

A Statistical Model to Describe Invariants Extracted from a 3-D Quadric Surface Patch and Its Applications in Region-Based Recognition

G.Y. Wang*, Z. Houkes†, P.P.L. Regtien†, M.J. Korsten†, G.R Ji*

(* Ocean University of Qingdao, P. R. of China)

(† University of Twente, The Netherlands, E-mail: Z.Houkes@el.utwente.nl)

Abstract

A statistical model, describing noise-disturbed invariants extracted from a surface patch of a range image, has been developed and applied to region based pose estimation and classification of 3D quadrics. The Mahalanobis distance, which yields the same results as a Bayesian classifier, is used for the classification of the surface patches. The results, compared with the Euclidean distance, appear to be much more reliable.

1 INTRODUCTION

Invariants have been explored in varieties and applied for object recognition [5][6][7][10]. In practice, there exists uncertainty in the extracted invariants because of noise disturbances in the data acquisition. At all events the models described by invariants can be easily distinguished from each other, the Euclidean distance is usually used as a measure in comparison for recognition [5]. However, if some of the registered models are close together in terms of invariants, the results of classification by the measure of the Euclidean distance might become unreliable. Although topics on uncertainty in parameter estimation have attracted attention from researchers [1][2][4][8], realization of an optimal recognition system still depends on the objective of its applications.

One of the applications of machine vision is the task of classifying man-made mechanical elements, in which case such man-made elements can be described with quadric and planar primitives. Usually the scene is segmented into a set of surface patches and the input of the recognition system is a region-based data set [1][3]. Since all objects share the same geometric description from a planar patch, even small differences among the geometric invariants of quadric patches might become dominant for the recognition. Moreover, retrieval of pose information is usually simultaneously required in the machine vision tasks.

Concerning the above applications of the underwater robot 3-D vision system [13], in this paper, the uncertainty of invariants extracted from a set of noisy 3-D data points of a quadric surface patch is studied in a statistical framework. A statistical model to describe the uncertainty of the estimated invariants is formulated. Based on this model, an optimal Bayesian-ruled classification system has been developed to classify quadric surface patches. In this work, it is assumed that objects have been described with quadric primitives. The input consists of 3-D noisy

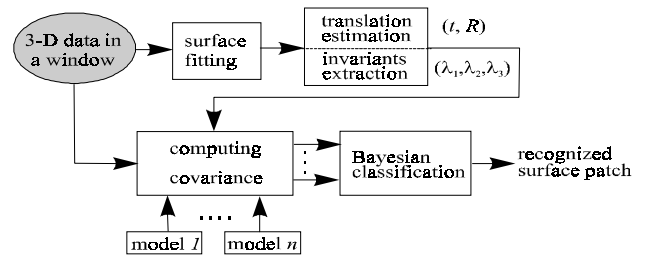


Fig. 1 A block diagram of the implementation of the region-based recognition system

surface points from one of those surface patches, i.e., it is a region-based data set.

Quadric and polynomial representation have been studied and applied for surface description and object recognition [3][6][9]. In case of a surface, expressed as an implicit quadric form $x^T Ax + v^T x + k = 0$, where $x = (x, y, z)^T$ is a surface point, the geometric invariants can be unified as the three eigenvalues of A , which geometric interpretation is explicit. In this paper, a feature descriptor of a quadric surface will be represented by the vector of these three eigenvalues.

At the first step (see Fig. 1), a normalization process is addressed for surface fitting, by which the estimates of the surface parameters in world coordinates can be transformed into the estimate of a standard form which is defined in model coordinates, combined with the retrieval of

the Euclidean transformation parameters. The approach of invariants extraction by normalization for surface fitting is discussed in Section 2.

The statistical characteristics of invariants, extracted from a quadric surface patch, are studied in Section 3. Based on a statistical model, describing the perturbances in the estimates of the invariants by a normally distributed pdf, an optimal region-based recognition system using a Bayesian-decision rule is implemented. The functional description of such a recognition system is illustrated in Fig.1. Experimental results with synthetic data are shown in Section 4. Finally, some remarkable elements and conclusions are summarized in Section 5.

