

Optimal similarity and rotation angle retrieval using Zernike moments

Jérôme Revaud, Guillaume Lavoué and Atilla Baskurt

July 16, 2007

Abstract

This paper presents a novel and robust algorithm that retrieves the rotation angle between two different patterns, together with their similarity degree. The result is optimal in the sense that it minimizes the euclidean distance between the two images. This approach is based on Zernike moments and constitutes a new way to compare two Zernike descriptors that improves standard Euclidean approaches.

1 Introduction

Because of the semantic gap, it still remains quite difficult for a computer-vision system to really understand the meaning of a real world image. A relevant way for understanding an image is to recognize the 3D objects which compose it, as well as their pose in the scene. The context of our work is the recognition of 3D objects present on a hand drawn storyboard of a computer-animated film. The issue is here to cross the gap between the 3D model and the cartoonist strokes. Formally, a single panel of the storyboard is a drawing that represents various objects, which in addition are separately known as 3D polygonal meshes in the database. As the storyboard is digitally sketched, the cartoonist creates a distinct layer for each object. In this paper, we focus on identifying precisely for each object its 3D viewing angle along with its rotation angle in the drawing plane (we restrict to the unoccluded case). To reach that goal, we index an important number of views for each 3D object using a global feature descriptor, like in [1, 2]. The objective is then, for each drawing, to determine the most similar view together with the corresponding planar rotation angle.

The choice of Zernike moments as global descriptor is well suited for our problematic. Indeed, Zernike moments are widely used in image processing because of both their robust and optimal ability to describe the information. Rotation invariant pattern recognition using the magnitude of Zernike moments has already been extensively studied [3, 4]. Moreover, a small number of Zernike moments characterizes effectively the global shape of a pattern, which makes them robust to deformations. Some optimized processing method were recently developed to compensate the high cost of Zernike moments computation [5–7].

A recent comparative study can be found in [8]. One Zernike moment contains two kinds of information: magnitude and phase, however, the usual way of comparing two Zernike descriptors in the context of 2D and 2D-3D indexing only considers the moments magnitudes (as it brings invariance to rotation). This loss of information can induce erroneous similarity measures when it deals with relatively symmetric object such as cars or characters.

Moreover, the phase information of Zernike moments can be useful to find the rotation angle between two patterns, together with their similarity. Two approaches have focused on the study of that rotation angle estimation problem. The first one was brought by Kim and Kim [9]. It proved to be very robust with respect to noise even for circular symmetric patterns. Nevertheless, the probabilistic model used to recover the rotation angle makes the approach to be disconnected from a concrete Euclidean distance between the images. The second method was proposed by Kanaya et al. [10] to estimate the direction of arrival (DOA) of a sound using antennas array. However, their approach is not really applicable to our images indexing issue. Besides, those two methods do not consider the possibility of recovering an angle between two different patterns (i.e. in our case, a 3D model view and a sketch).

To solve these issues, we have developed a general and rigorously founded approach for comparing two Zernike descriptors. Our algorithm enables to retrieve, in the same process, a similarity measure and an optimal rotation angle between two different patterns. The following section concisely presents Zernike moments. In section 3, we lean upon drawbacks of the conventional approach to build our method. The computational efficiency is also a constraint because the resulting comparator will be used within a matching process; We thus detail a fast implementation in section 4. Finally, we present experimental results and conclude in the last two sections.

2 Zernike moments

Complex Zernike functions constitute a set of orthogonal basis functions mapped over the unit circle. Zernike moments of a pattern are constructed by projecting it onto those functions. They share three main properties:

- The orthogonality: this property ensures that the contribution of each moment is unique and independent.
- The rotation invariance: the magnitude of Zernike moments is independent of any planar rotation of a pattern around its center of mass.
- The information compaction: low frequencies of a pattern are coded into the low order moments. As a result, relatively small descriptors are robust to noise or deformations.

