3D Chebyshev–Fourier Moments as an ATS Molecular Structure Representation

Satrya Fajri Pratama[1], Azah Kamilah Muda[1(✉)], Yun-Huoy Choo[1],
Ramon Carbó-Dorca[2] and Ajith Abraham[3]
[1]Computational Intelligence and Technologies (CIT) Research Group,
Center for Advanced Computing and Technologies, Fakulti Teknologi Maklumat dan Komunikasi,
Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
Email: {satrya, azah, huoy}@utem.edu.my
[2]Institut de Química Computacional i Catàlisi, Universitat de Girona, 17071 Girona, Catalonia, Spain
Email: ramoncarbodorca@gmail.com
[3]Machine Intelligence Research Labs (MIR Labs), Scientific Network for Innovation and Research
Excellence, Auburn, WA, USA
Email: ajith.abraham@ieee.org

Abstract—All countries are fighting the battle against substance addiction, most especially amphetamine-type stimulants (ATS). The ATS molecular structure profoundly dictates their identification process. However, due to the advent of new, sophisticated, and increasingly complex ATS molecular structures, the method is becoming more unreliable. The distinctive features of the molecular structure of the ATS need, therefore, to be precisely obtained. This paper formulates a novel 3D Chebyshev–Fourier molecular descriptor to represent the drug's molecular structure. The benchmarking of the proposed technique is using drug chemical structures obtained from the pihkal.info database for the ATS. Non-ATS drugs, by comparison, are arbitrarily selected from the ChemSpider database. The assessment indicates that the technique is qualified to be further explored and modified to be entirely consistent with the ATS detection domain in future works.

Index Terms—3D moments; ATS; Chebyshev–Fourier moments; molecular similarity; molecular descriptors.

I. Introduction

THE destructive potential of amphetamine-type stimulants (ATS) substance addiction is continually threatening every nation in this country. One of the critical reasons for the abuse of ATS is that a chemist may derive a novel and uncontrolled structure of ATS and produce it in clandestine laboratories, because it is mainly synthetic, provided with adequate technical knowledge [1, 2]. For law enforcement authorities, these creative designs proved to be a challenge, often leading to both false positive and negative outcomes, and often wrongly outlining the result of the criminal justice system [3-5].

Despite the noticeable need for effective methods for the detection and identification of ATS in the field of forensic toxicology, computer-assisted methods have only recently started to be used [6-8]. In the field of chemistry, the use of computer science breeds a new area, namely cheminformatics. In order to locate structurally related compounds, cheminformatics researchers use molecular similarity search [9, 10]. The effectiveness of the search for molecular similarity depends on the representation of molecular structures employed, also known as molecular descriptors [11, 12].

Formal molecular shapes and surface representations provide more detailed and chemically applicable details than simplistic molecular graphs and stereochemical bonding structures. It also offers a more detailed definition of the actual molecular recognition and interaction processes that result in a quantitative shape-activity relations (QShAR) domain [13]. The shape is an essential visual attribute and is one of the fundamental characteristics used to define the content of the image [14]. Therefore, searching for an image using the shape features poses many research challenges, because extracting the features that reflect and describe the shape is an arduous task [15] that also holds in the QShAR domain [13].

There are several types of molecular descriptors; the most widely used are 2D or topological molecular descriptors and 3D or geometric molecular descriptors. Geometric descriptors typically have more detail and discrimination power for related molecular structures and molecular configurations than for topological descriptors [16-18]. This study assumes that molecular similarity can also be used to detect similarity between unknown and reference compounds and can therefore be used to detect new brands of ATS based on known instances of ATS molecular structure [6, 7] drugs, due to the presence of aromatic phenyl ring and carbonyl side chains which gives its unique characteristics in all ATS analogues [1].

Recent developments in science have allowed the characterization of the physical molecular structure through molecular microscopy, such as the use of transmission electron microscopy (TEM), scanning tunneling microscopy (STM) and more recently, the non-contact atomic force microscope (nc-AFM) [19]. Single-molecule images obtained with nc-AFM are reminiscent of wire-frame chemical structures and also allow for the detection of variations in chemical bond-order [20]. The high-resolution images obtained indicate that the 3D model used for decades to represent the molecular structure is nearly identical to the physical molecules.

