Automated fish age estimation from otolith images using statistical learning

Ronan Fablet, Nicolas Le Josse

Automated fish age estimation from otolith images using statistical learning

Ronan Fablet [a] Nicolas Le Josse [a]

[a] *Ifremer/LASAA, BP 70, 29280 Plouzané,France, France*

Abstract

The acquisition of age and growth data is of key importance for fisheries research (assessment, marine ecology issues,...). Consequently, automating this task is of great interest. In this paper, we investigate the use of statistical learning techniques for fish age estimation. The core of this study lies in the definition of relevant image-related features. We rely on the computation of a 1D representation summing up the content of otolith images within a predefined area of interest. Features are then extracted from this non-stationary representation depicting the alternation of seasonal growth rings. Thus, fish age estimation can be viewed as a multi-class classification issue using statistical learning strategies. In particular, a procedure based on demodulation and remodulation of fish growth patterns is used to improve the generalization properties of the trained classifiers. The experimental evaluation is carried out over a dataset of 320 plaice otolith images from age groups 1 to 6. We analyze both the performances of several statistical classifiers, namely SVMs (Support Vector Machines) and neural networks), and the relevance of the proposed image-based feature sets. In addition, the combination of additional biological and shape features to the image-related ones is considered. We reach a rate of correct age estimation of 88% w.r.t. the expert ground truth. This demonstrates the relevance of the proposed approach for the automation of routine aging and for computer-assisted aging.

Key words: statistical learning, SVMs, neural nets, otolith images, automated fish aging, performance evaluation, plaice dataset

1 Problem statement and related work

The analysis of fish age-length keys is at the core of stock assessment and marine ecology issues. Acquiring fish age data is mainly performed by interpreting fish otoliths. These calcified structures are formed through an accretionary process, which results in alternated translucent and opaque rings, as illustrated by Fig.1 for a plaice otolith. Fish age is then determined by counting the number of translucent or opaque rings. This tedious task is achieved by expert readers, who typically interpret several thousand of otoliths a year. Computer-assisted tools are then sought to help in the acquisition of age-length keys and in the storage of otolith interpretation (Morison et al., 1998; Small and Hirschhorn, 1987; Ogor and Fablet, 2004).

[Figure 1 about here.]

As far as the automated estimation of fish age and growth is concerned, two broad categories of approach can be distinguished. One the one hand, 1D methods restrict their analysis to an intensity profile extracted along a given reading axis. The issue is then to count and detect its meaningful extrema. Different techniques have been proposed using for instance filtering techniques (Welleman and Storbeck, 1995) or time-frequency analysis (Lagardère and Traodec, 1997). The main drawback lies in the absence of 2D perception of ring continuity, being thus sensitive to local artifacts in otolith images. The evaluation carried out on plaice otolith sets shows

Email address: rfablet@ifremer.fr (Ronan Fablet).
URL: www.ifremer.fr/lasaa/rfablet (Ronan Fablet).

that these 1D approaches achieve only low performance (Guillaud et al., 2002). On the other hand, 2D ring segmentation has also been investigated: deformable templates deduced from the external shape of the otolith (Traodec et al., 2000), a graph framework (Rodin et al., 1996), locally deformable spline-based models (Benzinou et al., 1997) and multi-agent systems (Guillaud et al., 2002). Important improvements were obtained for fish age estimation (Guillaud et al., 2002) compared to 1D approaches.

[Figure 2 about here.]

All these 1D and 2D techniques simply performed age and growth estimation by validating all the detected rings. However, even for rather "easy-to-read" species, such as plaice, the interpretation of fish otoliths remains a complex task. Two main issues arise:

- The discrimination of actual growth rings from false rings or checks as illustrated by Fig.2. The presence of false rings is often due to biological or environmental events (sexual maturity, environmental stress,.....). Both ring strength (in terms of image contrast), ring length and inter-ring distances are exploited by expert readers, even if the definition of quantitative criteria is difficult;
- the high biological variability both in terms of ring appearance and shape, and of otolith growth, illustrated for plaice otolith growth by Fig.7. This implies that it remains difficult to define objective and quantitative criteria (in terms of ring appearance and in terms of ring growth increments) to discriminate actual growth rings from false rings.

These two main difficulties explain that only expert readers with an important experience for a given species and stock can provide reliable age estimate, since they have usually observed several ten thousands of otoliths before being consistent in

their otolith interpretation. This difficulty in defining well-formalized criteria for the interpretation of otolith rings has been emphasized by the results of a European workshop on plaice otolith reading held in 2003: inter-reader agreement rates vary from 40% to 95% depending on the readers' experience for fish samples from age groups one to six, whereas this rate is indeed comprised between 85% and 95% for expert readers. These results actually points out the complexity of the otolith interpretation even for "easy-to-read" species.

As far as the automation of fish age estimation is concerned, the first consequence is that we are required to go beyond the simple detection of growth rings to carry out the actual interpretation of otoliths from image-based features. Besides, to achieve satisfactory results, it is obvious that one needs to benefit somehow from the experience of expert readers to design a computer-based aging system. Statistical learning approaches are then particularly appealing. They are known to be capable of handling complex biological phenomena, and they allow the system to benefit from otolith samples interpreted by expert readers to infer the underlying aging rules. This issue was first considered in (Robertson and Morison, 1998) using neural approaches. Though promising, reported results show that features extracted from otolith images do not bring significant cues compared to biological characteristics (fish length, otolith weight, fish sex,...). This contradicts readers' experience, who mainly rely on otolith interpretation for age determination. In fact, biological features are considered only to check possible unrealistic age estimates. Theses results might be due to the use of unadapted image-based features: features formed by the real and imaginary parts of the Fourier coefficients of an intensity profile along a given reading axis were considered (Robertson and Morison, 1998). Such a frequency analysis is generally not suited to deal with otolith images for two main reasons. First, yearly otolith growth is not linear but rather exponential (see

Fig.3.a). This results in non-stationary image features in terms of frequency content (i.e., the frequency content of the analyzed signal is “time-dependent”). Second, the otolith is usually thicker around the center. Consequently, image intensities tend to become brighter close to the edge. This leads to the observation of a continuous trend, as illustrated by Fig.3.b. These two important characteristics of otolith image stress that Fourier coefficients should be used as relevant descriptors of the image content.