Mathematically, Zernike basis functions are defined with an order p and a repetition q over $D = \{(p, q) | 0 \leq p \leq \infty, |q| \leq p, |p - q| = \text{even}\}$.

$$Z_{pq} = \frac{p+1}{\pi} \int \int_{x^2+y^2 \leq 1} V_{pq}^*(x, y) f(x, y) \partial x \partial y \quad (1)$$

where V_{pq}^* denotes the complex conjugate of V_{pq} , defined as:

$$V_{pq}(\rho, \theta) = R_{pq}(\rho) \cdot e^{iq\theta} \quad \text{and} \quad R_{pq}(\rho) = \sum_{\substack{k=|q| \\ |p-k| \text{ even}}}^p \frac{(-1)^{\frac{p-k}{2}} \frac{p+k!}{2}}{\frac{p-k!}{2} \frac{k-q!}{2} \frac{k+q!}{2}} \rho^k \quad (2)$$

From eq. (1) and (2), Zernike moments of a pattern rotated by an angle α around its origin are given as :

$$Z_{pq}^\alpha = \frac{p+1}{\pi} \int_0^{2\pi} \int_0^1 V_{pq}^*(r, \theta) f(r, \theta + \alpha) r \partial r \partial \theta = Z_{pq} e^{iq\alpha} \quad (3)$$

Eq (3) proves the invariance of the magnitude of Zernike moments to rotation since $|Z_{pq} e^{iq\alpha}| = |Z_{pq}|$. Finally, Zernike moments allow the analyzed function to be reconstructed:

$$\tilde{f}(x, y) = \sum_{(p,q) \in D} Z_{pq} V_{pq}(x, y) \quad (4)$$

In the reminder of this paper, we will denote the usual comparator (i.e. invariant to rotation) of Zernike descriptors as the *classical* one. Formally, this comparator is nothing else than an Euclidean distance between the magnitudes of the moments:

$$distance = \sum_{(p,q) \in D} (|Z_{pq}| - |Z'_{pq}|)^2 \quad (5)$$

3 Similarity measure and rotation angle retrieval

In this section, we first describe the two main problems that arise from the classical approach in order to set our method off against these drawbacks. We then define in the next subsection our new algorithm for Zernike descriptors comparison, allowing to retrieve both an improved similarity and a generalized rotation angle between two different images.

3.1 The classical approach

One smart solution to our problem of retrieving the most similar 3D pose of a synthetic 3D object, along with its corresponding planar rotation angle, from a given hand-made sketch would be (see the upper panel of fig. 1):

1. to describe each views of the 3D model using Zernike moments (off line),
2. to compute Zernike moments of the sketched object,

3. to compare the descriptors two by two in a rotation invariant way, using eq. (5)
4. to consider the best match to be the correct 3D pose, and finally
5. to estimate the planar rotation angle using existing angle estimation approaches [9, 10].

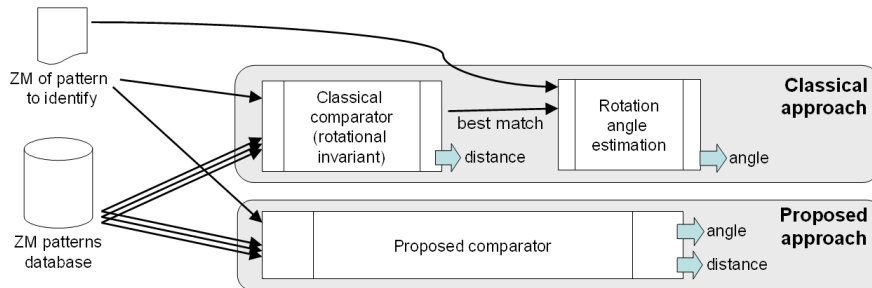


Figure 1: General plot of the conventional and the proposed approach. Whereas the classical approach has to first elect the best *rotation invariant* match to then estimate its rotation angle, our comparator unifies both similarity and rotation angle retrieval.

However, steps (3) and (5) can both induce erroneous results. To begin with, the classical comparator loses the phase information of Zernike moments. As a result, two different patterns can be mistaken to be the same ones if only their magnitudes are the same, like it is illustrated on fig. 2. In particular, since a pattern has the same Zernike moments magnitudes than its symmetric, two viewing angles of a 3D symmetric object present the same magnitudes and thus can not be differentiated by the classical Euclidean comparator; Fig. 3 presents such an example. As a consequence, this classical comparator may not enable a correct 3D pose retrieval – in particular when it deals with symmetric 3D objects.