Invariance to the labelling, molecular atoms numbering, and molecular translation and rotation is the necessary property of the molecular descriptor. Besides, it must also have a simple, algorithmically quantifiable description, and the values must be within the appropriate numerical range for the molecule set where applicable [21, 22]. Since the molecular descriptor is independent of the characteristics of the molecular representation, it is possible to consider the molecular form as an image and thus to apply the image processing to depict the shape of the molecular structure.

Various molecular structure shape representation techniques have been proposed in the QShAR domain itself [13, 23]. In the meantime, several shape representations techniques have been explored in the pattern recognition and image processing domain to extract features from the object shapes. A good shape descriptor should be able to find physically similar forms, similar to human beings when comparing object shapes. One of the applications of 2D and 3D image processing methods is moment invariants (MI) which can easily achieve these invariance properties. MI is a special case of moments (the scalar quantity used to describe a function and to capture its key features) [24]. The first use of MI as a 3D molecular descriptor is the 3D Zernike descriptor, although it was implemented to represent the molecular surface of the protein structure [25].

There are some benefits to MI-based molecular form representation compared to traditional representations. First, MI-based descriptors enable easy retrieval and comparison of molecular structures. Second, due to its invariance properties, molecular structures do not need to be aligned for comparison. Finally, the resolution of the interpretation of the molecular structures can be conveniently and naturally modified by adjusting the order of the shape descriptors [25, 26].

However, it is very difficult to derive MI from a moment of technique, as it must follow the invariance requirements where the representation of an object must be invariant while the object is undergoing translation, size, and rotation transformation [27]. The goal of this study is, therefore, only to suggest a 3D moment-based molecular representation technique instead of an MI-based one, due to the advantages of 3D molecular shape and its similarity to physical molecular structure. In contrast, the implementation of the MI-based molecular representation technique will be carried out extensively in future works, and this research will be used as the basis of the MI-based technique if the findings of this study are satisfactory.

Before the proposed technique is defined and the rationale of the findings obtained is provided, the current moment-based techniques, which will also be used as a comparison of the proposed technique, must be discussed first. Thus, the rest of this paper is structured as follows. The proposed technique is introduced, and the experimental setup explaining the data source collection and experimental design is described in Sections 2 and 3, respectively. Meanwhile, Sections 4 and 5 discusses the findings, hypotheses, and future works, respectively.

II. Proposed 3D Chebyshev–Fourier Moments

Moments are used to create a series of numbers that uniquely describe the global characteristics of an image and have been used in a variety of fields, from mechanics and statistics to pattern recognition and image analysis [28]. References [29] and [30] were pioneering the use of moments to quantify image features in image processing and pattern recognition.

A 2D image is regarded as piece-wise continuous real function $f(x, y)$ of two variables defined on a compact support $D \subset R \times R$ and having a finite nonzero integral, which by extension is also applicable to 3D images. Some of the existing and well-known moments, including orthogonal moments, are geometric [29], complex (proposed by [31] and later extended by [32]), Legendre, and Zernike moments, both are proposed by [33].

On the other hand, the Chebyshev–Fourier moments was first introduced by [34], which was motivated by the findings of [35], which have shown that a polynomial that is invariant for any rotation of axes about the origin must be of the form

$$V(r\cos\varphi, r\sin\varphi) = R_s(r)e^{-\hat{\imath}q\varphi}, \tag{1}$$

where $R_s(r)$ is a radial polynomial in $r = \sqrt{x^2 + y^2}$ of degree s and $e^{-\hat{\imath}q\varphi}$ is an angular part of the respective polynomial with $\hat{\imath} = \sqrt{-1}$ and $\varphi = \arctan\left(\frac{y}{x}\right)$. The area of orthogonality is a unit disk $\Omega = \{(x, y) | x^2 + y^2 \leq 1\}$. To calculate the moments, the image $f(x, y)$ must be scaled such that it is fully contained in Ω [36]. If the Chebyshev radial polynomial is used to construct an orthogonal function system to decompose an image, the orthogonal function system will be the invariant features of the picture.

