Experimental results of bispectral invariants discriminative power

Kubicki, Karol.; Kakarala, Ramakrishna.

2012

Kubicki, K.,& Kakarala, R. (2012). Experimental results of bispectral invariants discriminative power. Proceedings of SPIE - Three-Dimensional Image Processing (3DIP) and Applications II, 82900F.

https://hdl.handle.net/10356/99100

https://doi.org/10.1117/12.906526

Experimental results of bispectral invariants discriminative power

Karol Kubicki and Ramakrishna Kakarala

School of Computer Engineering, Nanyang Technological University, Singapore

ABSTRACT

One of the main tools in shape matching and pattern recognition are invariants. For three-dimensional data, rotation invariants comprise of two main kinds: moments and spherical harmonic magnitudes. Both are well examined and both suffer from certain limitations. In search for better performance, a new kind of spherical-harmonic invariants have been proposed recently, called bispectral invariants. They are well-established from theoretical point of view. They posses numerous beneficial properties and advantages over other invariants, include the ability to distinguish rotation from reflection, and the sensitivity to phase. However, insufficient research has been conducted to check their behavior in practice. In this paper, results are presented pertaining to the discriminative power of bispectral invariants. Objects from Princeton Shape Benchmark database are used for evaluation. It is shown that the bispectral invariants outperform power spectral invariants, but perform worse than other descriptors proposed in the literature such as SHELLS and SHD. The difference in performance is attributable to the implicit filtering used to compute the invariants.

Keywords: Bispectral invariants, object matching, pattern recognition

1. INTRODUCTION

Visual signals carry a huge amount of data. Extracting useful information out of them may seem to be an easy task from a human's perspective, but it is certainly not in case of computers. One of the priorities for computer vision is to recognize objects. Particular shapes, seen from different viewpoints, constitute different sets of data, but there exist certain properties that do not change. Among those properties are the so-called geometric primitives, such as the following: distance, area and volume. According to this observation, it is possible to calculate certain characteristics that will remain the same, regardless of object's position and orientation in space and will change for different objects. A function that assigns values to shapes in that manner is called an invariant.

Translation, scaling and rotation are three basic transformations with respect to which it is desired to obtain invariance. It is easy in case of first two. We can achieve translation and scale invariance by moving the centroid of a given object to the origin and than normalizing it with the longest distance of vertexes from the origin. It is more difficult with rotation, in particular three-dimensional (3-D) rotation as examined in this paper. One of the ideas is to find a canonical orientation and adjust it so that it lies in x,y,z directions. An example of such approach is finding the inertia ellipsoid (more details can be found in Galvez & Canton[1]). Another and most common approach is to simply use functions with the rotation invariance embedded, although it is not easy to find them. Generally, this method is preferred, because it is more robust with respect to noise.

Two main kinds of 3-D rotation invariants functions are moments and spherical invariants. The former are well analyzed[2] and used also in case of 2-dimensional object recognition. Three dimensional moments of order $(p+q+r)$ for a object described by a function f may be defined as:

$$M_{pqr} = \sum_x \sum_y x^p y^q z^r f(x,y,z) \; . \tag{1}$$

Then centroid coordinates are $\bar{x} = \frac{M_{100}}{M_{000}}$, $\bar{y} = \frac{M_{010}}{M_{000}}$, $\bar{z} = \frac{M_{001}}{M_{000}}$ and the central moments, which are translation invariant, are defined as follows:

$$\mu_{pqr} = \sum_{x,y,z} (x-\bar{x})^p (y-\bar{y})^q (z-\bar{z})^r f(x,y,z) \; . \tag{2}$$

Further author information:
K.K.: E-mail: karol.kubicki@uj.edu.pl
R.K.: E-mail: ramakrishna@ntu.edu.sg

Three-Dimensional Image Processing (3DIP) and Applications II, edited by Atilla M. Baskurt, Robert Sitnik,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8290, 82900F · © 2012 SPIE-IS&T
CCC code: 0277-786X/12/$18 · doi: 10.1117/12.906526

Scaling invariance may be obtained by normalization: $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\rho}}$, where $\rho = \frac{p+q}{2} + 1$. Classical examples[2] of 3-D rotation invariant moments are presented below.

$$I_1 = \eta_{200} + \eta_{020} + \eta_{002} \tag{3}$$

$$I_2 = \eta_{020}\eta_{002} + \eta_{200}\eta_{002} + \eta_{200}\eta_{020} \tag{4}$$

$$I_3 = \eta_{200}\eta_{020}\eta_{002} + 2\eta_{110}\eta_{101}\eta_{011} - \eta_{200}\eta_{011}^2 - \eta_{020}\eta_{101}^2 - \eta_{002}\eta_{110}^2. \tag{5}$$