Besides, in step (5) the existing methods for rotation angle estimation are based on the hypothesis that the two patterns are the *same* (that is, except some noise and the rotation). In our case, we have to compare different patterns since in one hand we have 3D model views and in the other hand we have drawings. In such situation, the existing angle estimators are quite ineffective (see an example in fig. 4).

To solve the first issue, the classical descriptor has to take moments phases information into account. One solution to the second issue is to define the optimal angle as that which minimizes the Euclidean distance between the first image and the rotated second one. We show in the following section that those two solutions can be expressed in a single framework (fig. 1, bottom) to prevent the before-mentioned errors to occur.

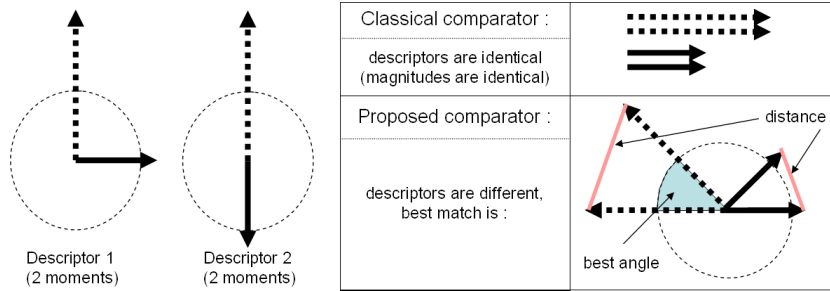


Figure 2: The classical comparator can not distinguish two descriptors that only differs by their phases, while the proposed comparator does.

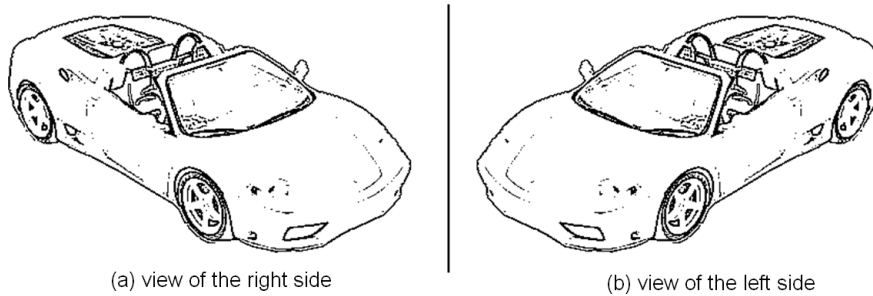


Figure 3: Edge images of a car seen from both side. The magnitude of the Zernike moments is the same for both views (a) and (b). Consequently, it is impossible for the classical approach to decide whether a given sketch represents either of these two views.

3.2 The proposed Zernike comparator

In this section, we describe a generalized method that retrieves both a similarity measure and a rotation angle between two different images. The angle is considered to be optimal since the Euclidean distance between the first image and the rotated second one is minimal.

Let I and I' be two different images, and $(I' * \mathfrak{R}_\phi)$ be the I' image rotated by ϕ . The Euclidean distance between I and $(I' * \mathfrak{R}_\phi)$ can be expressed as a function of the rotation angle ϕ as follows:

$$distance(I, I', \phi) = \sum_{x^2+y^2 \leq 1} \sum [I(x, y) - (I' * \mathfrak{R}_\phi)(x, y)]^2 \quad (6)$$

Thus our objective is to minimize this expression so as to determine the corresponding angle ϕ . Eq. (3) has shown that if the set of moments $\{Z_{pq}^{I'} | (p, q) \in D\}$ represents the I' image, then $\{Z_{pq}^{I'} e^{iq\phi} | (p, q) \in D\}$ represents $I' * \mathfrak{R}_\phi$. By replacing I and $(I' * \mathfrak{R}_\phi)$ in equation (6) by their exact Zernike reconstruction, we

