*Sensors & Transducers, Vol. 160, Issue 12, December 2013, pp. 660-666*

www.sensorsportal.com

**Sensors & Transducers**



*http://www.sensorsportal.com*

# A Coarse-to-Fine Keypoint Detection Method for 3D Model

**[1] Hui ZENG, [2] Han WU**

[1] School of Automation and Electrical Engineering University of Science and Technology Beijing, Beijing 100083, China

[2] China National Computer Products Quality Supervising Test Center, Beijing 100083, China

[1] E-mail: hzeng@163.com

*Received: 30 September 2013 /Accepted: 22 November 2013 /Published: 30 December 2013*

**Abstract:** This paper proposes a coarse-to-fine 3D keypoint detection method based on Principal Component Analysis and Harris operator. At first, the local neighborhood of each vertex is determined according to the conception of “ring”. Then the Principal Component Analysis method is performed on the local surface, and the ratio between the first two principal axes of the local neighboring surface is used for selecting candidate keypoints. Finally we compute the Hessian matrix of the local surface through paraboloid fitting, and the Harris operator is used to obtain final keypoint. Extensive experimental results have testified the effectiveness of the proposed method, and it is more robust to noise, especially to high level noise. *Copyright © 2013 IFSA.*

**Keywords:** 3D model, Keypoint detection, Principal Component Analysis, Hessian matrix, Harris operator.

## 1. Introduction

In recent years, with the rapid development of 3D model acquisition and processing technique, the 3D models have been widely used in virtual reality, human identification, video surveillance, electronic commerce, the protection of ancient buildings, medical diagnosis and intelligent robot, etc. As the 3D model generally requires a huge amount of data for representation in the raw data format, the computational complexity has become one of the bottlenecks preventing the development and application of 3D model related techniques [1-5]. So before performing analysis of the 3D model, some points are often labeled to represent the 3D model effectively. The traditional solution method for reducing computational cost was to performing a sparse sampling or random selection of the surface to reduce number of the 3D points to be processed [3, 4]. This kind of method can’t select the best keypoints in terms of their repeatability and distinctiveness, and some useful information may be lost. Then the results of subsequent registration or identification steps will be influenced. To represent the 3D model effectively, recent techniques began to detect some salient points, lines or surfaces according to their distinctiveness and uniqueness [5]. For a 3D object, different 3D model acquisition devices may obtain different 3D data with different scale, position, rotation angle and topology structure. Furthermore, the acquisition process of 3D models may exist partial occlusions. Any the above factors may increase the difficulty of 3D keypoint detection.

The aim of the keypoint detection method is selecting the salient points with highly repeatability according to their distinctiveness and uniqueness [6]. That is to say, for different 3D models of a same object, their 3D keypoint detection results should be similar, and the 3D keypoint detection results should be not sensitive to noise.

Mian et al. give a keypoint detection method based on Principal Component Analysis [7, 8]. They



Article number P_1671

define the local neighborhood of the vertex is the points in the sphere of radius $r$ centered at the vertex. For a vertex, a local surface is cropped depending on the value of radius $r$, and then it is rotated so that the normal of the point is aligned with the positive z-axis. The ratio between the first two principal axes of the local surface is used to detect the keypoint. Finally the keypoint quality is computed to refine the above detection result. Ho and Gibbins proposed a method for extraction salient local features based on the local curvature [9]. The points with high local curvature are selected as feature points. The vertex with higher curvature is selected as keypoint, but its local neighboring region may have uniform shape variation. Under this situation, it cannot provide much information about surface and it shouldn't be selected as keypoint. Castellani et al. proposed a detection method of salient points over several views of object [10]. They performed a series of Gaussian filtering on the 3D meshes, and the salient point was detected if its saliency value was a local maximum and it is higher than the 30 % of the global maximum. Here the saliency measure of the point is defined by the significant motion of the vertices along the direction perpendicular to their local surface. Gelfand et al. present a feature point selection algorithm for automatic alignment of two 3D shapes [11]. At first, they compute the integral volume descriptor of each point. Then the feature points are selected according to the uniqueness of the descriptor value at the point.

