presented in Section 4. Section 5 concludes the paper.

II. Wavelet Theory

The Discrete Wavelet Transform (DWT) decomposes a signal into a set of orthogonal components describing the signal variation across the scale [1]. The orthogonal components are generated by dilations and translations of a prototype function ψ, called mother wavelet.

In analogy with other function expansions, a function f is presented for each discrete coordinate t as a sum of a wavelet expansion up to certain scale J plus a residual term, that is:

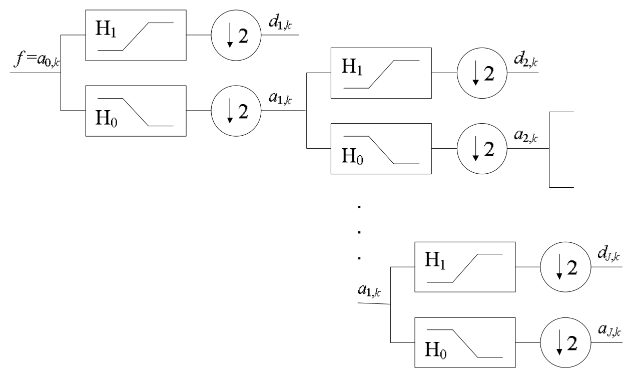


Fig. 1. Discrete wavelet transform tree.

$$f(t)=\sum_{j=1}^{J}\sum_{k=1}^{2^{-j}M} d_{jk}\psi_{jk}(t)+\sum_{k=1}^{2^{-J}M} a_{Jk}\phi_{Jk}(t) \qquad (1)$$

where ψ_{jk} and ϕ_{jk} denote wavelet and scaling function, respectively, the indexes j and k are for dilatation and translation, and a_{Jk} and d_{jk} are approximation and detail coefficients.

Wavelet decompositions and multiresolution concepts are closely related to filter bank theory. For this reason, it is helpful to view the scaling and wavelet function as a low pass and high pass filters, $\mathbf{H_0}$ and $\mathbf{H_1}$, respectively. The wavelet transform is applied to low pass results (approximations) as it is illustrated in Fig. 1.

The most popular form of conventional wavelet-based signal filtering [2], can be expressed by:

$$\begin{aligned}&\{\mathbf{A}^{(k)},\mathbf{D}^{(1)},\mathbf{D}^{(2)},\cdots,\mathbf{D}^{(k)}\}=\mathrm{DWT}(\mathbf{s}),\\&\mathbf{s}^{*}=\mathrm{IDWT}\left(f\left(\mathbf{A}^{(k)},\mathbf{h}^{(1)}.\times\mathbf{D}^{(1)},\mathbf{h}^{(2)}.\times\mathbf{D}^{(2)},\cdots,\mathbf{h}^{(k)}.\times\mathbf{D}^{(k)}\right)\right)\end{aligned} \qquad (2)$$

where $\mathbf{s}$ is input signal, $\mathbf{s}^*$ is filtered signal, $\mathbf{A}^{(k)}$ and $\mathbf{D}^{(k)}$ are approximation and detail coefficients at level k, respectively, f is a function of the modified detail and approximation coefficients, $.\times$ is element-by-element multiplying and

$$\mathbf{h}^{(k)}=\left[h_1^{(k)},h_2^{(k)},\cdots,h_j^{(k)}\right]^T \qquad (3)$$

are weighting coefficients of the corresponding detail coefficients at level k.

In case of conventional hard threshold filtering the weighting coefficients are

$$h_j^{(k)}(hard)=\begin{cases}1, & \text{if } \left|D_j^{(k)}\right|>\tau^{(k)}\\0, & \text{otherwise}\end{cases}, \qquad (4)$$

[1]Mitko Kostov, Mile Petkovski and Ilija Jolevski are with the Faculty of Technical Sciences, I.L.Ribar bb, 7000 Bitola, Macedonia, E-mails: mitko.kostov@uklo.edu.mk, mile.petkovski@uklo.edu.mk, ilija.jolevski@uklo.edu.mk.