A comprehensive review of generalized moments and moment invariants can be found in the book by Flusser.[2] A general theory is presented for constructing 2-dimensional moment invariants of any order (with respect to translation, scaling and rotation). It is achieved with two approaches: complex moments and normalization (which involves transforming image into a canonical position). Moment invariants can be also constructed with multiple integrals of geometric primitives like: distance, area and volume. This method is shown by Xu & Li.[3] Its main advantages are: intuitiveness and the fact that it can be easily generalized to higher dimensions. Object recognition via shape normalization is described by Galvez & Canton.[1] Firstly, moments of orders from 0 to 2 are used to move object to the origin, rotate it to canonical position and scale it. Secondly, moments of orders 2-5 are employed to count feature vector which serves for discrimination.

Spherical invariants are based on the properties of spherical harmonics. Any (square-integrable) function f defined on a two dimensional sphere can be decomposed into an infinite sum of harmonics:

$$f(\varphi,\theta) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} F_l^m Y_l^m(\varphi,\theta) \,, \tag{6}$$

where $Y_l^m(\varphi,\theta) = P_l^{|m|}(\cos\theta)e^{im\varphi}$ is spherical harmonic and $P_l^{|m|}$ is associated Legendre polynomial. The l subscript is called the frequency and $f_l = \sum_{m=-l}^{l} F_l^m Y_l^m$ is a frequency component of f. The crucial property of the aforementioned decomposition is that rotation of f doesn't change the norm of its frequency components.[2] Therefore, $P_l(f) = \|f_l\|$ is rotation invariant. Note that because the spherical harmonics are an orthonormal basis, we have that $\|f_l\| = \|F_l\|$, where F_l is the $2l+1$ dimensional vector

$$F_l = [F_l^{-l}, ..., F_l^0, ..., F_l^l] \tag{7}$$

Using the analogous terminology from signal processing, P_l will be called a power spectral invariant in the further part of this paper.

A very good introduction to spherical harmonics theory can be found in Shønefeld.[4] Essential mathematical background is provided. It includes: orthogonality in space of functions and associated Legendre polynomials. After introducing spherical harmonics and their properties, applications in computer graphics are discussed. Different approach to the topic is presented by Mohlenkamp.[5] The key element of this approach is the fact that spherical harmonics are derived in analogy to Fourier series on a circle.

Both moments and spherical invariants suffer from certain intrinsic limitations. Moment invariants are relatively hard to derive, and do not provide high discriminative power. For example, the invariants I_1, I_2, and I_3 cannot distinguish rotation from reflection. That limitation is also shared by power spectral spherical harmonic invariants. In pursuit of better performance, new kind of invariants has been developed. The trick is to use the idea of bispectrum. It is well known in statistical signal processing, but has not been used in computer vision until recently.[6] Such a bispectrum gave birth to invariants that retain information not only about magnitudes, but also phase.

To introduce bispectral invariants, a few definitions are needed. For $k,l \geq 0$ and $|k-l| \leq i \leq k+l$, let:

$$F_l \otimes F_k = [F_l^{-l}F_k, ..., F_l^0 F_k, ..., F_l^l F_k] \,, \quad \hat{F}_i = [0, ..., 0, F_i, 0, ..., 0] \,. \tag{8}$$

The number of zeros preceding F_i is n_p and zeros succeeding is n_s:

$$n_p = \sum_{q=|k-l|}^{i-1} (2q+1) \,, \qquad n_s = \sum_{q=i+1}^{k+l} (2q+1) \,. \tag{9}$$

Now, the i-th coordinate of (k,l)-th vector bispectrum for function f is:

$$b_f^{kl}(i) = [F_k \otimes F_l] C_{kl} \hat{F}_i^{\dagger} \,, \tag{10}$$

where the † symbol means conjugate transpose and C_{kl} is Clebsch-Gordan matrix (see the book by Hamermesh for details[7]).

Bispectral invariants theory is thoroughly discussed in a previous paper.[6] In fact, triple correlation and then bispectrum is defined for functions which domain is group $SO(3)$ (group of all rotations about the origin in $\mathbb{R}^3$). This is justified by the fact that for any function on sphere S^2 there exists corresponding function on $SO(3)$. Having those two, bispectrum vector invariant to rotation is defined for any function on S^2. Another important result, presented in the paper is the close relation between bispectrum invariants and moments. It turns out that monomials of a given order can be linearly transformed into spherical harmonics coefficients. This fact is more deeply investigated elsewhere.[8] As a result, a simple method for calculating bispectral invariants is obtained. The invariants come down to be a weighted inner product. A little more details are given in the Research Methodology section.

