Definition of colour object signatures based on Zernike moments

Camille Counilh, Jean-Christophe Burie, Pierre Courtellemont; Laboratoire L3i: Laboratoire Informatique, Image, Interaction Université de La Rochelle; FRANCE

Abstract

In this paper we present a method to define signatures for colour object based on the Zernike moments. The aim of these signatures is to characterize simultaneously the object shape and texture. We propose two kinds of signatures that are invariant to image scale, rotation and translation.

The first one is based on a marginal approach that processes each RGB component separately. The Zernike moments are computed considering the greyscale value of the pixels and the used method provides two maps of moments for each component, one with real values and one with imaginary values. An adapted comparison procedure is used to compare the moment maps of two distinct objects.

The second signature is based on a vectorial approach where each pixel is characterized by a 3D vector defined in the RGB cube. The Zernike moments are now computed considering the colour variation between two adjacent pixels. This second approach is faster and provides only 2 moment maps that represent the colour variations of the object.

These signatures have been used to track exotic fishes in an aquarium in the framework of the Aqu@theque program.

Introduction

The tracking of an object in a sequence of images needs a recognition phase. In order to associate a signature to a given object, we have focused our attention on works dealing with classification and indexing. This signature can be a scalar number, a vector or a matrix depending on the method used. Our principal objective is to identify a given object from its signature to track it through a sequence of images.

Many methods are based on histogram comparison [1], [2], [3]. They consist in estimating, for instance, the number of colours used in the image and their distribution. A lot of comparison criteria have been defined to compare images or regions of interest (ROI) [4]. These methods are usually fast; however they are not enough discriminating, because they don't permit to make the distinction between two different objects that exhibit similar colorimetric properties.

Another group of methods works on object shapes. They were generally used in the case of binary images, but their uses have been extended to greyscale images. Some of these methods use moments (geometric, Hu, Legendre …) [5], [6]. But, they have to be adapted in order to obtain a signature that is invariant to image scale, rotation and translation. Few methods, such as the Zernike moments [7] [8], the Radon transform [9] [10] or the Fourier-Mellin transform [5], exhibit naturally these invariance properties.

The purpose of our study is to work simultaneously on object shape and texture. The Zernike moments [11], [12], [13], [14] seem to be the appropriate method for several reasons.

1 – The used computation method is invariant to image scale and rotation.

2 – The computation of the moment at a particular point provides the translation invariance property.

3 – The projection of the pixel intensity on a set of orthogonal polynomials allows estimating the texture variation in a given direction.

However, this method is usually used with binary and greyscale images. So, a new approach based on Zernike moments, has been developed to define signatures for colour object. In this paper, we present a marginal approach and a vectorial one to compute signatures. These two methods have been applied to the tracking of fishes in an aquarium in the framework of the Aqu@theque Program.

Complex Zernike moments

Complex Zernike moments are constructed using a set of complex polynomials which form a complete orthogonal basis set defined on the unit disc $|r| \leq 1$. The Zernike moments of order p with repetition q are given by the following expression:

$$Z_{pq} = \frac{p+1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} V_{pq}(r,\theta)\, f(r,\theta)\, r\, dr\, d\theta \qquad (1)$$

where:

- $p = 0, 1, 2\ldots \infty$ defines the order.
- q is an integer subject to these two constraints: $0 \leq |q| \leq p$ and $p - |q|$ is an even number
- $f(r,\theta)$ is the image function being described (the object).
- $V_{pq}(r,\theta)$ is the complex-valued Zernike polynomial. It is expressed in polar coordinates as:

$V_{pq}(r,\theta) = R_{pq}(r)e^{jq\theta}$ where (r,θ) are defined over the unit disc.

