A method for the possible species discrimination of juvenile gadoids by broad-bandwidth backscattering spectra vs. angle of incidence

Bo Lundgren and J. Rasmus Nielsen


Measurements were made of the broad-bandwidth (80–220 kHz) acoustic backscattering from free-swimming juvenile gadoids at various orientations and positions in an acoustic beam, under controlled conditions. The experimental apparatus consisted of a stereo-video camera system, a broad-bandwidth echosounder and echo-processor system, a narrowband 120 kHz split-beam echosounder, a large tank, and a fishnet cage. The net cage was centred on the acoustic beams and was virtually transparent, both acoustically and optically. Accurate three-dimensional positions and angular orientations of individual fish were estimated from stereo-images captured synchronously when broad-bandwidth echoes were received from passing fish. Fish positions were also estimated from data collected with a synchronized split-beam echosounder. Software was developed for image analysis and modelling, including calibration, alignment of acoustic and optical-reference frames, and automatic position-fitting of fish models to manually marked fix-points on fish images. The software also performs Fourier spectrum analysis and pulse-shape analysis of broad-bandwidth echoes. Therefore, several measurement series on free-swimming juvenile gadoids were evaluated. The method and data may be used to improve the acoustic identification of fish species and sizes, and thereby improve investigations of spatial prey–predator relationships, and the accuracy and efficiency of acoustic surveys.

Keywords: broad bandwidth, gadoids, species discrimination, split-beam, stereo-image analysis, target strength, tracking.

Received 14 December 2006; accepted 7 January 2008; advance access publication 19 March 2008.

B. Lundgren: Danish Technical University, National Institute of Aquatic Resources, PO Box 101, North Sea Centre, DK-9850 Hirtshals, Denmark. J. R. Nielsen: Danish Technical University, National Institute of Aquatic Resources, Charlottenlund Castle, DK-2920 Charlottenlund, Denmark. Correspondence to J. R. Nielsen: tel: +45 33963381; fax: +45 33963333; rm@aqua.dtu.dk

Introduction

In fisheries research, acoustic and trawl surveys from research vessels are used to estimate the spatial distributions and biomasses of fish stocks (MacLennan and Simmonds, 2005; Kalikman and Yudanov, 2006). The method commonly applied uses data from echosounders operating at one or more discrete frequencies (e.g. 38 and 120 kHz), and estimates the species mix and their sizes from net catches (e.g. Nielsen et al., 2001; ICES, 2005a, b). The frequency dependence of acoustic backscatter from fish (see Horne, 2000; Korneliussen and Ona, 2002) could be better exploited to improve species-identification techniques. Echosounders operating over a broad frequency range (e.g. 30–200 kHz) and with a large sampling volume could provide more precise information about fish taxa or species and their size distributions in scientific investigations. Acoustic identification based on variation of the angular backscatter with frequency could, for example, ease the description of single fish prey and predator distribution patterns in the sea (Nilsson et al., 2003), but may also diminish the need for frequent trawling during routine surveys (Lundgren and Nielsen, 2002).

There are good reasons to believe that species-specific, spectral characteristics for direct recognition do exist (Demer et al., 1999). The fact that toothed whales use broadband backscatter when locating marine surroundings and prey (Teilman et al., 2002; Beedholm et al., 2006) supports the potential of the method. Multiple investigations regarding acoustic fish identification have been reported by Lebourges (1990a, b), Simmonds and Armstrong (1990), and Simmonds et al. (1996).

Previous investigations are of two types, either similar to ours, measuring the backscatter from a single fish, e.g. Lebourges (1990a), or directed towards measuring the backscatter from aggregations of fish, e.g. Simmonds and Armstrong (1990), Simmonds et al. (1996), Zakaria et al. (1996), and Rogers et al. (2004), as well as multifrequency measurements by Korneliussen and Ona (2002). Lebourges measured broad-bandwidth reflectivity of tethered trout (Salmo trutta fario) and sea perch (Morone labrax). The shape of the fish body was monitored with a two-camera system. Simmonds et al. (1996) made measurements of free-swimming fish aggregations, and used a single still camera to identify targets. Other researchers measured broad-bandwidth sound-scatter from in situ fish aggregations using modified ADCP transducers (Rogers et al., 2004). In these three examples, the methods used for acoustic-target identification include spectral comparisons, and discriminant function and neural-network analyses. Their measurements, though, give only limited information on the cause of the frequency dependence of the fish spectrum and at the angle at which the greatest differences are found.
Measurements on fish aggregations need to be made with a transducer with equal beam width for all relevant frequencies, and Simmonds and Copland (1986), Simmonds et al. (1996), and Zakharia et al. (1996) designed special transducers and transceivers to obtain this information, and found spectral features in their respective frequency ranges (27–54 and 20–80 kHz) that indicated that it would be possible to distinguish between some of the species included in the investigations with a certain probability.

