Autonomous agents for edge detection and continuity perception on otolith images

Anne Guillaud\textsuperscript{a,*}, Abdessalam Benzinou\textsuperscript{a}, Hervé Troade\textsuperscript{b}, Vincent Rodin\textsuperscript{a}, Jean Le Bihan\textsuperscript{a}

\textsuperscript{a}Laboratoire RESO, Ecole Nationale d’Ingénieurs de Brest, Technopôle Brest-Iroise, BP 30815, 29608 Brest Cedex, France
\textsuperscript{b}Laboratoire de Sclérochronologie des Animaux Aquatiques, Centre de Brest, Institut Français de Recherche pour l’Exploitation des Mers, BP 70, 29280 Plouzané, France

Accepted 4 July 2002

Abstract

An automatic method for edge detection on biological images (otolith images) using a multi-agent system is presented. One of the major problems encountered during an automatic contour detection is the lack of structure continuity perception. In this paper we present a new approach to perceive continuity based on a 2D reconstruction of closed contours using a multi-agent system. Each agent is provided with sensors on the image, which allow it to follow local intensity extremes. The purpose is to detect alternative light and dark concentric structures in an image. To improve the detection of these reactive agents, we have added high-level information about the shape of the contour. An application to fish otolith growth ring detection is presented in this paper. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Multi-agent system; Reactive agent; Roof edge detection; Otolith

1. Introduction

An otolith is a calcified structure located in the inner ear of the fishes. Otolith growth is an accretionary process, as for trees trunks (Fig. 1). The otolith structure is made of alternative opaque and translucent concentric rings. Each year one opaque ring and one translucent ring appear. The purpose of growth rings identification is to acquire data on age and growth of fish population. Such data are needed in a great number of biological and ecological studies and to improve stock management. Up to now, this analysis has been mainly limited to a ring count, processed by human readers on hundreds of thousands of otoliths per year. Therefore we are searching for automatic methods for ring detection or semi-automatic methods to help human interpretation. Ring detection on such images involves edge detection but also continuity perception.

Active contours are currently used to deal with problems such as edge detection and contour following [1,16]. They have inspired methods using deformable templates [24,25], which have already been applied on otolith images [21]. The shape of the otolith contour was modeled by a closed B-spline and reduced using a homothetic transform inside the otolith to detect the position of rings, by minimizing an energy function depending on the gray levels in the image. The contour shape was used as a template because of the shape memory phenomenon occurring during otolith growth. This method gave good results with young individuals, until age group 3, but beyond this age the percentage of wrong age estimation was quite important. This can be due to different reasons. For old individuals last rings are very thin (Fig. 1), and the energy function might be difficult to minimize in such conditions, so that the method will often under-estimate the number of rings. Moreover on old otoliths the similarity between central rings and external rings is less evident. Thus the shape of first rings will not be respected when represented by a homothetic transform of the contour shape. In Ref. [6] a deformable model (LDBBM) has been applied to otolith images. The model was initialized at the growth center of the otolith and then inflated by computing local forces based on the gray levels and on the global shape of the otolith. Detected rings shape was thus more accurate than with the previous method using otolith contour exact shape, but the method still encountered problems with old individuals, which last rings are very thin and poor contrasted. Therefore we needed to search for another method, which uses a more local perception of the
rings. This local perception requires to detect edges corresponding to rings location.

Edges are generally considered as discontinuities in the gray levels of an image, and are called step edges. Nevertheless one can find other types of edges in images, like roof edges or ridge edges. In this paper we will deal with roof edges, as we need to find local extremes of the image for our application. We could search for step edges in such images, which would give the position of the boundary between a dark and a light ring. But in some images this boundary is fuzzy, and the real transition position is difficult to find, whereas one can have a more accurate position of the light and dark rings by searching for roofs and valleys on the image [5].

In Ref. [10] a way to evaluate a good edge detector is defined by using some criteria: the detector must give a good localization of the edge and a low number of bad detected points (pixels detected that do not correspond to edges pixels or edges pixels not detected). The optimal filters defined in Ref. [10] were adapted in Ref. [26] for roof and ridge detection and extended to the 2D case. This filter was applied on synthetic images featuring lines in an image containing gaussian noise. The results show a good localization but an important over-detection of the edges, if no threshold is applied. In Ref. [15] a cubic polynomial in the two variables: row and column, was determined, which coefficients are computed using masks on the neighborhood of the pixels, to detect ridges and valleys in images. This idea is inspired from forms used in statistical regression problems. To identify a ridge or a valley, all second partial derivative of the polynomial at the origin, which is the central pixel are computed. Then a zero crossing of the first derivative in the direction α is searched to determine the edge position. The method needs an important amount of computations and thresholds and the coefficients of the polynomial have to be tuned. In Ref. [20] a survey is proposed concerning methods to detect skeletons on line images which could be used to detect edges, with an application to handwritten characters recognition. For these methods the image has to be first binarized, which can be easy for this type of application, where the gray levels are quite uniform within a handwritten character. Nevertheless this step is risky in several cases. For example, in the case of otoliths, a low threshold would erase the contrast between the last rings that are light, and a high threshold would erase the first rings that are darker.