$R_{pq}(r)$ is the orthogonal radial polynomial, defined as:

$$R_{pq}(r) = \sum_{k=0}^{(p-|q|/2)} (-1)^k \frac{(p-k)!}{k!((p+|q|)/2-k)!((p-|q|)/2-k)!} r^{p-2k}$$

As an example, the first six radial polynomials are:

$R_{00}(r) = 1$	$R_{11}(r) = r$
$R_{20}(r) = 2r^2 - 1$	$R_{22}(r) = r^2$
$R_{31}(r) = 3r^3 - 2r$	$R_{33}(r) = r^3$

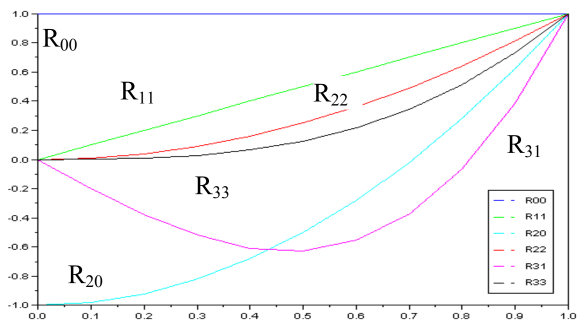


Figure 1: First radial polynomials

Figure 1 shows the response of the first radial polynomials. We can notice they are all different, and thus, allow a characterisation of the function image with "different points of view".

For a discrete image, the equation (1) becomes:

$$Z_{pq} = \lambda_Z(p,N)\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} R_{pq}(r_{kl})e^{-jq\theta_{kl}} f(r_{kl},\theta_{kl}) \qquad (2)$$

$$\lambda_Z(p,N) = \frac{p+1}{(N-1)^2}$$

where $f(r_{kl},\theta_{kl})$ represents the grey level value of a pixel whose the co-ordinates are expressed in polar co-ordinates system.

To compute the Zernike moments of an image $f(r,\theta)$, the image (ROI) is first mapped to the inscribe unit disc using polar co-ordinates, where the centre of the image (or ROI) is the origin of the unit disc (*cf.* figure 2). Those pixels falling outside the unit disc are not used in the calculation.

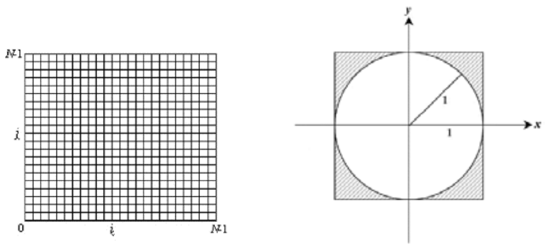


Image of N X N pixels
1 puts = 1 pixel

Unit disc inscribes in the image whose centre is the centre of the image

Figure 2: Discrete calculation of the Zernike moments

This method provides naturally invariance to translation and scale change. Now, if the image is rotated through angle φ, the relationship between Z'_{pq} and Z_{pq} is

$$Z'_{pq} = Z_{pq}.e^{-iq\varphi} \qquad (3)$$

Then $|Z_{pq}|$, the magnitude of the Zernike moment is a rotation invariant feature of the underlying image (or ROI).

Signature definition from the Zernike moments

For an order *p* with repetition *q*, the computation result of Zernike moments is a complex value from which it is possible to calculate a scalar value (magnitude). It appears clearly that a signature based on only one value is not sufficient.

For example, the figure 3 shows 2 test images representing a circle with the same coloured slices. Each slice has the same number of pixels, so the distribution of a given colour is the same in the 2 images. The only difference between image 1 and image 2 is the location of the red (1), green (2), blue (3), cyan (4), magenta (5) and white (6) slices.