Our approach involved measuring ex situ broad-bandwidth acoustic backscattering from free-swimming individual juvenile gadoids at various orientations and positions within acoustic beams. Measurements were made on various sizes of small cod (Gadus morhua), whiting (Merlangius merlangus), and saithe (Pollachius virens). Characteristics of the reflected echosounder pulses are explored for possible methods of species recognition, and we suggest a methodology for discrimination, describe the experimental design, the data-collection system, and the calibration methods, as well as data analysis and modelling methodology.

Methods

The experimental set-up is an extension of the system described in Nielsen and Lundgren (1999). It consists of a stereo pair of video cameras, a broad-bandwidth (80–220 kHz) echosounder, a 120 kHz split-beam echosounder (Simrad EY500), and a large outdoor experimental tank (Figures 1 and 2). During measurements, a fish was kept centred in the acoustic beams using a net cage. Accurate three-dimensional positions and angular orientations of the fish were obtained by tracking the fish on stereo-video images, recorded when the broad-bandwidth echosounder received fish echoes (Lundgren et al., 2001). The trigger signals were emitted when the echo level from a passing fish was above a selectable threshold (usually 25 mV) and corresponding roughly to the echo level from a −60 dB target in the centre of the beam at ∼100 kHz within a selected distance interval (2.8–3.8 m).

General design and set-up

Tank, cage, and fish

A large outdoor tank situated at the North Sea Centre, DTU-Aqua, Hirtshals, Denmark, was used for the experiments (Figure 1). The tank, 20 m in diameter and 5 m deep, was filled with 2000 m³ of seawater taken directly from intakes on the Skagerrak seashore. A small laboratory with an observation window (0.5 m × 1.5 m) was placed on one side of the tank.

The fish cage (Figure 1) was constructed from light nylon mesh attached to a 2 m diameter frame constructed of light plastic tube. The frame was positioned outside the main lobes of the transducers at the measurement depth. The transparency of the net was advantageous when tracking fish optically (Lundgren et al., 2001). The fish were caught in the wild and acclimatized to the water temperature and cage in the measurement tank before being measured (see Data sampling below).

Instrumentation

The experimental set-up (Figures 1 and 2) consisted of three computers, the EY500, the broad-bandwidth echosounder, and a dual-camera video system. The customized broad-bandwidth sounder comprised an echo-processor (SignalData), power and signal amplifiers (Apex and Reson V1000 preamplifier, respectively), and two transducers (Reson TC2130). The sampling rates of the transmitter and receiver were both 526.3158 kHz. A pulse detector on the receiver was used to trigger the EY500 and camera systems via a signal generator (HP8111A), configured as a variable pulse-delay generator.

The broad-bandwidth transmitter generated constant amplitude pulses (~0.6 ms) with a linear-frequency sweep from 80 to 220 kHz. The pulses were amplified and sent to the transmitting transducer. The repetition rate was variable, but was set to ~3–4 pulses s⁻¹. The broad-bandwidth receiver digitized continuously, retaining the last 1.4 ms in a temporary buffer. When a sufficiently strong echo was received in the selected depth interval, the receiver was triggered and the buffered data were recorded. The stereo-video images were concurrently triggered. The trigger to the EY500 was delayed by 15 ms to avoid interference with the broad-bandwidth system (Figure 2). Each ping of the broad-bandwidth data contained 512 samples of the echo signal, the time and date of the trigger event, the time corresponding to the sample number corresponding to the trigger event.

The image-acquisition part of the system consisted of a computer with two frame-grabber boards (Data Translation DT3152), with resolutions of 656 × 472 pixels × 8 bit, and two cameras (Sony XC55 Progressive Scan CCD) with 659 × 494 pixels, and 256 grey levels (Figure 2). The 12-mm lenses had fixed apertures. A dynamic range of about 18 dB was achieved with automatic gain control (AGC).

The two video cameras were positioned at right angles to each other, 3 m from the centre of the acoustic beams, and 3.5 m below the transducers (Figure 1). Measurements were made in the far-field of the 120 kHz transducer (see Nielsen and Lundgren, 1999). The apparatus and software are shown in Figure 2 and detailed in Lundgren et al. (2001).

Calibration

Optical calibration under water

The theory and procedures used in camera calibration were based on the principles given by Tsai (1986, 1987), who basically assumed a pinhole-camera model and included second-order corrections for lens aberrations. The air–window–water interface of the camera assembly (camera and case; Figure 3) acted as an
extra lens, changing the apparent opening angle, aberration and focal width of the camera assembly.