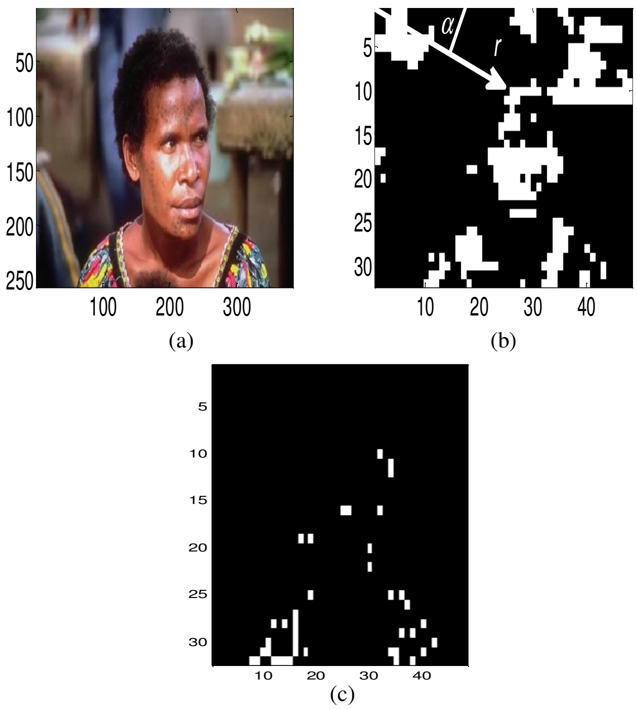


Fig. 2. (a) Image with resolution 384x256; (b) and (c) The most important wavelet approximation/detail coefficients at level 3.

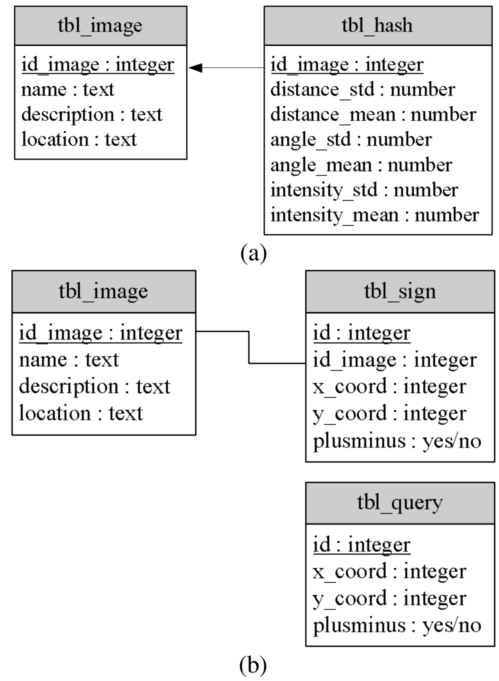


Fig. 3. Relation schemas in the images database.

while for the soft threshold filtering they are

$$h_j^{(k)}(soft) = \begin{cases} 1 - \dfrac{\tau^{(k)} \operatorname{sgn}\left(D_j^{(k)}\right)}{D_j^{(k)}}, & \text{if } \left|D_j^{(k)}\right| > \tau^{(k)} \\ 0, & \text{otherwise} \end{cases}, \quad (5)$$

where $\tau^{(k)}$ is user specified threshold for the *k*-th level details.

III. PSEUDO-HASH

The main idea is to search for a particular image-query in a large images database on the basis of a small piece of information calculated from the images themselves [3]; in the following text it is referred as pseudo-hash information. By comparing the similarity between the images' pseudo-hashes stored in the database and the image-query's pseudo-hash, few images (candidates) are selected to be considered visually if some of them correspond to the image-query.

Pseudohash is based on the wavelet transform. Namely, the wavelet transform tends to concentrate the energy of a signal into a small number of coefficients, while a large number of coefficients have small energy. By applying a threshold given with (4) the most important wavelet coefficients are selected and used in creating pseudo-hash of an image.

A. Pseudo-hash based on approximation coefficients values

RGB images are converted to YCbCr colour space, where Y is the luminance (intensity) component and Cb (blue chrominance) and Cr (red chrominance) are the blue-difference and red-difference chroma components, respectively. The Y components are taken into consideration and wavelet transform is applied. The most important approximation coefficients at a certain level are selected and for these coefficients, three variables are defined: distance, angle and intensity, as it is illustrated in Fig. 2b. An image with resolution 384x256 is shown in Fig. 2a and the most important wavelet approximation coefficients calculated with haar wavelet at level 3, are shown in Fig. 2b. For the three variables (distance, angle and intensity) both mean value and standard deviation are calculated and they compose the pseudohash information. Hence, an image pseudohash is consisted from six values: mean value and standard deviation for distance, angle and intensity of the non-zero pixels.