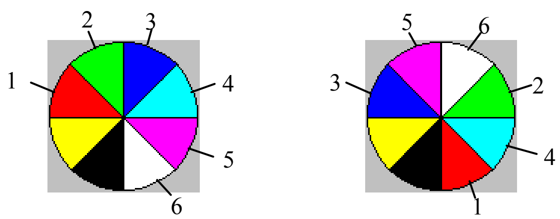


Image 1 *Image 2*

Figure 3: Two test images

	Z_{00}	Z_{11}	Z_{22}	Z_{20}	Z_{33}
Red component					
Image 1	293,16	262,14	187,15	454,90	119,52
Image 2	293,16	262,14	187,15	454,90	119,52
Green component					
Image 1	293,16	262,14	187,15	454,90	119,52
Image 2	293,16	262,14	187,15	454,90	119,52
Blue component					
Image 1	293,16	262,14	187,15	454,90	119,52
Image 2	293,16	262,14	187,15	454,90	119,52

Table 1: The computation result of "classical" Zernike moments

Table 1 shows the magnitude of the first Zernike moments for each component. The computed values are equal; the images will be detected as being the same one.

In order to enhance the discriminating capacity, we have defined a "moment map" which is obtained from the decomposition of the computation of the Zernike moments.

For an order p with repetition q, a value for the angle θ is imposed. Then, it is possible to calculate the following expression:

$$\int_{r=0}^{1} V_{pq}(r,\theta)f(r,\theta)rdr \qquad (4)$$

This term represents the sum of pixel intensities weighted by the Zernike polynomials calculated on the radius length of the unit disc. The result is then stored in a two dimensions map (*cf.* figure 4). The abscissa represents the angle θ, ranging from 0 to 2π. The ordinate represents the Zernike moments: Z_{00}, Z_{11}… Z_{pq}… Z_{mn}, where m and n are respectively the maximum order and repetition chosen to define the signature. The line number of the obtained signature depends on the number of moments used to characterize the object.

This computation method provides two maps: one for real values and another one for imaginary values for all computed moments (*cf.* figure 4). These maps define the signature of the studied object.

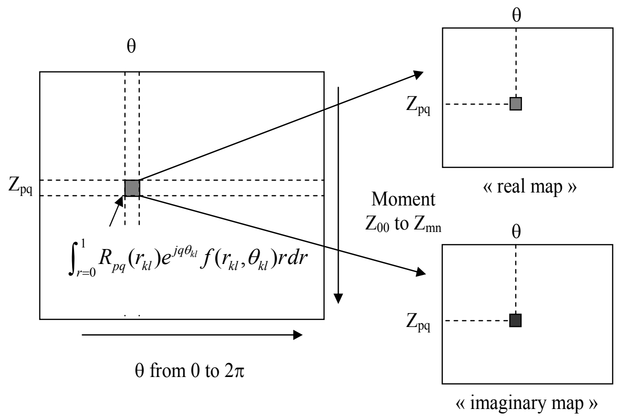


Figure 4: "moment maps"

Signature based on marginal approach

In order to generalise this method to colour images, firstly, we have used a marginal strategy. Two "moment maps" are defined, as described above, for each colour component: red, green and blue. So, 6 "moment maps" have to be computed to characterize an object.

To allow visual comparison of the signatures, a colour "map" is obtained by mixing the three "real maps" (corresponding to red, green and blue component) to create a coloured image (composed of a red, green and blue channel). The same method is used for "imaginary maps". Finally, we obtained two coloured representations of the signature of our studied object. But the interpretation is easier with the magnitude map, computed from the real and imaginary maps.

For the images of figure 3, the following "moment maps" (*cf.* figures 5, 6) are obtained with the marginal approach.