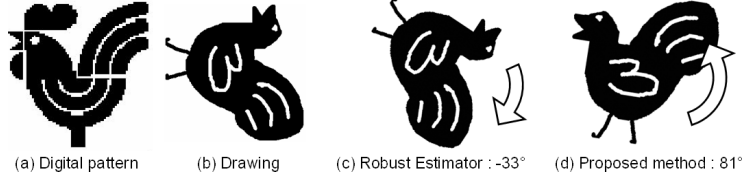


Figure 4: Rotation angle retrieval between (a) a digital pattern and (b) its rotated sketch. The estimation is erroneous for the robust estimator from [9] (c) as it deals with different pictures, while the proposed approach (d) retrieves a correct angle.

obtain:

$$\begin{aligned}
distance(I, I', \phi) &= \sum_{x^2+y^2 \leq 1} \sum \left[\sum_{(p,q) \in D} V_{pq}(x, y) (Z_{pq}^I - Z_{pq}^{I'} e^{iq\phi}) \right]^2 \\
&= \sum_{x^2+y^2 \leq 1} \sum \left[\sum_{(p,q),(u,v) \in D^2} V_{pq}(x, y) (Z_{pq}^I - Z_{pq}^{I'} e^{iq\phi}) V_{uv}(x, y) (Z_{uv}^I - Z_{uv}^{I'} e^{iv\phi}) \right] \\
&= \sum_{(p,q),(u,v) \in D^2} (Z_{pq}^I - Z_{pq}^{I'} e^{iq\phi}) (Z_{uv}^I - Z_{uv}^{I'} e^{iv\phi}) \left[\sum_{x^2+y^2 \leq 1} V_{pq}(x, y) V_{uv}(x, y) \right] \\
&= \sum_{(p,q) \in D} \langle V_{pq}, V_{p-q} \rangle \left| Z_{pq}^I - Z_{pq}^{I'} e^{iq\phi} \right|^2 \tag{7}
\end{aligned}$$

where Z_{pq}^I and $Z_{pq}^{I'}$ represent Zernike moments of images I and I' , respectively, and $\langle V_{pq}, V_{uv} \rangle$ denotes the scalar product of two Zernike basis functions over the unit disc. Thanks to the orthogonality of the basis, this product is null except for the case where $(p, q) = (u, -v)$. In that case, it can be simplified into:

$$\langle V_{pq}, V_{p-q} \rangle = \langle V_{pq}, V_{pq}^* \rangle = \frac{\pi}{p+1} \tag{8}$$

At first glance, eq. (7) is not trivial to minimize but it can be rewrote as follows:

$$\begin{aligned}
distance(I, I', \phi) &= \sum_{(p,q) \in D} \sum \frac{\pi}{p+1} \left| Z_{pq}^I - Z_{pq}^{I'} e^{iq\phi} \right|^2 \\
&= \sum_{(p,q) \in D} \sum \frac{\pi}{p+1} \left| Z_{pq}^I \right|^2 \left| \frac{Z_{pq}^{I'}}{Z_{pq}^I} e^{i(q\phi + [Z_{pq}^{I'}] - [Z_{pq}^I])} - 1 \right|^2 \\
&= \sum_{(p,q) \in D} \sum \frac{\pi}{p+1} \left(\left| Z_{pq}^I \right|^2 + \left| Z_{pq}^{I'} \right|^2 - 2 \left| Z_{pq}^I Z_{pq}^{I'} \right| \cos(q\phi + [Z_{pq}^{I'}] - [Z_{pq}^I]) \right) \tag{9}
\end{aligned}$$

where $|Z_{pq}|$ and $[Z_{pq}]$ denote respectively the modulus and the argument of Z_{pq} . Formula (9) points out that the only parameter whose the distance depends is ϕ , which in addition is only present into the cosine functions. As a consequence, the search for optimal distance and angle will result in minimizing a sum of cosines.

To conclude with, equation (9) describes a generalized manner to compare two Zernike descriptors which enables the solving of both issues mentioned in section 3.1 : the similarity measure is improved and it depends on a single parameter ϕ which is the optimal rotation angle. The next section focuses on resolving this minimization so as to insure a low complexity and a fast computing time.

