An entropy-based approach for shape description

1st Vittoria Bruni
University of Rome La Sapienza/
IAC - CNR
Rome, Italy
vittoria.bruni@sbai.uniroma1.it

2nd Lorenzo Della Cioppa
University of Rome La Sapienza
Dept. of SBAI
Rome, Italy
lorenzo.dellacioppa@sbai.uniroma1.it

3rd Domenico Vitulano
National Research Council
IAC
Rome, Italy
d.vitulano@iac.cnr.it

***Abstract*—In this paper an automatic method for the selection of those Fourier descriptors which better correlate a 2D shape contour is presented. To this aim, shape description has been modeled as a non linear approximation problem and a strict relationship between transform entropy and the sorted version of the transformed analysed boundary is derived. As a result, Fourier descriptors are selected in a hierarchical way and the minimum number of coefficients able to give a nearly optimal shape boundary representation is automatically derived. The latter maximizes an entropic interpretation of a complexity-based similarity measure, i.e. the normalized information distance. Preliminary experimental results show that the proposed method is able to provide a compact and computationally effective description of shape boundary which guarantees a nearly optimal matching with the original one.**

***Index Terms*—Shape representation, Fourier descriptors, non linear approximation, differential entropy, normalized information distance (NID)**

I. Introduction

In the last years, automatic contour-based object analysis and recognition is receiving an even more increasing interest, especially due to the rapid increase of multimedia information. The huge amount of visual information requires effective content description for storage, retrieval and processing purposes, just to mention some applications. Shape is one of the key features for describing an object, along with texture and color. A shape descriptor can be defined as a set of numbers which gives an accurate, compact and discriminative shape representation, allowing straightforward comparisons. A shape descriptor is also required to be low computationally demanding. Shape descriptors can be broadly classified in: volume-based and contour-based. The former involve all pixels of the object; geometric [7] and Zernike moments [11] belong to this class. Contour based descriptors only use shape boundary, as, for example, curvature scale space [10], [14], active shape models [3], hierarchical skeletons [18] and Fourier descriptors [15], [20]. The latter can be considered global descriptors, since they describe the whole contour; on the contrary, local descriptors focus on its local relevant features. Local descriptors play a fundamental role in the search of correspondences between objects, feature detection, registration, segmentation and labeling; global descriptors are involved in object retrieval, classification, recognition and clustering. Fourier descriptors are the coefficients of the discrete Fourier transform of some representation of a closed contour, often denoted by signature. The shape signature is a one dimensional function which is derived from shape boundary coordinates, such as complex coordinates, curvature, centroid distance, cumulative angular function, and so on. Independently of their definition or the adopted signature [5], [7]–[9], [17], [19], the main aim in defining effective descriptors is to have invariance properties under transformations like translation, scaling, and rotation [7]. However, the recent literature has shown the need of defining descriptors able to have both local and global properties, especially for applications in fields like medicine, biology, cultural heritage [1], [18]. That is why methods for combining properly such kind of descriptors have been recently proposed [16]. In this paper, we look at the shape description as a non linear approximation problem. Among all the existing descriptors, we focus on Fourier descriptors, as they allow a straightforward presentation of the proposed method. Fourier description commonly consists of the expansion of the shape signature in a Fourier basis and retaining the first N coefficients, resulting in a linear approximation problem [13]. As the literature on this topic proved, this kind of approximation cannot be effective in the sense of sparsity or compactness, since the N vectors are selected a priori, without any information about the function under study. On the contrary, non linear approximation selects the N elements that better correlate the signal, i.e. its main features, meaning that even very high frequency content may be selected if relevant for the representation. Non linear approximation is thus able to guarantee compactness of the representation and faster convergence to the original signal with respect to L^2 norm. As a result, an hierarchical definition of the set of Fourier descriptors allows us to describe the shape through its own main features, i.e. the most representative frequency content, and to enable its recognition and characterization from very few descriptors. The selection of the best number of coefficients able to completely represent the shape, i.e. the one providing a nearly zero approximation error, is a well known but not trivial task in approximation theory. In fact, it corresponds to select a proper threshold to apply to the sorted moduli of the Fourier coefficients. In this paper we propose to select this threshold in an automatic way by defining it as the one maximizing the distance between two competing sets of Fourier coefficients. More precisely, a formal relation between the differential entropy of the transform and its sorted version is derived; on the basis of this formal result, an entropic