A database that keeps pseudohash information for the images contains a few relations with their schemas given in Fig. 3a. In the relation tbl_image, the attribute id_image is the primary key. This relation contains description of the images: name, description and location (if the images are picture files in the file system). The relation tbl_hash contains pseudohash information for each image. The primary key, id_image, at the same time is foreign key that takes its values from the primary key of the relation tbl_image.

Next, the algorithm is extended to search for circularly shifted images by involving Fourier transform, which is time/space invariant. Instead of calculating the wavelet coefficients from the Y component of the original images, the Fourier transform is applied to the Y components of all the images which eliminates the space information, then magnitudes of the Fourier coefficients are calculated, and at the end the inverse Fourier transform is applied. These modified images are used to calculate wavelet coefficients and

Fig. 4. Part of images database with 1000 images used for experiments.

pseudohash information as it is described in the basic algorithm.

B. Pseudo-hash based on coefficients coordinates

RGB images are first converted to YCbCr colour space and the pseudo-hashes are calculated from the most important wavelet detail coefficients of the Y components at a certain level. Then the wavelet detail coefficients are filtered in order to keep only the most important coefficients (Fig. 2c). Both the position and sign of the coefficients make up the pseudo-hash information for an image: a set of triplets (x, y, s), where (x, y) denotes a coefficient position, and s is the coefficient sign (+/–).

A database that keeps images information can contain a few relations with their schemas given in Fig. 3b. The relation *tbl_image* contains information like name, description and location of the images (if the images are picture files in the file system). In this relation, the primary key is the attribute *id_image*. The relation *tbl_sign* contains the pseudo-hash information for all the images. The primary key is *id*, while *id_image* is foreign key that takes its values from the primary key of the relation *tbl_image*.

The relation *tbl_query* contains the pseudo-hash calculated from the image-query by applying the same algorithm. A list of images candidates can be obtained by comparing the matching results from the tables *tbl_query* and *tbl_sign*. The image with a maximum number of matching points (positions and signs of the most important wavelet coefficients) is likely the image that is looked for.

IV. Experimental Results

In this Section, our experimental results are explained. The experiments for fast searching are carried out with 1000 images which pseudo-hashes are stored in Microsoft Access 2007 databases with schemas given in Fig. 2. Some of these images are shown in Fig. 4. The database contains a lot of similar images with people, animals, landscapes, objects, etc. The database does not contain the images themselves; the images are picture files in the file system.

A. Pseudo-hash based on approximation coefficients values

All these images are converted in YCbCr colour space, and the Fourier transform is applied over their Y components. Next, inverse Fourier transform is applied to the magnitudes of the Fourier coefficients. The obtained images from these successive operations of Fourier transform and inverse Fourier transform over the image from Fig. 2a are shown in Fig. 5a and Fig. 5b.

The haar wavelet transform in three levels is applied over the images obtained with inverse Fourier transform. The most important 20% pixels from the wavelet approximation coefficients at the third level are kept (Fig. 2b). It can be noticed that resolution of the wavelet coefficients at this level is 48x32, which means that only a few coefficients are taken into consideration for the calculation the pseudohash. For the image in Fig. 2, the number of non-zero pixels after applying wavelet threshold is only 307. These coefficients are normalized so the maximum intensity is 255.

For the selected wavelet coefficients, mean value and standard deviation are calculated for the variables defined with distance, angle and intensity and they are stored in the relation tbl_hash (Fig. 3a).

Table I shows pseudohash data for a part of the images from our database. The all attributes domains are integer in order to save space. The value of the attribute id_image for the image from Fig. 2a is 1.

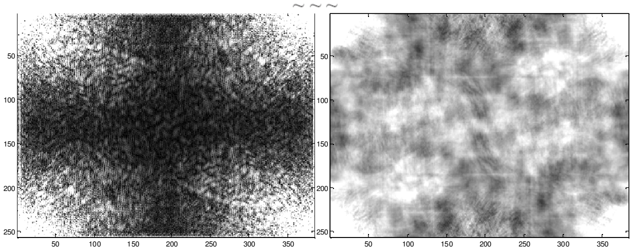
In the querying process, the same algorithm is applied over an image-query and a simple SQL SELECT statement is used to select the image from the database which pseudohash values corresponds to pseudohash data calculated from the image-query.

