TEM Journal. Volume 6, Issue 1, Pages 143-146, ISSN 2217-8309, DOI: 10.18421/TEM61-21, February 2017.

MRA Based Efficient Database Storing and Fast Querying Technique

Mitko Kostov [1], Elena Kotevska [1], Metodija Atanasovski [1]

[1] *Faculty of Technical Sciences, St. Kliment Ohridski University, Bitola, Macedonia*

Abstract **– In this paper we consider a specific way of organizing 1D signals or 2D image databases, such that a more efficient storage and faster querying is achieved. A multiresolution technique of data processing is used in order of saving the most significant processed data.**

Keywords **– multiresolution, wavelet coefficients, threshold, database.**

1. Introduction

The property of wavelets to localize both time and frequency makes them very suitable for analysis of non-stationary signals [1]. They are an excellent tool for feature extraction, signal and image compression, edge detection and compression.

This paper considers a practical implementation of wavelet transformation for content-based signal/image retrieval from a database. Instead the signals/images itself, we place their wavelet coefficients in the database and this gives us a basis for creating an application for fast querying the database. Actually, content-based image retrieval (CBIR) emerged in 1980s [2]. This approach for searching an image database in which a query is expressed as a low-resolution image is known as query by content [3, 4, 5]. The paper is organized as follows. After the introduction, in section 2 we give the basic definitions of wavelet transform. Section 3 describes the organization of database and Section 4 presents our experimental results.

DOI: 10.18421/TEM61-21
https://dx.doi.org/10.18421/TEM61-21

Corresponding author: Mitko Kostov, Faculty of Technical Sciences, St. Kliment Ohridski University, Bitola, Macedonia
Email: mitko.kostov@uklo.edu.mk

2. Discrete Wavelet Transformation

2.1. *Definition*

Discrete wavelet transform (DWT) decomposes the signal into a set of orthogonal components describing the signal variation across the scale [6]. The orthogonal components are generated by dilations and translations of a prototype function, called the mother wavelet.

$$\psi_{i,k}(t) = 2^{-i/2}\psi(t/2^i - k), \quad k,i \in Z \tag{1}$$

The last equation means that the mother function is dilated for by integer value i and translated by an integer value k. A function f for each discrete coordinate t can be presented as a sum of a wavelet expansion up to certain scale J plus a residual term

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

The estimate of d_{jk} and a_{Jk} can be achieved via iterative algorithm for decomposition using two complementary filters h_0 (low-pass) и h_1 (high-pass) [7]. This is illustrated in Figure 1. for 1D DWT and 2D DWT.

2.2. *Wavelet Shrinkage*

Having in mind that DWT has a tendency to concentrate the energy of a signal into few coefficients, while large number of coefficients have small SNR, the wavelet coefficients d_{jk} are modified by

$$\hat{d}_{jk} = d_{jk} \cdot h_k \tag{3}$$

where filter h_k describes "hard" or "soft" threshold filtering, known also as wavelet shrinkage [8]:

$$h_{jk}^{(\text{hard})} = \begin{cases} 1, & \text{ако } |d_{jk}| \geq \tau_j \\ 0, & \text{ако } |d_{jk}| < \tau_j \end{cases} \quad \text{or}$$

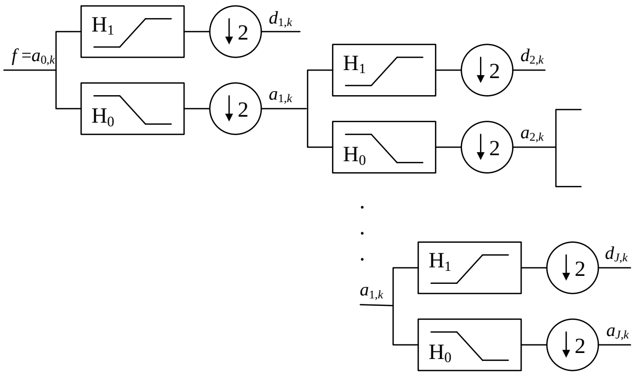


Figure 1. Discrete wavelet transform tree.

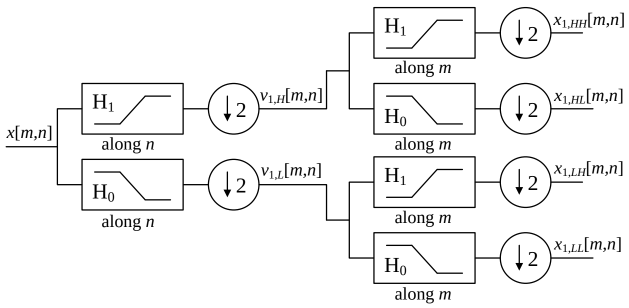


Figure 2. Schematic diagram of 2D wavelet transform.

$$h_{jk}^{(\mathrm{soft})} = \begin{cases} 1 - \dfrac{\tau_j \operatorname{sgn}(d_{jk})}{d_{jk}}, & \text{ако } |d_{jk}| \geq \tau_j \\ 0, & \text{ако } |d_{jk}| < \tau_j \end{cases} \tag{4}$$

2.3. Wavelet Packets

Wavelet packet transform (WPT) is a kind of extension of DWT providing a more flexible tool for time-frequency data analysis. WPT can be applied to both approximation and details.