ISBN 978-90-827970-1-5 © EURASIP 2018

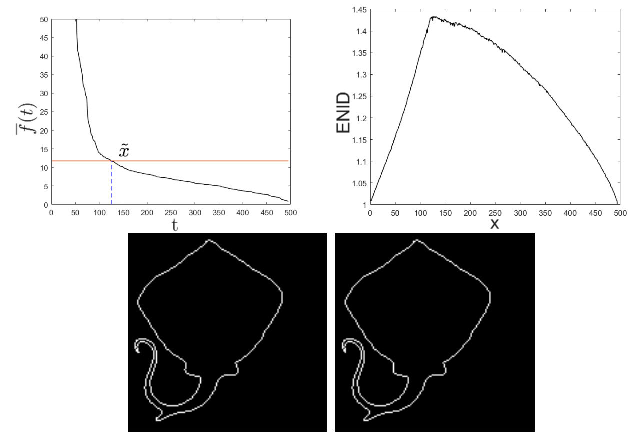


Fig. 1. **Top:** Plot of sorted absolute value of Fourier coefficients (*left*) of the complex cartesian coordinates of the shape at the bottomleft of the figure; plot of ENID as defined in (4) (*right*); the horizontal line indicates the threshold value corresponding to maximum point of ENID. **Bottom:** Original shape (*left*); recovered shape using the proposed method (*right*) — 25% of sorted Fourier coefficients have been used.

interpretation of the normalized information distance (NID) [2] is given. The hierarchical Fourier description is then defined as the set of sorted (in decreasing order) moduli of the Fourier coefficients such that it maximizes the entropic version of NID when computed with respect to the remaining set of coefficients. The proposed method is a first approach to coefficients selection considering the problem in the framework of information distance and Kolmogorov complexity. Preliminary experimental results show that the proposed maximization procedure is able to provide a nearly optimal and compact shape representation which shows to have promising results in retrieval-based applications. Indeed, the proposed method is not limited to Fourier descriptors but can be understood in the more general sense of selecting the most representative coefficients in any hierarchical shape description.

The remainder of the paper is the following. Next section presents the proposed model and the relative algorithm. Preliminary experimental results in shape representation are then presented in Section III along with some concluding remarks.

II. The proposed model

The proposed model is based on the following strategy. The 2D shape contour is given in terms of Fourier descriptors. Since we are interested in catching the shape main features, Fourier descriptors are sorted. It is worth observing that sorted Fourier descriptors preserve the invariance properties of the classical Fourier descriptors. As Fig. 1 shows, sorted Fourier descriptors are a composition of two consecutive decreasing signals: the first one, whose domain is $[0, \tilde{x}]$, usually is strongly decreasing, while the second one ($(\tilde{x}, \bar{x}]$, where $\bar{x}$ is the total length of the signal) is slightly decreasing. We are obviously interested in the first signal, composed of those Fourier coefficients that convey the greater amount of contour information. An entropic version of the Normalized Information Distance (NID) [2], [12] is then proposed to automatically find out $\tilde{x}$. Entropic NID is then maximized as explained in the next section. Finally, it is worth outlining that a new form of differential entropy is proposed in Prop. 1. The advantage of using this form is that it is calculated in terms of a sorted signal — that, in turn, is linked to its cumulative distribution function.

A. Mathematical formulation

We give the definition of (continuous-time) sorting of a function f. This definition is similar to the one in [6].