[Figure 3 about here.]

In this paper, we also investigate statistical learning for fish aging applications, but with an emphasis on taking into account these two key features of otolith images. We focus on the definition of new image-based features. A semi-local approach is presented in Section 2 to characterize the 2D image content within a given area of interest. Besides, we exploit an *a priori* knowledge on otolith growth to perform an appropriate filtering of the extracted intensity-based representation with a view to enhancing the structures of interest (i.e., image valleys and ridges). We finally present two types of image-based features using either frequency or peak information. Section 3 outlines the application of neural networks and Support Vector Machines (SVMs) to the considered multi-class classification issue. In particular, a procedure targeted at increasing the generalization properties of the classifiers is presented. The experimental evaluation reported in Section 4 is carried out on a dataset of 320 otolith images. We evaluate the performances of several statistical classifiers and the relevance of the proposed image-based feature sets as well as their combination to biological and geometrical descriptors. Section 5 provides concluding remarks and outlines for future work.

2 Image-related features

The core of our study consists in extracting relevant features to characterize the content of otolith images, in terms of alternation of translucent and opaque rings. We investigate the use of two different feature sets: frequency-based and peak-based features. The key components of our scheme are a semi-local template-based framework to compute a 1D representation of the 2D image content within a given area of interest, and a growth-adapted filtering to cope with non-stationary features of otolith images.

2.1 From the 2D image content to a 1D representation

Otolith interpretation is usually performed by experts along a predefined reading axis, as shown by Fig.1 for a plaice otolith. Therefore, the automated analysis of the content of otolith images can be restricted to a region around the main reading axis. Within the considered region of interest, we aim at summing up the associated 2D image content to a 1D representation depicting the alternation of the translucent and opaque rings. This more compact representation will ease the extraction of relevant features.

[Figure 4 about here.]

To this end, we use a template-based approach inspired from (Traodec et al., 2000) as illustrated by Fig.4. More precisely, given the otolith center 0 and an angular sector $\mathcal{S}$, the following 1D signal $s_{\mathcal{S}}$ is computed:

$$s_{\mathcal{S}}(\alpha) = med(I(O + \alpha * T_{\mathcal{S}})) \quad (1)$$

where $med()$ is the median operator, $T_{\mathcal{S}}$ is the template model used within $\mathcal{S}$, I the image intensity function, and α a scaling factor between 0 and 1. As in (Traodec et al., 2000), the template model $T_{\mathcal{S}}$ is given by the sampled external otolith shape within $\mathcal{S}$. $s_{\mathcal{S}}(\alpha)$ is then the median intensity value along the template $T_{\mathcal{S}}$ scaled by α w.r.t. the otolith center O.

Fig.5.b depicts the signal $s_{\mathcal{S}}$ extracted for the otolith image given in Fig.1. Since the scaling of the external otolith shape w.r.t. the otolith center provides locally a good approximation of the shape of the seasonal rings (Traodec et al., 2000), $s_{\mathcal{S}}$ depicts oscillations corresponding to the alternation of translucent and opaque ring. It then provides a representation of the image content within $\mathcal{S}$, and features describing these oscillations will supply relevant information for fish aging.

2.2 Growth-adapted filtering

As previously stressed, the signal $s_{\mathcal{S}}$ representation involves a frequency modulation associated with the non-linear otolith growth. Therefore, a straightforward use of the Fourier transform as in (Robertson and Morison, 1998) is not appropriate. In addition, to extract meaningful image features, noise artifacts as well as the continuous trend of $s_{\mathcal{S}}$ should be removed. Since the relevant frequency content of the signal $s_{\mathcal{S}}$ varies with the distance to the otolith center, we need to locally adapt the scale at which the signal is analyzed to perform this pre-processing step.

As a consequence, we propose a growth-adapted filtering technique. We first perform a demodulation step w.r.t. a mean *a priori* growth model. Then, noise artifacts and the continuous trend are filtered out using Gaussian filters applied to the demodulated signal. A growth-adapted filtering is thus introduced to analyze $s_{\mathcal{S}}$.

Given a mean *a priori* growth model $L = \Phi(t)$, with L the distance to the otolith center, and t the time variable in years, $s_{\mathcal{S}}$ is initially demodulated w.r.t. Φ:

$$s_{\mathcal{S}}^{DM}(t) = s_{\mathcal{S}}(\Phi(t)), \tag{2}$$

The mean growth model $L = \Phi(t)$ is computed from a sample set of growth patterns estimated on otoliths by an expert. Though non-stationary, the demodulated signal $s_{\mathcal{S}}^{DM}$ involves a more compact frequency content, what makes easier its analysis.

To remove the continuous component $s_{\mathcal{S}}^{T}$ from $s_{\mathcal{S}}^{DM}$, $s_{\mathcal{S}}^{T}$ is estimated using a convolution to a Gaussian kernel g_{σ_T} with a large variance σ_T^2: $s_{\mathcal{S}}^{T} = g_{\sigma_T} * s_{\mathcal{S}}^{DM}$. High-frequency noise can also be filtered out using a Gaussian kernel with a low variance σ_G, such that the growth-adapted filtering comes to consider the filtered signal $s_{\mathcal{S}}^{G}$ computed as:

$$s_{\mathcal{S}}^{G}(t) = g_{\sigma_G} * \left[s_{\mathcal{S}}^{DM} - s_{\mathcal{S}}^{T} \right] \tag{3}$$

Fig.5 illustrates the proposed growth-adapted filtering: Fig.5.a depicts the mean *a priori* growth model used for demodulation purposes, Fig.5.b the signal $s_{\mathcal{S}}$ extracted for the otolith image given in Fig.1 and Fig.5.c its demodulation version $s_{\mathcal{S}}^{DM}$ (–), the continuous component $s_{\mathcal{S}}^{DM}$ (-.) and the filtered signal $s_{\mathcal{S}}^{G}$. This example has been processed with the parameter setting detailed in Section 4. Compared to the original 1D representation $s_{\mathcal{S}}$, the filtered version $s_{\mathcal{S}}^{G}$ provides all useful information for aging, while making further analysis easier.