2 INVARIANTS EXTRACTION BY NORMALIZATION IN SURFACE FITTING

To describe the geometric invariants of a quadric surface patch, the quadric is represented in model coordinates, using the standard form, by

$$x_s^T A_s x_s + 1 = 0 \quad (1)$$

with A_s a diagonal matrix. The invariants to be extracted are just the diagonal elements of A_s .

Given a set of 3-D points measured from an arbitrary posed surface patch in world coordinates, firstly the surface fitting is carried out. By applying coordinate transformations, the fitted quadric surface patch is expressed in standard form (1). After this so-called normalization process, the invariants are computed from the standard form as an estimate and put in a feature vector $q = (\lambda_1, \lambda_2, \lambda_3)^T$ for classification.

2.1 Extraction of Invariants by normalization

A quadric surface observed in world coordinates is expressed in general form by

$$x^T A x + v^T x + k = 0 \quad (2)$$

with $k = const$ the constraint. The normalization is implemented with two steps as below.

By applying the coordinate translation $x = x' + t$, the expression (1) becomes

$$x'^T A x' + (2t^T A + v^T)x' + t^T A t + v^T t + k = 0 \quad (3)$$

The surface is described in the coordinates of x' as

$$x'^T A x' + k' = 0 \quad (4)$$

$$t = -\frac{1}{2} A^{-1} v \quad (5)$$

$$k' = t^T A t + v^T t + k \quad (6)$$

A rotation $x' = R x_s$ is used to diagonalize the matrix A of (4), which result can be used to compute eigenvalues. Substituting the expression for x' in (4), dividing (4) by k' and substituting t of (5) into the resulting expression, the diagonal matrix A_s in the standard form of (1) is derived from

$$A_s = \frac{R^T A R}{k - \frac{1}{4} v^T A^{-1} v} \quad (7)$$

In the normalization process, the translation t and the rotation R are also estimated.

2.2 Applying gradient weighted least-square fitting as MLE solution

The *gradient weighted least-squares* fitting [12] is applied to obtain the estimates of the parameters of A and v in world coordinates. This method will be reformulated in terms of *maximum likelihood estimation* (MLE). The advantage in using MLE solution is two-fold:

- (1) The solution of a MLE in one parameter space can be obtained with the MLE in another parameter space by parameter transformation, assuming the transformation is a one-to-one map [14]. Therefore, among different coordinates, the estimates \hat{A} and \hat{v} of A and v , obtained from surface fitting, can be transformed by their dependency on the pose transformation. Therefore the estimate of invariants by (7) is an MLE solution and independent of the measuring coordinates, also the estimate of translation and rotation are MLE.
- (2) The asymptotic distribution of the invariants, obtained with MLE, is normal. This suggests that a statistical model with a normal *pdf* to describe the estimate of the invariants is reasonable.

Given a surface function $h(x, \theta) = 0$, with $h(\cdot)$ the left part of (2) and θ the vector with the surface parameters described as $\theta = (a_{11}, a_{22}, a_{33}, a_{12}, a_{13}, a_{23}, v_1, v_2, v_3)$, the cost function for the gradient weighted least-square fitting is:

$$C(\cdot) = \sum_{i=1}^n w_i h_i^2 = \sum_{i=1}^n \frac{h_i^2}{\|\nabla h_i\|^2} \quad (8)$$

where $h_i = h(x_i, \theta)$ and w_i is the so-called ‘‘gradient weight’’. The Taylor expansion of h_i with respect to x_i at the true surface point x_{i0} and the true parameter θ_0 yields the first order approximation described by $h_i = h(x_i, \theta_0) \approx (\nabla h_{i0})^T \Delta x_i = \|\nabla h_{i0}\| n_i \cdot \Delta x_i$, where n_i is the unit vector of the gradient at x_{i0} and Δx_i is the noise. Defining a random variable $d = n \cdot \Delta x$, which in fact is the projected distance Δx of x with respect to its original x_0 , the cost function of (8) can be expressed as the sum of d_i^2 over the input points. Supposing Δx is Gaussian distributed with $E[\Delta x] = 0$ and $Cov[\Delta x] = \sigma^2 I$, then d is also Gaussian distributed with $E[d] = 0$ and $Var[d] = \sigma^2$. Therefore, under first order approximation, the estimate of θ can be interpreted as an MLE with the

joint likelihood function of the probability density function of $\{d_i\}$.