Def. 1 (Sorting of f): Let $f(t) : [0, \bar{x}] \to \mathbb{R}$ be a Riemann-measurable, limited function and let $M_f(y)$ be the cumulative distribution function of f, that is, $M_f(y) = \text{meas}(\{t : f(t) < y\})$, where the measure used is the Riemann measure. Then the sorting $\bar{f}$ of f is the inverse of its cumulative distribution function:

$$\bar{f}(t) = (M_f)^{-1}(t). \tag{1}$$

A new form of differential entropy [4], written in terms of sorted function, can be then given.

Prop. 1: Let $f(t) : [0, \bar{x}] \to \mathbb{R}$ be a continuous, differentiable function. Then the following identity holds:

$$E_{[0,\bar{x}]} = \frac{1}{\bar{x}} \int_0^{\bar{x}} \ln(\bar{x}\bar{f}'(s))ds, \tag{2}$$

where $E_{[0,x]}$ is the differential entropy of f and $\bar{f}(t)$ is the continuous-time sorting of f.
Proof is in Appendix.

It is easy to generalize (2) to an arbitrary interval $[x_1, x_2] \subset [0, \bar{x}]$, i.e. $E_{[x_1,x_2]}$. The latter depends on the sorting of the restriction of f to the interval $[x_1, x_2]$.

Before defining a new form of entropic NID in terms of $E_{[x_1,x_2]}$, we give just a short description of the classical Normalized Information Distance (NID), defined in [12]. It measures the similarity of two data objects on the base of the more general concept of Kolmogorov complexity. Given a data object a, the Kolmogorov complexity $K(a)$ of a is the length of the shortest computer program that outputs a. More precisely, NID is defined as

$$NID(a, b) = \frac{max(K(a|b), K(b|a))}{max(K(a), K(b))}, \tag{3}$$

where $K(a|b)$ is the generalized conditional Kolmogorov complexity and represents the length of the shortest computer program which outputs a given b. NID is a universal metric; in addition, if two objects are similar with respect to a given similarity metric, then they also are similar with respect to NID. Even though NID is non-computable, it can be well approximated by compression-based distances, like the Normalized Compression Distance (NCD) [2]. The latter compares the compressed version of the two objects. More precisely, it measures how good the lossless compression algorithm C can use the information given in one string to better compress the other one and viceversa. Hence, NCD is a computable version of NID in the sense that the lossless compressor C is used for approximating the Kolmogorov complexity K. NCD

ISBN 978-90-827970-1-5

is able to capture the dominant common information in all pairwise comparisons and already gave high standard results in hierarchical clustering [2]. The entropic NID, denoted by ENID, is then defined:

Def. 2 (Entropic NID):

$$ENID(x) = \frac{\bar{x}E_{[0,\bar{x}]} - \min(C_1(x), C_2(x))}{\max(C_1(x), C_2(x))}, \quad (4)$$

where $C_1(x) = xE_{[0,x]}$ and $C_2(x) = (\bar{x} - x)E_{[x,\bar{x}]}$.
It is possible to prove that ENID has a unique maximum, whose argument $\tilde{x}$ is the solution we are looking for. In particular:

Prop. 2: Let $f(t) : [0, \bar{x}] \to \mathbb{R}$ such that $f \in C^1([0, \bar{x}])$, $f'(t) < 0$, $|f'(x)| > e^{-1}/x$ and $|f'(x)| > e^{-1}/(\bar{x} - x)$ in $[\varepsilon, \bar{x} - \varepsilon]$. Then

$$i) \quad \exists!\, \tilde{x} \in [\varepsilon, \bar{x} - \varepsilon] : \tilde{x} = argmax_x ENID(X)$$

$$ii) \quad ENID(x) \geq 1. \quad (5)$$

Taking into account that the first interval ($[0, \tilde{x}]$) and the second one ($(\tilde{x}, \bar{x}]$) can be seen as produced by two different information sources, it is straightforward to see that the difference of the entropy of these two messages has a maximum when $\tilde{x}$ correctly separates them — avoiding mixtures.