4 Efficient minimum search

In this section, we describe how to efficiently extract the global minimum of eq. (9) by restricting the search using Nyquist-Shannon sampling theorem. Assuming that the Zernike moments of each image are known until the N^{th} order, the sum initially comprises $O(N^2)$ cosine terms. Nonetheless, the equation can be simplified by removing the constant terms and by aggregating the cosine terms that have the same frequency, as follows:

$$A_1 \cos(q\phi + B_1) + A_2 \cos(q\phi + B_2) = |C| \cos(q\phi + [C]) \quad (10)$$

where C is a complex that worths $A_1 e^{iB_1} + A_2 e^{iB_2}$. Then, eq. (9) can be equivalently expressed as a sum of N cosines :

$$f_N(\phi) = \sum_{q=1}^N A_q \cos(q\phi + B_q) \quad \text{with } (A_q, B_q) \in \mathbb{R}^2 \quad (11)$$

4.1 Restricting the search of a global minimum

One can notice that $f_N(\phi)$ is a 2π -periodic function. Usually, the general technique for finding the minimum of a periodic function is a gradient descent. Indeed, $f_N(\phi)$'s first and second derivative are easy to compute. However, such functions generally have many local minima whereas our approach requires to find the global one. One expensive solution is then to find every local minima and maxima with the gradient descent method by following the function from $\phi = 0$ to 2π .

However, $f_N(\phi)$ owns a discrete Fourier spectrum bounded by a maximal frequency $f_N^{MAX} = N/2\pi$. Hence, $f_N(\phi)$ have at most N local maxima and N local minima in $[0, 2\pi[$. Moreover, the Nyquist-Shannon sampling theorem teaches us that the function can not change substantially between two consecutive sampling points taken at the Nyquist frequency $F = 1/T = N/\pi$. The minimal distance between two consecutive minima is thus bounded by π/N . The initial starting points for finding every possible minima with a gradient descent can thus be equally scattered in $2N$ points. Moreover, by cutting f_N into

$4N$ intervals, we ensure that only one minimum *or* one maximum is present in each interval (see fig. 5).

4.2 Optimized minima retrieval

Our approach for optimizing the gradient descent takes advantage of the previously formulated properties. f_N is sampled by $4N$ points equally spread between $[0, 2\pi[$: $\{x_n = n\pi/2N \mid 0 \leq n < 4N\}$. We compute f_N 's differential, denoted as f'_N , for each of those points. Section 4.1 ensures that if and only if a minimum is present between two consecutive points $[x_n, x_{n+1}]$, then $f'_N(x_n)$ is negative and $f'_N(x_{n+1})$ is positive (see fig. 5). Moreover, those differential values enable to approximate the abscissa of the local minimum. Indeed, by approximating f_N between $[x_n, x_{n+1}]$ as a second degree polynomial, then the minimum abscissa can be evaluated at:

$$x_{minimum} = \frac{x_{n+1}f'_N(x_n) - x_nf'_N(x_{n+1})}{f'_N(x_n) - f'_N(x_{n+1})} = x_n - \frac{\pi}{2N} \frac{f'_N(x_n)}{f'_N(x_n) - f'_N(x_{n+1})} \quad (12)$$

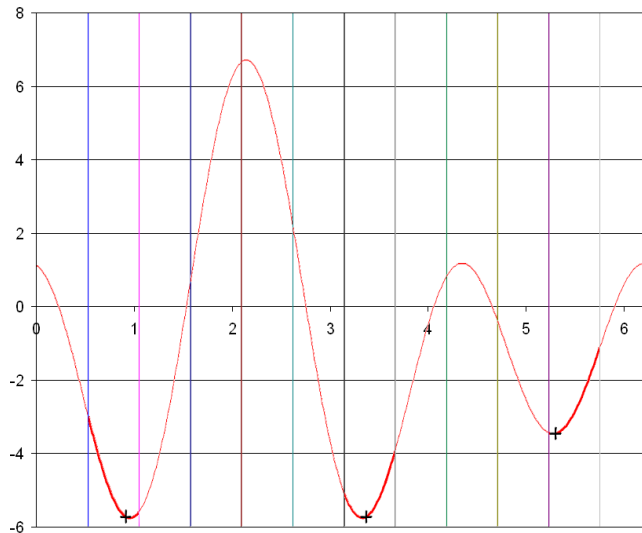


Figure 5: A random function f_3 cut into $4N = 12$ intervals. Each interval contains at most one local minimum. When a minimum exists in the interval, the left derivative is negative and the right derivative is positive. Positions of the second order approximation are represented by crosses.