For these reasons more adaptive methods were required to deal with the problem of ring detection on otolith images. In this paper we present a method of image segmentation using autonomous agents located on the image to detect edges. In Section 2 we will introduce multi-agent systems and their interest in computer vision. In Section 3 we will describe the characteristics of agents used in our application and their ability to detect edges on synthetic images, and then on otolith images. To improve the detection on otolith images when rings are discontinuous or bad contrasted, we have added high level information available for the agents orientation, which is the shape of the contour of the otolith, as explained in Section 4. Performances of the method are presented and discussed in Section 5.

2. A multi-agent system for edge detection

Multi-agent systems foundations can be found in different research themes, such as Distributed Artificial Intelligence or Artificial Life [12,23]. An agent is an entity, which can be virtual or physically embodied, evolving in an environment, which can contain other agents. Agents are generally autonomous, which means that they do not need external intervention to act according to the data that they perceive [18]. A multi-agent system is composed of an environment, passive objects situated in this environment, at least two agents which can act on these objects, and relations between all entities of the system.

Recently multi-agent systems have been used in computer vision [2,3,7,8,17]. These agents achieve quite simple actions, but by sharing their results, their work can bring to a more complex process. In Ref. [17] autonomous agents were used to detect homogeneous regions in brain scan images. Each agent can achieve tests on pixels around it in a circular neighborhood, as computing the variance or the mean gray level. If it finds that its neighborhood satisfies the conditions to be a region, the central pixel will be marked and new agents will be generated to grow the region. If the agent does not recognize a region, it will move to another place. This method is well adapted for brain scan images, because of regions characteristics regularity for tumors, sane parts, etc. In Refs. [8,9] a multi-agent system is used to segment cytological images. One type of agent is defined that could be adapted to the research of four different features in cytological images, as the nucleus or the background. The discrimination is also mainly based on the regions characteristics. On the opposite if we consider that otolith images can be divided in two types of regions: dark rings and light rings, it will be almost impossible to find
some statistical characteristics that would always fit to each type of region. For example, as the mean gray level increases from the nucleus to the edge, the mean value of a dark ring near the edge can be superior to the mean value of a light ring near the nucleus (Fig. 1). In Refs. [3,4] a multi-agent system is proposed to detect concentric rings that can be found in natural objects such as tree trunks. Each agent can move around in its environment which is a grayscale image; its two square-shaped sensors on the pixels of the image allow it to follow light rings (light agents) or dark rings (dark agents) by moving in the direction of the lighter pixels for light agents (resp. darker pixels for dark agents). If the agents have gone over a loop, they can validate their path as a ring. This method presents several advantages: agents are able to detect very quickly circular structures (a few seconds). Agents have also a local vision of rings, which favors a good precision in edge detection. Moreover agents can restitute ring continuity by the way they move in the image. This property is interesting for an application on otolith images featuring concentric rings. Nevertheless agents may encounter problems to detect very textured rings, or to find all loops for old individuals, with very thin and poor contrasted rings (Fig. 1). This problem was also observed with other methods dedicated to otoliths image processing ([6,19,21]). Human experts take into account the global otolith shape in order to give the most accurate position of discontinuous rings. To simulate human perception, we have added high-level information about the otolith shape available for helping the agent’s decision. In the following sections we will explain the interest of our approach, inspired by Ref. [3], to improve the detection of otolith growth rings. In the following sections we will describe the agents in their original version, which goal is to follow and detect edges, and how we have adapted them to the problem of ring detection on otolith images.

3. Agents description

The multi-agent system suggested in Ref. [3] is composed of reactive agents. Agents are called reactive when their behavior consists in a response to external stimuli, such as gray level intensities in an image. The agents are first placed at random on the image, then they are guided by their sensors (Fig. 2). Each agent has one unit sensor which allows it to locate itself on the image, and gives its central position \((x, y)\). Two square-shaped sensors made up of unit sensors are located in front of the agent and distant one from the other (Sensor 1 and Sensor 2). They return the sum of the grayscale levels on the part of the image where they are located. By comparing the gray level intensities of its two sensors, the agent is able to detect which sensor indicates the direction of growing intensities or decreasing intensities.