There exists a website associated with this research available at: `http://langust.x10.mx/project.htm`. Matlab codes and all the obtained outcomes can be found there for all the experiments conducted.

1.1 RESEARCH OBJECTIVE

Bispectral invariants are well-established from theoretical point of view. Yet, little research has been conducted on how do they behave in practice. It is proven that they posses numerous advantages:[6]

- They can distinguish between whole object rotation and rotation of different frequencies of spherical components,
- They can discriminate between rotation with and without reflection,
- They are more robust to noise in shape recognition than spherical magnitude invariants,
- They are fast to compute.

Bearing those in mind, bispectral invariants seem to be a very promising tool. Though, there is still a question how good they are at shape matching and distinguishing between different class of objects. Generally, it would be best if the bispectral invariant values were close within one category and distant between different categories.

Another issue to investigate is how does the discriminative power of bispectral invariants change when we use more of them (adding those of higher orders). Is it possible to approach any arbitrary level of performance with respectively high order or does there exist a limit? After the aforementioned issues are addressed, the range of applications for bispectral invariants in 3-D imaging increases.

2. RESEARCH METHODOLOGY

The research was conducted using the Princeton Shape Benchmark (PSB) database.[9] It is a publicly available database of 3-D polygonal models (see `http://shape.cs.princeton.edu/benchmark/`). The total amount of 1814 objects is provided. The PSB objects were gathered from World Wide Web and categorized into several classifications. First, they were split into two disjoint groups: train and test, 907 objects each. Within each of those, 4 classifications has been proposed. Base classification is the most detailed one. Next three have been constructed by merging the classes from the previous one so that coarser partitioning is obtained. In this paper, the base classification of the train group will be most prominent. It will be referred to as train1 classification.

It is worth mentioning that classification files are written in a simple language described in the documentation provided by PSB. Thanks to that, new classifications can be easily constructed and added.

The PSB database provides also tools for evaluation of performance of matching algorithms. Those tools were used to generate most of the results presented in this paper. Before they will be briefly explained, one remark is needed. For each object in the database, bispectral invariant vector was calculated. With those vectors, distance between each pair of objects was computed. Having a matrix of distances it is possible to say for a given object which other are the closest and how many of them are from the same class of objects.

Precision-recall plot shows the relation between 'precision' (vertical axis) and 'recall' (horizontal axis) value for a given object. Where:

$$\text{precision(OBJ)} = \frac{\text{objects from the same class as OBJ within first K closest objects}}{\text{K}}, \tag{11}$$

$$\text{recall(OBJ)} = \frac{\text{objects from the same class as OBJ within first K closest objects}}{\text{number of objects in the class that OBJ belongs to}} . \quad (12)$$

The curve is parameterized by K which changes from 1 to the number of models in the classification. The best outcome is a horizontal line at the level of 1 (for any number of models retrieved, all of them are from the same class as query). Having multiple plots, the one which is the highest is the best. The plots presented in the next section are generated as an average value over all objects from target classification.

We use the following statistics to evaluate the performance of bispectral invariants: Nearest Neighbor (NN), first tier (1stTier), second tier (2ndTier) and E-measure. Their detailed description can be found in Shilane *et al.*[9] It is enough to know that for all of them the higher the value the better and the maximal value is 1. Thanks to that they are easily comparable.

Tier images for a given classification may be generated. They will not be presented in this paper, but can be seen on the project website. It is useful tool to check within which classes the matching was most successful. Namely, those classes are the ones which have bright, class-size block of pixels attached to the diagonal at their respective level. Best match results can be also seen on this projects website. They are easy to interpret, but not to convenient for getting a general picture. For each object in a given classification a web page is generated where first 90 closest model matches are presented. Blue frame means that the model is from the same class as the query and red frame the opposite.

Each object in PSB database is stored as coordinates of vertexes and numbers of vertexes that constitute faces. All but one of the results presented in this paper were obtained using only the vertexes. The only different result was obtained to check that it was justified to skip the information about faces.