For our application, the gradient descent algorithm does not need to be iterated to reach a high precision since this simple approximation (represented as crosses in fig. 5) is precise enough for our purposes (cf. experimental results in 5.1). Finally, the computational complexity of our approach for the distance minimization using the approximation is $O(N^2)$. Our approach has then the same computational complexity than the classical Zernike comparator (eq. (5)).

5 Experimental results

This section includes experimental results that demonstrate the efficiency of our minimum search, a comparison with existing angle estimation approaches and some visual examples within our sketched object recognition application.

5.1 Efficient minimum search

In order to demonstrate the efficiency of our minimization approach (see section 4), a set of 30,000 random functions $f_N(\phi)$ was created. For each of 3 different bandwidths $N = \{6, 12, 24\}$, there were 10,000 functions with random A_p and B_p . Firstly, we have computed for each function the exact solution ϕ and the approximation $\tilde{\phi}$ using respectively an exhaustive gradient descent and the proposed optimization. Secondly, we have compared both results and distinguished two types of situation:

1. The global minimum is found : $\phi \approx \tilde{\phi}$,
2. Another minimum is found.

The second case derives from situations where the function admits more than one solution. In term of Zernike description, when there are $n - 1$ secondary minima as small as the global one, that concretely means that the pattern is n -fold symmetric. For instance, a circular symmetric pattern can be rotated in an equivalent way by two different angles $\{\alpha, \alpha + 180^\circ\}$. Consequently, what truly determines the rightness of the retrieved angle is the vertical error on the depth of the approximated minimum.

		$\phi \approx \tilde{\phi}$	$\phi \neq \tilde{\phi}$
$N = 6$	Occurrences	9,986	14
	RMS error on $\tilde{\phi}$	0.81°	107.6°
	RMS error on \tilde{f}_N	0.197%	0.537%
$N = 12$	Occurrences	9,991	9
	RMS error on $\tilde{\phi}$	0.36°	119.4°
	RMS error on \tilde{f}_N	0.122%	0.472%
$N = 24$	Occurrences	9,986	14
	RMS error on $\tilde{\phi}$	0.16°	98.8°
	RMS error on \tilde{f}_N	0.083%	0.215%

Table 1: The average RMS errors corresponding to the approximation of the minimum position for various N .

Table 1 details the RMS errors of angle $\tilde{\phi}$ and depth of the minimum $f_N(\tilde{\phi}) = \tilde{f}_N$ for both situations. The proportion of occurrences of the second case (the retrieved minimum is not the global one) is comprised below 0.15%. In the first case, the RMS errors of the angle and the minimum depth do not run over 0.81° and 0.2%, respectively, even for $N = 6$ (when the approximation is

with the most severe noise yields less than 1.9° , while it lightly overtakes 2.3° using the robust estimator. On average, our estimator is about 20% better than the Kim and Kim approach [9].

		Noise-Free	SNR=20dB	SNR=10dB
Robust Estimator Method	Overall average	0.21°	0.33°	0.76°
	Average of the worst 15%	0.49°	0.71°	1.69°
	Average of the worst 10%	0.56°	0.80°	1.93°
	Average of the worst 5%	0.67°	0.95°	2.34°
Proposed Method	Overall average	0.17°	0.30°	0.58°
	Average of the worst 15%	0.37°	0.64°	1.32°
	Average of the worst 10%	0.42°	0.73°	1.51°
	Average of the worst 5%	0.52°	0.89°	1.85°

Table 2: The average RMS error of rotation angle using the robust estimator method and the proposed method.