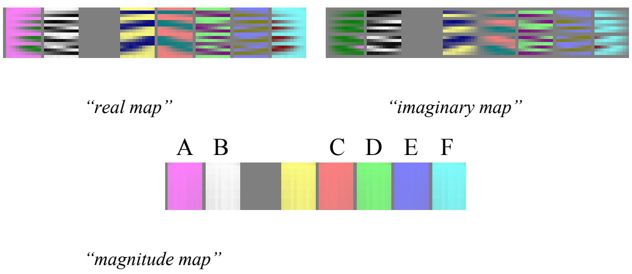


Figure 5: The signature of Image 1

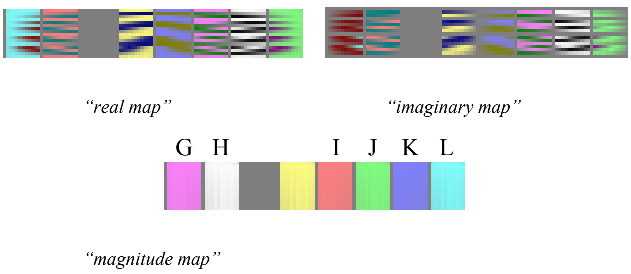


Figure 6: The signature of Image 2

A visual comparison of the magnitude maps allows immediately concluding that the two objects are different. In this example, each coloured part of the signature corresponds to a coloured slice of the circle. The part A (figure 5) corresponds to the slice 5 (figure 3: image 1), and the part G (figure 6) corresponds to the slice 4 (figure 3: image 2). The "colour" of each element of the signature depends on the mixing of the value of Zernike moments computed for the red, green and blue components. For the green slice (2), the red and blue components are equal to zero and their Zernike moments are also equal to zero. It is different for the green component, this is why the part D (figure 5) appears green. The differences between the parts A and G, but also between B (resp. C, D, E, F) and H (resp. I, J, K, L) allow distinguishing the objects.

Figure 7 shows another example. The object (image 1) has been rotated (180 degrees), and scaled (double size) to obtain the object of image 2. Figure 8 shows the signature of image 1 and the 9 the one of image 2. The yellow colour is marked by the item (1) and the red colour with the item (2). The comparison between the maps of the figure 8 and 9 demonstrates that a rotation of the object involves a shifting of the "magnitude map".

To compare objectively the two signatures, we need to define a comparison procedure that takes into account the shifting effect of the signature when a rotation is applied to an object.

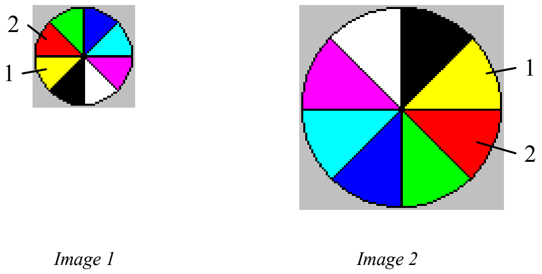


Figure 7: Two test images. Image 2 has undergone a rotation of 180° and a scaling (double).

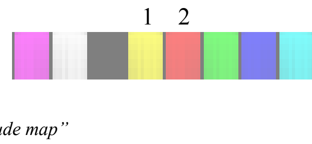


Figure 8: The signature of Image 1

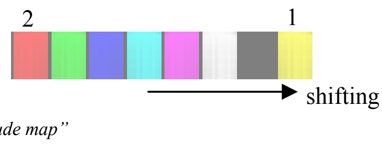


Figure 9: The signature of Image 2

The comparison between the "moment maps" of two images (or ROI) is based on the χ^2 test. But, to take into account the shifting effect. The first signature is fixed and is compared to each circular permutation of the second one. The minimum value is kept as the similarity index between the two signatures.

The figure 10 shows an example of an image and the two maps obtained for the fish marked by a white disc.

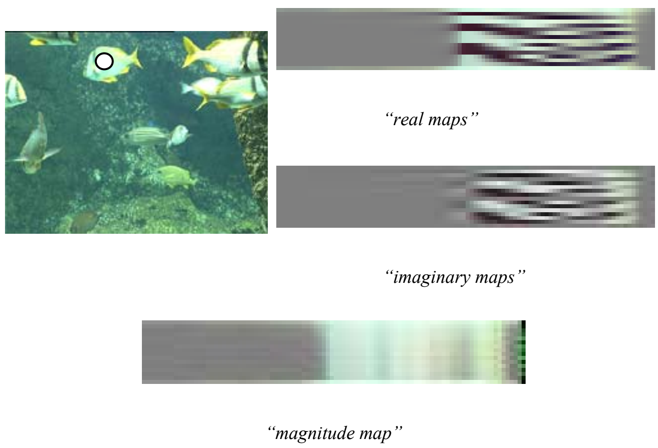


Figure 10: Visual signature of the fish marked by a white disc obtained with the marginal approach