[Figure 5 about here.]

2.3 Frequency-based representation

The frequency content of the filtered signal $s_{\mathcal{S}}^{G}$ might offer a relevant descriptor of the alternation of ridge and valley rings within $\mathcal{S}$. Rather than computing the Fourier coefficients of $s_{\mathcal{S}}^{G}$, we resort to the DCT (Discrete Cosinus Transform) coefficients of $s_{\mathcal{S}}^{G}$. The DCT coefficients are computed as the decomposition of the signal onto a basis formed by cosinus functions. Compared to the Fourier transform, it has been proven to lead to more compact representation (Mandyam et al., 1997). Given the DCT coefficients of $s_{\mathcal{S}}^{G}$, we consider the N_{DCT} first DCT coefficients to form the feature vector (typically, $N_{DCT} = 200$).

2.4 Peak-based representation

The analysis of the protocols followed by experts to interpret otoliths leads us to define a second kind of features. More precisely, experts do not rely on the analysis of the image frequency content, but rather on the analysis of the relative positions of the likely translucent and opaque rings. Consequently, we define a new feature vector from the positions of likely growth rings.

Within the angular sector of interest $\mathcal{S}$, growth rings correspond to extrema of the signal $s_{\mathcal{S}}^{G}$. We first compute the zero-crossings of the first-order derivative of the signal $s_{\mathcal{S}}^{G}$ to determine the positions $\{t_1, ..., t_n\}$ of its local maxima. The proposed peak-based representation $s_{\mathcal{S}}^{PB}$ is then defined by:

$$s_{\mathcal{S}}^{PB}(t) = min\left\{\rho(t - t_k)\right\}_{k\in\{1,...,n\}} \tag{4}$$

where ρ is a distance kernel. In practice, we use $\rho(u) = 2atan(\alpha \cdot u)/\pi$, with α a scaling parameter. The resulting signal $s_{\mathcal{S}}^{PB}$ only depends on maxima positions.

In particular, compared to s_S^G, s_S^{PB} is invariant to the variability of intensity ranges over images.

[Figure 6 about here.]

Fig.6 supplies an example of the peak-based representation for the otolith image depicted in Fig.1. If the extracted maxima refer to the actual translucent rings, the proposed peak-based representation can be viewed as a continuous version of the associated growth pattern. The computation of this peak-based representation is also similar in spirit to the computation of signed surface to characterize 2D shapes (Leventon et al., 2000). The vector formed by the values of the signal s_S^{PB} is called hereafter the peak-based feature vector.

3 Statistical learning applied to fish aging

Given the feature vectors extracted from otolith images, fish age estimation is regarded as a pattern classification issue. Given an otolith image, we indeed aim at mapping the associated feature vector x^n to an age class y^n. To cope with this multi-class classification, we assume that a reference set of labeled examples is provided to achieve statistical learning. We focus on two particular techniques: neural networks (Bishop, 1995) and SVMs (Scholkopf and Smola, 2002).

3.1 Feature vectors

In addition to the feature types introduced in the previous Section, we might also exploit biological and geometrical features. Our aim will be to evaluate the relative relevance of each feature set w.r.t. the classification task.

The biological features are mainly the length, the sex and the catch date. If available, other information such as the sexual maturity or the otolith weight could also be considered. Although these features might be relevant for fish aging issues (Robertson and Morison, 1998), growth variability among fish, as illustrated by Fig.7, is usually too high to supply accurate age estimates using only such biological information.

Geometrical features computed from the external shape of the otolith might also be considered (M. Cardinale and Mosegaard, 2004). We typically compute the perimeter, the surface, and the rate between the length along the two principal axis. This type of feature is obviously not sufficient alone to infer fish age. But, we are interested in evaluating whether or not it can bring additional cues in combination to other feature types to improve fish age estimation.

3.2 Neural networks

Neural networks have been widely used for pattern recognition issues (Bishop, 1995). They are aimed at optimizing the interconnection weights between neurons to minimize the mean square error. Here, a back propagation neural network with one hidden layer is chosen. The decision function is a sigmoid. The number of neurons of the hidden layer has been empirically optimized, and set to fifty. In our experiments, we use the Netlab Matlab toolbox (Nabney, 2001).

3.3 SVM

Kernel-based techniques such as SVMs (Scholkopf and Smola, 2002) are among the most efficient approaches for pattern recognition. The key idea of SVMS, which

are maximum margin classifiers, is to map the original feature space to a higher dimensional space. The selected mapping Ψ has to verify that the dot product between two points in the new feature space $\Psi(x_1) \cdot \Psi(x_2)$ can be rewritten as a kernel function $K(x_1, x_2)$. The binary SVM classification then resorts to determine the hyperplane separating the two classes within the feature space associated to the kernel function. We let the reader refer to (Scholkopf and Smola, 2002) for further details.

In our study, a Gaussian kernel $K(x_1, x_2) = exp(-(x_2 - x_2)^2/2\sigma^2)$ was experimentally proven to be the most efficient compared to polynomial and sigmoid kernels. The binary SVM classification is extended to a multi-class classification using a one vs. all strategy. We use the libSVM package (Chang and Lin, 2003) to test for two different SVM techniques: C-SVM and nu-SVM (Chang and Lin, 2003; Scholkopf and Smola, 2002).

3.4 Training issues

Given a training set $\mathcal{T}$ of otolith images and the associated feature vectors $\{x^k\}_{t \in \mathcal{T}}$ (x^k possibly refers to frequency-based, peak-based, biological or shape features, or a combination of these features), we first apply a Principal Component Analysis (PCA) (Bishop, 1995) to the image-related feature vectors to retain the first N_{PCA} components. Statistical learning is then performed within the resulting feature space.