B. The Algorithm

Let $(x(\tau), y(\tau))$ the parametrized cartesian coordinates of the analysed shape contours — τ is the parameter.

- Construct the complex signal $s(t) = x(t) + iy(t)$
- Compute the discrete Fourier Transform of $s(t)$, and denote it by f
- Sort $|f|$, the absolute value of f, in descending order. Let $\bar{f}(t)$ be the sorted signal and $\bar{x}$ its length
- $\forall\, x \in [0, \bar{x}]$, compute $ENID(x)$ as in (4)
- Select $\tilde{x} = argmax_x ENID(x)$
- Invert the discrete transform using the selected $\tilde{x}$ frequencies while putting the others equal to zero and get the approximation of s.

III. Experimental results and Conclusions

The proposed method has been tested on different shapes. In this section some results will be presented. The first test is oriented to fix the accuracy of shape description given by a hierarchical representation of Fourier coefficients with respect to the classical method. Even though the result in terms of MSE is theoretically predictable, this test aims at showing that few sorted coefficients carry information related to corners or some geometrical structures of the analysed shape allowing a straightforward characterization of the single shape or the class it belongs to. This property plays a fundamental role in monitoring based applications as it can be concerned to some pathologies at a very beginning state — fissures or cracks in building of historical importance in cultural heritage or spicular mass of various kinds of cancer in medicine. As Fig. 2 shows, the use of the first six Fourier coefficients allows us to recognize the particular and distinctive geometric profile of the analyzed shape. The same information is not gathered using the same number of traditional Fourier descriptors; the first 10 Fourier coefficients have to be retained for getting comparable information. Fig. 3 shows similar results for two different shapes. It is worth observing that, the linear selection of 6 Fourier descriptors provides similar profiles, making impossible to distinguish the two different shapes; on the contrary, hierarchical Fourier coefficients are able to provide this information. The second test is oriented to show that the optimal number of coefficients that has been selected using the proposed method, provides all the information necessary to reconstruct the shape. The reconstructed images are in fact very close to the original one — the discarded information is really not informative, at least from the visual point of view. Fig. 1 shows the original shape and the one reconstructed using the output of the maximization procedure presented in the previous section. As it can be observed, there are not significant visual differences between the two images. The same figure depicts the entropic NID and the threshold value applied to the coefficients of the transform; as it can be observes, 25% of coefficients allows us to recover the shape with high precision — it corresponds to apply a threshold value equal to 11.76 to the absolute value of Fourier coefficients. Table I contains some results obtained for some selected images having different features; the percentage of retained coefficients and the reconstruction errors, measured in terms of mean square error (MSE) between original image and reconstructed one, are given. As it can be observed, less than 30% of coefficients allow us to recover the images with very high precision. In addition, for more regular shapes, the error is nearly zero. Table II contains the results obtained for shapes

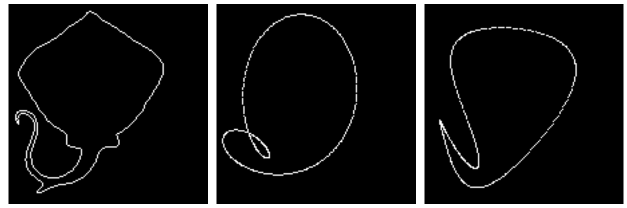


Fig. 2. Original shape (*left*). Recovered shape using the first 6 Fourier descriptors (*middle*). Recovered shape using the first 6 sorted Fourier descriptors (*right*).