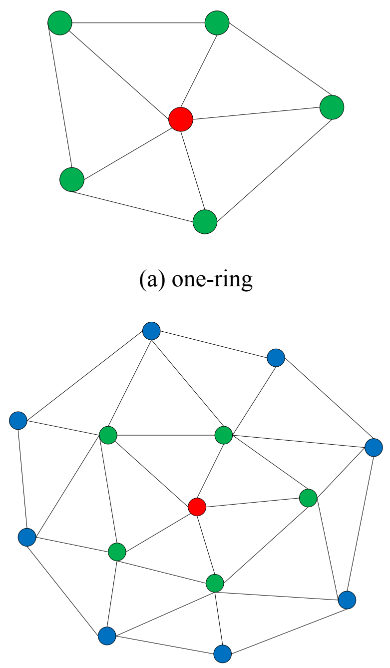
Inspired by the work described above, this paper presents a novel coarse-to-fine keypoint detection method. At first, detect the candidate keypoint using PCA based method. This step is similar to the approach proposed by Mian, and our improvement is to use the conception of "ring" to determine the neighborhood of the vertex. Compared with the neighborhood determination method based on the sphere, the computation burden of the ring based method is smaller. Then the Hessian matrix and Harris operator is used to refine the detection result. Extensive experiments have performed to testify the effectiveness of the proposed method.

The rest of the paper is organized as follows. In section 2, the Candidate keypoint detection based on Principal Component Analysis is introduced. Then fine keypoint detection based on Hessian matrix and Harris operator is presented in Section 3. Section 4 reports the experimental results and some concluding remarks are listed in Section 5.

## 2. Candidate Keypoint Detection Based on Principal Component Analysis

### 2.1. Local Neighborhood Definition

The local neighborhood of the vertex was often chosen depending on the radius of the sphere that is centered at the vertex. That is to say, it can be determined according to Euclidean distance between the vertex and its surrounding surface points. If the distance between the vertex and its surrounding surface point is below the predefined threshold, this surrounding surface point is specified as the neighboring point of the vertex. As the computation burden of this kind of method is huge, we use the conception of "ring" to define the local neighborhood of the surface point in this paper [9]. For example, if the local neighborhood of the vertex is defined by one ring, it contains the vertexes that lie in the first ring of the vertex. As shown in Fig. 1(a), the five surrounding green points are the local "one-ring" neighborhood of the red vertex. If the local neighborhood of the vertex is defined by two rings, it contains the vertexes that lie in the first and second ring of the vertex. The first ring contains all the direct neighboring surface points of the vertex and the second ring contains all the direct neighboring surface points of the points in the first ring, and so on. As shown in Fig. 1(b), the eight blue points are the "one-ring" neighborhood of five green points, and they are the second ring neighborhood of the red vertex. The blue points and the green points constitute the local "two-ring" neighborhood of the red vertex.



(a) one-ring

(b) two-ring

**Fig. 1.** The example of local "ring" neighborhood.

### 2.2. Candidate Keypoint Detection

In this paper, we use the principle of Principal Component Analysis (PCA) to detect the candidate keypoint. PCA is a useful statistical technique that has been widely used for applications such as dimensionality reduction, lossy data compression, feature extraction, image compression, and pattern recognition [12, 13]. It can highlight the similarities and differences of the original data, and it can reduce the number of dimensions with much loss of information. PCA can be defined as orthogonal projection of the data onto a lower dimensional uncorrelated linear space, called principal subspace. The projected data is called principal components. This projective transformation is defined in such a way that the first principal component has the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to the preceding components.

Let $V_i$ ( $i = 1, 2, \cdots, n$ ) be the *i*-th vertex of the 3D mode, and $P_j^{(i)} = [x_j^{(i)}, y_j^{(i)}, z_j^{(i)}]^T$ ( $j = 1, 2, \cdots, n_i$ ) be the points in the local neighborhood of the vertex $V_i$. The mean of the vertex $V_i$ and its neighboring points is given by

$$P_m^{(i)} = \frac{1}{n_i} \sum_{j=1}^{n_i} P_j^{(i)} \tag{1}$$

The covariance matrix of the vertex $V_i$ and its neighboring points is given by

$$C_i = \frac{1}{n_i} \sum_{j=1}^{n_i} (P_j^{(i)} - P_m^{(i)})(P_j^{(i)} - P_m^{(i)})^T \tag{2}$$