In next section performance results will be presented pertaining to bispectral, power spectral and covariance invariants. Now the methods used to calculate them will be presented using the following notation[8](matlab scripts can be found on the research website). Let OBJ be a $Nx3$ matrix of vertexes and $[\mu_x, \mu_y, \mu_z]$ be a mean row from OBJ (a centroid). A vector of central moments of order l of length $n_l = \frac{1}{2}(l+1)(l+2)$ is:

$$\mathcal{M}_l = [\mu_{l00}, ..., \mu_{0l0}, ..., \mu_{00l}] \quad (13)$$

There exists such a matrix [8] $\mathcal{A}_l$ that $F_l = \mathcal{M}_l \mathcal{A}_l$, where F_l is the vector of Fourier coefficients defined in eq. (8). When we substitute this equation into (10) we get:

$$b^{kl}_{OBJ}(i) = [(\mathcal{M}_l \mathcal{A}_l) \otimes (\mathcal{M}_k \mathcal{A}_k)] C_{lk} (\widehat{\mathcal{M}_i \mathcal{A}_i})^\dagger . \quad (14)$$

Now, the expression $(\widehat{\mathcal{M}_i \mathcal{A}_i})^\dagger$ can be rewritten as $S_i (\mathcal{M}_i \mathcal{A}_i)^\dagger$, where S_i is a matrix of dimensions $(n_l n_k)$xn_i. It contains only zeros, except for the $2i+1$ dimensional identity matrix embedded after t_i rows of zeros, where $t_0 = 0$ and $t_i = \sum_{q=|l-k|}^{i-1} (2q+1)$ for $i > 0$. Having that and using the following property of Kronecker product: $(ab) \otimes (cd) = (a \otimes c)(b \otimes d)$ we obtain:

$$b^{kl}_{OBJ}(i) = (\mathcal{M}_l \otimes \mathcal{M}_k) \left[(\mathcal{A}_l \otimes \mathcal{A}_k) C_{lk} S_i \mathcal{A}_i^\dagger \right] \mathcal{M}_i^T . \quad (15)$$

The symbol ^{T} in the above equation is used instead of † to underline the fact that $\mathcal{M}_i$ coordinates are real. Denoting $B^i_{lk} := (\mathcal{A}_l \otimes \mathcal{A}_k) C_{lk} S_i \mathcal{A}_i^\dagger$ (a $(n_l n_k)$xn_i-dimensional matrix that can be precomputed) we finally get:

$$b^{lk}_{OBJ}(i) = (\mathcal{M}_l \otimes \mathcal{M}_k) B^i_{lk} \mathcal{M}_i^T . \quad (16)$$

The above equation was employed to compute the bispectral invariants vectors that were used to obtain the results presented in this paper. The term 'long (bispectral) invariants vector of order k' (LBIV) will be used to refer to a vector containing all the bispectral invariants of orders 3-k:

$$\text{LBIV of order } k = [\text{bisp\_inv\_vect\_order\_3 bisp\_inv\_vect\_ord\_4 ... bisp\_inv\_vect\_ord\_}k] . \quad (17)$$

Power spectral invariant of orders 1-11 were calculated accordingly to the following equation[8](l denotes order, OBJ_l denotes l-th frequency component of OBJ):

$$P_l(OBJ) = \|OBJ_l\| = F_l F_l^T = \mathcal{M}_l \mathcal{A}_l \mathcal{A}_l^\dagger \mathcal{M}_l^T . \quad (18)$$

The values obtained in this way were put into 1x11-dimensional vector, which was the basis for getting the distances between objects (euclidean norm of difference of two vectors).

For a baseline comparison, the covariance invariants vector was counted in four steps. First, the centroid of *OBJ* was moved to the origin. Then, all coordinates were scaled by $max_{(x,y,z)\in OBJ}\|(x,y,z)\|$. For such object, 3x3 covariance matrix $cov(OBJ)$ was computed (the three random variables are x-,y- and z- coordinates of vertexes). Finally, covariance invariant vector is vector of eigenvalues of $cov(OBJ)$ in ascending order; such a vector is clearly invariant to 3-D rotation of the points in OBJ.

3. DATA ANALYSIS AND RESULTS

In this section results pertaining to discriminative power of bispectral invariants will be presented. Our first result, in Figure 1, shows a precision-recall plot for long bispectral invariants vectors of orders 3, 4 and 8. Orders 5-7 and 9-14 were skipped so that the Figure is more legible. Their graphs are not significantly different from the one for order 8. This observation suggest that increasing the order of bispectral invariants beyond 8 does not result in any increase of discriminative power. Another issue is that presented graphs are far from ideal which is horizontal line at the level of 1. Comparing Figure 1 to Figure 2 from[9] it seems that bispectral invariants perform worse than SHELLS which is the weakest descriptor from PSB.

Analogous conclusions can be drawn from Table 1 (for each given descriptor, the closer the value to 1 the better). Increase in discriminative power can be observed when using LBIV of order up to 8, after which the values of all statistics fluctuate about a constant level. This leads to quite surprising conclusion that no new information is introduced with bispectral invariants of higher orders. That is why in further part of this section only LBIV of order 8 will be used for comparisons with other descriptors (it also justifies why it was chosen to be shown in Figure 1).