This method has been used to track fishes in an aquarium and provides good results. But two problems have been observed. Firstly, it's time consuming since a significant number of maps have to be computed (one for each colour component). Secondly, the three components aren't correlated any more, since they are processed separately. So a second approach has been defined.

Signature based on vectorial approach

In this approach, each pixel is characterized by a vector whose the components are defined from its colour components (red, green, blue) defined in the RGB cube. The principle used to compute the signature is the same as above except that we don't consider anymore the pixel intensity (greyscale value), represented by the term f(r,θ) in equation (4). This term is replaced by an expression that estimates the colour variation (*Cv*) between two adjacent pixels along the radius of the unit circle. This expression is given by the following equation:

$$Cv = \alpha\left(\frac{1-\cos\varphi}{2}\right) + \beta\frac{2}{\sqrt{3}}\left|\left\|\overrightarrow{V_{pixel1}}\right\| - \left\|\overrightarrow{V_{pixel2}}\right\|\right| \quad (5)$$

where:

- $\overrightarrow{V_{pixel1}}$ and $\overrightarrow{V_{pixel2}}$ represent the vectors which characterize the colour of two adjacent pixels along the radius of the unit circle.

Note that the origin of these vectors is the centre O' of cube RGB as shown on figure 11 and their extremity correspond to the colour of the two adjacent pixels.

- φ is the angle formed by these two vectors.
- The parameters α and β are used to weight each term of this expression with the constraint $\alpha + \beta = 1$.

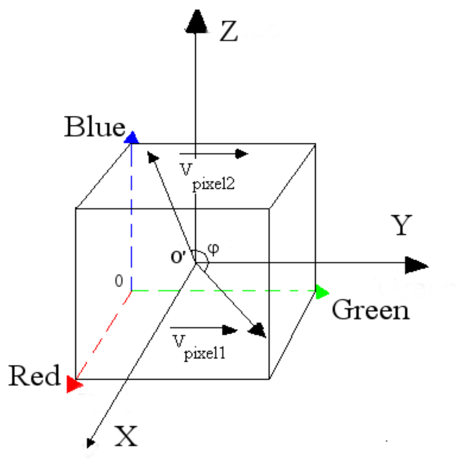
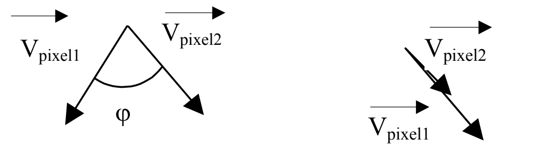


Figure 11: RGB Cube.

When the second term (difference of the norms) tends towards zero, the first term of the expression (equation 5) is used to differentiate them thanks to the angle φ formed by these two vectors (*cf.* figure 12a). The maximum value of the first term is obtained for opposite vector (φ = 180) and the minimum value for vectors with the same direction (φ = 0).

Conversely, when the angle φ tends towards zero, the second term of the expression (eq. 5) allows differentiating the two vectors (*cf.* figure 12b). The maximum value of the second term is obtained when $\overrightarrow{V_{pixel\,1}}$ is a null vector (middle grey "colour") and $\overrightarrow{V_{pixel\,2}}$ is the longer vector (colour corresponding to extremities of the cube such as red, green, white …). The minimum value is equal to zero when the magnitudes of the vectors are equals.

If two adjacent pixels exhibit a similar colour, the term Cv tends toward zero. On the other hand, more the colour is different more the term Cv tends toward one.