To improve the robustness of the learning stage, we virtually increase the number of elements within the training set $\mathcal{T}$ as follows. Thanks to the interpretation provided by the expert, any otolith image k within the training set is associated with

a growth pattern $L = \Phi_k(t)$ (i.e., the growth pattern plots the distance from the center to the edge as a function of the age). Using these growth data, we apply a demodulation/remodulation procedure. Given $k \in \mathcal{T}$, the template-based signal s_S^k (cf. Eq.1) is demodulated w.r.t. the growth model $L = \Phi_k(t)$ and remodulated w.r.t. growth models randomly chosen within $\{\Phi_n\}_{n\in\mathcal{T}}$. This scheme generates new virtual signal, which are realistic in terms of growth information.

The computation of image-based feature vectors is then applied to the combination of original and virtual signals to create a larger training set. The remodulation is typically performed with five growth models. Thus, the size of the training set is multiplied by a factor six. This is expected to improve the generalization behavior of the trained classifiers.

4 Experiments

4.1 Training and test sets

[Figure 7 about here.]

The experiments are carried out for a database of 320 images of plaice otoliths from age group 1 to 6. The ground truth for fish age estimation is provided by the interpretation of an expert in plaice otolith readings. Due to the lower number of samples in age groups 5 and 6, we merge both groups to end up with a five-class issue: age groups 1 to 4 and 5+. This grouping can also be motivated by the fact that age groups 4 and less usually represent more than 85% of the commercial landings. The age distribution of this otolith set is given in Fig.7.a. As an illustration, Fig.8 depicts four sample images for each age group. We also outline with Fig.7.b the

biological variability within the processed dataset, which makes difficult the aging issue. More precisely, this plot proves that there is a high overlapping in the distance from the otolith center to the first three translucent rings. This feature makes unfeasible an accurate *a priori* prediction of the ring positions and leads to a complex multi-class classification issue.

[Figure 8 about here.]

The training set is formed by selecting the two third of the images of each age class, while the remaining images are used for the test set. For each image, image-related features are extracted for five non-overlapping angular sectors of width $\pi/50$ around the main reading axis. Angular sectors, for which information required to interpret the otolith is missing due to ring discontinuities or local artifacts, are discarded from the training set (in practice, about 20% of the training data were discarded), but no such checking is carried out over the test set. We resort to training and test sets including respectively 800 and 600 elements. Using the virtual growth-based signal generation, the training set finally comprises 6000 elements.

In the reported experiments, we use the following parameter setting. The growth-adapted filtering is performed with $\sigma_T = 1.5$ and $\sigma_T = 0.05$. For frequency-based representation, we use $N_{DCT} = 200$ DCT coefficients. The peak-based representation is computed with the parameter α of the kernel distance ρ set to 0.1. When applying PCA, we keep $N_{PCA} = 40$ coefficients. Besides, the hyper-parameters of SVMs and neural nets have been set by cross-validation.

4.2 Classification results over the test set

We evaluate the performances of the proposed approach using different feature sets and classifiers. The performances are analyzed in terms of mean correct classification rate over the test set w.r.t. the five age groups. The overall evaluation shows that the best results (86%) are obtained by C-SVMs (Scholkopf and Smola, 2002) combined to the peak-based features and the remodulation procedure within the training set. Thus, this configuration is used as the reference in the subsequent.

[Figure 9 about here.]

We first compare in Fig.9 four statistical classifiers: C-SVMS and nu-SVMs with a Gaussian kernel, neural nets (NN) and linear C-SVMs. These experiments stress the complexity of the considered classification task since linear techniques achieve only poor results. Otherwise, neural nets are also significantly outperformed by the two SVM techniques, while C-SVMs are only slightly better than nu-SVMs.

The second evaluation reported in Fig.10 is concerned with the comparison of different feature types using C-SVMs: peak-based features (PB), PB features combined to the remodulation scheme within the training set (Rem+PB), Frequency-based features combined to the remodulation scheme (Rem+FB), Fourier coefficients of the template-based signal $s_{\mathcal{S}}$ as used in (Robertson and Morison, 1998) (S&R), biological features (B) and geometrical features (G). These results indicate that biological and geometrical features are far too simple to provide alone relevant cues to infer fish age. The influence of the growth-adapted filtering is also demonstrated by the rather low performance (69%) achieved using the frequency features from the template-based signal as in (Robertson and Morison, 1998). Let us stress than in (Robertson and Morison, 1998) frequency features are computed for a 1D

intensity profile straightforwardly extracted along the considered reading axis, what should even lower the classification performances given that we use here the proposed template-based semi-local scheme. Besides, the peak-based representation is shown to outperform the frequency-based features (86% versus 80%). Finally, the remodulation scheme within the training set actually improves the generalization properties of the trained C-SVM classifier to result in an overall correct classification of 86% compared to 82.6% without the remodulation step.

[Figure 10 about here.]

[Figure 11 about here.]

The last experiments evaluate the impact of the combination of the peak-based features to biological and geometrical features. Surprisingly as reported in Fig.11.c, this does not bring any improvement compared to PB features alone. Tab.1 provides the confusion matrix of the classification results for the PB features alone and their combination to the geometrical features. The confusion matrix provides the percentage of samples from any class i assigned to any class j (for instance in Tab.1, 88.3% of the samples from class 1 are assigned to class 1, 10% to class 2, 0.8% to class 3 and 0.8% to class 5+). It stresses that this combination tends to decrease the mean correct classification rate, while the mean-square error is slightly lowered (0.54 for the combination of geometrical and PB features and 0.57 for the PB features alone). This is certainly due to the high overlapping in terms of length and shape features between the different age groups, which decreases the discriminant power of these features.

[Table 1 about here.]

4.3 Extension to image classification

From this overall evaluation over the test set, the best configuration is to combine C-SVMs to PB features and to exploit the remodulation procedure. The reported classification results refer to the estimation of fish age from the analysis of one angular sector. Consequently, one can deduce a more robust scheme by combining age estimations for several angular sectors around the main reading axis. Exploiting age estimations for five non-overlapping angular sectors around the main reading axis as detailed previously in the parameter setting, a straightforward voting procedure is used to infer fish age from one otolith image. Tab.2 reports the confusion matrix of the resulting image-based classification. The mean correct classification rate is 88%, and only age group 4 has a classification rate just below 80%.