It was marked before that best matches results can be seen on the project website. It is quite common among the objects that within their closest matches some diametrically different shapes are present (for example the closest match to aeroplane fighter jet is leopard). This is a striking phenomenon. It may indicate that there exists a problem with the way in which the objects data is used to calculate the invariants. It's conceivable that too much information about the object is lost when only vertexes are used. To remind, the vertexes of a given shape are stored in Nx3 matrix which is used to calculate moments that are later transformed into bispectral invariants.

Lets move on to comparing bispectral invariants with other descriptors. Table 2 shows that bispectral invariants perform better than power spectral ones. It was expected, because they retain more information. What is surprising though is the fact that covariance matrix produced results at the same level as bispectral invariants. Covariance matrix was supposed to be the lower boundary for performance in shape recognition. It is in fact the simplest possible invariant that retains only information about object orientation in space.

The same inference may be drawn from precision-recall plot presented in Figure 2. For comparison, also SHELLS descriptor from[9] is presented. It is the worst performing tool in PSB, but it still turns out to be better than bispectral invariants. It is important to note that there is a sizable difference in performance of power spectral invariants (ps) and SHD (spherical harmonic descriptor) presented in PSB[9] (see Table 2). Those two are actually the same tools, but calculated in different ways. This confirms former presumption that way in which bispectral invariants are counted may be key factor to their poor performance.

It seems intuitive that for a given invariant the more discriminative power it has, the more varied values it will attain over a large collection of models. That is the reason why for each coordinate of LBIV of order 11 standard deviation was computed (over the whole PSB database). Then a new vector was constructed containing only those elements of LBIV that have standard deviation greater than 1. That vector consisted of 35 coordinates whereas the whole LBIV contains 67. Performance of such a vector (sig_bi - 'significant' bispectral invariants) is shown in Figure 3 and Table 3. As expected, it is almost the same as performance of LBIV of order 8. Two conclusions can be drawn from this. First, it is possible to use significantly smaller vector of invariants without loss of discriminative power. Second, there exist a strong relationship between standard deviation of a given bispectral invariant and its effectiveness in shape matching.

On the other hand, according to data, the higher the order of bispectral invariant vector the greater standard deviation can be found within its coordinates. Earlier we have established that bispectral invariants of order higher than 8 do not improve the performance. Bearing those two in mind, it is apparent that the higher standard deviation doesn't mean better

discriminative power. But it turns out that there exits certain standard deviation level which can serve as significance test for a given coordinate of long bispectral invariant vector. To confirm that proposition another test was conducted. Each coordinate of LBIV of order 14 was used as a separate matching descriptor (108 of them in total). Then, the second Tier statistic was computed and with respect to it the coordinates were sorted. It was conducted with both train1 and base test classification, so two sets of data were generated. In both cases, about the first half of coordinates had standard deviation greater than 0.1 with the opposite result for second half. The data can be seen on this project website.

Reflection of a given object means complex conjugation of its bispectral invariant. The question arises if this is what we want in terms of object matching. On one hand, this means that bispectral invariants retain a lot of information and will be able to distinguish between greater spectrum of objects. But on the other hand, reflections are really similar (like left and right shoe). So it is desired to have really close invariants for them. Complex conjugates could fail to meet that condition (as we calculate the distance as norm of their difference). This is the reasoning behind the following experiment. For each object, its bispectral invariant was transformed according to formula: $x + i \cdot y \longmapsto x + i \cdot |y|$. Thanks to this, the complex conjugate effect was turned off. The performance of modified invariants was exactly the same as before transformation (all the statistics and plots were equal). It turned out that imaginary parts of bispectral invariants are tiny compared to real parts and therefor insignificant. This is confirmed by the figures in Table 4. Presented statistics pertain to the quantity $r(object)$ computed at every model from PSB:

$$r(object) = \frac{\max |\Im(bi(object))|}{\max |\Re(bi(object))|} \tag{19}$$

Where $bi(object)$ is LBIV of order 14 of a given object, $\Im(v)$ and $\Re(v)$ are vectors with respectively imaginary and real parts of coordinates of vector v.

More direct proof was obtained for insignificance of aforementioned imaginary parts. It turned out that the performance remained unchanged when distances between objects were calculated using only real parts of bispectral invariants. What is more, the performance got worse when the imaginary parts were scaled by the factor of 1000 and 10000. After scaling by factor of 100, the difference was negligible. The performance mentioned here was measured with the NN, 1stTier, 2ndTier and Emeasure statistics. To sum up, imaginary parts of bispectral invariants can be used to distinguish between reflections of models, but are of no importance in object matching.