(a) $\|V_{pixel1}\| = \|V_{pixel2}\|$ *and* $\varphi \neq 0$ *(b)* $\|V_{pixel1}\| \neq \|V_{pixel2}\|$ *and* $\varphi \approx 0$

Figure 12: Comparison of the vectors

With this approach, the signature of an object is only characterized by two "moments maps" and is less time consuming than the previous approach since the three components (red, green and blue) are processed simultaneously. The "moment maps" give a description of the colour variations of an object (ROI). Since the signature has the same properties as above, the same comparison procedure is used.

Figure 13 shows an example of an image and the two maps obtained for the fish marked by a white disc. To allow visual comparison and interpretation, the maps are presented like grey level images.

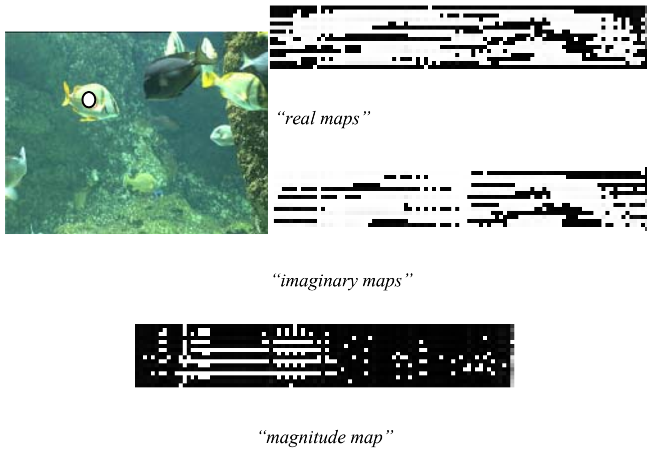


Figure 13: Signature of the selected fish with the vectorial approach.

Application

This work is carried out in the framework of the Aqu@theque project [15] in collaboration with the "Aquarium of La Rochelle". These two approaches have been applied to track exotic fishes in an aquarium.

To test the validity of the signature based on Zernike moments, we have used the following process. First, we select a fish in the first image of the sequence and its signature is computed (*cf.* figure 14a). Then, the signature of all the fishes detected in the following image is defined. All the signatures are compared to the one of the selected fish. The chosen fish is the one whose the signature is the most similar to the signature of the selected fish (*cf.* figure 14b).

In the first sequence (*cf.* figure 14), the attention have been focused on the fish marked by a white disc. During the sequence, the direction of the fish change and can be considered as a rotation. However, the algorithm shows its robustness to rotation since it successes to track the fish.

(a) First image: $t = t_0$ (b) Image at $t = t_1 > t_0$

Figure 14: Tracking of a fish which has undergone a rotation.

The second sequence (*cf.* figure 15) shows the tracking of a fish whose the direction and the size change (modification of the distance between the camera and the fish). The results prove that the signature is also invariant to image scale changes.

(a) First image: t=t0 (b) Image at t=t1 > t0

(c) Image at t=t2 > t1

Figure 15: Tracking of a fish which has undergone many rotations and scale changes.

The aim of this last experiment (*cf.* figure 16) is to track a "blue tang" fish. During the sequence, some illumination variations modify the appearance of the fish. In spite of the luminosity change, the signature is robust enough to track the fish (*cf.* figure 17).

(a) First image: $t = t_0$ (b) Image at $t = t_1 > t_0$

Figure 16: Tracking of the "blue tang" fish.

Conclusion

In this paper, we have presented a method based on Zernike moments to define signatures for colour objects. The invariance properties to translation, rotation and scale change ensure to be able to track selected objects if their moves can be decomposed in simple geometrical transformations.

The vectorial approach allows characterising the colour texture of an object since the moments give a description of colour variations.