Then we perform PCA on the covariance matrix $C_i$. The eigenvalues of the matrix $C_i$ are denoted by $\lambda_1^{(i)}$, $\lambda_2^{(i)}$ and $\lambda_3^{(i)}$ ( $\lambda_1^{(i)} \ge \lambda_2^{(i)} \ge \lambda_3^{(i)}$ ), and the corresponding eigenvectors of $\lambda_1^{(i)}$, $\lambda_2^{(i)}$ and $\lambda_3^{(i)}$ are denoted by $e_1^{(i)}$, $e_2^{(i)}$ and $e_3^{(i)}$. Then we have

$$C_i e_k^{(i)} = \lambda_k^{(i)} e_k^{(i)} \qquad k = 1, 2, 3 \tag{3}$$

Here the matrix $C_i$ is a symmetric and positive semi-definite matrix, so the eigenvalue $\lambda_1^{(i)}$, $\lambda_2^{(i)}$ and $\lambda_3^{(i)}$ are real-valued [14]. The eigenvector $e_1^{(i)}$, $e_2^{(i)}$ and $e_3^{(i)}$ are orthogonal with each other, corresponding to the principal components of the point set $P_j^{(i)}$ and the vertex $V_i$.

Let the matrix of eigenvectors be $E = [e_1^{(i)} \quad e_2^{(i)} \quad e_3^{(i)}]$, and the 3D points can be aligned with its principal axed using

$$Pe_j^{(i)} = E(P_j^{(i)} - P_m^{(i)}), \tag{4}$$

where $Pe_j^{(i)} = [xe_j^{(i)}, ye_j^{(i)}, ze_j^{(i)}]^T$. We define the maximum value of $xe_j^{(i)}$ ( $j = 1, 2, \cdots, n_i$ ) is $xe_{\max}^{(i)}$ and the minimum value of $xe_j^{(i)}$ is $xe_{\min}^{(i)}$, and the maximum value of $ye_j^{(i)}$ is $ye_{\max}^{(i)}$ and the minimum value of $ye_j^{(i)}$ is $ye_{\min}^{(i)}$. Then the ratio between the first two principal axes of the local neighboring surface is

$$\delta = \frac{xe_{\max}^{(i)} - xe_{\min}^{(i)}}{ye_{\max}^{(i)} - ye_{\min}^{(i)}} \tag{5}$$

Because $\lambda_1^{(i)} \ge \lambda_2^{(i)} \ge \lambda_3^{(i)}$, the variance of $xe_j^{(i)}$ is highest than $ye_j^{(i)}$ and $ze_j^{(i)}$, and the variance of $ye_j^{(i)}$ is higher than $ze_j^{(i)}$. So $xe_{\max}^{(i)} - xe_{\min}^{(i)} > ye_{\max}^{(i)} - ye_{\min}^{(i)}$, that is to say $\delta \ge 1$.

For symmetrical surfaces, the value of $\delta$ will be equal to 1 and for unsymmetrical surfaces, the value of $\delta$ will be greater than 1 [8]. In this paper, we detect the candidate keypoint according to the value of $\delta$. For a vertex, if its corresponding value of $\delta$ is larger than the threshold $t$, then the vertex is chosen as the candidate keypoint. Here the parameter $t$ can be chosen by experience. In this paper, we set $t$ to 1.05.

## 3. Fine Keypoint Detection Based on Harris Operator

### 3.1. Harris Operator

Harris operator is firstly proposed by Harris and Stephens for image corner detection [15]. It is robust to illumination and image noise, so it has been widely used in image matching, motion detection, video tracking, 3D reconstruction and object recognition [16]. The Harris operator uses the differential local auto-correlation value with respect to different directions.

Without loss of generality, let the grayscale image be $I$, and an image patch over the area $(x, y)$ is shifted by $(u, v)$. The change of intensity for the shift is measured by

$$E(u,v) = \sum_{x,y} w(x,y) \left[ I(x+u, y+v) - I(x,y) \right]^2, \tag{6}$$

where $E(u,v)$ is the weighted sum of squared differences (SSD), and $w(x,y)$ is the Gaussian window function centered on $(x,y)$. From Equation (6) we can see that, the value of $E(u,v)$ is determined by the variance of gray value between these two patches. For nearly constant patches, $E(u,v)$ is near 0. For distinctive patches, $E(u,v)$ is larger.