The definition of the shifted Chebyshev polynomial of the second kind is

$$U_s^*(r) = U_s(2r - 1), \tag{2}$$

where the function $U_s(2r - 1)$ is the Chebyshev polynomial of the second kind defined as

$$U_s(\cos\varphi) = \frac{\sin(s + 1)\,\varphi}{\sin\varphi}, \tag{3}$$

with

$$\begin{aligned}\sin(s + 1)\,\varphi = \sin\varphi \Big[&(2\cos\varphi)^s \\ &- \binom{s-1}{1}(2\cos\varphi)^{s-2} \\ &+ \binom{s-2}{2}(2\cos\varphi)^{s-4} - \cdots\Big].\end{aligned} \tag{4}$$

Let $\cos\varphi = 2r - 1$, then

$$\begin{aligned}U_s^*(r) &= U_s(\cos\varphi) \\ &= \sum_{k=0}^{\frac{s+2}{2}} (-1)^k \frac{(s-k)!}{k!\,(s-2k)!}[2(2r-1)]^{s-2k}.\end{aligned} \tag{5}$$

The shifted Chebyshev polynomial of the second kind with the weighting function is orthogonal in the interval [0,1] as

$$w(r) = \sqrt{r - r^2}, \tag{6}$$

which satisfy the following orthogonality property to the weight function

$$\int_0^1 U_p^*(r)U_q^*(r)w(r)\,\mathrm{d}r = \frac{\pi}{8}\delta_{pq}, \tag{7}$$

where δ_{pq} is the Kronecker delta, $\delta_{pq} = 1$ if $p = q$ and $\delta_{pq} = 0$ otherwise.

From Equations (5) and (7), the radial polynomial is

$$R_s(r) = \sqrt{\frac{8}{\pi r}w(r)}U_s^*(r) \tag{8}$$

then the orthogonality property becomes

$$\int_0^1 R_p(r)R_q(r)r\,\mathrm{d}r = \delta_{pq}, \tag{9}$$

Thus, the Chebyshev–Fourier moments can be defined as

$$\Psi_{pq} = \int_0^{2\pi}\int_0^{1} R_p(r)e^{-iq\varphi}f(r,\varphi)r\,\mathrm{d}r\,\mathrm{d}\varphi. \tag{10}$$

This study proposes the extension of Chebyshev–Fourier moments for 3D images. The proposed 3D Chebyshev–Fourier moments is adopting the generalization of *n*-dimensional moments on a sphere [36, 37], and defined as

$$\Psi_{nl}^{m} = \int_0^{2\pi}\int_0^{\pi}\int_0^{1} R_s(\varrho)\overline{Y_l^m(\theta,\varphi)}f(\varrho,\theta,\varphi)\varrho\,\mathrm{d}\varrho\,\mathrm{d}\theta\,\mathrm{d}\varphi, \tag{11}$$

where $R_s(\varrho)$ is the radial polynomial given in Equation (8) with $\varrho = \sqrt{x^2+y^2+z^2}$, $\theta = \arctan\left(\frac{y}{x}\right)$, and $\varphi = \arccos\left(\frac{z}{\varrho}\right)$. The $Y_l^m(\theta,\varphi)$ is the spherical harmonics [38] defined as

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)\,e^{-iq\varphi}, \tag{12}$$

where P_l^m is an associated Legendre function defined as

$$P_l^m(a) = (-1)^m(1-a^2)^{\frac{m}{2}}\left(\frac{\mathrm{d}}{\mathrm{d}a}\right)^m L_l(a), \tag{13}$$

and $L_l(a)$ is a Legendre polynomial defined as

$$L_s(a) = \sum_{k=0}^{s} c_{k,s}a^k = \frac{(-1)^s}{2^s s!}\left(\frac{\mathrm{d}}{\mathrm{d}a}\right)^s[(1-a^2)^s], \tag{14}$$

which can also be written as

$$L_s(a) = \sum_{k=0}^{\left\lfloor\frac{s}{2}\right\rfloor} D_{sk}a^{s-2k}, \tag{15}$$

with

$$D_{sk} = (-1)^k\frac{(2s-2k)!}{2^s k!\,(s-k)!\,(s-2k)!}. \tag{16}$$

The recursive relation of Legendre polynomials, $L_s(a)$, is given as

$$L_s(a) = \frac{(2s-1)aL_{s-1}(a)-(s-1)L_{s-2}(a)}{s}, \tag{17}$$

where $L_0(a) = 1$, $L_1(a) = a$ and $s > 1$.