The next experiment pertains to class matching ability of bispectral invariants. Having two disjoint classifications of objects (907 models each) with over a dozen of common classes a following test can be run. Within arbitrary classification, average invariant can be computed for each class of models. Then, a question arises: how well does this average match the average from other classification? So for each of common classes, the classes from other classification were sorted with respect to distance and the position of respective class in this sequence was stored. This position is our random variable. Sample statistics counted from train1 and base test classification perspective are presented in Table 5. The lower the median/mean the better, because it means that the respective class from other classification was closer (and matching better). It is easy to notice that LBIV of order 4 produces really good results. It is pretty surprising considering all the previous results. It might be interesting to check if this result holds in a greater amount of experiments. Also, in case of both classifications, LBIV of orders 6 and 7 are equally bad and worst than all the others.

In all the aforementioned experiments, calculations were made using only the vertexes of each object. It is reasonable to ask if the use of information about the faces could improve the performance of bispectral invariants. That is why the following test was run. Having an arbitrary object, a centroid of each of its faces has been computed and added as a new vertex. In that way, a new extended list of vertexes was generated and using it bispectral invariants were calculated. The results of their performance can be seen in Table 6. Presented figures prove that there is no significant difference if new vertexes are added. The conclusion is that initial nets of vertexes are dense enough to contain sufficient information about objects.

It is possible that one aspect of the method of calculations adopted in Shilane *et al*[10] could be used to enhance the performance of any rotation invariant: specifically, this is the idea of spherical shells. Namely, any object after moving to the origin and size normalization, can be divided into layers by concentric spheres. Then for each layer, invariant can be calculated and the outcome concatenated. As a result, new invariant is obtained. This is a kind of implicit filtering that can explain why SHD performs much better than PS, though they are using the same underlying spherical harmonic magnitude.

4. CONCLUSIONS AND FUTURE WORK

In this paper experimental results are presented pertaining to discriminative power of bispectral invariants. According to the figures shown, they perform worse than expected witth the given method of calculation (15). Their effectiveness using (15) is no better then that of a covariance matrix invariant which is one of the simplest and limited ones. It also turned out that there is no significant increase in performance when bispectral invariants of orders higher than 8 are used. What is more, a sizable amount of long bispectral invariant vector coordinates can be dropped without loss to the results. They are easy to find as standard deviation may be employed as a criterion. Another finding is that imaginary parts of bispectral invariants have no influence on their behavior in object matching (although they can be used to distinguish between reflections). The method used for calculating bispectral invariants may be responsible for their bad performance. It would be interesting to check whether the filtering approach adopted in Shilane *et al*[10] will work as well with bispectral invariants as it did with spherical harmonic magnitude invariants. This is a good starting point for future work.

In the experiments with bispectral invariants, it was necessary to use overlapping layers to retain rotation invariance. This has additional advantage. Overlapping layers make the new invariant change values when different layers of object are rotated by different angle. With this approach, new invariant is obtained with two degrees of freedom (number of layers and overlapping factor). It is possible that there exist such a combination of those two that could improve a performance of a given initial invariant (that was the case with low order bispectral invariants). This may be an interesting idea to investigate.

REFERENCES

[1] Galvez, J. M. and Canton, M., "Normalization and shape recognition of three-dimensional objects by 3d moments," *Pattern Recognition* **26**, 667–681 (1993).

[2] Flusser, J., Suk, T., and Zitová, B., [*Moments and Moment Invariants in Pattern Recognition*], Wiley & Sons Ltd., Chichester (2009).

[3] Xu, D. and Li, H., "Geometric moment invariants," *Pattern Recognition* **41**, 240–249 (2008).

[4] Schønefeld, V., "Spherical harmonics." `heim.c-otto.de/~volker/prosem_paper.pdf` (2005). [Online; accessed 03-October-2011].

[5] Mohlenkamp, M., "A users guide to spherical harmonics." `www.ohio.edu/people/mohlenka/research/uguide.pdf` (2010). [Online; accessed 03-October-2011].

[6] Kakarala, R. and Mao, D., "A theory of phase-sensitive rotation invariance with spherical harmonic and moment-based representations," in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], 105–112 (2010).

[7] Hamermesh, M., [*Group theory and its application to physical problems*], Dover (1962).

[8] Kakarala, R., Kaliamoorthi, P., and Li, W., "Viewpoint invariants from three-dimensional data: the role of reflection in human activity understanding," in [*IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*], 57–62 (2011).

[9] Shilane, P., Min, P., Kazhdan, M., and Funkhouser, T., "The princeton shape benchmark," in [*Shape Modeling International*], 167–178 (2004).

[10] Kazhdan, M. M., Funkhouser, T. A., and Rusinkiewicz, S., "Rotation invariant spherical harmonic representation of 3d shape descriptors," in [*Symposium on Geometry Processing*], 156–165 (2003).