Taylor expansion can be used to compute the shifted patch $I(x+u, y+v)$. Then $E(u,v)$ can be computed by

$$E(u,v) \cong \begin{bmatrix} u & v \end{bmatrix} \bullet M \bullet \begin{bmatrix} u \\ v \end{bmatrix} \quad (7)$$

Here $M$ is 2×2 matrix computed from image derivatives:

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \quad (8)$$

where $I_x$ and $I_y$ are the partial derivatives of $I$.

By analyzing the eigenvalues of matrix $E$, the interest point should have two large eigenvalues. Because the computation cost of the eigenvalues is expensive, Harris and Stephens proposed the following measure of corner response:

$$R = \det(M) - k[trace(M)]^2, \quad (9)$$

where $\det M = \lambda_1 \lambda_2$, $traceM = \lambda_1 + \lambda_2$ and $k$ is an empirically determined constant.

Because the parameterization of 3D model is a very complex problem, the topology representations of 3D surface are not unique. The above 2D image based computation method of Harris operator cannot be directly extended to the 3D field.

### 3.2. 3D Fine Keypoint Detection

In this section, we introduce a 3D keypoint detection method using Harris operator. For each detected candidate keypoint, we firstly build its principal coordinate system $O - e_1^{(i)} e_2^{(i)} e_3^{(i)}$ and transform the coordinates of its local neighboring points into its principal coordinate system. Then the vector $e_1^{(i)}$, $e_2^{(i)}$ and $e_3^{(i)}$ become $[1 \;\; 0 \;\; 0]^T$, $[0 \;\; 1 \;\; 0]^T$ and $[0 \;\; 0 \;\; 1]^T$ in its principal coordinate system respectively. As the vector $e_3^{(i)}$ is the eigenvector with the lowest associated eigenvalue, the distribution of rotated neighboring points is close to the plane $e_1^{(i)} O e_2^{(i)}$. That is to say, the rotated points exhibit a good spread in the plane $e_1^{(i)} O e_2^{(i)}$.

Then we translate the local neighboring points so that the vertex coincides with the origin of the principal coordinate system $O - e_1^{(i)} e_2^{(i)} e_3^{(i)}$. A least-square method is performed to fit a quadratic patch to the local neighborhood of the vertex. The form of the fitted paraboloid is

$$f(x,y) = ax^2 + by^2 + cxy + dx + ey + f \quad (10)$$

The Hessian matrix of the local patch can be written as

$$M = \begin{bmatrix} 2a & c \\ c & 2b \end{bmatrix} \quad (11)$$

Then the Harris operator can be computed from the Hessian matrix of the local patch and used to select fine keypoints from the candidate keypoints [17]. The Harris operator is defined as

$$R_{key} = \det(M) - k[trace(M)]^2, \quad (12)$$

where k is set to 0.05. For a 3D candidate keypoint, if $H_{key}$ is larger than threshold $T$, it will be selected as final keypoint. Here the parameter $T$ can be chosen by experience.

## 4. The Proposed Coarse-to-Fine Method

Given a 3D model, the proposed coarse-to-fine keypoint detection method is summarized as follows:

1) Determine the local neighborhood of each vertex, and then perform PCA on each local neighboring surface.

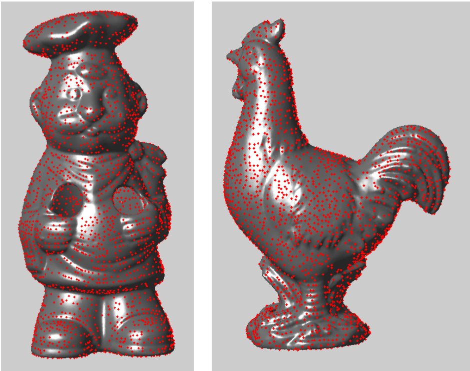
2) For each vertex, compute its corresponding measure $\delta$ using Equation (5), and determine the candidate keypoints using the threshold $t$.