A dark agent tries to move where the values returned by the sensors are minimal, in order to detect dark rings. A light agent does the opposite, in order to detect light rings. The greater the difference between the two sensors, the more the agent will deviate. The direction of the agent is indicated by the angle \(d\) (Fig. 2), which is the angle formed between the horizontal direction and the median direction between the two sensors of the agent. The difference between the new direction \(d_{n+1}\) (direction \(d\) at step \(n + 1\)) and the previous one \(d_n\) (direction \(d\) at step \(n\)) is proportional to this gray level difference. Eq. (1) explains how the agent computes its new orientation.

\[
d_{n+1} = d_n + \frac{\sum_{i=1}^{l} \sum_{j=1}^{l} I_1(i, j) - \sum_{i=1}^{l} \sum_{j=1}^{l} I_2(i, j)}{P} \times S
\]  

Different variables are present in this equation: \(d_n\) is the agent direction in radians at step \(n\), \(I_1(i, j)\) (resp. \(I_2(i, j)\)) is the intensity of a pixel of sensor 1 (resp. sensor 2) with coordinates \((i, j)\) relative to a corner of the square sensor. \(P\) is a coefficient transferring a gray level difference to an angle (here in radians). \(S\) is an integer which value can be +1 or -1. If the agent searches for light rings (resp. dark rings) \(s = +1\) (resp. \(s = -1\)); this sign allows the agent to move in the direction of the sensor giving the extreme values it is searching for.

The coordinates of an agent at step \(n + 1\) depend on its coordinates at step \(n\) as explained in Eq. (2).

\[
x_{n+1} = x_n + \cos(d_{n+1}), \quad y_{n+1} = y_n + \sin(d_{n+1})
\]

\(x_{n+1}\) and \(y_{n+1}\) are the coordinates of the central position \((x, y)\) of the agent at step \(n + 1\) (Fig. 2).

This behavior allows the agent to remain on an edge and follow it step by step, by always trying to keep the edge direction. As these agents seemed interesting for our application, we have found necessary to first analyze the performances of these agents as roof edge detectors, which had not been made previously. For this analysis we will use
synthetic images featuring roof edges and additive gaussian noise. Roof edges are modeled by creating images with vertical lines of constant gray levels decreasing linearly on both sides of the roof top, before noise addition (Fig. 3).

3.1. Edge detection on synthetic images

As the behavior of dark agents is equivalent to the one of light agents, in this section we will only evaluate light agents behavior. The agents behavior, and consequently their ability as edge detectors, depends on four parameters (Fig. 2):

- the distance \( L \) between the central position of the agent and the sensors,
- the angle \( \theta \) separating the orientation of the two sensors,
- the width \( l \) of the square sensors,
- the coefficient \( P \) that divides the difference between the two sensors to compute the deviation.

If the agents have sensors of \( l \times l \) square pixels (Fig. 2), they act as if they were preprocessing the image with a mean filter of size \( l \times l \) before computing their movement. As a mean filter is not always the more accurate filter to preprocess an image, we will now deal with sensors of one square pixel, so that the problem of image preprocessing can be treated separately. To simplify the parameterization, \( \theta \) will also be fixed at the value of \( \frac{\pi}{4} \), which ensures a good directivity of the agents and a sufficient distance between the two sensors. This distance can then be set by varying parameter \( L \).\[14\]

According to the type of edge (roof width, roof slope, noise level), the parameters \( L \) and \( P \) can be tuned to improve edge detection. A method of parameterization is proposed in Ref. [14]. We will here give some examples of the performance of the agents as edge detectors. Their performance can be evaluated using Canny criteria [10]: good detection, good localization, low multiple responses, to compare the ideal edge position and the detected edge. In order to create the image of the detection achieved by the agents, their path is recorded during the processing. The path consists in the coordinates of the central position of the agent (Fig. 2). Each agent records its path in a common image. The initial gray level of this image is zero; the gray level of each pixel is incremented (arbitrarily of 2 gray levels) each time an agent goes on it. As the agents are first placed at random on the image, some pixels will be marked in the background, that are not edge points, even if the edge is completely detected. Nevertheless the non-edge pixels will be less intense than real edge pixels, as the agents are guided by the edge and try to follow it. Fig. 5 illustrates the histogram of the agents’ path image. The original image features a roof edge of amplitude 60 gray levels, of width 3 pixels, altered by a gaussian noise of standard deviation 2 gray levels (Fig. 4(a)). The path of the agents on this image is illustrated in Fig. 4(b).

As we can see in Fig. 5 the histogram of an image representing the agents path features a first lobe in the low gray levels and one peak at high gray level, which corresponds to the pixels intensity on the edge. We can therefore use an automatic threshold to separate the interesting peak from the other intensities. In this method, defined in Ref. [13], a threshold is searched in the gray levels to define two optimal classes in the histogram, so that the variance of the group of pixels in each class is minimized. The result of this operation is presented in Fig. 4(c). This binary image can then be compared with the real edge position using the Canny criteria.

We will here give some results on synthetic images (as in Fig. 4(a)) featuring a roof edge with noise which standard deviation is increasing. The roof width is 3 pixels (minimum value to obtain a roof), the slope is 15 gray levels/pixel, the gaussian noise standard deviation \( \sigma \) increases from 0 to 12. We can see in Fig. 6 that until \( \sigma = 10 \), the detection error remain quite low; for \( \sigma = 12 \), we obtain an important over-detection, which is due to the binarization step. The histogram of agents path in this case does not comprise two lobes, as the intensity of pixels on the edge position does not feature a peak in the histogram, due to the important noise level.