In order to validate the signature, some experiments have been done to track fishes in an aquarium. The first tests are encouraging but a certain sensitivity of the signatures has been detected when the centre of the unit disc is slightly moved. When it appears the tracking can fail. So, we are working now to reduce this sensibility by fixing more precisely the centre of the unit disc.

In the case of our application, we have also to take care of the deformable aspect of the objects and the different views with which the fishes can be seen (front view, side view …). Since its aspect evolves during the sequence, we are also working on a dynamic version of the signature which adapts itself to the tracked object.

References

[1] Tae-Wong Yoo and Il-Seok Oh, A fast algorithm for tracking human faces based on chromatic histograms, Pattern Recognition Letters, Volume 20, Issue 10, Octobre 1999, Pages 967-978.

[2] Z.-N. Li, O.R. ZaTiane, Z. Tauber, Illumination invariance and object model in content-based image and video retrieval, J. Visual Commun. Image Representation 10 (3) (1999) 219–244.

[3] M.J. Swain, D.H. Ballard, Color indexing, Int. J. Comput. Vision 7 (1) (1991), pp. 11–32.

[4] Sameer Antania, Rangachar Kasturia; ., Ramesh Jainb, A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video, Pattern Recognition, 35, (2002), pp.945-965.

[5] G.Y. Yang, H.Z. Shu, C. Toumoulin, G.N. Han and L.M. Luo, Efficient Legendre moment computation for grey level images, Pattern Recognition, Vol. 39, Issue 1, Jan uary 2006, Pages 74-80

[6] MK Hu. Visual pattern recognition by moment invariants,. Ieee Trans. on Information Theory, 8, pp. 179-187, 1962.

[7] F. Zernike, Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode. Physica, 1:pp. 689–704, 1934.

[8] M.R. Teague, Image analysis via the general theory of moments, J. of Opt. Soc. of America 70, pg. 920–930, (1980).

[9] J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. Berichte Sächsische Akademie der Wissenschaften, Leipzig, Mathematisch-Physikalische Klasse, 69:262–277, 1917.

[10] S. Tabbone, L. Wendling and J.-P. Salmon. A new shape descriptor defined on the Radon transform. Computer Vision and Image Understanding, 102(1):42--51, April 2006.

[11] C. Kan and M.D. Srinath, Invariant character recognition with Zernike and orthogonal Fourier-Mellin moments, Pattern Recognition, vol n°35, p 143-154, 2002.

[12] Khotanzad A. and Hong Y.H., Invariant Image Recognition by Zernike Moments, IEEE Trans. on PAMI, 12 (5). 489-497 may 1990.

[13] C.W Chong, P. Raveendran and R. Mukudan, Translation invariants of Zernike moments, Pattern Recognition, vol. n°36, p1765 - 1773, 2003.

[14] C.W Chong, P. Raveendran and R. Mukudan, A comparative analysis of algorithms for fast computation of Zernike Moments, Pattern Recognition, vol. n°36, p 731-742, mars 2003.

[15] D. Semani, T. Bouwmans, C. Frélicot and P. Courtellemont., Automatic Fish Recognition in Interactive Live Videos*,* Int.

Workshop on Interactive Video between Research challenges and Industrial Applications (IVRCIA), XIV : 94-99, 2002.

Biography

Camille COUNILH received her MS in Computer Science from Bordeaux University (2002). She is now following a PhD degree in Computer Science at the University of La Rochelle (France) in the Department of Computer Science (L3i). Her current research is mainly focused on Colour Image Processing and Image sequence processing. Jean-Christophe Burie and Pierre Courtellemont are respectively assistant professor and professor at La Rochelle University.