3) For each candidate keypoint, compute its corresponding Hessian matrix and the response of Harris operator.

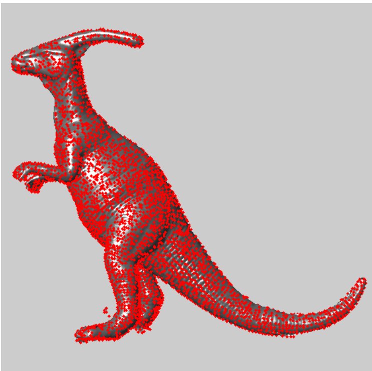
4) Detect the final keypoints of the 3D model. For a candidate keypoint, if its response of Harris operator is larger than the threshold $T$, it is labeled final keypoint.

## 5. Experimental Results

In this paper, our experiments are carried out on the four models download from Mian's website [18, 19]. Fig. 2 and Fig. 3 show the experimental 3D models named as "Chef", "Chicken", "Parasaurolophus" and "T-rex" and their detection results. In Fig. 2, the red dots denote the extracted candidate keypoints using our proposed method. In Fig. 3, the yellow dots denote the extracted final keypoints using our proposed method. From Fig. 2 and Fig. 3 we can see that the number of the final keypoints is significantly smaller than the number of the candidate keypoints, and the distribution of the final keypoints is relative uniform and most of the vertices with higher local shape variation are selected as final keypoints. Table 1 shows the number of vertices, the number of candidate keypoints and the number of final keypoints for each 3D model. Compared with the number of vertices, the number of final keypoints is significantly smaller than the number of vertices and the number of candidate keypoints.



(a) Chef (b) Chicken

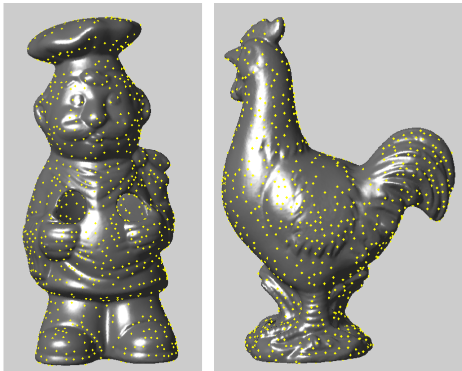


(c) Parasaurolophus

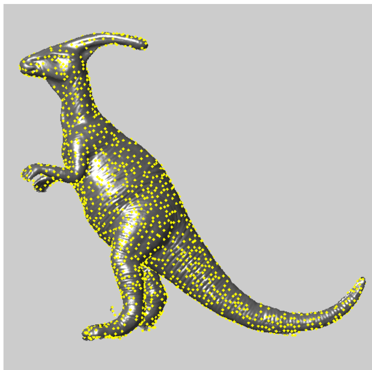
(d) T-rex

**Fig. 2.** The four 3D models and their candidate keypoints.

In order to testify the effectiveness of the proposed coarse-to-fine method, we compared it with PCA based method and Harris operator based method. Here we investigate the robustness of the new proposed method by comparing the respectability of the three methods. Gaussian noise with mean 0 and standard deviation ranging from 0 to 0.12 is added to the coordinates of the 3D models. Fig. 4(a), Fig. 4(b), Fig. 4(c) and Fig. 4(d) are the results of the repeatability of the four 3D models. "PCA+Harris" denotes our proposed coarse-to-fine method. "PCA" denotes the method based on PCA, which uses the ratio between the first two principal axes of the local neighboring surface to detect keypoint. "Harris" denotes the method based on Hessian matrix and Harris operator, which uses the value of Harris operator to detect keypoint. From Fig. 3 we can see that the repeatability of the method "PCA+Harris" is higher than the method "PCA" and "Harris", and the repeatability of the method "Harris" is higher than the method "PCA". Compared with the other two methods, the method "PCA+Harris" is more robust to noise. Especially when the level of the noise is high, the performance of the method "PCA+Harris" is more stable.