For the images presented in Fig. 6 (results of detection for an image as in Fig. 4(a) described above, with noise standard deviation increasing from 0 to 12), the percentage of wrong pixels PWP has been computed. This criteria is used to summarize the three Canny criteria, represented here by \( c_1, c_2, c_3 \) (cf. Eq. (3)),

![Fig. 4](image-url)

Fig. 4.

![Fig. 5](image-url)

Fig. 5.
and is computed as written below.

\[
P_{\text{WD}} = \frac{\text{number of marked pixels that are not edge pixels} + \text{number of non-detected edge pixels}}{\text{total number of pixels in the image}},
\]

\[
c_1 = \frac{\text{number of detected pixels that are not edge pixels}}{\text{number of non-edge pixels}},
\]

\[
c_2 = \frac{\text{number of non-detected edge pixels}}{\text{number of edge pixels}},
\]

\[
c_3 = \text{mean distance between the detected edge pixel and the real edge pixel position}
\]

The criterium PWP takes into account errors due to under-detection but also over-detection of edges, which allows to evaluate globally the quality of detection using a single variable.

We can see in Table 1 that PWP is very low until \( \sigma = 11 \) (inferior to 0.23\%). Until \( \sigma = 11 \) the criterium \( c_3 \) is also low with a mean distance between detected edges and real edges inferior to 0.05 pixels.

As explained in Ref. [14], parameter \( P \) needs to be increased when noise standard deviation increases, in order to optimize edge detection. Nevertheless the detection is also correct for superior values of \( P \), which guarantees a robust edge detection. To test the robustness of the detection, we have evaluated the criteria in a wide range of \( P \) values for different images with additive gaussian noise of standard deviation 2. On the graph in Fig. 7 we can observe that after a minimum value for \( P \) (\( P = 130 \)), the detection keeps stable on a wide range, which allows to detect different types of edges with the same parameters. To illustrate this property the criterion used for the graph in Fig. 7 is the percentage of wrong pixels PWP, as in Eq. (3).

When roofs are larger and noisy, increasing \( L \) can improve the detection. This point is developed in Ref. [14], and can be illustrated in Fig. 8. The slope of this roof is 1.52 gray levels/pixel and the width is 73 pixels. We have then added gaussian noise of standard deviation 2. This image simulates the type of roofs we can find in otolith images. In Fig. 8(b) with \( L = 2 \) we have an important under-detection of the edge (Table 2), whereas with \( L = 6 \) we have no under-detection but a little over-detection (Fig. 8(c)).

We have analyzed the properties of the agents as roof edge detectors, we will now study how they can be used to detect growth rings on otolith images.

3.2. Ring detection on otolith images

In Section 3.1, agents as roof edge detectors have been evaluated; now another property concerning their behavior

<table>
<thead>
<tr>
<th>( \sigma )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L )</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>( \theta )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
<td>( \Pi/4 )</td>
</tr>
<tr>
<td>( P )</td>
<td>33</td>
<td>35</td>
<td>38</td>
<td>40</td>
<td>48</td>
<td>55</td>
<td>63</td>
<td>74</td>
<td>90</td>
<td>80</td>
<td>118</td>
<td>143</td>
<td>193</td>
</tr>
<tr>
<td>PWP in %</td>
<td>0.07</td>
<td>0.07</td>
<td>0.03</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.13</td>
<td>0.15</td>
<td>0.17</td>
<td>0.05</td>
<td>0.23</td>
<td>33.5</td>
</tr>
<tr>
<td>( c_1 ) in %</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>34.4</td>
</tr>
<tr>
<td>( c_2 ) in %</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( c_3 ) in pixels</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>
will be described. The agents are indeed able to validate a closed loop, by recording their path and detecting if they find again their initial position. If an agent finds one contour, it will write it on a common image, so that any agent that will detect that its central position is on an already known contour will be killed. This behavior prevents from redundancy. Therefore in the case of closed contours, no threshold is necessary to binarize the agents path, as only loops will be validated. This property is interesting to process otolith images, as they are composed of concentric rings. Agents can therefore detect continuous rings in a few seconds, but as we can see in Fig. 9 they encounter problems to detect discontinuous rings. For example in Fig. 9 as the external dark ring is not well contrasted in the left part of the image the agents have bifurcated on the internal dark ring when they have met a dark blob in the upper part of the image. And when the dark ring is written in the common image, light agents are killed when they meet this contour. Fig. 9(a) illustrates this fact, in this image the gray levels have been increased (resp. decreased) each time a light agent (resp. dark agent) has gone on it.

This drawback is particularly due to the fact that agents have only a local perception of the image. We have tried to give them more global information about the shape of the contour to simulate human perception of the continuity of the rings. We present a new way to exploit agents behavior by using an a priori knowledge about otolith growth.