Compared to previous work using statistical learning for fish aging (Robertson and Morison, 1998), the reported classification rates are much greater due to the definition of more relevant image-based features. Besides, as reported in (Guillaud et al., 2002), previously proposed 1D and 2D techniques are both outperformed by the proposed statistical learning scheme with respectively 50% and 80% of correct classification. In addition, compared to 2D approaches which require about one minute per image, the SVM classification from PB features just needs a few seconds on a 2.4GHz PC in a Matlab implementation.

[Table 2 about here.]

Inter-expert agreement rates for plaice otolith readings were recently evaluated between 85% and 95% during the European workshop held in 2003. Thus, our automated scheme shows performances comparable to expert ones, what demonstrates its relevance for an application to routine aging.

5 Discussion and future work

We have dealt with automated fish aging from otolith images using statistical learning. We have in particular defined a new set of image-based features to characterize the alternation of opaque and translucent growth rings within otolith images, and developed a remodulation scheme to improve the generalization properties of the trained classifiers. The evaluation carried out on a dataset of 320 plaice otolith shows that the best performances are obtained using C-SVMs combined to the proposed peak-based representation and to the remodulation step before training. With an overall correct classification rate raising 88%, the proposed scheme outperforms previous work and can be favorably compared to inter-expert agreement rates comprised between 85% and 95%.

Considering the proposed scheme as an alternate aging procedure to the acquisition of fish age data by expert readers, different issues arise. First of all, a fully automated scheme requires to combine the technique proposed in this paper to the automation of the acquisition of otolith images and of the detection of the otolith nucleus as proposed in (Cao et al., 2004). For routine aging, the cost-benefits analysis of such automated system need to balance between the accuracy reached by the system compared to readers and the gain in terms of processing time. As stressed by our results, for an “easy-to-read” species, such as plaice, the potential gain is huge, since the accuracy levels of the automated system should be in the range of the inter-expert agreement rates, while being greater than those observed for in-training readers. Besides, even if the proposed automated system consider only the younger age groups (up to 6), these age groups represent up to 85% of the sampled data for the considered species and stock. More generally, while the number of age groups to be considered needs to be kept rather low to apply statistical learn-

ing techniques, the use of the proposed scheme for the most important age groups should at least highly reduced the work to be done by expert readers. Finally, let us stress that in any case the use of statistical learning techniques implies to provide the system with interpreted data. Consequently, expert readers will still be required to interpret at least a representative subset of otoliths to update the trained model and to prevent errors due to important changes over time in ring appearance and in otolith growth patterns.

The proposed age estimation technique is not only an alternative to expert readings. It can also be used in a number of different ways to assist the readers for routine aging. For instance, it could provide a low-cost solution for a second reading of otolith samples to evaluate confidence measures of the age estimation. Such measures are of great interest for assessment models (S.A.Reeves, 2003). In addition, the proposed system also provides an automated tool to detect very unlikely age estimates by comparing automatical and reader age estimates. Thus, the readers would be automatically asked for a second reading for unlikely samples. Besides, interactive tools could also be designed for training purposes. These different potential applications stress that the proposed system is not limited to the automation of fish age estimation. More importantly, it can provide a series of new tools for computer-assisted aging in the context of quality assurance and quality control (Morison et al., 1998).

The application of our approach to other species with more difficult otoliths and longer-lived species needs to be investigated. The proposed system comprised two main steps: the feature extraction step leading to the definition of the peak-based image features, and the interpretation step relying on statistical learning. Let us point out that the proposed approach is generic in the sense that the exploited peak-based representation of the image content within a region of interest can be de-

fined from the outputs of any ring detection algorithm (Guillaud et al., 2002; Rodin et al., 1996). As a consequence, for difficult species, for which the proposed local template-based scheme might not be suited, our future work will focus in designing appropriate ring detection techniques. The application of our scheme to longer-lived species might be an issue given the large number of age groups to be considered. A possible solution would be to exploit statistical learning as regression tools rather than classification tools (Scholkopf and Smola, 2002).

Acknoledgments

The authors wish to thank André Ogor from for his help in the acquisition of image data, and Marie-Line Manten who provides the expert interpretation of the image of plaice otoliths.

References

Benzinou, A., Traodec, H., Bihan, J. L., and Rodin, V. 1997. The locally deformable B-Bubble model: an application to growth ring detection on fish otoliths. In *Proc. of th 10th Scand. Conf. on Image Analysis, SCIA'97*, pages 181–187, Lappeenranta.

Bishop, C. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press. 2

Cao, F., Bouthemy, P., and Fablet, R. 2004. Automatic morphological detection of otolith nucleus. In *Proc. of 17th Int. Conf. on Pattern Recognition, ICPR'04*, Cambridge, UK.

M. Cardinale, P. Doering-Arjes, M. K. and Mosegaard, H. 2004. Effects of

sex, stock, and environment on the shape of known-age Atlantic cod (Gadus morhua) otoliths. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(2):158–167.

Guillaud, A., Benzinou, A., Troadec, H., Rodin, V., and Bihan, J. L. 2002. Autonomous agents for edge detection and continuity perception on otolith images. *Image and Vision Computing*, 20(13-14):955–968.

Lagardère, F. and Traodec, H. 1997. Age estimation in common sole *solea solea* larvae: validation of daily increments and evaluation of a pattern recognition technique. In *Marine Ecology Progress Series*, volume 155, pages 223–237.

Leventon, M., Grimson, E., and Faugeras, O. 2000. Statistical shape influence in geodesic active contours. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'2000*, pages 316–323, Hilton Head Island.

LibSVM, 2003. LibSVM: a library for support vector machines. www.csie.ntu.edu.tw/cjlin.

Mandyam, G., Ahmed, N., and Magotra, N. (1997). Lossless image compression using the discrete cosine transform. *Jal of Vis. Comm. and Im. Repr.*, 8(1).

Morison, A., Robertson, S., and Smith, D. 1998. An integrated system for production fish aging: image analysis and quality insurance. *North American Journal of Fisheries Management*, 18:587–598.