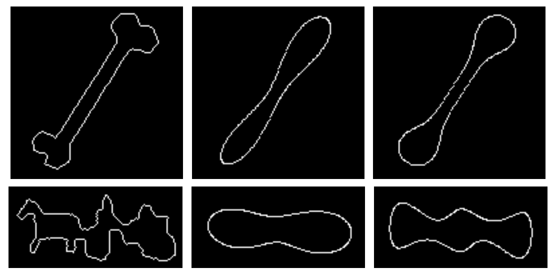
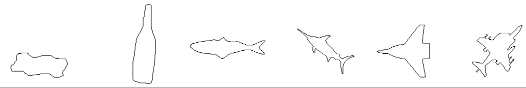
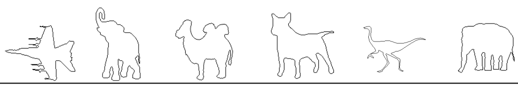


Fig. 3. Original shape (*left*). Recovered shape using the first 6 Fourier descriptors (*middle*). Recovered shape using the first 6 sorted Fourier descriptors (*right*). Two different shapes are shown.

ISBN 978-90-827970-1-5



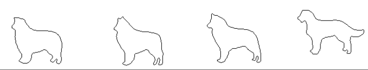
%	27.84	24.62	29.26	26.77	26.84	27.21
	$(\frac{71}{255})$	$(\frac{81}{329})$	$(\frac{91}{311})$	$(\frac{83}{310})$	$(\frac{84}{313})$	$(\frac{117}{430})$
MSE	0.0000	0.0000	0.0000	0.0009	0.0001	0.0021



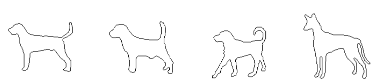
%	29.78	27.18	29.47	26.37	28.06	29.22
	$(\frac{187}{628})$	$(\frac{159}{585})$	$(\frac{160}{543})$	$(\frac{164}{622})$	$(\frac{211}{752})$	$(\frac{173}{592})$
MSE	0.0020	0.0009	0.0028	0.0002	0.0003	0.0006

TABLE I

RESULTS IN TERMS OF % OF RETAINED FOURIER COEFFICIENTS AND APPROXIMATION ERROR MEASURED IN TERMS OF MSE (4 DECIMAL DIGITS) FOR DIFFERENT SHAPES. IN THE BRACKETS: NO. OF SELECTED COEFFICIENTS VS SIGNAL LENGTH.



%	24.46	27.60	27.40	25.78
	(80/327)	(93/337)	(97/354)	(99/384)
MSE HFD	0.0011	0.0014	0.0009	0.0007
MSE FD	0.0027	0.0020	0.0012	0.0025



%	27.73	29.39	27.73	26.85
	(127/458)	(134/456)	(160/577)	(189/704)
MSE HFD	0.0007	0.0005	0.0019	0.0006
MSE FD	0.0012	0.0009	0.0049	0.0028

TABLE II

RESULTS IN TERMS % (NO. OF COEFFICIENTS/SIGNAL LENGTH) OF RETAINED FOURIER COEFFICIENTS AND APPROXIMATION ERROR MEASURED IN TERMS OF MSE (4 DECIMAL DIGITS) FOR SHAPES IN THE SAME CLASS.

belonging to the same class. As it can be observed, the number of selected coefficients is consistent with the complexity of the shape; however, the percentage of selected frequencies is less than 30%. The same table contains the approximation error provided by selecting the same number of Fourier descriptors for increasing frequency value (classical Fourier descriptors). As it can be observed, reconstruction errors are higher in this case. In addition, the more detailed the shape contour, the better the precision provided by the proposed method. The last test aims at a preliminary evaluation of the discriminative properties of the proposed set of descriptors. Fig. 4 contains the first 5 shapes in the analyzed database closest to the shape in Fig. 1 — as it can be observed, 3 of them are in the same class and the remaining 2, although not in the same class, have high frequency content similar to the original image.

Fig. 4. Images in the database closest to the shape in Fig. 1. For each shape the optimal number of Fourier descriptors provided by the proposed method have been considered.

The proposed method is computationally advantageous as the estimation of the optimal number of Fourier coefficients requires simple and fast operations — 0.02 secs are required on average in a not optimized MatLab implementation.