4. Integration of an a priori knowledge

The otolith, as other calcified structures (scales, fin rays, vertebrae) or trees, is composed of nearly concentric rings. Therefore the shape of rings is quite parallel to the global shape of the otolith. Otolith growth starts from its center, which is called nucleus. By comparing the direction of the agent turning around this nucleus with the direction of the contour of the otolith, a decision can be taken to recognize whether its path is correct or not. To do that information concerning the shape of the otolith and the position of its nucleus need to be recorded. This point is based on the perception of the rings achieved by a human reader. Two approaches have been used to control the agents orientation explained in Fig. 10.

4.1. Description of the three first common steps

Three steps are common for the two approaches, which allow to acquire high level knowledge for further processing (Fig. 10).

During step 1 the contour of the otolith is detected and its coordinates are recorded. This step is quite easy because the background mean gray level is very inferior to the mean gray level of the otolith; this contrast is set during the image acquisition using a microscope and a video camera.

During step 2, the position of the nucleus is automatically detected. To find the nucleus, the first white ring, which is the innermost white ring; is searched by agents as they are described in Section 3.2. In most cases the first white ring is well contrasted and agents are able to validate it. So the ring center of gravity, weighted by the gray levels of the ring and

\[
\begin{array}{|c|c|c|}
\hline
\text{i} & \text{Fig. 7(b)} & \text{Fig. 7(c)} \\
\hline
L & 2 & 6 \\
\hline
\text{c}_{1} \text{ in } \% & 0.38 & 0.56 \\
\text{c}_{2} \text{ in } \% & 23.86 & 0 \\
\text{c}_{3} \text{ in pixels} & 0.33 & 0 \\
\hline
\end{array}
\]
corresponding to \((\text{grav}_x, \text{grav}_y)\) is searched as in Eq. (4). In this computation dark points have more weight than light points, as points are ponderated with the complement to 255 (maximal intensity) of their intensity. The nucleus is considered by experts as the central point in the darker area inside the first ring. Thus the minimum gray level pixel is searched in a little neighborhood around \((\text{grav}_x, \text{grav}_y)\) and estimated as the nucleus.

\[
\text{grav}_x = \frac{\sum_{(x,y) \in \text{ring}\_\text{surface}} x \times (255 - I(x,y))}{\text{ring\_surface} \times \sum_{(x,y) \in \text{ring}\_\text{surface}} 255 - I(x,y)},
\]

\[
\text{grav}_y = \frac{\sum_{(x,y) \in \text{ring}\_\text{surface}} y \times (255 - I(x,y))}{\text{ring\_surface} \times \sum_{(x,y) \in \text{ring}\_\text{surface}} 255 - I(x,y)}.
\]

During step 3, two images are created which will be available for the agents. One image will record information concerning the local orientations of the contour. Knowing the position of the nucleus and the co-ordinates of points constituting the contour of the otolith, the orientation is computed as explained below:

\[
\text{angle} = \arctan(\frac{y_2 - y_1}{x_2 - x_1})
\]

In Eq. (5) \((x_1,y_1)\) and \((x_2,y_2)\) are the co-ordinates of two close points of the contour. This information is recorded on a single image (Fig. 11(a)); this image will be used by agents when they are placed once more on the otolith image during steps 4a and 4b. The agents will be able to read the intensity of the pixel, which is relative to the orientation of the contour in the angular sector where they are located.

Another image will be created, which will contain information about the relative localization of each pixel in the otolith image Fig. 11(b). As otolith growth is an accretionary process starting from the nucleus, the shape of last rings is more similar to the one of the contour than the shape of first rings. Thus agents need to take into account this knowledge, to be strongly influenced by the otolith shape near the contour, and only slightly influenced by the otolith shape near the nucleus. This knowledge will be represented by an image which gray levels increase from the nucleus to the edge, as we can see in Fig. 11(b). As the area around the main growth axis is generally well contrasted, agents are kept free in this zone, which means that the influence of the contour is null as the gray level in this zone is zero.

We will now present two approaches that have been developed to detect the rings using these new sources of knowledge.
4.2. First approach: directed agents

In this approach, the agents are placed on the image, as explained in step 4a in Fig. 10 and in Section 4.1. They have now two more sensors, which localization is the same as the central position (Fig. 2):

- one sensor on the orientation image (Fig. 11(a))
- one sensor on the localization image (Fig. 11(b))

At every step they make, the orientation computed with the gray levels of their sensors is compared with the orientation of the contour of the otolith in the area where they are located. From both sources of knowledge they will compute their new orientation $d_{\text{agent}}$ as explained in the equation below:

$$d_{\text{agent}} = d_{\text{local}} + \alpha(d_{\text{contour}} - d_{\text{local}})$$

In this equation, three variables are present: $d_{\text{local}}$ is equivalent to the orientation computed by agents using their sensors on the gray levels of the otolith image as in Eq. (1). $d_{\text{contour}}$ is the orientation of the contour in the angular sector where the agent is located, which is proportional to the gray level in Fig. 11(a). $\alpha$ is a correction strength, which value is in the interval $[0,1]$. $\alpha$ is equal to 0 near the nucleus and equal to 1 near the otolith edge. This value is proportional to the gray level measured by agents in image Fig. 11(b).