The set of Legendre polynomials $L_s(a)$ forms a complete orthogonal basis set on the interval $[-1, 1]$

$$\int_{-1}^{1} L_p(a)L_q(a)\,\mathrm{d}a = \frac{2}{2p+1}\delta_{pq}. \tag{18}$$

Spherical harmonics are orthonormal on the surface of the unit sphere per the relation

$$\int_0^{\pi}\int_0^{2\pi} Y_l^m(\theta,\varphi)Y_{l'}^{m'}(\theta,\varphi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = \delta_{ll'}\delta^{mm'}. \tag{19}$$

The proposed 3D Chebyshev–Fourier moments are implemented in discrete domain for digital images, and thus Equation (11) becomes

$$\Psi_{nl}^{m} = \sum_{\varrho=0}^{1}\sum_{\theta=0}^{\pi}\sum_{\varphi=0}^{2\pi} R_n(\varrho)\overline{Y_l^m(\theta,\varphi)}f(\varrho,\theta,\varphi)\varrho^2\sin\theta, \tag{20}$$

and can be implemented in Cartesian coordinates as

$$\Psi_{nl}^{m} = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\sum_{z=0}^{N-1} R_n(x,y,z)\overline{Y_l^m(x,y,z)}f(x,y,z). \tag{21}$$

since $\varrho^2\sin\theta$ is the Jacobian of the transformation of Cartesian to spherical coordinates ϱ,θ,φ, and after substituting $\varrho = \sqrt{x^2+y^2+z^2}$. In the next section, the performance of the existing and proposed techniques on ATS and the non-ATS dataset is revealed.

III. Experimental Setup

Pursuant to the objective set out in the above section, a comparative empirical analysis must be planned and performed thoroughly and rigorously. A detailed overview of the experimental method can be found in this section.

A. Dataset Collection

This section describes the process of converting the molecular structure of ATS into 2D and 3D computational data representation as illustrated in [39]. The ATS dataset used in this analysis is extracted from [1] which consists of 60 ATS molecular structure, while 60 non-ATS drug molecular structures are arbitrarily acquired from [40] to match the number of ATS. After the generation of the voxel data, the 3D geometric, complex, Legendre, Zernike, and Chebyshev–Fourier moments are measured up to the 8th order, which produces 165 features. While the features of 3D geometric and Legendre moments are real numbers, 3D complex, Zernike, and Chebyshev–Fourier moments on the other hand are complex numbers.

Therefore, these complex numbers need to be converted into real numbers, so most of the pattern recognition tasks can only handle real numbers. Reference [41] suggested a method consisting of four techniques to represent a complex number as a real number, and the Cartesian bit interleaved was found to be the best representation technique. The value of the zeroth-order moments of ecstasy for each 3D moment represented using the Cartesian bit interleaved is shown in Table I.

TABLE I
Cartesian Bit Interleaved Values of Zeroth-order Moments of Ecstasy for Each 3D Moments

3D moments	Original number	Represented number
Geometric	306425	4254572170020069956 7041133799352041472
Complex	16130711836.218561	4257684755048437479 8153183560267891362
Legendre	0.000285380519926548	1417517392444323061 8113893434503725056
Zernike	7708.229987404831	4253810814878636215 5157822007265511528
Chebyshev–Fourier	9344.919014397588	4253854169438291160 4971099286388998786

B. Operational Procedure

The conventional structure of the tasks of pattern recognition, which are pre-processing, feature extraction and classification, will be used in this article. This paper will therefore compare the efficiency of the existing and proposed 3D moments. All extracted instances were tested using the training and testing dataset mentioned earlier for processing time, memory usage, intra-and inter-class variation, and classification of the drug molecular structure using the leave-one-out classification model, all of which were completed 50 times.

The quartile dispersion coefficient (QCD) of the normalized median absolute deviation (NMAD) is used to explicate the quality of the characteristics of each moment technique in terms of intra-and inter-class variance. Intra-and inter-class variance is a common option to calculate the similarity or dissimilarity of a representation technique [42]. QCD calculates dispersion and is used to make comparisons between and within data sets [43], and it is defined as

$$QCD_i = \frac{Q3_i - Q1_i}{Q3_i + Q1_i}, \tag{22}$$

where $Q1_i$ and $Q3_i$ are the first and third quartile of the ith feature set, respectively. Meanwhile the median absolute deviation (MAD) is a stable alternative to the standard deviation since it does not impact the outliers [44]. However, the MAD may differ in various cases, so it should be normalized to the original function in order to ensure consistency with different data, such that

$$\mathrm{NMAD}_i = \frac{\mathrm{MAD}_i}{x_i}. \tag{23}$$

In this analysis, intra-class variance is classified as NMAD QCD for the ith feature of a molecular structure compared to intra-class molecular structures, and inter-class variance is defined as NMAD QCD for the ith feature compared to inter-class molecular structures.