FIGURES AND TABLES

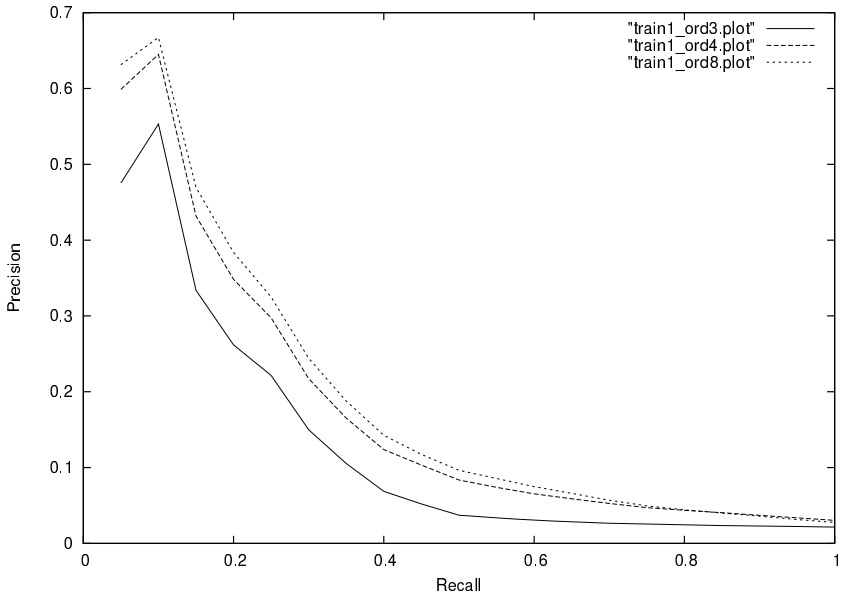

Figure 1. Precision-recall plot for bispectral invariants of orders 3,4 and 8, train1 classification.

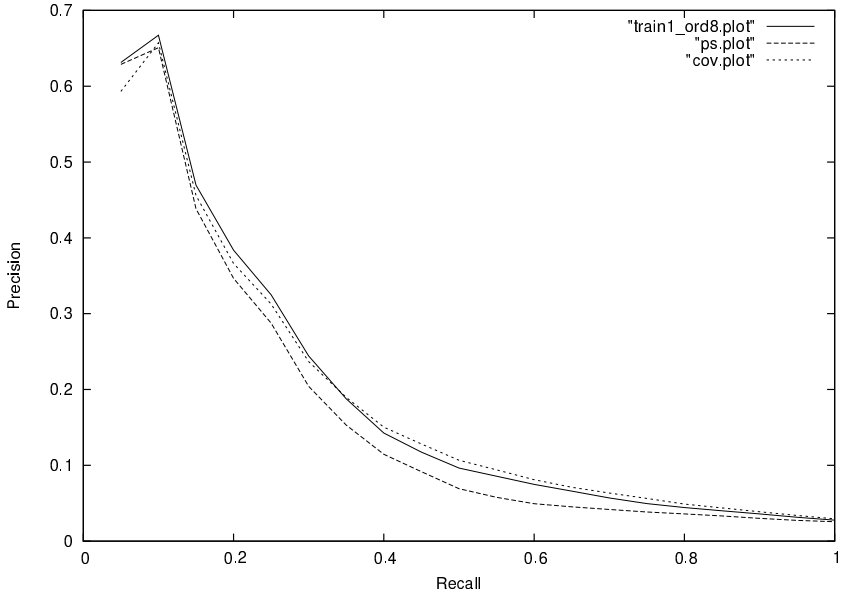

Figure 2. Precision-recall plot for power spectral, bispectral (order 8) and covariance matrix invariants.