Thus if images present problems as in Fig. 9 this new behavior will allow agents to correct their path and avoid wrong trajectories to form rings which global shape is inspired from the otolith contour shape. Fig. 12 illustrates the improvement obtained with this approach, on the same otolith image as in Fig. 9.

In Fig. 12 we can see that all the rings have been detected, which is essential to estimate the right age for an otolith. Nevertheless we can observe that the localization of the rings does not correspond precisely with the edges position, because when agents are constrained to follow the contour, their trajectory may be different from the real shape of the ring. Other drawbacks appear in Fig. 13. On image (a) the shape of the contour is irregular and the real shape of inside rings is quite different from agents forced trajectory. On image (b) rings shape is more regular but many rings are still missing. On these rings the information is concentrated on an area around the main growth axis. Agents may have covered correctly this area but as these rings cannot be well distinguished on other parts of the otolith, the probability of the agent to find again its original position on this type of ring is very low. In this case the partial edge detection achieved by agents is lost.

To avoid these drawbacks, a new approach has been developed to collect as much information as possible, by leaving the agents free on the image and compare their path with the shape of the contour a posteriori. The alternation of light and dark rings will also be taken into account, which was not the case here.

4.3. Second approach: ring reconstruction using agents path (reconstructed path)

4.3.1. Agents path segmentation

In this approach the three first steps are identical as for the first approach (Section 4.1). Then agents are placed once more on the image to detect the rings (step 4b). Agents parameters are selected to fit with ring size: knowing fish specie, otolith size, image scale, and a general growth pattern, ring size can be roughly predicted, so agents parameters are set bigger in the central otolith area and smaller in the external part of big otoliths. Agents compare their orientation computed as in Eq. (1) with the one of the contour, do not achieve any correction (in opposition with step 4a), but record their path on an image if the difference between their direction and the direction of the contour is

![Fig. 12](image1.png)

![Fig. 13](image2.png)
inferior to $\Pi/4$ in this area of the image (Fig. 14(a)). This value is chosen arbitrarily but allows an adaptation on each image, since angular discrepancy between agents path and contour orientation gives a qualitative evaluation of agents work, whatever the size of the otolith is. Another important information is the occurrence of the agents on a pixel, as explained in Section 3.1. This element is also recorded on another image (Fig. 14(b)). Every step an agent makes, the gray level of the pixel corresponding to its present position is incremented on this image. On images Fig. 14(a)–(c), only light agents work is represented, for more visibility; dark agents path would be similar.

Using the images created before (Fig. 14(a) and (b)), agents paths that very likely fit to the rings position are selected, which are the most frequented and best oriented traces. To obtain the image in Fig. 14(c) an automatic thresholding in Fig. 14(b) as explained in Section 3.1 is used and then only the pixels present in both the resulting image and Fig. 14(a) are kept.

Fig. 14(c) gives a rings segmentation which will be used to reconstruct the rings. On this image we find again the almost entire first rings that could have been detected by initial agents (as in Fig. 9), and we can also find a trace of the bigger rings mainly around the main growth axis.

4.3.2. Ring reconstruction using agents path segmentation

As the first ring is almost always detected by basic agents (ignorant of the shape of the contour), its shape will be used to reconstruct the following rings. According to the biological growth process of the otolith starting from the nucleus, a ring is close to the geometric inflation of the previous ring. So the first ring shape will be inflated until an optimal correspondence occurs between the inflated shape and the ring segmented by agents visible in Fig. 14(c). This correspondence is evaluated by maximizing the correlation between the inflated shape and the segmented ring. The transform is made by translating each point of the previous ring using a particular vector. As we have seen in step 1 (Section 4.1), the otolith external contour is detected and its coordinates are recorded in an order following the shape of the otolith. For each point of the external contour, with coordinates $(X_n, Y_n)$ a vector is determined which direction is the direction of a line going through the nucleus and this point, as illustrated in Fig. 15. The first white ring detected by agents is sampled by taking one point in the direction of each vector $V_n$. This first ring is thus divided in points which number is equal to the number of points of the external edge.

Then each point of the first ring is translated using the corresponding vector. These vectors do not have the same modulus. This modulus depends on the otolith axis length in a particular direction, as explained in Eq. (7). This condition means that the maximal modulus will be one pixel in the direction of the main growth axis, and the different modulus of the other vectors will allow the transform to be influenced by the global shape of the otolith, as the speed of points will be proportional to the length of the axis in its sector. 

$$|V_n| = \frac{\text{axis_length_in_direction}_n}{\text{greater_axis_length}} \quad (7)$$

The shape of the previous ring is inflated step by step by translating each point with the associated vector, until the shape reaches the optimal position of the next ring, which is obtained by counting the number of common pixels between the inflated shape and the non-zero pixels present in Fig. 14(c). The shape is then adjusted to obtain the best fitting with the agents’ path. Adjustment is executed for each point in a direction corresponding to the vector associated and in a neighborhood which width is limited by the distance between the inflated shape and the previous ring. This means that if one point of the ring is not over a non-zero pixel present in Fig. 14(c), the nearest non-zero pixel in Fig. 14(c) will be used instead of the initial point position to adjust the shape. After a smoothing step this new ring will be used to search the next ring with the same method. Light and dark rings are searched alternatively until the shape reaches the otolith edge. The result of the method on an eight year old individual is presented in Fig. 16.