During camera calibration, the fish cage was removed and the mount for the acoustic transducers was used to support a 1 × 1 m white polyethylene plate supported by aluminium backing. On the surface of this calibration plate were 100 needles with dark, spherical heads (diameter 2.5 mm) arranged in a 10 × 10 cm matrix grid. Images were recorded with the plate at ∼45° to both cameras, and then at ∼90° to each of the two cameras (Figure 4).

The pixel positions corresponding to the needle heads in a calibration image pair were evaluated both with standard image-particle detection software (GlobalLab) and manually with the fix-point marking routine described below, with similar results. The results from the 45° exposure run were processed together with information about the camera parameters (pixel-size and pixel-aspect ratio) by Tsai’s (1986, 1987) procedure for a co-planar calibration grid. The procedure is available in a C-language software package distributed on the internet (Willson, 1995). The results from the 90° runs were used to check the pixel-aspect ratio parameters. The calibration procedure estimated the apparent lens focal width, the lens distortion of the camera, and the three-dimensional position (distance and angles) of the calibration plate relative to the optical axis of the lens. From them, new constants \((k_x, k_y, \kappa)\) were deduced, which provided the relationships (1) and (2) between any three-dimensional point \((x_o, y_o, z_o)\) on the object, and the corresponding two-dimensional image point expressed in pixels \((x_i, y_i)\):

\[
x_i = k_x x_o / z_o + \kappa \tan^2(v_o) \tag{1}
\]

\[
y_i = k_y y_o / z_o + \kappa \tan^2(v_o) \tag{2}
\]

where \(\tan^2(v_o) = (x_o^2 + y_o^2) / z_o^2\).

For both the image and the object coordinates, the origin was the optical axis of the lens, and \(v_o\) was the angle between the optical axis and a particular ray through the apparent pinhole lens. The calibration software was further developed to introduce the camera-parameter and calibration-plate data into the Tsai-calibration routines, to update the three-dimensional to two-dimensional coordinate-transformation constants, and to display the calibration data. The software also contained a model simulating a line grid corresponding to the needle-head grid on the calibration plate. This line grid could be overlaid and aligned

---

**Figure 2.** Schematic connection diagram for the combined acoustic and video system.

**Figure 3.** Geometry of the apparent change of focal width of the camera through refraction in the window–water interface of the camera case.

**Figure 4.** Positions of cameras and calibration plate during the calibration. Fine dashed and solid lines, opening angles of cameras in air and water, respectively. Bold black line, calibration plate during the main calibration. Bold grey lines, calibration plate during supplementary runs.
with the needle-head spots on the calibration-plate images, to check the calibration constants.

**Calibration of acoustic systems**

Shortly before the experiments, the broad-bandwidth transducers were factory-calibrated (Reson) with calibrated, reference hydrophones (Figure 5). The transmit and receive sensitivities on the acoustic axis were measured from 80 to >240 kHz in 1-kHz bands. Variations of the sensitivities with angle were measured at 100, 150, and 200 kHz in two perpendicular directions, and were represented by four-degree polynomials fitted to the measurement points (Figure 5b). Finally, sensitivity variations at other frequencies were calculated by representing the frequency variations of the second- and fourth-order coefficient by second-degree polynomials. Additionally, both the split-beam and broad-bandwidth echosounders were calibrated in the tank using a standard copper sphere of 30.5 mm diameter in the following manner:

(i) The calibration sphere was placed in the acoustic beam. Broad-bandwidth, split-beam, and video data were recorded with the sphere placed in two different positions.

(ii) The broad-bandwidth echo-pulses were selected and Fourier transformed.

(iii) The initial transducer sensitivities were applied at each frequency.

(iv) A TVG-correction referring all echo levels to 3 m distance was applied.

(v) The sphere positions obtained from the split-beam echosounder and video-tracking were aligned and used to correct the amplitude for angular-sensitivity variation of the transducers at each frequency. Aligned acoustic positions were relative to the optical centre. Like the procedure for fish described above, the software fitted a sphere model to manually marked sphere images to obtain the spatial positions.

(vi) The measured amplitude spectra were compared with theory (Figure 5a). The theoretical spectra were calculated with software (D. MacLennan, pers. comm.), using equations described in MacLennan (1981, 1982) and in MacLennan and Simmonds (2005).

(vii) Calibration corrections were calculated at frequencies of local maxima (Figure 5a), and corrections at other frequencies were estimated from linear interpolations between values at the local maxima (Figure 5a). This technique minimized potential biases attributable to measurements in the amplitude valleys having lower signal-to-noise ratios (Figure 5a). The difference in corrections between the theoretically calculated and the measured target strength were 1.00, 3.30, 2.40, –1.30, and 3.50 dB for frequencies 94, 121, 145, 170, and 196 kHz, respectively.