(a) Chef (b) Chicken



(c) Parasaurolophus

(d) T-rex

**Fig. 3.** The four 3D models and their final keypoints.

**Table 1.** The number of vertices and the number of keypoints.

| Model | Number of vertices | Number of candidate keypoints | Number of final keypoints |
|---|---|---|---|
| Chef | 176920 | 4638 | 2017 |
| Chicken | 135142 | 3987 | 1734 |
| Parasaurolophus | 184933 | 7512 | 3228 |
| T-rex | 176508 | 5286 | 2283 |

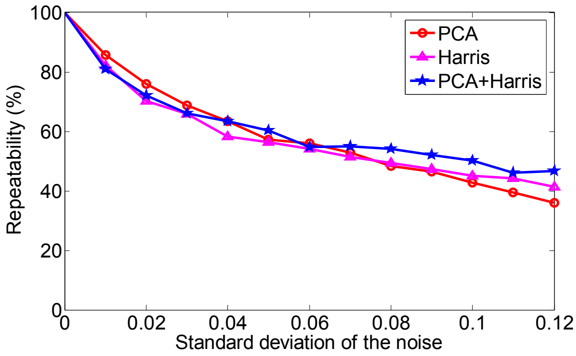



**Fig. 4 (a).** The repeatability of keypoints detected at different noise level: Chef.

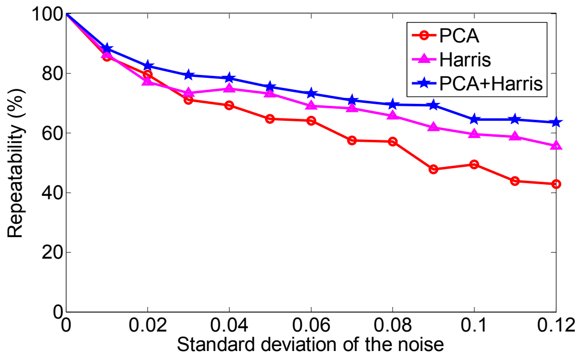



**Fig. 4 (b).** The repeatability of keypoints detected at different noise level: Chicken.

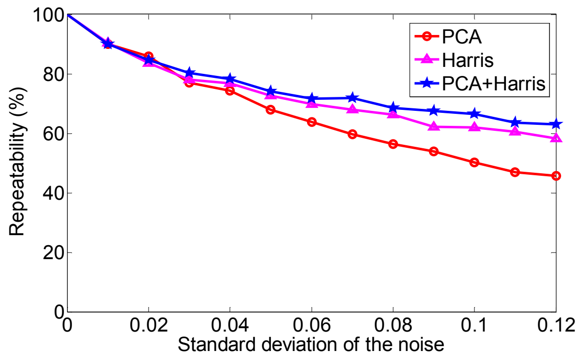



**Fig. 4 (c).** The repeatability of keypoints detected at different noise level: Parasaurolophus.

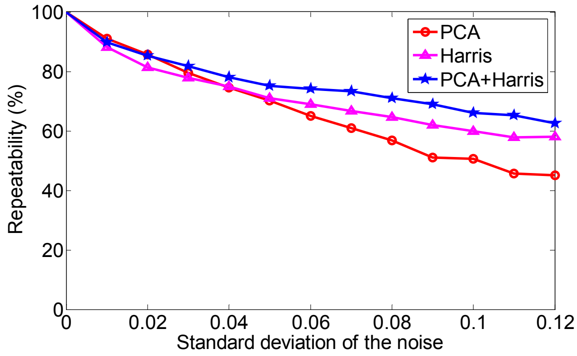



**Fig. 4 (d).** The repeatability of keypoints detected at different noise level: T-rex.

## 6. Conclusion

In this paper, a coarse-to-fine method for 3D keypoint detection is proposed. The local neighborhood of each vertex is determined and PCA is performed on the neighborhood, and then the ratio between the first two principal axes of the local neighboring surface is used to select the candidate keypoint. Finally for each candidate keypoint, its corresponding Hessian matrix is computed and the final keypoint is determined using Harris operator. This paper use the conception of “ring” to obtain the neighborhood of the vertex, and this approach is simple and effective. The effectiveness of the proposed coarse-to-fine keypoint detection method is validated by extensive experiments. Our future work will focus on researching multiscale keypoint detection method based on PCA and Harris operator, the description method of the detected keypoints etc.