3. Efficient Storage

In this section we present a way of more efficient storage of signals/images in a single database. Instead of storing the original signals in time or frequency domain, in the database we store their wavelet coefficients calculated at different scales. WT has a tendency of grouping the coefficients in two groups: one smaller group of coefficients with great energy and one larger group with coefficients with smaller energy[9]. Moreover, if we apply a threshold given in (4), then in the new multiresolution version of the signal, a large number of pixels will have zero value and as such need not to be stored in the database. By adjusting the threshold we can control the number of nonzero pixels.

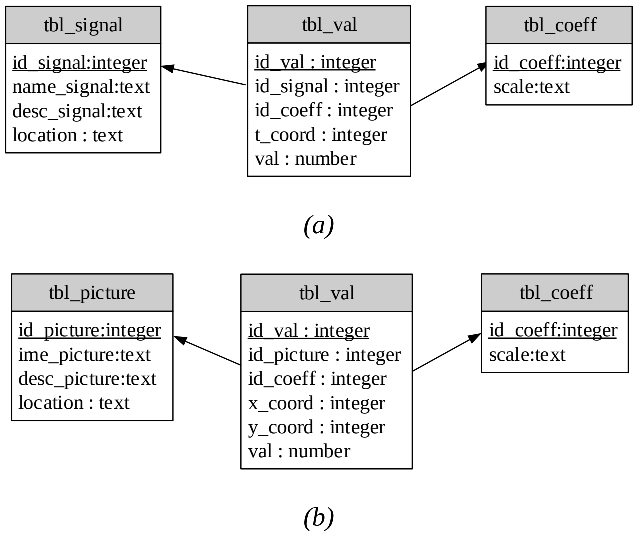


Figure 3. Relations schemes in a database for storing: (a) 1D signal, (b) 2D picture.

3.1. Organization of the database

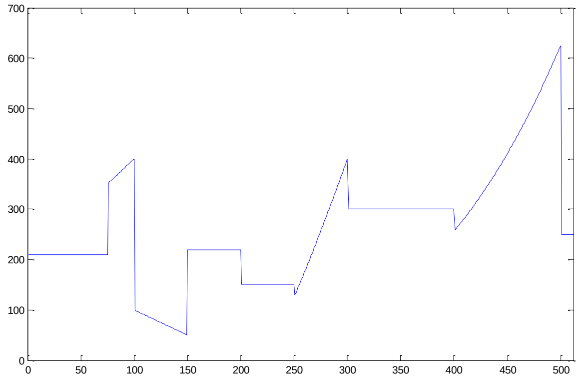
The structure of the proposed database for storing a 1D signal would include several relations with schemes given in Figure 3a.

In the relation tbl_signal, the primary key is the attribute id_signal and the relation includes description of the signals. The relation tbl_coeff with primary key id_coeff includes the types of used wavelet coefficients (approximation coefficients of scale 1,2,… and detail coefficients of scale 1, 2, ...). The relation tbl_val contains the values of nonzero wavelet coefficients of signals, resulting from threshold application. The primary key in this relation is the attribute id_val, while attributes id_signal and id_val are foreign keys pointing to the primary keys from relations tbl_coeff and tbl_val. In the case of database for storing 2D images, the database would be of a structure shown in the Figure 3b.

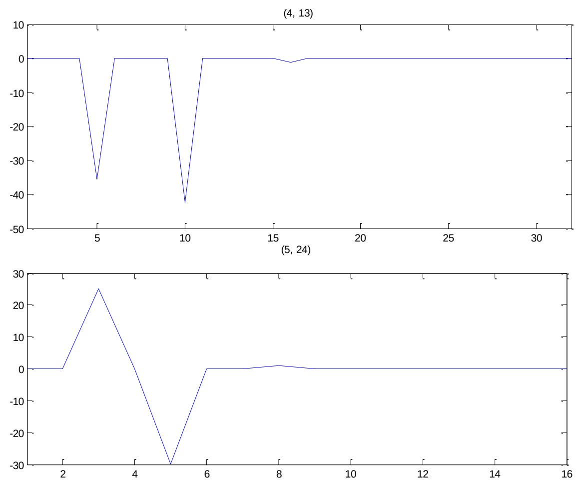
The advantage of such database organization is saving the memory space, but also the possibility for fast data querying. There are two ways of using this type of organization of a database in order to achieve fast querying as it follows.