**Figure 5.** (a) Data from the calibration of the broadband echosounder with a 30.5 mm copper sphere. Filled triangles and small filled squares: measured sphere spectra at two different positions corrected for factory calibration, position in the beam and distance from the transducer; Large filled squares: selected peak values used for the interpolation; Solid line: running mean over 2 kHz of the theoretically calculated TS-spectrum (according to D. N. MacLennan, pers. comm.); x-x-x: difference between the average of the two measurements and the theoretical values at the same frequency; +–+–+: interpolated difference; Open circles: average measured spectrum after application of the interpolated difference correction. The corrections have been limited to the frequency range 85 – 208 kHz, where the SNR is acceptable. (b) Transmit lobe pattern of the TC2130 broadband transducer. One-way transmit sensitivity in dB plotted against angle at three different frequencies. Squares, triangles, and circles represent the values measured, and the solid lines represent the fourth-order polynomials fitted to the measurements.
The EY500 was calibrated using a program that logs target-strength detections of the standard sphere and estimates beam parameters and calibrated-system gains (Lobe, Simrad).

**Data sampling**

Measurement sessions lasted from <1 h to several hours. Most measurement sessions began with recordings of the calibration sphere at a few positions in the acoustic beam. Then, the fish were transferred one at a time from the storage tank to the net cage, and kept there for periods of half-a-day up to a maximum of ~2 d. Before this, the top of the cage was treated with soap solution, and the empty cage put in place for a while to minimize bubble accumulation. Measurements of the fish began after a few hours to avoid stress behaviour. Sessions with more than one fish in the cage were attempted, but abandoned because the fish tended to school, making it impossible or difficult to separate single-fish echoes. Recording runs were then started and allowed to continue until between 300 and 3000 image pairs with accompanying acoustic data had been collected. Between runs, the cage was carefully raised so that its topside was just above the surface, then lowered again. This removed possible accumulated bubbles.

**Data processing and development of special processing and analysis software**

The three dataseries were post-synchronized to a precision of ± 0.1 s, utilizing the fact that pauses of variable length occurred at irregular intervals between regularly sampled data.

After synchronization, a number of groups of data corresponding to fish passes through the acoustic beams were identified and selected for further analysis within each recording run. These tracks had (i) overlapping broad-bandwidth, split-beam, and image-pair data, (ii) single-target detections from at least three consecutive pings, (iii) a maximum of one missing ping in the sequence, and (iv) a total acoustic split-beam angle detection of <3.6°. To display and visualize the data, a program was developed (ImageAnalyze; based on ImageGrab) to pair, store, and process images with the acoustic data (Figure 6). It allowed contrast enhancement of the images and inspection of the post-synchronization. Finally, to obtain the three-dimensional positions of a fish from the video images, it was necessary to identify and recognize some fix-points on the fish that were visible on both images. The program allowed manual marking, visualization, and storing of fix-points, which were later used for automatic estimation of the position and orientation of the fish.

**Creation of models for individual orientation**

For each dataseries, a line-grid model was created to act as a visual aid in the tracking of a fish. The grid points were used to construct a line drawing of the model that could be moved around on top of each of the images in a pair, using the coordinate-conversion parameters obtained during the optical calibration described above. A number of fix-points corresponding to eyes, mouth, and fin details were enhanced to aid the operator in placing the model accurately on top of a fish image. One point, in front of the dorsal fin, was defined as the position of the fish model. A magnification factor was included to make it easy to adapt the model to a new fish of similar shape but different size. The model outline could be modified by dragging the definition points to fit a new fish in a separate window (Figure 6). The model also included an option to bend the model sideways to approximate a sub-carangiform swimming mode (see Blake, 1983; Webb, 2002). Adapting the model to swimming movements clearly improved the fit between model and images. The transformation [Equations (3)–(8) below] of the centreline of the original fish model was used (Figure 7):

\[
x(0) = a_0 + \theta_1 \times \sin(q_0 + q_1),
\]

\[
dx(z_0) = a_0 \times a(z_0) \times \sin(p(z_0) + q(z_0)),
\]

\[
x(s) = S_c \times \sum_{z_0=0}^{t} dx(z_0) \times dz_0,
\]

\[
z(0) = 0,
\]

\[
dx(z_0) = \sqrt{1 - dx^2},
\]

\[
z(s) = S_c \times \sum_{z_0=0}^{t} dz(z_0) \times dz_0,
\]