In addition, if the image-query is a circularly shifted version of an image in the database (Fig. 6a), it has the same pseudohash data as the original image due to the using of the

TABLE I
PSEUDO-HASH DATA ACCORDING TO III.A

id_ image	dist_ std	dist_ mean	ang_ std	ang_ mean	int_ std	int_ mean
0	16	32	27	50	24	11
1	14	33	28	39	22	10
2	15	33	28	37	21	9
3	16	32	24	57	26	12
4	14	32	25	33	25	11
5	17	33	28	41	22	10
6	15	33	27	39	19	8
7	17	32	27	36	22	9
8	18	33	30	44	24	11
9	14	32	26	39	20	8
10	15	31	24	35	22	10
11	15	33	29	53	24	11
12	17	33	27	39	21	9
13	15	32	28	48	19	8
14	17	32	28	38	28	13
15	18	33	30	44	23	10

~ ~ ~



(a) (b)

Fig. 5. (a) Magnitude of the Fourier transforms over the images in Fig. 1a and Fig. 6a; (b) Inverse Fourier transform from Fig. 5a.

TABLE II
RESULTS OF SEARCHING AN IMAGE ACCORDING TO III.B

[1]	[77]
[216]	[14]
[790]	[9]
[747]	[8]
[142]	[8]
[134]	[8]
[126]	[8]
[705]	[8]
[704]	[7]
[118]	[7]
[113]	[7]
[803]	[7]
[833]	[7]
[764]	[7]
[56]	[7]
[766]	[7]
[190]	[7]

(a) (b)

Fig. 6. (a) Circularly shifted version of the image from Fig. 2a; (b) Image with missing pixels.

Fourier transform. Moreover, if the image-query misses some pixels or parts of the image-query miss as it is shown in Fig. 6b, its pseudohash data differs slightly from the pseudohash of the original. The original image from the database still can be selected by loosening the criterion in the sql statement and involving the operator 'between' in the 'where' clause to select a range of data between two values. The calculated pseudohash data for the image from Fig. 6b is (13 33 28 39 22 10), what is very similar to the pseudohash for the image with id_image=1 (Fig. 2). The missing pixels in the image in Fig. 6b are not estimated.

B. Pseudo-hash based on coefficients coordinates

All these images are converted in YCbCr colour space. The haar wavelet transform in three levels is applied over the Y components. The most important 5% pixels from the wavelet detail coefficients at the third level frequency region (3, 15) are kept (Fig. 2c). Similar results can be obtained by taking into consideration coefficients from other frequency region or from all the frequency regions.

Next, as a query-image we use the image shown in Fig. 2a, which image resolution is 256x384. The database already contains pseudo-hash information for this image in the table *tbl_sign* and the $id = 1$. Its pseudo-hash is calculated from the coefficients shown in Fig. 2c.

Next step is comparing the image-query pseudo-hash with the pseudo-hashes in the database. A simple SQL SELECT statement is used to select all the images from the database which pseudo-hash values correspond to pseudo-hash data calculated from the image-query. The result of matching points between the tables *tbl_sign* and *tbl_query* is shown in the Table II. It can be seen that the number of matching points (matched coefficients positions and signs) is highest for the image with $id = 1$ (77 matched points). The nearest image-candidate is the image with $id = 216$ (14 matched points).

V. CONCLUSION

In this paper two algorithms for fast searching in an images database are presented. The searches are based on calculating pseudo-hash information for all the images in the database. For this reason the wavelet transform is exploited in order to select the most important wavelet coefficients, which are used to calculate the pseudo-hashes. This can be used to control the motion of a robot depending on the visual information extracted from the images captured by one or several cameras.

REFERENCES

[1] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley-Cambrige Press, 1996.

[2] D. L. Donoho, "Wavelet Thresholding and W.V.D.: A 10-minute Tour", *Int. Conf. on Wavelets and Applications*, Toulouse, France, June 1992.

[3] C. E. Jacobs, A. Finkelstein, D. H. Salesin, "Fast multiresolution image querying", *SIGGRAPH '95 Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*.