Laboratoire L3i: Laboratoire Informatique, Image, Interaction
Université de La Rochelle
Pôle Sciences et Technologie
17042 La Rochelle Cedex 1 - FRANCE
camille.counilh@univ-lr.fr

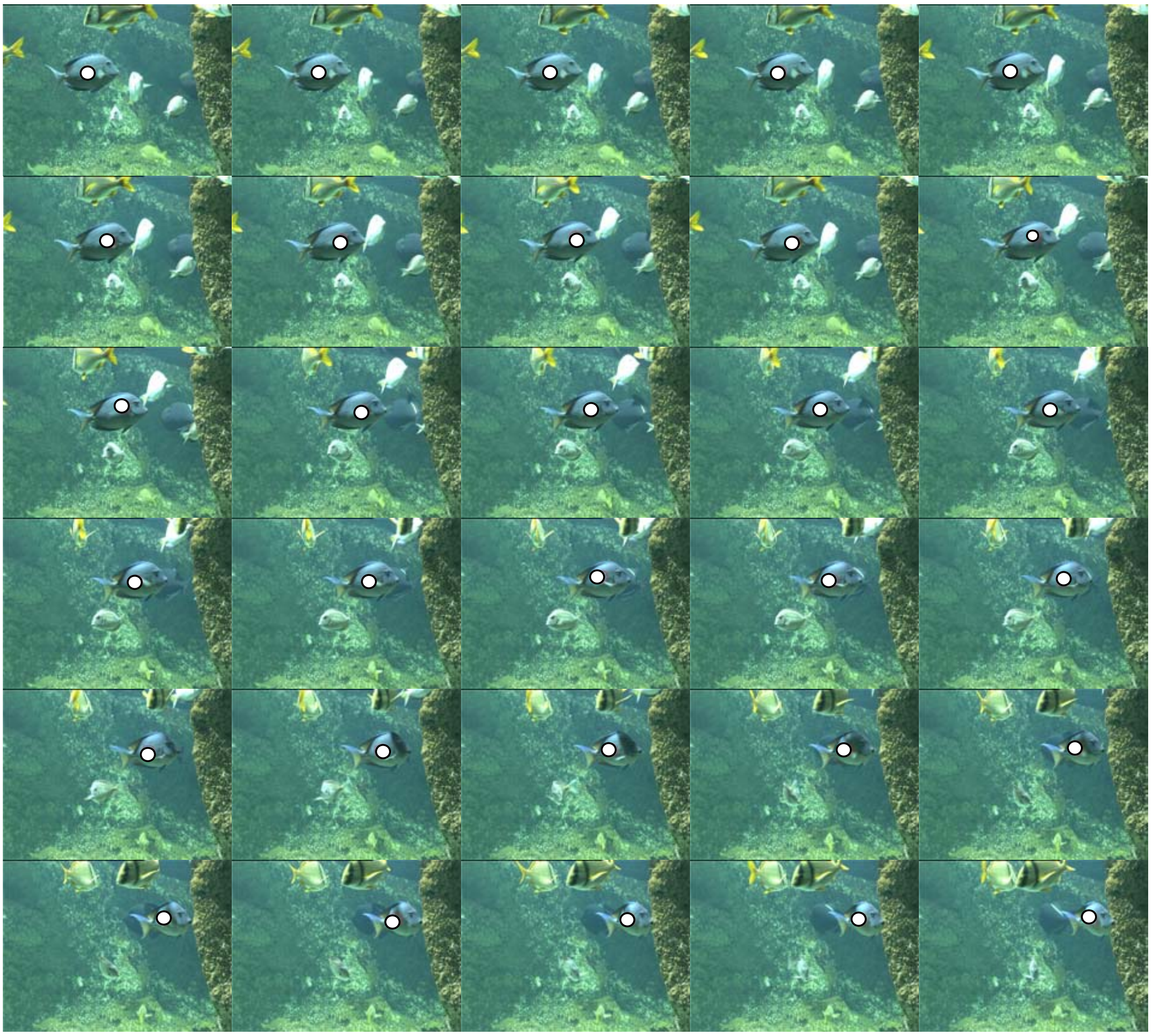


Figure 17: Tracking of the "blue tang" fish during a sequence