On the other hand, the features are evaluated in terms of classification accuracy against the well-known Random Forest (RF) classifier [45] from WEKA Machine Learning Package [46]. RF is used in this study because previous studies conducted by [47-51] found that RF is the most appropriate for molecular structure data. In this analysis, the number of trees used by RF is 165, equal to the number of attributes of all 3D moments.

IV. Results and Discussion

Current and proposed 3D moments will be evaluated numerically in this section to determine their merit and consistency in molecular structure representation. Table II displays the average processing time, the memory usage, the intra-class variance ratio relative to the total number of features, and the average classification accuracy of 50 executions.

TABLE II
Average of Processing Time, Memory Usage, and Intra-class Variance Ratio of 3D Moments

3D moments	Processing Time (ns/voxel)	Memory Usage (byte/voxel)	Intra-class Variance Ratio	Classification Accuracy
Geometric	13	419	92.12%	63.42%
Complex	39	841	77.58%	60.43%
Legendre	16	1195	67.88%	73.50%
Zernike	59	4405	64.24%	71.58%
Chebyshev–Fourier	474	2426	87.88%	72.30%

The outcomes presented in Table II show that despite 3D Chebyshev–Fourier performs slowest and second-highest memory usage among other 3D moments, it ranks in second for intra-class variance ratio and classification accuracy. The slow performance high memory usage of 3D Chebyshev–Fourier is attributed to its polynomial computation which is time consuming. Since the classification accuracy is the main consideration of this study, the results should undergo statistical validation.

The results of the classification accuracy should be tested for normality before conducting the statistical validation. If the results are normally distributed, parametric tests, such as ANOVA [52], may be used to verify the accuracy of the classification, otherwise non-parametric tests should be used instead. The normality of the classification accuracy is tested in this analysis using the Shapiro–Wilk test of normality [53]. The outcome of the test of normality is shown in Table III, which shows that the precision of the classification for all 3D moments is normally distributed $(p > 0.05)$, except for geometric moments.

TABLE III
Tests of Normality Results

3D moments	Statistic	df	Sig. (p)
Geometric	0.933	50	0.007
Complex	0.978	50	0.480
Legendre	0.973	50	0.305
Zernike	0.975	50	0.353
Chebyshev–Fourier	0.97	50	0.240

Therefore, the assumption of ANOVA has been violated and the Kruskal–Wallis *H* tests must be used instead. The Kruskal–Wallis *H* test results for the classification accuracy are shown in Table IV, with the post-hoc results using multiple Mann–Whitney *U* test statistic are shown in Table V. It should be noted that the post-hoc test results shown in Table V only show the comparison between 3D Chebyshev–Fourier moments against other 3D moments, also known as planned contrast, and thus statistical tests conducted is to validate the proposed technique.

TABLE IV
Kruskal–Wallis *H* Test Results for Classification Accuracy

Kruskal-Wallis H	*df*	Asymp. Sig.
192.932	4	0

TABLE V
Post-hoc Test Results using Multiple Mann–Whitney *U* Tests Againts 3D Chebyshev–Fourier Moments

Opposing 3D moments	Mann-Whitney *U*	Wilcoxon *W*	*Z*	Asymp. Sig. (2-tailed)
Geometric	1.5	1276.5	-8.635	< .001
Complex	0	1275	-8.639	< .001
Legendre	806	2081	-3.092	0.002
Zernike	973	2248	-1.926	0.54

Based on the results shown in Table IV, there are a statistically significant effect of classification accuracy $[H(4) = 192.932, p = 0]$. The degree of freedom (df) of this test is the number of groups minus one, which in this case is 5

groups, and thus the $df = 4$. Post-hoc comparisons using multiple Mann–Whitney U test on each pair of groups and adjusting the p value with the Bonferroni method to avoid Type I error [54] shown in Table V indicated that there is a statistically significant difference between the classification accuracy of 3D Chebyshev–Fourier moments and other 3D moments techniques ($p < p_{Bonferroni} = 0.005$), with one exception to 3D Zernike moments ($U = 973, p = 0.54$).

Despite providing a second highest performance after 3D Legendre moments and without significant difference to 3D Zernike moments in third place, this study nevertheless proposes a new 3D moments technique and shows that the proposed 3D Chebyshev–Fourier possesses certain potentials to be explored in the future, most notably on its invariance properties.