where \(0 < s < 100, dz_0 = 1, p(z_0) = 2\pi z_0/100, \) and \(z_0\) is the coordinate along the centreline of the original model. Amplitude and phase in Equation (4) vary along the centreline according to Equations (9)–(11) below:

\[
a(z_0) = a_2, q(z_0) = q_0 \quad \text{for} \quad 0 < z_0 \leq z_1,
\]

\[
a(z_0) = a_2 + a_3 \times \exp(k_1(z_0 - z_1)), \quad q(z_0) = q_0 \quad \text{for} \quad z_1 < z_0 > z_2,
\]

\[
a(z_0) = a_2 + a_3 \times \exp(k_1(z_0 - z_1)) + a_4 \times \exp(k_2(z_0 - z_2)),
\]

\[
q(z_0) = q_1 + q_2 \quad \text{for} \quad z_2 < z_0 < 100.
\]

Scl is a scale factor corresponding to the actual size of the fish.

The constants \(f_1 = 1\) (number of body waves per body length), \(a_1 = -25\) (relative amplitude of snout movement), \(q_1 = 0.3\pi\) (start phase for snout movement), \(a_2 = 0.1\) (amplitude factor for basic body wave), \(z_1 = 40\) (start of exponential amplitude growth as a percentage of body length), \(a_2 = 0.1\) (amplitude factor for exponential growth), \(k_1 = 0.1\) (growth factor for exponential growth), \(z_2 = 90\) (start of the tailfin as a percentage of body length), \(a_3 = 0.1\) (amplitude factor for tailfin movement), \(k_1 = 0.1\) (growth factor for tailfin movement), and \(q_2 = 0.2\pi\) (phase difference for tailfin movement) were predefined constants in the program, and \(a_0\) (amplitude) and \(q_0\) (phase) were modified interactively from the main window or the model window. The constants were initially adjusted empirically to give a reasonable visual fit between model and images in most of the image material used. The model did not include options for vertical bending, which occurred relatively infrequently.

**Three-dimensional position, orientation, and bending of fish from video image pairs**

Whenever model images coincided with fish images in a pair on the screen, it was adjudged that position, orientation, and bending in the model were also the position, orientation, and
bending of the fish at the time of exposure. Position \((x, y, z)\) and angular orientation \((\text{pitch, tilt, yaw, and roll})\) in the reference-coordinate system defined by the stereo-image calibration, as well as the bending parameters \((\text{amplitude and phase})\) and magnification factor, appeared on the screen. The magnification factor is the size of the modelled fish relative to the basic model fish.

**Manual and automatic fitting of the fish model to fix-points**

The position, orientation, relative size \((\text{magnification})\), and bending of the model fish were initially adjusted interactively until the converted fish model seemed to overlap with the fish on both images correctly (manual fitting). The fitting process was aided by matching the positions of the fix-points on the fish (blue crosses) and the model (white dots).

When the fitting was satisfactory, the data describing the model fish shape \((\text{model points})\), the three-dimensional spatial position, orientation, relative size, and the bending parameters, together with the image-pair number and the image recording time, i.e. the trigger time obtained from the image file name, were recorded in an intermediate result file together with the target-position data \((\text{depth, athwartship, and alongship angles})\) from the split-beam echosounder.

To speed up the process of obtaining fish positions, an automatic, iterative fitting process with steps similar to the manual trial-and-error fitting process was designed to minimize the...
mean distance between the manually entered fix-points and the model fix-points on the images. Automatic fitting included the x, y, and z coordinates and the pitch and yaw angles, but not the roll angle or the bending parameters. However, the last two parameters could be interactively updated, whereas the automatic fitting of the other parameters was active.

As the cameras were approximately at right angles and equidistant from the fish, the iterative process was initialized by setting the model pitch-and-yaw angles equal to estimated fish pitch-and-yaw angles calculated by directly using the pixel distances between the available fix-point pairs on the fish images. On the first image, the model-fish position was initialized to (0, 0, 0), otherwise the position obtained in the last iteration process was used. The initial pixel positions of the model fix-points on the images were calculated using Equations (1) and (2) (see the Optical calibration section). Then, the model position was incremented a small amount beginning with the y-coordinate (up or down), and new positions for the model fix-points on the images were calculated. A first estimate of the actual vertical position of the fish was then calculated by assuming that the ratio between the two distances (fish position to initial fish-model position, and incremented fish model position to initial fish model position) in the camera coordinate system was the same as the ratio between the corresponding distances in the image-pixel coordinates, i.e.:

\[
(y_t - y_{\text{mo}})/(y_{\text{md}} - y_{\text{mo}}) = \text{AvDist}(fxp_{\text{md}}, fxp_{\text{imo}})/\text{AvDist}(fxp_{\text{imd}}, fxp_{\text{imo}}),
\]