The method of “bias-corrected” estimation [8][12], combined with estimation of σ^2 , was improved and subsequently applied, by which reliable “bias-corrected” results were obtained.

3 INVARIANTS ESTIMATION WITH A STATISTICAL MODEL

In this section the uncertainty in the invariants is discussed within a statistical perturbation framework. The following analysis gives rise to a statistical description of the estimated invariants by a *normally distributed* probability density function resulting in a formulation for the computation of the covariance matrix.

3.1 Perturbation of the invariants in measurement

Given the estimates of the surface parameters \hat{A} and \hat{v} in measuring coordinates, according to the normalization process, the estimate of the invariants is extracted by computing the eigenvalues of the matrix

$$\hat{A}'_s = \frac{\hat{A}}{k - \frac{1}{4}\hat{v}^T \hat{A}^{-1} \hat{v}} \quad (9)$$

The above expression is obtained from (7) by rotating the diagonal matrix A_s to get $A'_s = R^{-T} A_s R^{-1}$. Denoting A'_s as the noise-free matrix and $\Delta A_s = \hat{A}'_s - A'_s$ as the so-called *perturbation* matrix, the eigenvalues of A'_s and \hat{A}'_s respectively, are denoted by λ and $\hat{\lambda}$. Assuming the added noise to be under the level at which the first order approximation is practical, according to perturbation theory [11], the perturbation of $\hat{\lambda}$ can be expressed as $\Delta\lambda = \hat{\lambda} - \lambda \approx e^T \Delta A'_s e$, where e is the eigenvector of A'_s associated with λ . Because $\hat{\lambda}$ is independent of the measuring coordinates in the sense of MLE, it will be convenient to use the model coordinates for analysis of the perturbation, in which case $e_1 = (1,0,0)^T$, $e_2 = (0,1,0)^T$, $e_3 = (0,0,1)^T$ and consequently it is

$$\Delta\lambda_i = \Delta a'_{s,ii} = \hat{a}'_{s,ii} - a'_{s,ii}, \quad \text{with } i \in \{1,2,3\} \quad (10)$$

where $a'_{s,ii}$ are the elements of A'_s in model coordinates, $\hat{a}'_{s,ij}$ is obtained by (9), while the \hat{A} and \hat{v} are estimated in measuring coordinates.

Next, we will discuss the first and second order moments of $\hat{\lambda}$, which will be based on the statistical model stating the normal distribution of $\hat{\lambda}$. Especially, the computation of the *covariance matrix* of the feature vector, defined as $q = (\lambda_1, \lambda_2, \lambda_3)^T$, is formulated.

3.2 Formulating the covariance matrix

Supposing the parameters \hat{A} and \hat{v} are obtained in model coordinates, the elements of \hat{v} will be small deviations about zero, because their true values are zero. Since the constraint $k = c$ (c is a constant) is imposed in surface fitting, neglecting the second order term of \hat{v} in the right side of (9), we have

$$\hat{A}'_s = \frac{\hat{A}}{c} \quad (11)$$

Because the parameters in surface fitting are estimated with a reliable “bias-corrected” method and the parameter transformation is linear among different coordinates, so

\hat{A}'_s in (11) can be treated as an unbiased estimate and

according to (10), $\hat{\lambda}_i$ is also unbiased, i.e., we have

$E[\hat{q}] = q$, where q is the true vector describing the model invariants.