## Acknowledgements

This paper is supported by the National Natural Science Foundation of China under the Grant No. 61005009, No. 61175059 and No. 61170116, and the Fundamental Research Funds for the Central Universities under the Grant No. FRF-TP-12-100A.

## References

[1]. R. J. Campbell and P. J. Flynn, A survey of free-form object representation and recognition techniques, *Computer Vision and Image Understanding*, Vol. 81, Issue 2, 2001, pp. 166-210.

[2]. Ming Dong, Yanhua Chen, Salient region detection and feature extraction in 3D visual data, in *Proceedings of the IEEE International Conference on Image Processing*, San Diego, CA, USA, 2008.

[3]. Erdem Akagündüz, Ilkay Ulusoy, Scale and orientation invariant 3D interest point extraction using HK curvatures, in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, September 2009, pp. 697-702.

[4]. Timothy C. Faltemier, Kevin W. Bowyer, Patrich J. Flynn, A region ensemble for 3-D face recognition, *IEEE Transactions on Information Forensics and Security*, Vol. 3, Issue 1, March 2008, pp. 62-73.

[5]. J. J. Koenderink and A. J. van Doorn, Surface shape and curvature scales, *Image and Vision Computing*, Vol. 10, October 1992, pp. 557-565.

[6]. C. Dorai and A.K. Jain, Cosmos – a representation scheme for 3d free-form objects, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, Issue 10, October 1997, pp. 1115–1130.

[7]. Ajmal Mian, M. Bennamoun and R. Owens, Keypoint detection and local feature matching for textured 3D face recognition, *International Journal of Computer Vision*, Vol. 79, No. 1, 2008, pp. 1-12.

[8]. Ajmal Mian, M. Bennamoun and R. Owens, On the repeatability and quality of keypoints for local feature-based 3D object retrieval from cluttered scenes, *International Journal of Computer Vision*, Vol. 89, No. 2-3, 2010, pp. 348-361.

[9]. Huy Tho ho and Danny Gibbins, Multi-scale feature extraction for 3D models using local surface curvature, in *Proceedings of the IEEE Conference on Digital Image Computing: Techniques and Applications*, 2008, pp. 16-23.

[10]. U. Castellani, M. Cristani, S. Fantoni, V. Murino, Sparse points matching by combining 3D mesh saliency with statistical descriptors, *Eurographics*, Vol. 27, No. 2, 2008, pp. 643-652.

[11]. N. Gelfand, N. Mitra, L. Guibas, H. Pottmann, Robust global registration, in *Proceedings of the Eurographics Symposium on Geometry Processing*, 2005, pp. 197-206.

[12]. H. Abdi, L. J. Williams, Principal component analysis, *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 2, Issue 4, 2010, pp. 433–459.

[13]. I. T. Jolliffe, Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed., *Springer*, NY, XXIX, 2002, 489 p.

[14]. Mark Pauly, Markus Gross, Leif P. Kobbelt, Efficient simplification of point-sampled surfaces, in *Proceedings of the Conference on Visualization*, 2002, pp. 163-170.

[15]. C. Harris and M. Stephens, A combined corner and edge detector, in *Proceedings of the 4th Alvey Vision Conference*, 1988, pp. 147–151.

[16]. Miroslav Trajković, Mark Hedley, Fast corner detection, *Image and Vision Computing*, Vol. 16, Issue 2, 1998, pp. 75-87.

[17]. L. Sipiran, B. Bustos, A robust 3D interest points detector based on Harris operator, in *Proceedings of the Eurographics Workshop on 3D Object Retrieval*, 2010, pp. 7-14.

[18]. Ajmal S. Mian, M. Bennamoun and R. Owens, A novel representation and feature matching algorithm for automatic pairwise registration of range images, *International Journal of Computer Vision*, Vol. 66, No. 1, 2006, pp. 19-40.

[19]. Ajmal Mian, M. Bennamoun and R. Owens, 3D model-based object recognition and segmentation in cluttered scenes, *IEEE Transactions in Pattern Analysis and Machine Intelligence*, Vol. 28, No. 10, 2006, pp. 1584-1601.

---