or

\[
y_t = y_{\text{mo}} + (y_{\text{md}} - y_{\text{mo}}) \times \text{AvDist}(fxp_{\text{md}}, fxp_{\text{imo}})/\text{AvDist}(fxp_{\text{imd}}, fxp_{\text{imo}}),
\]

where \(y_t\) is the estimate of the \(y\)-coordinate of the fish, \(y_{\text{mo}}\) the \(y\)-coordinate for the original model position, and \(y_{\text{md}}\) is the incremented position. The function \(\text{AvDist}(fxp_1, fxp_2)\) calculated a weighted mean pixel distance between two image fix-point sets. The name \(fxp_{\text{imo}}\) represents the set of model fix-points in the images with the model in the initial position, \(fxp_{\text{imd}}\) with the model in the incremented position, and \(fxp_{\text{md}}\) the set of fix-points on the fish images. The magnitude and sign of the weights depended on which variable (\(y, x, z, \text{pitch}, \text{or yaw}\)) was incremented. The model position was changed to the newly estimated position and the iteration loop repeated until \(\text{AvDist}(fxp_{\text{md}}, fxp_{\text{imo}})\) was \(<1\) (pixel). A similar iteration loop was done for each of the other variables: \(x, z, \text{pitch}, \text{and yaw, in that order. The whole set of iteration loops were repeated until the }<1\text{-pixel condition was fulfilled for all variables simultaneously. Finally, for each image pair, the automatic fitting process was checked manually.}

Fish tracking and alignment of optical and acoustic tracks

The actual position of the fish in the horizontal \((x-z)\) plane as defined by the model position, and the split-beam data were plotted in the tracking window (Figure 6d). The position data (in the intermediate file) closest in time to the then displayed image pair were plotted as track lines. The program also included an option to move the origin and rotation of the acoustic-reference coordinate system interactively, to align the optically and acoustically measured tracks in the best possible way (Figure 6e). The alignment parameters were stored in the intermediate data file mentioned above, whenever it was updated.

Display of the echo-amplitude of the acoustic broad-bandwidth data

The interval of \(~1.4\) ms of broad-bandwidth pulse data recorded was displayed as a plot of the echo-amplitude against time along a vertical axis (Figure 6). From start to end of the fish-echo pulse, the pulse was highlighted by a different plot colour. The start was defined by the trig-sample number and a pulse-start adjustment parameter, and the fish-echo-pulse length was assumed to be equal to the transmit-pulse length (\(~0.6\) ms). The depth calculated from the delay time between start of the transmit pulse and start of the echo pulse (sound speed 1478 m s\(^{-1}\) at a salinity of 32 psu and a temperature of 8°C) was also displayed on the screen. The pulse-start adjustment parameter, which compensates for the delay between the actual pulse-start time and the trig time, was stored in the intermediate output file when next it was updated. The software allowed for interactive modification of both pulse-start and -length parameters.

Fourier spectrum analysis of the broad-bandwidth echo pulse

The actual fish pulse data as defined by the pulse-start and -length parameters (default 300 samples plus some margin) were extracted to a 1024-point buffer, and the rest of the buffer was zeroed. The first five and the last five samples were multiplied with a Gaussian-shaped function to diminish possible ringing caused by sharp pulse edges. The data in the buffer were then converted to a spectrum by a 1024-point Fourier transformation. The real and imaginary components of the spectrum were combined into a power spectrum.

Display of uncorrected and corrected Fourier spectra

The part of the spectrum corresponding to frequencies between 80 and 220 kHz was plotted as power in dB against frequency in kHz (Figure 6c). A curve representing the power spectrum compensated by the transducer-calibration factors was also displayed. Using the lobe-shaped pattern of the broad-bandwidth transducers, an amplitude correction was obtained by assuming that the
actual target position (angular distance from beam axis) in the broad-bandwidth beam was the same as obtained by the target tracking with the split-beam echosounder or with the optical tracking after alignment. Small errors in the target angle will not affect the correction for the broad-bandwidth hydrophones significantly, because the main lobe is much wider and flatter for all frequencies than the split-beam main lobe. The lobe pattern was derived by interpolation as described above. The curve of the compensated power spectrum was also displayed.