Future work will be devoted to extensive comparative studies with the state of the art methods. The main aim will be the use of the proposed entropy-based procedure for the automatic selection of the best number of significant descriptors, independently of the adopted shape signature.

IV. APPENDIX

Proof of Prop. 1 Given a probability distribution function $m(y)$ with support S, the differential entropy is

$$E = -\int_S m(y)\ln(m(y))dy. \tag{6}$$

The distribution of a function f is the derivative of the cumulative distribution function given in **Def.** 1, normalized by x. By setting $y = \bar{f}(t)$ in (6) and since $\frac{df^{-1}}{dt} = \frac{1}{\frac{df}{dt}}$, we have (2).

Proof of Prop. 2 Since f is monotonically decreasing, then $\bar{f}(x) = f(\bar{x}-x)$ and $\bar{f}'(x) = |f'(\bar{x}-x)|$. By setting $u = \bar{x} - s$, we obtain $E_{[0,x]} = x^{-1}\int_0^x \ln(x|f'(u)|)du$ and $E_{[x,\bar{x}]} = (\bar{x}-x)^{-1}\int_x^{\bar{x}} \ln((\bar{x}-x)|f'(u)|)du$. Hence $C_1'(x) = \ln(ex|f'(x)|)$ and $C_2'(x) = -\ln(e(\bar{x}-x)|f'(x)|)$. Since $C_1'(x) > 0$ and $C_2'(x) < 0\ \forall\, x \in [\varepsilon, \bar{x}-\varepsilon]$, $C_1(x)$ and $C_2(x)$ are continuous by definition, $C_1(0) = 0 = C_2(\bar{x})$ and $C_1(\bar{x}) = \bar{x}E_{[0,\bar{x}]} = C_2(0)$, there exists a point $\tilde{x}$ such that $C_1(\tilde{x}) = C_2(\tilde{x})$. Hence, for $x \in [\varepsilon, \bar{x}-\varepsilon]/\{\tilde{x}\}$, it holds:

$$\frac{d}{dx}\max(C_1, C_2)(x) = C_2'(x)\chi_{[0,\tilde{x})}(x) + C_1'(x)\chi_{(\tilde{x},\bar{x}]}(x) \tag{7}$$

$$\frac{d}{dx}\min(C_1, C_2)(x) = C_1'(x)\chi_{[0,\tilde{x})}(x) + C_2'(x)\chi_{(\tilde{x},\bar{x}]}(x) \tag{8}$$

Let us split (4) into $I_1(x) = \bar{x}E_{[0,\bar{x}]}/\max(C_1(x), C_2(x))$ and $I_2(x) = -\min(C_1(x), C_2(x))/\max(C_1(x), C_2(x))$. Using (7), $I_1(x)$ is increasing before $\tilde{x}$ and decreasing afterward, while using (8) we have $I_2'(x) = -\left[\left(\frac{C_1(x)}{C_2(x)}\right)'\chi_{[0,\tilde{x})}(x) + \left(\frac{C_2(x)}{C_1(x)}\right)'\chi_{(\tilde{x},\bar{x}]}(x)\right]$, and $(C_1(x)/C_2(x))' = -(C_1(x)/C_2(x))^2(C_2(x)/C_1(x))'$. Since $(C_1(x)/C_2(x))' < 0 \Leftrightarrow C_1'(x)/C_2'(x) < 0 < C_1(x)/C_2(x)$, $I_2(x)$ is increasing before $\tilde{x}$ and decreasing afterward, proving the proposition.

With regard to the second part, it is easy to prove that

$$xE_{[0,x]} + (\bar{x}-x)E_{[x,\bar{x}]} \leq \bar{x}E_{[0,\bar{x}]}. \tag{9}$$

Then we have

$$\bar{x}E_{[0,\bar{x}]} \geq \min(C_1(x), C_2(x)) + \max(C_1(x), C_2(x)). \tag{10}$$

Hence, (5) follows.