V. Conclusion

A new 3D moments technique to represent ATS molecular structure has been proposed and the extensive comparative study to the existing 3D moments has been presented in this paper, namely 3D Chebyshev–Fourier moments. Despite the experiments have shown that the proposed technique performs better compared to most of existing 3D moments in terms of processing time and memory usage, but nonetheless it has shown potential in intra- and inter-class variance, and more importantly, classification accuracy. Thus, this study serves as a basis towards a better 3D molecular structure representation, especially on using continuous orthogonal moments defined on a sphere.

Hence, future works to extend the proposed technique so that it has invariance properties, as well as better representing the molecular structure based on this preliminary study are required. The proposed feature extraction technique will be further validated in the future works using specifically-tailored classifiers for drug shape representation. Furthermore, ATS molecular structure data from National Poison Centre, Malaysia, will also be used as additional dataset in the future works.

References

[1] United Nations Office on Drugs and Crime, "Recommended Methods for the Identification and Analysis of Amphetamine, Methamphetamine and Their Ring-substituted Analogues in Seized Materials," Sales No. E.06.XI.1, UNODC, 2006.

[2] United Nations Office on Drugs and Crime, "World Drug Report 2016," Sales No. E.16.XI.7, UNODC, 2016.

[3] P. L. Cary, "Designer Drugs: What Drug Court Practitioners Need To Know," *Drug Court Practitioner Fact Sheet,* vol. IX, no. 2, pp. 1-13, 2014.

[4] M. J. Swortwood, "Comprehensive Forensic Toxicological Analysis of Designer Drugs," Dissertation, Florida International University, Florida, USA, 2013.

[5] M. C. F. Smith, "But what of designer drugs?," *Advances in Psychiatric Treatment,* vol. 17, no. 2, pp. 158-158, 2011.

[6] M. D. Krasowski, and S. Ekins, "Using cheminformatics to predict cross reactivity of "designer drugs" to their currently available immunoassays," *J Cheminform,* vol. 6, no. 1, pp. 22, 2014.

[7] M. Petrie, K. L. Lynch, S. Ekins, J. S. Chang, R. J. Goetz, A. H. Wu, and M. D. Krasowski, "Cross-reactivity studies and predictive modeling of "Bath Salts" and other amphetamine-type stimulants with amphetamine screening immunoassays," *Clin Toxicol (Phila),* vol. 51, no. 2, pp. 83-91, Feb, 2013.

[8] A. Amine, Z. Elberrichi, M. Simonet, and A. Rahmouni, "A Hybrid Approach Based on Self-Organizing Neural Networks and the K-Nearest Neighbors Method to Study Molecular Similarity," *International Journal of Chemoinformatics and Chemical Engineering,* vol. 1, no. 1, pp. 75-95, 2011.

[9] Y. C. Martin, J. L. Kofron, and L. M. Traphagen, "Do structurally similar molecules have similar biological activity?," *J Med Chem,* vol. 45, no. 19, pp. 4350-8, Sep 12, 2002.

[10] A. Bender, and R. C. Glen, "Molecular similarity: a key technique in molecular informatics," *Org Biomol Chem,* vol. 2, no. 22, pp. 3204-18, Nov 21, 2004.

[11] A. Bender, "Studies on Molecular Similarity," Dissertation, Darwin College, University of Cambridge, Cambridge, UK, 2005.

[12] V. Consonni, and R. Todeschini, "Molecular Descriptors," *Recent Advances in QSAR Studies: Methods and Applications*, T. Puzyn, J. Leszczynski and T. M. Cronin, eds., pp. 29-102, Dordrecht, Netherlands: Springer Netherlands, 2010.

[13] P. G. Mezey, "Shape-Similarity Measures for Molecular Bodies: A Three-Dimensional Topological Approach to Quantitative Shape-Activity Relations," *Journal of Chemical Information and Computer Sciences,* vol. 32, no. 6, pp. 650-656, 1992.

[14] D. Zhang, and G. Lu, "Shape-based Image Retrieval using Generic Fourier Descriptor," *Signal Processing: Image Communication,* vol. 17, no. 10, pp. 825-848, 2002.

[15] A. K. Muda, "Authorship Invarianceness for Writer Identification Using Invariant Discretization and Modified Immune Classifier," Dissertation, Universiti Teknologi Malaysia, Johor, Malaysia, 2009.