5. Results and discussion

The method has been tested on 116 samples of plaice otolith images, which age had been estimated by human experts. Therefore results illustrate discrepancy or adequacy with human interpretation. The number of dark and light rings on images gives the number of seasons the fish has lived, and knowing the season where the fish has been
caught, the experts can estimate the fish age in years. The number of an age group corresponds to the age of the fishes in this group. Age groups are comprised between 1 and 13, with approximately 15 samples per age group from 2 to 8 year old individuals. Results for age groups 1, 9, 10, 11, 12, 13 are less significant since they are not composed of more than four individuals per group.

5.1. Automatic nucleus detection

With a resolution of $512 \times 512$ pixels, the error (distance between the nucleus given by a human expert and the nucleus detected automatically) was inferior to 10 pixels excepted on 10 samples over 116. Fig. 17 illustrates these results. The mean error was 6 pixels and the standard deviation was 10 pixels.

In Fig. 17 we can observe that the most frequent error between human localization and automatic nucleus detection is 3 pixels. This error is not dramatic, since a human expert committed a mean error of 3.55 pixels for this set of images, when localizing the nucleus on the same otolith but at different dates. For only a few individuals the nucleus error is quite important. This can be observed on images where rings are discontinuous or very bad contrasted; therefore the first white ring has not been detected by agents that have no information concerning the contour shape, to correct their path whenever they meet an obstacle during this step. Then if the smaller detected ring is not the first one, its center of gravity may not correspond with the first ring center of gravity (Fig. 18(a)).

In Fig. 18 the cross with vertical and horizontal bars is the detected nucleus, the other crosses are the center of gravity weighted by gray levels inside the ring (on the left on both images) and the geometric center of the ring (on the right on both images). To improve the reliability of automatic nucleus detection, a cooperation could be set with another method, which consists in searching for the principal inertia axis and the darkest area near the middle of this axis.

5.2. Age estimation with the first approach (directed agents)

For age group 1–3, age estimation is quite satisfying with, respectively, 50, 80, and 46.7% of good age estimation (Fig. 19(a)). Then the performances decrease. This is due to the fact that when fish gets older the new rings appearing are only visible on an area centered around the main growth axis, then even if the agents have properly followed these rings in this area they may have encountered problems to follow it on the rest of the otolith, and so they have very little chance to meet again their original position in order to validate a closed contour. Nevertheless as we can see in Fig. 19(b), the percentage of rings properly detected has increased compared to rings detected by free agents (as they are initially described in Ref. [3]), as fanciful loops have been avoided.

5.3. Age estimation with the second approach (reconstructed path)

With this approach the percentage of good age estimation has strongly increased compared to the first approach, with, respectively, 100, 92, and 93% for age group 1–3. The mean of good age estimation for all age groups is 61%. Compared to previous work on plaice [22] which was based on signal envelope detection of integrated profiles,
performances were improved by our study. A comparison with plaice age groups from 2 to 4 shows a 50% error for Wellemans’s work [22] against 15% for the proposed method with a resolution of 512×512 pixels.

One drawback of this method is that the first ring needs to be correct to ensure a good reconstruction of next rings. In Fig. 21(a) only the second ring has been detected during step 2, so that otolith age will be under-estimated; in Fig. 21(b) the first ring shape overlaps the second light ring, and this error creates a bulge and its repercussion on next rings detection. To tackle this problem the first detected ring could be evaluated by comparing its position with edges detected as in Fig. 14(c), and if necessary by reducing its shape to find whether a smaller ring exists with a similar shape in Fig. 14(c).

Nevertheless ring reconstruction is not very sensitive to nucleus detection error. If we simulate an error on the nucleus position on the same image as in Fig. 16, results keep almost constant.

Fig. 22 illustrates the fact that ring reconstruction is not very influenced by little errors on nucleus localization. On the two images on the first line, the real nucleus position has been translated of 10 pixels on the right or on the left. On the next line the nucleus has been translated vertically and also horizontally of 10 pixels in both directions. On the last line the nucleus has been translated vertically of 10 pixels up and down. We can observe that rings shape are a little different but the same number of rings has been estimated, and their position are globally similar.

This method is more time consuming than the first approach presented (Directed Agents), with approximately 15 s for the first approach against 1 min with the second approach (Reconstructed Path), with 512×512 pixels images processed on a workstation (195 MHz). Nevertheless this processing time is inferior to previous methods, with for example 2–4 min with deformable templates method [21].