To compute the covariance matrix of \hat{q} , first we have to formulate the covariance matrix of the estimated surface parameter vector $\hat{\theta}$, defined as $\hat{\theta} = (\hat{a}_{11}, \hat{a}_{22}, \hat{a}_{33}, \hat{a}_{12}, \hat{a}_{13}, \hat{a}_{23}, \hat{v}_1, \hat{v}_2, \hat{v}_3)^T$. With surface fitting, the estimate $\hat{\theta}$ is obtained by minimizing the cost function of (8), i.e., $\frac{\partial}{\partial \theta} (\sum_{i=1}^n d_i^2) = 0$. Applying the linearization,

solving this equation, we get $\Delta\theta = M_{\theta\theta}^{-1} \sum_{i=1}^n M_{i,x\theta} \Delta x_i$, with

$$M_{\theta\theta} = \sum_{i=1}^n M_{i,\theta\theta}, \quad M_{i,x\theta} = \frac{\partial}{\partial \theta} d^2(\theta_0, x_{i0}), \quad M_{i,\theta\theta} = \frac{\partial^2}{\partial \theta^2} d^2(\theta_0, x_{i0})$$

It should be pointed out that such a linear approximation is based on the assumptions made with respect to the noise level and the use of the “bias-corrected” estimation of $\hat{\theta}$. Now the covariance matrix of $\hat{\theta}$ can be expressed by

$$Q_\theta = E[\Delta\theta\Delta\theta^T] = \sigma^2 M_{\theta\theta}^{-1} (\sum_{i=1}^n M_{i,x\theta} M_{i,x\theta}^T) M_{\theta\theta}^{-T} \quad (12)$$

Combining (10) and (11), the covariance matrix Q_q of \hat{q} can be directly obtained as the upper left 3×3 sub-matrix of the 9×9 matrix Q_θ , which is formulated in model coordinates. Since the transformation between the world coordinates and the model coordinates is simultaneously estimated through normalization in surface fitting, the input points can be transformed into model coordinates with the estimates of R and t to compute Q_θ with (12). Therefore, Q_q is adaptively computable. The variance σ^2 of noise the is, as already indicated in section 2, estimated in surface fitting [12].

Based on the analysis above, a recognition system that uses the invariants as input for classification of quadric surface patches can be implemented in the sense of Bayesian classification. The recognition system consists of the following steps:

1. Carry out surface fitting with region-based input in world coordinates.
2. Extract eigenvalues by normalization, incorporating estimates of rotation and translation.
3. Compute the covariance matrix Q_q for all surface models with the current input simultaneously.
4. Apply the *Bayesian* rule to classify the input surface patches.

The experiments on the recognition system are illustrated in the next section.

4 EXPERIMENTAL RESULTS

The input of our recognition system consists of a set of 3-D points generated from a synthetic quadric. The data-points were sampled within a $2 \times 2 \text{cm}^2$ window at the sampling distance of 0.1cm. The sampled surface patch is shown in fig. 2 as a grey-level image.

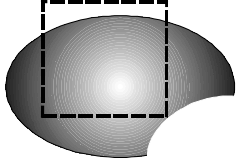


Fig.2 The region-based input from a quadric patch

Gaussian noise was added to the three coordinates of each point. The number of points in the input data is 400. The experiment comprises the steps mentioned in 3.B. We constructed a database with 3 selected elliptical surface primitives, which are close together in feature space. A model in the database is represented by its feature vector q , denoted as m_k with $k \in \{1,2,3\}$. These models are $m_1 = [-0.2, -0.4, -0.6]^T$, $m_2 = [-0.25, -0.4, -0.55]^T$, $m_3 = [-0.3, -0.4, -0.5]^T$ (13)

4.1 Bayesian classification

The input is generated using one of the modeled surfaces with added noise. The estimate of the feature vector is denoted by s . Because s is normally-distributed, assuming the *a priori* probability of the models and the risks be equal, then the Bayesian-decision criteria can be stated as:

Classify s as m_i , when it satisfies the condition

$$(s - m_i)^T Q_i^{-1} (s - m_i) \leq (s - m_j)^T Q_j^{-1} (s - m_j), \quad (14)$$

$$\forall j \neq i, \text{ with } i, j \in \{1,2,3\}$$

For the conditions which hold in our experiments, the *Bayesian* classifier yields the same results as the *Mahalanobis distance* classifier. In the following paragraphs, their difference was ignored.