5.3 Rotation angle retrieval between different patterns, application to 3D sketched object recognition

Since the rightness of the rotation angle retrieval between different pictures is quite hard to define, we have substituted a visual comparison instead, within the framework of our application: recognition of the pose of a 3D object present on a sketched storyboard. For that aim, several views of the 3D object (a 3D polygonal mesh) were taken every 20° intervals on latitude and longitude, then corresponding edge images were constructed and described using Zernike moments. On the other hand, the sketched patterns were also preprocessed (skeletonization, noise removal) in order to facilitate the recognition. Then, the most similar views (corresponding to specific 3D poses) along with their planar rotation angles were computed for each of three drawing sketches using the classical approach (Zernike descriptor matching and then rotation angle retrieval using [9]) and the proposed approach.

For each trailer sketch, the three best retrievals are presented in table 7. The quality of the first two sketches is quite good whereas the last one (last row) was degraded on purpose to test robustness to noise. Our method provides very stable results even with the most severe degradation. One can also notice that the first sketch looks like a union of the first two views retrieved by our method. On the contrary, the classical approach is unstable for the second and third sketches as it retrieves symmetric views for the third best results. In the case of the low quality sketch, the robust angle estimator [9] fails at retrieving the rotation angle even when the viewing angle is the correct one.

Even if the subjectivity takes a big place in this evaluation, it is clear that in the second and third sketch cases the proposed approach well evaluate the views whereas the classical method fails at retrieving the correct 3D pose or the planar rotation angle.

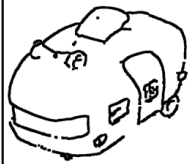




















Sketch		First	Second	Third
	Classical Method			
	Proposed Method			
	Classical Method			
	Proposed Method			
	Classical Method			
	Proposed Method			

Figure 7: Comparative statement of the classical and the proposed approach with for each method the first, second and third best retrieved 3D pose. Each image was rotated by the angle computed with the corresponding method.

6 Conclusion

We have presented a new efficient comparator for Zernike descriptors, which improves the similarity evaluation while additionally providing the rotation angle between the corresponding patterns. Moreover our approach owns the same $O(\text{order}^2)$ complexity than the classical Euclidean distance between Zernike descriptors. In terms of rotation angle estimation, experiments have shown a 20% improvement compared to the robust estimator from Kim and Kim [9]. This works is particularly suited for 3D indexing and pose retrieval.

Acknowledgment

We would like to thank *Pinka* Studio (<http://www.pinka-prod.com/>) which partially supports this work and has provided the storyboard and the 3D model of the caravan.

References

- [1] T. Filali Ansary, M. Daoudi, J.-P. Vandeborre, A bayesian 3D search engine using adaptive views clustering, *IEEE Transactions on Multimedia* 9 (1) (2007) 78–88.
- [2] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, D. Jacobs, A search engine for 3d models, *ACM Trans. Graph.* 22 (1) (2003) 83–105.
- [3] A. Khotanzad, Y. H. Hong, Invariant image recognition by zernike moments, *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (5) (1990) 489–497.
- [4] A. Khontanzad, Y. H. Hong, Rotation invariant image recognition using features selected via a systematic method, *Pattern Recogn.* 23 (10) (1990) 1089–1101.
- [5] C.-Y. Wee, R. Paramesran, Efficient computation of radial moment functions using symmetrical property, *Pattern Recogn.* 39 (11) (2006) 2036–2046.
- [6] C.-Y. Wee, R. Paramesran, On the computational aspects of zernike moments, *Image Vision Comput.* 25 (6) (2007) 967–980.
- [7] G. Amayeh, A. Erol, G. Bebis, M. Nicolescu, Accurate and efficient computation of high order zernike moments., in: *ISVC, 2005*, pp. 462–469.
- [8] C.-W. Chong, P. Raveendran, R. Mukundan, A comparative analysis of algorithms for fast computation of zernike moments., *Pattern Recognition* 36 (3) (2003) 731–742.
- [9] W.-Y. Kim, Y.-S. Kim, Robust rotation angle estimator, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21 (8) (1999) 768–773.
- [10] N. Kanaya, Y. Iiguni, H. Maeda, 2-d doa estimation method using zernike moments, *Signal Process.* 82 (3) (2002) 521–526.