Another limit for this method could be the tuning of the agents parameters to adapt to the ring size, which is not easy because of the variability of samples. The method presented in Section 4.3 is also sensitive to the good detection of the first ring for further reconstruction of other rings.

5.4. Discussion

In Table 3, the percentage of good age estimation obtained with the two methods presented in this paper is compared with methods previously developed. The 1D method is a mono-dimensional method based on extraction of radials starting from the nucleus and going to the external edge to search extremes in these radials [22]. This method
cannot take into account structures continuity and the results are lower than those obtained with the methods we propose.

The Template method uses the shape of the external otolith edge reduced by an homothetic transform centered on the nucleus, in order to search the position of growth rings [21]. Their position is determined by minimizing an energy function, which is computed with the gray levels of the image. This method can only give an approximate shape to rings and presents quite low results for age groups from 5 to 8, in comparison with the Reconstructed Path method.

The graph method uses a polar coordinates transform to extract radials starting from the nucleus and going to the edge in the original image to display them one under the other in the final image [19]. Peaks, corresponding to light rings and valleys corresponding to dark rings are then extracted using morphological transforms. Objects are then labeled and closest objects are connected to reconstruct rings. This method gives results similar to the Reconstructed Path method, but cannot be applied to otoliths having several nuclei because of the polar coordinates transform (Fig. 23). Moreover it is more time consuming.

The Deformable Model method presented in Ref. [6], inspired by Ref. [11] cannot be easily compared with the methods we propose because it has not been evaluated on a sample of test images. However, this method may encounter problems to detect thin rings, which are visible around the main growth axis for old individuals, as the energy function computed to inflate the model, uses gray levels all around the otolith.

Processing time for estimating age with Directed Agents and Reconstructed Path on otolith images may be longer than the time necessary for a human person to estimate age. Nevertheless a human person is not able to draw precisely by hand the position of the rings, or this would be a very long and tedious task. Moreover the same person, at different times, or two different people, can make different estimations for the same image.

Ring detection on otolith images is a very complex problem, as the number and the shape of structures to detect in an image is a priori unknown. Thus it is essential to use high-level knowledge to improve this detection. For this reason it may be difficult to compare the methods we propose in this paper with more general methods presented in the literature. However, agents present the advantage to perceive locally roof edges and structures continuity in the same time. Agents can also cooperate by sharing their work on the same image. The use of snake curves thrown on the image instead of agents could be suggested. Nevertheless snakes detect step edges, and in our application detection of peaks and valleys is more appropriate to localize growth rings. Moreover if agents are initialized on the image in an area where no ring is visible, they are able to move from this
area, which would be more difficult for a snake, if its energy function doesn’t allow it to evolve.

6. Conclusion

We have developed a method to perceive continuity of contours in textured, noisy and low contrast images, using autonomous agents. We have applied it to growth rings detection on fish otoliths and used the shape of the otolith edge as high-level information. Two methods were presented in this aim; the first one consists in constraining the agents to keep parallel to the external otolith contour, the second one consists in reconstructing rings using correct agents paths. In our method, detected rings shape is not completely constrained to fit with otolith contour shape, in opposition with deformable templates methods [21], as agents have a certain freedom to follow the position of real extremes in the image. Therefore rings detected have a more realistic shape. The agents can also adapt locally their parameters, which allows to detect simultaneously large rings near the nucleus and thin rings near the contour. Age estimation results on old individuals are indeed encouraging (Fig. 20).

Plaice otolith images were used to develop automatic age estimation methods as they are quite simple compared to other species. Nevertheless we hope to be able to adapt the method to more complex species (Fig. 23), which can have several nucleus and less regular shapes.

We have also proposed an automated method to detect the nucleus, which was previously interactively placed by a human operator. Ring detection results are encouraging but improvements have to be achieved with the adaptation of the agents to get more precision about thin rings appearing near the otolith edge, mainly for old individuals. In the case of plaice otolith images, the choice of the nucleus as the homothetic transform center is relevant, but it should be recanted for other types of fishes, which otolith growth are less regular.

References


<table>
<thead>
<tr>
<th>Method ID</th>
<th>Template method</th>
<th>Directed agents</th>
<th>Graph method</th>
<th>Computing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 110</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–3 Yrs</td>
<td>R = 90% AM = 1</td>
<td>R = 65% AM = 2</td>
<td>R = 87% AM = 3</td>
<td></td>
</tr>
<tr>
<td>3–5 Yrs</td>
<td>R = 90% AM = 2</td>
<td>R = 95% AM = 3</td>
<td>R = 65% AM = 2</td>
<td></td>
</tr>
<tr>
<td>5–8 Yrs</td>
<td>R = 90% AM = 3</td>
<td>R = 95% AM = 4</td>
<td>R = 65% AM = 3</td>
<td></td>
</tr>
</tbody>
</table>

R: percentage of good age estimation, AM: maximal error amplitude.