A typical example of the scenario in our experiments is an input of the classifier generated by the model m_2

Table1. Two instances of classification results. Both inputs were generated from model m_2 . The Bayesian rule was carried out by the measure of Mahalanobis distance. To give a comparison, results of an Euclidean distance classifier are also listed. Results of surface fitting are listed below. The noise level is $\sigma=0.05$. A match is indicated with “+” and a non-match with “-”.

Model for match	m_1	m_2	m_3
Mahalanobis distance to each model	44.84	0.8794	61.04
Classification by Mahalanobis distance	-	+	-
Euclidean distance to each model	0.007748	0.009319	0.02089
Classification by Euclidean distance	+	-	-

(a) Input $s = [-0.2701, -0.4392, -0.6358]^T$

Model for match	m_1	m_2	m_3
Mahalanobis distance to each model	45.36	0.5659	50.33
Classification by Mahalanobis distance	-	+	-
Euclidean distance to each model	0.01365	0.003400	0.003148
Classification by Euclidean distance	-	-	+

(b) Input $s = [-0.2447, -0.3939, -0.4922]^T$

True parameters in world coordinate $\theta_0 = [7.989, 7.273, 10.21, -1.564, 2.896, 4.510, -28.96, -45.09, -204.2, 1]^T \times 10^{-3}$

Estimate in surface fitting : (a) $\hat{\theta} = [8.365, 7.132, 10.35, -0.584, 2.054, 2.273, -39.12, -44.96, -205.4, 1]^T \times 10^{-3}$

(b) $\hat{\theta} = [9.149, 7.882, 10.03, -0.677, 1.307, 2.357, -25.82, -47.13, -202.7, 1]^T \times 10^{-3}$

and a noise level set to $\sigma = 0.05 \text{cm}$. The covariance matrix Q_i , with $i \in \{1,2,3\}$ was simultaneously computed by incorporating the input with the i th model respectively. The experiment consists of 60 trials.

Table 1 shows the classification results for two instances. For a comparison, Mahalanobis distance and Euclidean distance are both listed. The total results are listed in Table 2. The experimental results clearly demonstrates, what was already intuitively expected, that the Euclidean-distance based classification is unreliable and might be entirely corrupted in the cases that models are close together, such as is the case in the scenario involved in the experiment. However, according to the statistical model we proposed, the classification based on the Mahalanobis distance yields highly reliable results.

4.2 Hypothesis testing

Assuming only one model of m_1 in the database, given an input in feature vector, it is required to judge whether it belongs to m_1 or not. This is the case concerned in our experiments to apply the hypothesis testing. For those samples generated from m_1 , the quantitative normal pdf. can be evaluated by comparing the theoretical “significance level” and the experimental result. Given an input s , the hypotheses are stated as $H_0 : s = m_1$ and $H_1 : s \neq m_1$. All 120 samples were generated from the primitive described by m_1 . For each input, either H_0 or H_1 was accepted by comparing the Mahala-

Table2. Total results of classification. All inputs were generated from m_2 under the noise level $\sigma=0.05$.

Model for match	Classification by Mahalanobis distance	
	$\sigma=0.05$	$\sigma=0.02$
\bar{m}_1	3.3%	1.7%
\bar{m}_2	97%	99%
\bar{m}_3	0	0
Total error	3.3%	1.7%

Table3. Results of hypothesis testing. Samples were generated from m_1 . For a given α , the account of acceptance of hypothesis H_1 were expressed by a ratio. The results under noise level $\sigma=0.05$ and $\sigma=0.02$ were both listed.