Nabney, I. 2001. *Netlab: Algorithms for Pattern Recognition*. Springer.

S.A.Reeves 2003. A simulation study of the implications of age-reading errors for stock assessment and management advice. *ICES Jal of Marine Science*, 60:314–328.

Robertson, S. and Morison, A. 1998. Development of an artificial neural network for automated age estimation. Technical Report 98/105, Marine and Freshwater Resources Institute.

Rodin, V., Traodec, H., Pontual, H. D., Bihan, J. L., Benzinou, A., and Tisseau, J.

1996. Growth ring detection on fish otoliths by a graph construction. In *Proc. of 3rd Int. Conf. on Image Processing, ICIP'96*, volume 2, pages 685–688, Lausanne.

Scholkopf, B. and Smola, A. J. 2002. *Learning With Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press.

Small, G. and Hirschhorn, G. 1987. Computer-assisted age and growth pattern recognition of fish scales using a digitizing tablet. In R.C. Summerfelt, G. H., editor, *Age and Growth of Fish*, pages 397–410. ISU Press.

TNPC. TNPC software: digital processing of calcified structures. www.ifremer.fr/lasaa.

Traodec, H., Benzinou, A., Rodin, V., and Bihan, J. L. 2000. Use of deformable templates for otolith 2D growth ring detection by digital image processing. *Journal of Fisheries Research*, 46(1-3):155–163.

Welleman, H. and Storbeck, F. 1995. Automatic ageing of Plaice otoliths by means of image analysis. In Secor, D., Dean, J., and Campana, S., editors, *Recent developments in Fish Otolith Research*, pages 271–282. Univ. of South Carolina Press.

List of Figures

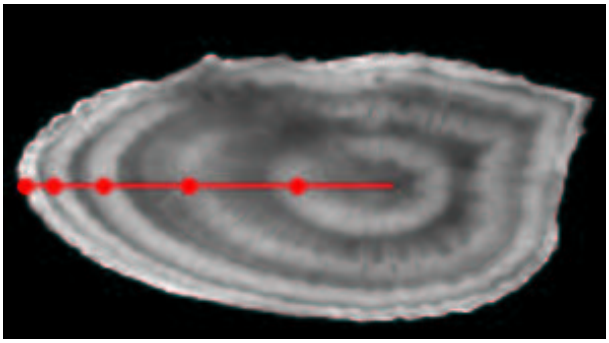

Figure 1. *Illustration of plaice otolith interpretation for a 4 year old individual. The otolith image has been acquired under a binocular with transmitted light. The interpretation of the translucent rings is displayed by the markers set on the radial drawn on the main reading axis from the center to the edge.*

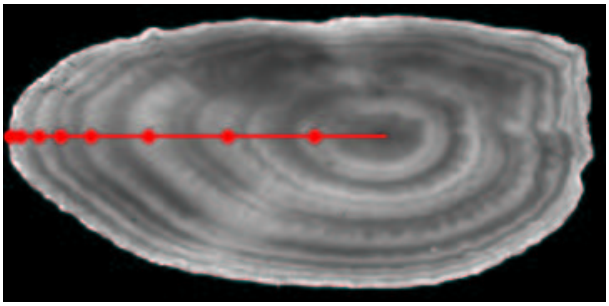

Figure 2. *Difficulties in otolith ring interpretation: this otolith presents false rings between the first and second rings, and between the second and the third ones. Besides, due to the non-linearity of the associated growth pattern, the last rings tend to be closer and hard to distinguish. The positions of validated rings is depicted by markers set on the radial drawn from the otolith center to the edge along the main reading axis.*

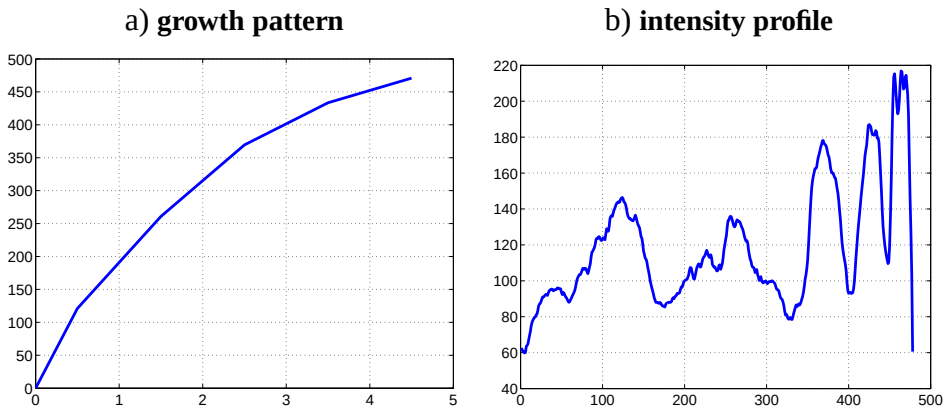

Figure 3. *Illustration of non-stationary features within otolith image for the example reported in Fig.1: a) growth pattern (distance from the center to the edge along the main reading axis as a function of the age), b) intensity profile extracted along the main reading axis from the center to the edge depicting a frequency modulation and a continuous trend.*

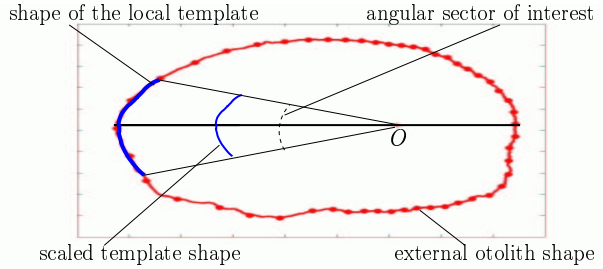

Figure 4. *Template-based approach aimed at extracting a 1D representation of the image content within a region of interest.*