ISBN 978-90-827970-1-5 © EURASIP 2018

REFERENCES

[1] V. Bruni, D. De Canditiis, D. Vitulano, Time-scale energy based analysis of contours of real-world shapes, Mathematics and Computers in Simulation, Elsevier, vol. 12, pp. 2891-2907, 2012

[2] R. Cilibrasi, P. M. B. Vitanyi, Clustering by compression, IEEE Transactions on Information Theory, vol. 51, no. 4, pp. 15231545, 2005.

[3] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, Active shape models — their training and application, Computer Vision and Image Understanding, vol. 61, no. 1, pp.38-59, 1995

[4] T. M. Cover, J. A Thomas, Elements of Information Theory, John Wiley sons, 1991.

[5] C. Dalitz, C. Brandt, S. Goebbels, D. Kolanus, Fourier Descriptors for Broken Shapes, EURASIP Journal on Advances in Signal Processing, vol. 2013:161, 2013

[6] P. J. S. G. Ferreira, Sorting Continuous-Time Signals: Analog Median and Median-Type Filters, IEEE Transactions on Signal Processing, Volume: 49, Issue: 11, Nov 2001

[7] R. Gonzalez, R. Woods, Digital Image Processing, 3rd ed. Prentice-Hall, New Jersey, 2008

[8] R.X. Hu, W. Jia, H. Ling, Y. Zhao, J. Gui, Angular pattern and binary angula pattern for shape retrieval, IEEE Trans. Image Process., vol. 23, issue 3, pp. 1118-1127, March 2014.

[9] H. Ling, X. Yang, and L. J. Latecki. Balancing deformability and discriminability for shape matching, Proceedings of the 11th European conference on computer vision conference on Computer vision: Part III (ECCV'10), K. Daniilidis, P. Maragos, and N. Paragios (Eds.). Springer-Verlag, Berlin, Heidelberg, 411-424, 2010.

[10] A. Jalba, M. Wilkinson, J. Roerdink, Shape representation and recognition through morphologicalcurvature scale spaces., IEEE Trans. Image Process., vol. 15, no. 2, pp. 331-341, 2006

[11] A. Khotanzad, Y.H. Hong, Invariant image recognition by Zernike moments, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, pp. 489-497, 1990

[12] M. Li, X. Chen, X. Li, B. Ma, and P. M. B. Vitanyi, The similarity metric, IEEE Trans. Inf. Theory, vol. 50, no. 12, pp. 32503264, 2004.

[13] S. Mallat, A wavelet tour of signal processing, Academic Press, 1998.

[14] F. Mokhtarian, S. Abbasi, J. Kittler, Robust and efficient shape indexing through curvature scale space, British Machine Vision Conference, pp. 53-62, 1996

[15] E. Persoon, K. Fu, Shape discrimination using Fourier descriptors, IEEE Trans. Syst. Man Cybern., vol. 7, no. 3, pp. 170.179, 1977

[16] M. A. Savelonas, I. Pratikakis, and K. Sfikas, Partial 3D Object Retrieval combining Local Shape Descriptors with Global Fisher Vectors, Proc. of Eurographics Workshop on 3D Object Retrieval, 2015

[17] E. Sokic, S. Konjicija, Novel fourier descriptor based on complex coordinates shape signature, Proc. of the 12th International Workshop on Content-Based Multimedia Indexing (CBMI), 2014

[18] C. Yang, O. Tiebe, K. Shirahama, M. Grzegorzek, Object matching with hierarchical skeletons, Pattern Recognition, vol. 55, issue C, pp. 183-197, 2016

[19] D. Zhang, G. Lu, Shape-based image retrieval using generic Fourier descriptor, Signal Processing: Image Commun., vol. 17, no. 10, pp. 825-848, 2002

[20] D. Zhang, G. Lu, Study and evaluation of different Fourier methods for image retrieval, Image and Vision Computing, vol. 23, pp. 33-49, 2005

ISBN 978-90-827970-1-5 © EURASIP 2018