Output data
Each measurement series was stored in a hard-disk file containing a record for each image pair with its number and name (time), the species name, a track number, and the image-pair number within the track. For each image pair with an accepted model fit, i.e. where fish position was well-defined by fix-points directly or could be determined by interpolation between the two adjacent image pairs, the position, angle, and bending data of the fish as obtained by model fitting were also stored. If an acoustic position was available from the split-beam echosounder, it was stored together with ping number, time, and measured target strength (compensated and uncompensated). Accordingly, broad-bandwidth data were stored as record (ping) number, ping date and time, and pulse data. Finally, auxiliary data, such as the alignment parameters between the optical and acoustic positions, the adjustment parameter to align measured split-beam, the broadband depths, and a data-status byte, were also stored. The pulse data consisted of two parts. The first was 19 samples of the pulse amplitude as a function of time; each sampled the mean square amplitude of 16 samples of original data, i.e. 304 samples starting at the adjusted trigger point. The second was 19 samples of the Fourier spectrum of the pulse; each sampled the root mean square of 15 samples of the original power spectrum corresponding to frequencies of ~80–220 kHz. The data-status byte indicated which data types were available for each image pair.

Apparent acoustic tilt angle of a fish
The calibration plate used for optical calibration as described earlier was aligned with its vertical axis parallel to the axis of the acoustic beam. The acoustic tilt angle of a fish, defined as the angle between an acoustic ray from the centre of the transducer and the y-axis of the fish (see Figure 7), is equal to the tilt angle when the fish is positioned along the fish y-axis and the projection of the acoustic ray on the vertical (y-z) plane of the fish according to Equation (14) (see Discussion for motivation):

\[
\cos(90 - \theta_{tt}) = \cos(\theta_{yw}) \times \cos(\theta_{tt}) + \sin(\theta_{yaw}) - \sin(\theta_{tt}) \times \cos(\theta_{yw}) \times
\cos(\theta_{tt}) \times \sin(\theta_{vlong}) + \sin(\theta_{tt}) \times \cos(\theta_{vot})\nu_{out}
= \sqrt{v_{atlow}^2 + v_{along}^2},
\]

\[\text{(14)}\]

where \(\theta_{tt}\) is the apparent tilt angle, \(tt\) the tilt angle, \(yw\) the yaw angle, and \(v_{atlow}\) and \(v_{along}\) are the two angles defining the angular distance of the target from the acoustic beam.

Results
The results are based on five measurement series, as summarized in Table 1. Figure 8 shows the power-spectrum data of the series as the relative variations in dB plotted against the estimated acoustic tilt angle of the fish. This incidence angle depends on the position of the fish relative to the beam axis and the swimming direction (yaw) of the fish, as defined above.

Patterns of the backscattering vs. tilt angle vary with frequency for all species and size groups (Figure 8). These directivity patterns resemble some of the patterns obtained in the modelling work presented by Clay and Horne (1994) and Jech and Horne (2001). Those authors used the Kirchhoff–Ray mode (KRM) theory and measurements of anatomical structure obtained from X-ray photographs to construct models that explain theoretically the physiological and anatomical features in fish directivity patterns.

Here, a simple physical model using length alone was fitted to the data to identify species dependence in frequency-dependent directivities:

\[A_{model} = \log_{10}(\frac{(\sin(C_1f \sin(\theta_{tt})))}{(C_2f \sin(\theta_{tt}))})^2 + K_1(\sin(C_1f \sin(\theta_{tt})))/(C_2f \sin(\theta_{tt}))^2 + K_2) + K_3(f),\]

\[\text{(15)}\]

where \(\theta_{tt}\) is the incidence angle according to Equation (14), and \(f\) the frequency. The values of the parameters \(C_1\), \(K_1\), and \(K_2\) were selected manually to give the best possible fit visually to the log-amplitude data in Figure 8a–e, so giving one set of parameters for each species and size. In each of these cases, the parameter \(K_3(f)\) was determined for each of the 18 frequencies in such a way that the mean values of the log-amplitude data and \(A_{model}\) over a selected angle interval were equal. The constant \(C_1\) is proportional to the selected characteristic length in the model.

As both \(K_1\) and \(K_2\) are much smaller than 1, the square of the Sinc function appears to explain most of the backscattering directivity patterns (see Denbigh, 1998). Despite the quasi-independent parameters, similarly shaped curves resulted. Consequently, these data could not be used for unambiguous discrimination of species or size. Moreover, visual scrutiny of the plots of amplitude against tilt angle (Figure 8) also failed to identify unambiguous distinctive patterns. For cod, however, there seems to be a distinct pattern in the backscattering for angles of incidence from 20° to 40°. For cod of both size groups, there is only one main lobe (−10° to −5°) and smooth side lobes (20°–40°), compared with the other species. Whiting also have a strong main lobe, but have distinct side lobes, especially the first-order side lobes. There were too few observations to describe the main lobes for