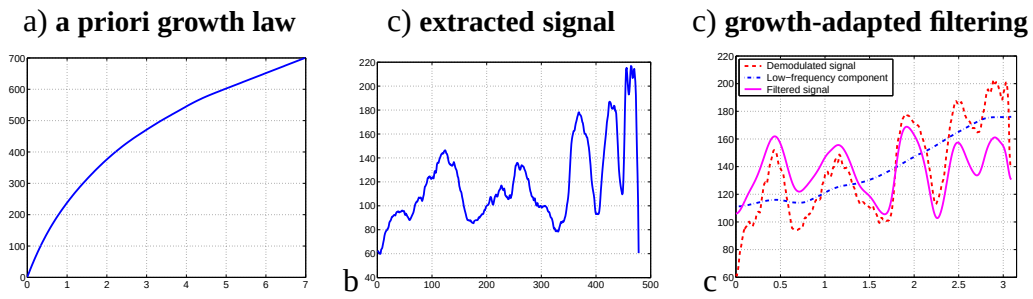

Figure 5. *Illustration of the growth-adapted filtering of the signal $s_{\mathcal{S}}$ for the otolith image reported in Fig.1: a) mean growth model, b) signal $s_{\mathcal{S}}$, c) the demodulated signal $s_{\mathcal{S}}^{DM}$ (–), its continuous component $s_{\mathcal{S}}^{DM}$ (-.) and the filtered signal $s_{\mathcal{S}}^{G}$ (-). For visualization issues, $s_{\mathcal{S}}^{G}$ has been shifted from 0 to the mean value of $s_{\mathcal{S}}^{DM}$.*

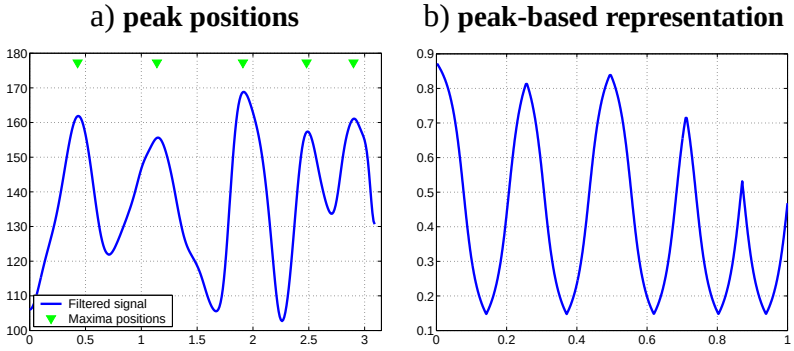

Figure 6. *Illustration of the peak-based representation for the otolith image depicted in 1: a) filtered signal* $s_{\mathcal{S}}^{G}$ *and the extracted maxima positions (^ markers), b) associated peak-based representation* $s_{\mathcal{S}}^{PB}$.

a) **age composition** b) **growth variability**

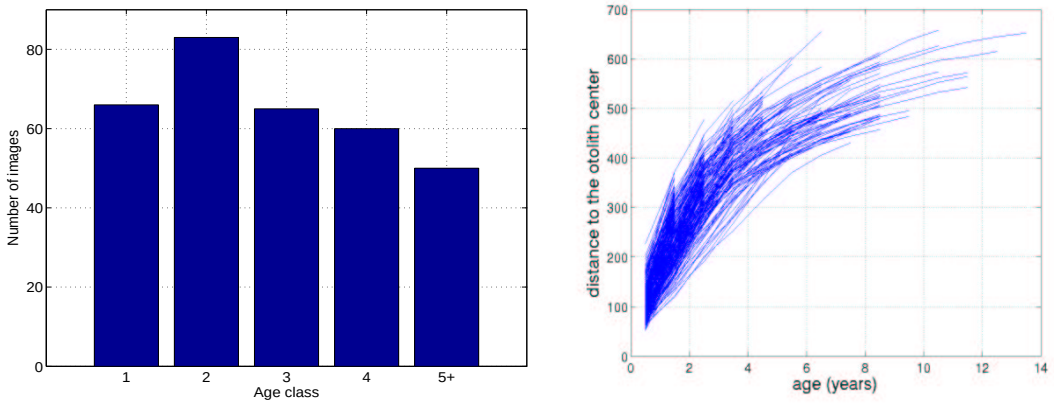

Figure 7. *Plaice otolith dataset: a) age distribution of the database of 320 plaice otoliths, b) illustration of growth variability within the considered dataset. We display the growth patterns (distance from the otolith center to the center as a function of the age) for the 320 plaice otoliths.*