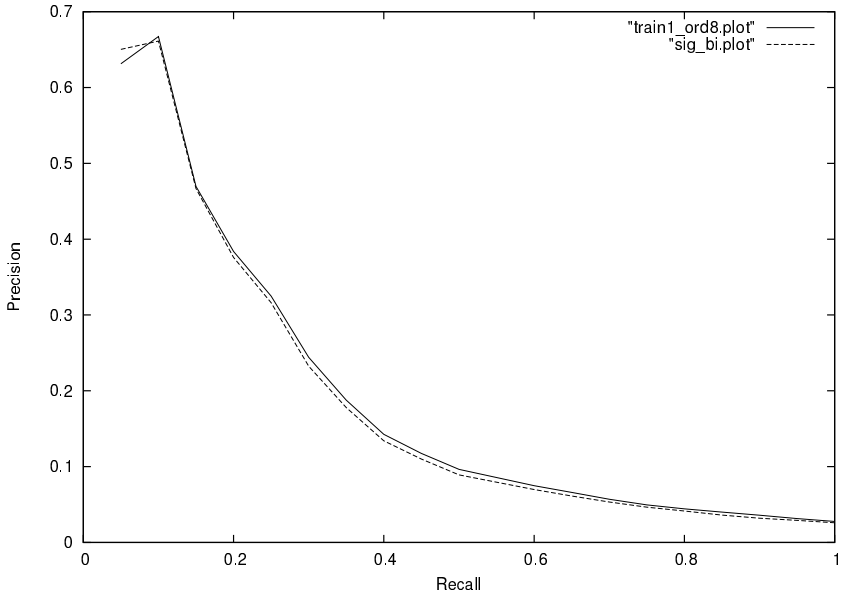

Figure 3. Precision-recall plot for subvector of LBIV of order 11 containing only those coordinates that have standard deviation greater than 1 (sig_bi.plot). Plot of bispectral invariants of order 8 presented for comparison. Generated with train1 classification.

Table 1. Statistics for LBIV's of orders 3-14, train1 classification

Order	NN	1st Tier	2nd Tier	E-measure
3	0.042	0.030	0.052	0.033
4	0.132	0.072	0.122	0.074
5	0.151	0.076	0.130	0.077
6	0.162	0.084	0.137	0.084
7	0.166	0.084	0.136	0.083
8	0.175	0.087	0.136	0.085
9	0.172	0.088	0.136	0.084
10	0.175	0.086	0.135	0.082
11	0.174	0.085	0.135	0.081
12	0.176	0.086	0.134	0.081
13	0.180	0.086	0.133	0.080
14	0.182	0.085	0.134	0.080

Table 2. Statistics for power spectral invariant (PS), LBIV of order 8, covariance matrix invariant (COV). SHELLS and SHD are taken from.[9] First three were computed with train1 classification. SHELLS and SHD with test classification.

Order	NN	1st Tier	2nd Tier	E-measure
PS	0.147	0.067	0.108	0.066
LBIV ord8	0.175	0.087	0.136	0.085
COV	0.165	0.088	0.146	0.087
SHELLS	0.227	0.111	0.173	0.102
SHD	0.556	0.309	0.411	0.241

Table 3. Statistics for LBIV of order 8 and subvector of LBIV of order 11 containing only those coordinates that have standard deviation greater than 1 (SIG_BI). Computed with train1 classification.

Order	NN	1st Tier	2nd Tier	E-measure
LBIV ord8	0.175	0.087	0.136	0.085
SIG_BI	0.180	0.081	0.131	0.079

Table 4. Statistics for the ratio between the biggest absolute values of imaginary and real parts of coordinates of LBIV of order 14, calculated for each object from PSB.

min	mean	max	std	quantile(0.75)	quantile(0.9)	quantile(0.95)
0	0.0012	0.1785	0.0066	0.000418	0.0020	0.0050

Table 5. Class matching results obtained with LBIV of orders 3-11. IQR - interquartile range - is the difference between third and first quartile. It measures the dispersion of data and is resistant to extreme values in contrast to standard deviation.

Train1 classification perspective (maximal value 92)									
Order:	3	4	5	6	7	8	9	10	11
min	1	1	1	1	1	1	1	1	1
median	32	2	29	43	43	39	32	32	32
max	58	32	40	56	56	50	47	47	47
IQR	25.25	3.25	33.75	46.75	42.75	37.75	34.5	34.5	33.75
mean	31.33	4.1	22.57	31.95	31.19	29.62	27.05	26.48	26.71
std	17.93	6.66	15.85	22.13	21.01	19.05	17.39	17.01	17.1
Base test classification perspective (maximal value 90)									
Order:	3	4	5	6	7	8	9	10	11
min	1	1	1	1	1	1	1	1	1
median	27	4	19	29	29	19	15	14	15
max	58	63	77	77	77	66	64	55	55
IQR	37	4.25	27	39.5	39.5	38.5	19.75	20.5	20
mean	28.48	6.52	18.00	27.38	27.48	23.57	17.24	16.52	18.00
std	19.44	13.24	19.24	21.48	21.37	19.87	14.55	13.46	15.38

Table 6. Statistics for bispectral invariants calculated in normal way (n'Number') and with the additional vertexes obtained as centroids of faces (f'Number'). Train1 classification was used.

Order	NN	1st Tier	2nd Tier	E-measure
n4	0.132	0.072	0.122	0.074
f4	0.144	0.073	0.121	0.071
n6	0.162	0.084	0.137	0.084
f6	0.162	0.084	0.135	0.080
n8	0.175	0.087	0.136	0.085
f8	0.175	0.090	0.135	0.083