3.2. Fast querying by wavelet coefficients comparison

This approach compares a given number of randomly chosen approximation wavelet coefficients from a given scale with the corresponding wavelet coefficients of the image we are querying. This



(a)



(b)

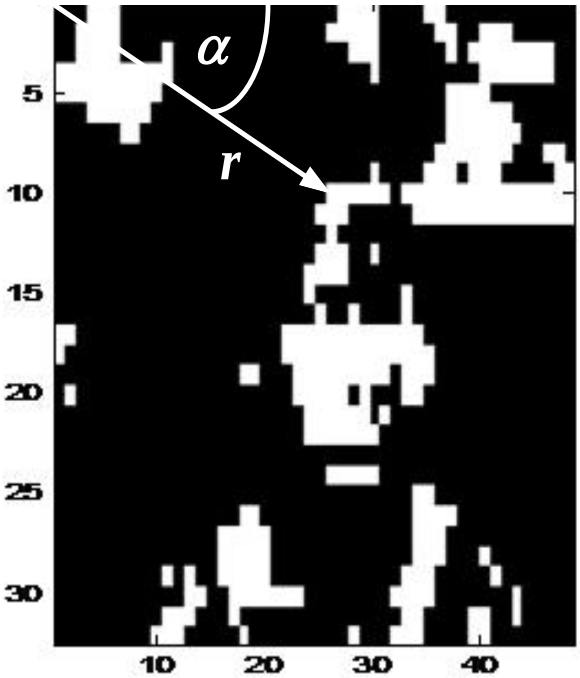
Figure. 4 (a) 1D Test function; (b) Wavelet packet decomposition at fourth and fifth level.

would result in elimination of large number of images from database and selecting a small number of candidates for the next iteration with from a finer scale. The procedure is applied over and over, until we have a few images that can be visually checked.

3.3. Fast querying by calculating a pseudohash values

This approach calculates a pseudohash sequence of values achieved from the most significant wavelet coefficients of a certain approximation scale *n*. To achieve this for every kept nonzero pixel three variables can be defined: distance, angle and intensity. Then the dispersion and/or mean value for each variables should be calculated. This would result in creation of one more relation, for example tbl_hash (id_picture : integer, dist_std : number, dist_mean : number, ang_std : number, ang_mean : number, val_std : number, val_mean : number).

(a)



(b)

Figure. 5 (a) 1D Test function; (b) Wavelet packet decomposition at fourth and fifth level.

4. Experimental Results

In order to demonstrate the proposed efficient storing of signals in a database, we chose 1D test function from Figure 4. It consists of few segments. We applied WPT by using the Haar wavelet in fivescales. The larger number of details wavelet coefficients have values very close to zero. This is illustrated in Figure 4. Applying the threshold defined by (4) we can keep the smaller part of coefficients (Figure 4b), meaning we can store data more efficiently. The experiments for fast querying in a database are made on one database including 1000 pictures. One of these pictures is given in Figure 3a. For each of these pictures converted in YCbCr colour space, haar WT in three levels is applied over the Y luminance component. We keep only 20% of the most significant pixels from the achieved approximation wavelet coefficients at third level. Approximation wavelet coefficients at third level of the Figure 5a are given in Figure 5b. It can be noticed that the resolution of wavelet coefficients at this level is 48x32, i.e., which means that only a few coefficients are taken into consideration for the calculation of the pseudohash. For this rough approximation for each picture we calculated the mean value and standard deviation of variables defined by distance, angle and intensity of the pixels kept. They are stored in the relation tbl_hash. In the querying process, the same algorithm is applied over an image-query and its pseudohash data is calculated from its Y component. 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.

5. Conclusion

In this paper we proposed a way for an efficient organization of signals/pictures in a database in order not only to save memory space, but also to enable fast data querying. For this purpose, a multiresolution technique is used in order to decompose signals, and to calculate pseudohash from the defined variables.

References

[1]. Vetterli M, Kovacevic J. *Wavelets and subband coding*. Prentice-Hall; 1995.

[2]. Liu Y, Zhang D, Lu G, Ma WY. “A survey of content-based image retrieval with high-level semantics”,*Pattern recognition* 40 no. 1,pp. 262-82, Jan. 2007.

[3]. Barber, R., W. Equitz, W. Flickner, W. Niblack, D. Petkovic, P. Yanker. "Efficient query by image content for very large image databases",*Compcon Spring'93, Digest of Papers,* pp. 17-19. IEEE, 1993.

[4]. 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*.

[5]. Wang, J.Z., Wiederhold, G., Firschein, O. and Wei, S.X., "Content-based image indexing and searching using Daubechies' wavelets",*International Journal on Digital Libraries 1.4*,pp. 311-328, 1998.

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

[7]. P.P.Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice – Hall (1992).

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

[9]. DeVore, Ronald A., Björn Jawerth, and Bradley J. Lucier. "Image compression through wavelet transform coding." *IEEE Transactions on information theory* 38.2 pp 719-746, 1992.