Prediction		$\alpha=0.01$	$\alpha=0.05$	$\alpha=0.1$	$\alpha=0.2$
		Result	$\sigma=0.05$	0.032	0.13
$\sigma=0.02$	0.020		0.080	0.21	0.36

nobis distance with a threshold associated with the significance level. For different significance levels α , the results are listed in Table 3 for noise levels $\sigma=0.05$ and $\sigma=0.02$.

From the results obtained with samples of m_1 , we can see that the actual results of α are larger than the theoretical prediction, which means that the actual uncertainty of the invariants is greater than that of the statistical description, But the errors were controlled. We believe that the normal pdf described in this paper is a practical approximation for analytical use.

5 CONCLUSIONS

This paper contributes a study on invariant feature extraction and recognition based on a statistical model. The geometric invariants are extracted through a normalization process, by which the rotation and translation between the model coordinates and the world coordinates can be estimated at the same time. The covariance matrix of the feature vector is formulated, which will also be useful in applications with pose retrieval. An analytical description for

the estimation of the invariants of quadric surfaces is given, which is combined with a normally distributed disturbance model. The way to compute the covariance matrix simultaneously with the inputs was presented. Based on this model, the Bayesian classification has been applied to the region-based recognition of a synthetic quadric surface patch. The issued problem and our statistical solution are especially significant for a region-based recognition system, which is believed to be important for the application of machine vision at a data-driven level.

REFERENCES

- [1] R. Bolle and D. Cooper, “Bayesian recognition of local 3D shape by approximation image intensity functions with quadric polynomials”, *IEEE Tr. PAMI*, vol. 16, no. 4, pp. 418-429, July, 1984.
- [2] R. Bolle and D. Cooper, “On optimally combining pieces of information, with applications to estimating 3D complex-object position from range data”, *IEEE Tr. PAMI*, vol. 8, no. 5, pp.619-638, Sept., 1986.
- [3] O. D. Faugras and M. Hebert, “The representation, recognition and locating of 3-D objects”, *Int. J. Robotics Res.*, vol. 5, no.3, pp.27-52, 1986.
- [4] W. Forstner, “Reliability analysis of parameter estimation in linear models with application to measurement problems in computer vision”, *CVGIP*, vol. 40, pp.273-310, 1987.
- [5] David Forsyth, et al, “Invariant Descriptors for 3-D Object Recognition and Pose”, *IEEE Tr. PAMI*, vol. 13, no. 10, pp. 971-990, Oct. 1991.
- [6] D. Keren, J. Subrahmonia and D.B.Cooper, ”Robust object recognition based on implicit algebraic curves and surfaces”, *IEEE Conf. On CVPR*, pp. 791-794, Champaign, June 1992.
- [7] D. Keren, “Some new invariants in computer vision”, *IEEE Trans. PAMI*, pp.1143-1149, Nov. 1994.
- [8] Kenichi Kanatani, “Statistical analysis of geometric computation”, *CVGIP: Image Understanding*, vol.59, no. 3, pp. 286-306,1994.
- [9] G. Taubin, “Estimation of planar curves, surfaces and non-planar space curves defined by implicit equation with applications to edge and range image segmentation”, *IEEE Trans. PAMI*, vol. 13, pp. 1115-1138, Nov., 1991.
- [10] G. Taubin and D.B. Cooper, “2D and 3D object recognition and positioning system based on moment invariants”, *Geometric Invariance in Machine Vision*, J. Mundy and Zisserman, eds., pp. 375-397, 1992.
- [11] G.W. Stewart, An introduction of matrix computation, Science & Technology, Shanghai, 1980.
- [12] Z. Zhang, “Parameter estimation techniques: a tutorial with application to conic fitting”, *Image and vision Computing*, vol. 15, pp.59-76, 1997.
- [13] G. X. Zheng, et al, “Laser difference-frequency scanning 3-D vision sensing system for underwater robot”, *Proc. of ROV’91*, pp. 90-96, Florida, 1991.
- [14] K.T. Fang, Practical multi-variables statistical analysis, East China Normal University, 1989.