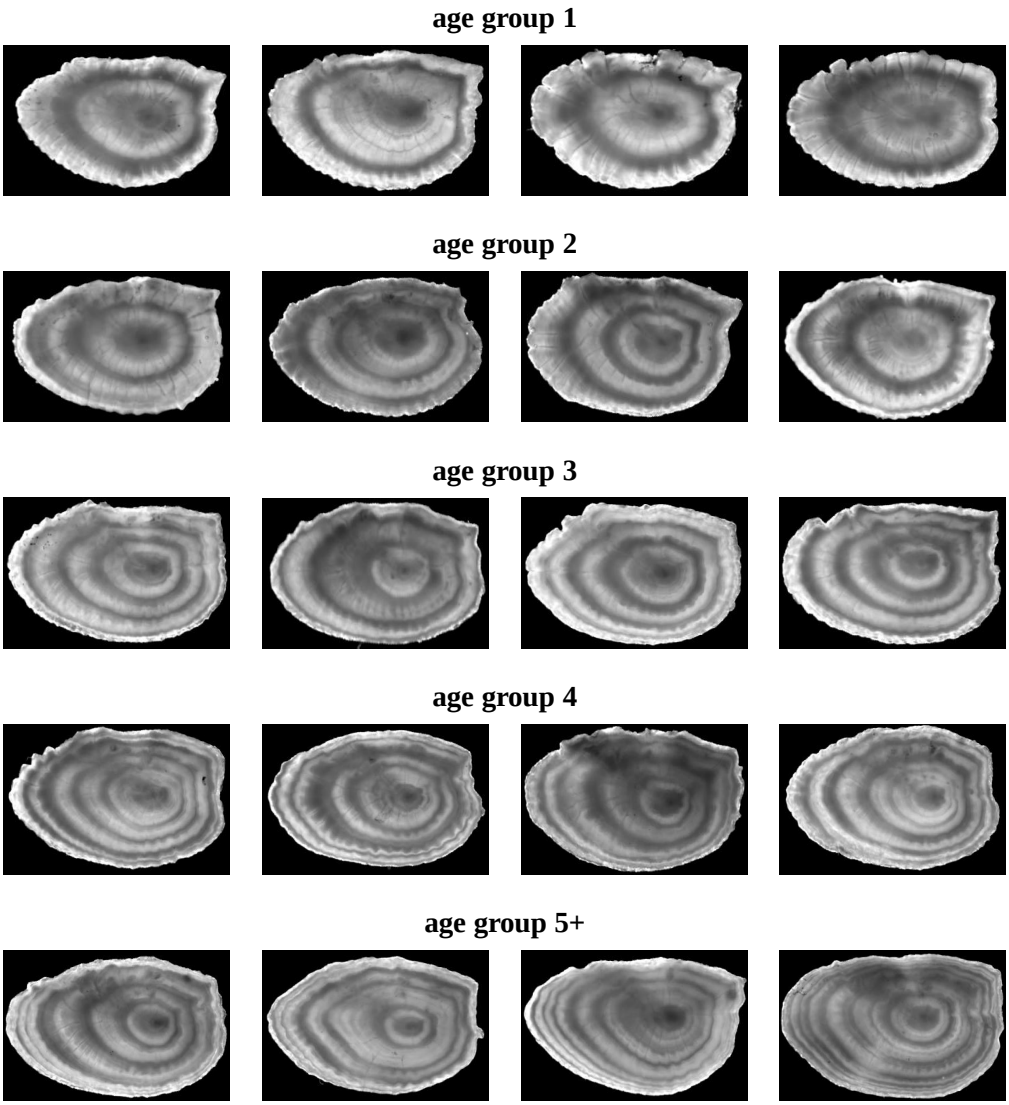

Figure 8. Sample images within the processed dataset of plaice otolith images.

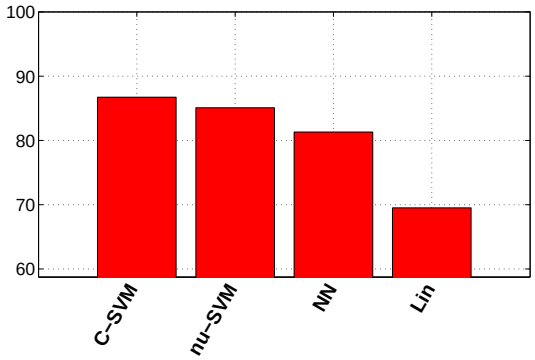

Figure 9. *Comparison of different classifiers using the peak-based representation and the remodulation scheme: C-SVMs, nu-SVMs, neural nets (NN) and linear SVMs (Lin).*

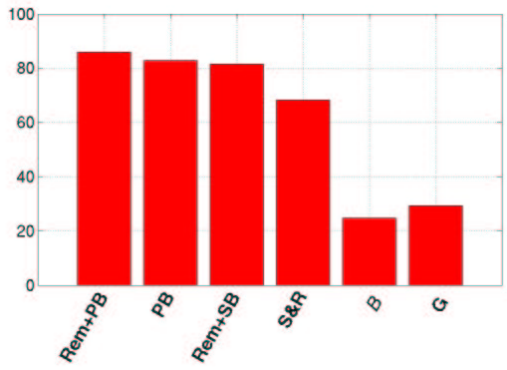

Figure 10. *Comparison of different feature types using C-SVMs: peak-based features (PB), PB features combined to the remodulation scheme within the training set (Rem+PB), Frequency-based features combined to the remodulation scheme (Rem+FB), real and imaginary parts of the Fourier coefficients computed for the template-based signal* s_S *as in (Robertson and Morison, 1998) (S&R), biological features (B) and geometrical features (G).*

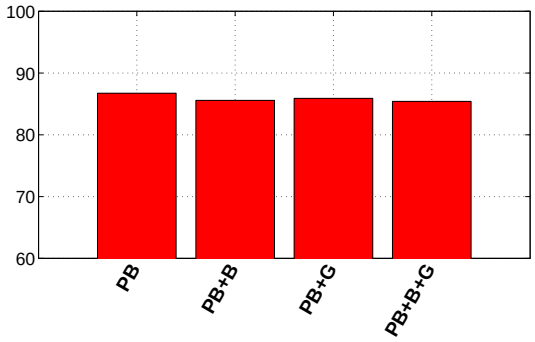

Figure 11. *Combination of peak-based features (PB) to biological (B) and Geometrical (G) features using C-SVMs and the remodulation step.*

List of Tables

	1	2	3	4	5+	overall rate of errors in [-1,+1]
1	**88.3**	**10**	**0.8**		**0.8**	**98.3**
	88.3	*10*	*0.8*		*0.8*	*98.3*
2		**92.3**	**4.7**		**2.9**	**97**
		90	*7*	*2.9*		*97*
3		**4.1**	**90**	**5.3**		**100**
		4.1	*92.5*		*3.3*	*96.6*
4		**4.3**	**14.7**	**75.6**	**5.2**	**95.6**
		5.2	*14.7*	*73.9*	*6*	*94.7*
5+	**1.1**		**7**	**8.2**	**83.5**	**91.7**
		1.1	*7*	*10.5*	*81.1*	*91.7*

Table 1
Confusion matrix of the results of fish aging over the test set using C-SVMs and the remodulation step for two feature types: peak-based features alone (bold) and the combination of peak-based and geometrical features (italic): for instance, 88.3% of the samples from class 1 are assigned to class 1, 10% to class 2, 0.8% to class 3 and 0.8% to class 5+.

	1	2	3	4	5+	overall rate of errors in [-1,+1]
1	91.6	8.3				100
2		91.1	5.8		2.9	97
3		4.1	91.6	4.1		100
4		4.3	13	78.2	4.3	95.5
5+			5.8	5.8	88.2	94.1

Table 2
Confusion matrix of the results of fish aging performed for each image of the test set by combining the classifications of five non-overlapping angular sectors along the main reading axis. One should read the table as follows.) table