[16] A. Axenopoulos, P. Daras, G. Papadopoulos, and E. N. Houstis, "A shape descriptor for fast complementarity matching in molecular docking," *IEEE/ACM Trans Comput Biol Bioinform,* vol. 8, no. 6, pp. 1441-57, Nov-Dec, 2011.

[17] E. Estrada, "Generalized Graph Matrix, Graph Geometry, Quantum Chemistry, and Optimal Description of Physicochemical Properties," *The Journal of Physical Chemistry A,* vol. 107, no. 38, pp. 7482-7489, 2003.

[18] S. Kortagere, M. D. Krasowski, and S. Ekins, "The importance of discerning shape in molecular pharmacology," *Trends Pharmacol Sci,* vol. 30, no. 3, pp. 138-47, Mar, 2009.

[19] D. G. de Oteyza, P. Gorman, Y. C. Chen, S. Wickenburg, A. Riss, D. J. Mowbray, G. Etkin, Z. Pedramrazi, H. Z. Tsai, A. Rubio, M. F. Crommie, and F. R. Fischer, "Direct imaging of covalent bond structure in single-molecule chemical reactions," *Science,* vol. 340, no. 6139, pp. 1434-7, Jun 21, 2013.

[20] L. Gross, F. Mohn, N. Moll, B. Schuler, A. Criado, E. Guitian, D. Pena, A. Gourdon, and G. Meyer, "Bond-order discrimination by atomic force microscopy," *Science,* vol. 337, no. 6100, pp. 1326-9, Sep 14, 2012.

[21] V. Consonni, and R. Todeschini. "Basic Requirements for Valid Molecular Descriptors," 28 January, 2016; http://www.moleculardescriptors.eu/tutorials/T3_moleculardescriptors_requirements.pdf.

[22] M. Randić, "Molecular bonding profiles," *Journal of Mathematical Chemistry,* vol. 19, no. 3, pp. 375-392, 1996.

[23] P. G. Mezey, "Theorems on Molecular Shape-Similarity Descriptors: External T-Plasters and Interior T-Aggregates," *Journal of Chemical Information and Computer Sciences,* vol. 36, no. 6, pp. 1076-1081, 1996.

[24] Y. Sun, W. Liu, and Y. Wang, "United Moment Invariants for Shape Discrimination." pp. 88-93.

[25] D. Kihara, L. Sael, R. Chikhi, and J. Esquivel-Rodriguez, "Molecular Surface Representation Using 3D Zernike Descriptors for Protein Shape Comparison and Docking," *Current Protein and Peptide Science,* vol. 12, no. 6, pp. 520-530, 2011.

[26] L. Sael, B. Li, D. La, Y. Fang, K. Ramani, R. Rustamov, and D. Kihara, "Fast protein tertiary structure retrieval based on global surface shape similarity," *Proteins,* vol. 72, no. 4, pp. 1259-73, Sep, 2008.

[27] D. Xu, and H. Li, "Geometric Moment Invariants," *Pattern Recognition,* vol. 41, no. 1, pp. 240-249, 2008.

[28] S. X. Liao, "Image Analysis by Moment," Dissertation, University of Manitoba, Manitoba, Canada, 1993.

[29] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory,* vol. 8, no. 2, pp. 179-187, 1962.

[30] F. L. Alt, "Digital Pattern Recognition by Moments," *Journal of the ACM,* vol. 9, no. 2, pp. 240-258, 1962.

[31] Y. S. Abu-Mostafa, and D. Psaltis, "Recognitive Aspects of Moment Invariants," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. PAMI-6, no. 6, pp. 698-706, 1984.

[32] J. Flusser, "On the Independence of Rotation Moment Invariants," *Pattern Recognition,* vol. 33, no. 9, pp. 1405-1410, 2000.

[33] M. R. Teague, "Image analysis via the general theory of moments*," *Journal of the Optical Society of America,* vol. 70, no. 8, pp. 920, 1980/08/01, 1980.

[34] Z. Ping, R. Wu, and Y. Sheng, "Image description with Chebyshev-Fourier moments," *J Opt Soc Am A Opt Image Sci Vis,* vol. 19, no. 9, pp. 1748-54, Sep, 2002.

[35] A. B. Bhatia, and E. Wolf, "On the circle polynomials of Zernike and related orthogonal sets," *Mathematical Proceedings of the Cambridge Philosophical Society,* vol. 50, no. 1, pp. 40-48, 1954/001/001, 1954.
[36] J. Flusser, T. Suk, and B. Zitová, *2D and 3D Image Analysis by Moments*, West Sussex, UK: John Wiley and Sons, Ltd, 2016.
[37] J. Flusser, T. Suk, and B. Zitová, *Moments and Moment Invariants in Pattern Recognition*, West Sussex, UK: John Wiley and Sons, Ltd, 2009.
[38] N. Canterakis, "3D Zernike Moments and Zernike Affine Invariants for 3D Image Analysis and Recognition." pp. 85-93.
[39] S. F. Pratama, A. K. Muda, Y.-H. Choo, R. Carbó-Dorca, and A. Abraham, "Preparation of Translated, Scaled, and Rotated ATS Drugs 3D Molecular Structure for the Validation of 3D Moment Invariants-based Molecular Descriptors," *International Journal of Computer Information Systems and Industrial Management Applications,* vol. 10, pp. 57-67, 2018.
[40] Royal Society of Chemistry. "ChemSpider Database," 23 January, 2016; http://www.chemspider.com/.
[41] S. F. Pratama, A. K. Muda, and Y.-H. Choo, "Arbitrarily Substantial Number Representation for Complex Number," *Journal of Telecommunication, Electronic and Computer Engineering,* vol. 10, no. 1-7, pp. 23-26, 2 April 2018, 2018.
[42] Z. He, X. Youb, and Y.-Y. Tang, "Writer Identification Using Global Wavelet-based Features," *Neurocomputing,* vol. 71, no. 10–12, pp. 1831-1841, 2008.
[43] D. G. Bonett, "Confidence interval for a coefficient of quartile variation," *Computational Statistics & Data Analysis,* vol. 50, no. 11, pp. 2953-2957, 7/20/, 2006.
[44] P. J. Rousseeuw, and C. Croux, "Alternatives to the Median Absolute Deviation," *Journal of the American Statistical Association,* vol. 88, no. 424, pp. 1273-1283, 1993/12/01, 1993.
[45] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, no. 1, pp. 5-32, 2001.
[46] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," *SIGKDD Explorations,* vol. 11, no. 1, pp. 10-18, 2009.
[47] S. F. Pratama, A. K. Muda, Y.-H. Choo, and A. Abraham, "Exact Computation of 3D Geometric Moment Invariants for ATS Drugs Identification," *Innovations in Bio-Inspired Computing and Applications*, Advances in Intelligent Systems and Computing V. Snášel, A. Abraham, P. Krömer, M. Pant and A. K. Muda, eds., pp. 347-358, Cham, Switzerland: Springer International Publishing, 2016.
[48] S. F. Pratama, A. K. Muda, Y.-H. Choo, and A. Abraham, "3D Geometric Moment Invariants for ATS Drugs Identification: A More Precise Approximation," *Proceedings of the 16th International Conference on Hybrid Intelligent Systems (HIS 2016)*, A. Abraham, A. Haqiq, A. M. Alimi, G. Mezzour, N. Rokbani and A. K. Muda, eds., pp. 124-133, Cham, Switzerland: Springer International Publishing, 2017.
[49] S. F. Pratama, A. K. Muda, Y.-H. Choo, J. Flusser, and A. Abraham, "ATS drugs molecular structure representation using refined 3D geometric moment invariants," *Journal of Mathematical Chemistry,* vol. 55, no. 10, pp. 1951-1963, June 28, 2017.
[50] S. F. Pratama, N. A. Muda, and F. Salim, "Representing ATS Drugs Molecular Structure Using 3D Orthogonal Fourier–Mellin Moments," *International Journal of Computer Information Systems and Industrial Management Applications,* vol. 9, pp. 135-144, 2017.
[51] S. F. Pratama, A. K. Muda, Y.-H. Choo, R. Carbó-Dorca, and A. Abraham, "Using 3D Hahn Moments as A Computational Representation of ATS Drugs Molecular Structure," *International Journal on Advanced Science, Engineering and Information Technology,* vol. accepted, 2018.
[52] R. A. Fisher, *Statistical Methods for Research Workers*, 14th ed., Connecticut, USA: Hafner Pub. Co., 1970.
[53] S. S. Shapiro, and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika,* vol. 52, no. 3-4, pp. 591-611, December 1, 1965, 1965.
[54] D. Flynn, *Student Guide to SPSS*, New York, USA: Barnard College, 2010.