<table>
<thead>
<tr>
<th>Fish species</th>
<th>Fish length (mm)</th>
<th>Fish weight (g)</th>
<th>Number of accepted pings/ image pairs</th>
<th>Tilt range(°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whiting</td>
<td>140</td>
<td>23</td>
<td>387</td>
<td>−13–+34</td>
</tr>
<tr>
<td>Whiting</td>
<td>149</td>
<td>27</td>
<td>93</td>
<td>−8–+13(+35)</td>
</tr>
<tr>
<td>Saithe</td>
<td>199</td>
<td>80</td>
<td>675</td>
<td>−9–+32</td>
</tr>
<tr>
<td>Cod</td>
<td>125</td>
<td>14</td>
<td>300</td>
<td>−(10)–3–+36(+45)</td>
</tr>
<tr>
<td>Cod</td>
<td>160</td>
<td>37</td>
<td>146</td>
<td>−17–+11</td>
</tr>
</tbody>
</table>

The number of unaccepted echoes using the described selection criteria was several thousands, especially for saithe. The tilt indicates the range of high-density data, and the numbers in parenthesis indicate range extensions with lower density data.

Table 1. Summary of data obtained from the experiments.
A method for the possible discrimination of juvenile gadoids

Figure 8. Backscatter power-spectrum data as a function of pitch and frequency from the five measurement series on single, free-swimming, juvenile gadoids. (a) Small whiting, (b) larger whiting, (c) saithe, (d) small cod, and (e) larger cod. The figure shows the collection of instantaneous spectra of the echopulses received from the fish. Each spectrum consists of 17 points vertically at the corresponding incidence angle, and each point represents the energy averaged over a frequency interval of 7.8 kHz. The equidistant centre frequencies range from 83.3 to 208.3 kHz. For clarity, the plots of the different frequencies have been separated by 10 dB, with the lowest frequency at the bottom. Incidence angle is the pitch angle corrected for fish position and yaw (swimming direction). The solid lines represent the functions fitted to the data points (see text).
saithe. However, distinct patterns do appear in the side lobes of saithe. In contrast to the other species, saithe have distinct second-order side lobes, with distinct nulls between the first- and second-order lobes.

Statistical discriminant function analyses (DFA) based on a principal component analysis (PCA) was also looked at in terms of its potential to distinguish species- and size-specific patterns in the backscattering. The PCA and DFA analyses were carried out with the R-functions princomp (PCA) and lda (linear discrimination) in the R-2.3.1 package (Ithaka and Gentleman, 1996; Venables and Ripley, 1996). Variables included in the analyses were fish species and length/weight (individual fish), tested with respect to dependence of fish position and orientation, i.e. acoustic incidence, pitch and yaw, for acoustic broad-bandwidth amplitudes at different frequency intervals with midpoints of 83, 91, 99, 107, 115, 122, 130, 138, 146, 153, 161, 169, 177, 185, 192, 200, 208, and 216 kHz within the measured broad-bandwidth frequency range of 80–220 kHz.

The PCA analyses were performed for each fish to compare principal components between individuals. They revealed that when yaw is included, the first principal component was solely dependent on this variable for all fish. This means that the other variables were independent of yaw. Accordingly, yaw was removed as a describing variable. As acoustic incidence angle is a direct function of yaw and pitch, pitch was also removed from the analysis, so only the acoustic incidence angle was retained. Therefore, PCAs were used to explore relationships between acoustic incidence and acoustic reflection (amplitudes) for each fish. The second round of PCAs revealed that the first principal component for all individual fish included all descriptive variables, with approximately equal weight. The first principal component was dominant for all fish, and the second included in general half of all the variables, though no common patterns could be found.

The quadratic DFA demands underlying normally distributed amplitudes (Venables and Ripley, 1996), but a Shapiro–Wilks test showed the distribution to be significantly different from normal. Accordingly, a series of linear DFAs was tested to identify the best model to describe the variability in data. The general linear DFA model used with different modifications was of the form

$$ldl = lda(species \sim length \times \text{AcuIncid} \times \text{Amp}_F83 \times \text{Amp}_F91 \times \text{Amp}_F99 \times \text{Amp}_F107 \times \text{Amp}_F115 \times \text{Amp}_F122 \times \text{Amp}_F130 \times \text{Amp}_F138 \times \text{Amp}_F146 \times \text{Amp}_F153 \times \text{Amp}_F161 \times \text{Amp}_F169 \times \text{Amp}_F177 \times \text{Amp}_F185 \times \text{Amp}_F192 \times \text{Amp}_F200 \times \text{Amp}_F208; \text{(16)}$$

where ldl is the linear discrimination line, lda the linear discriminant analysis, AcuIncid the acoustic incidence, and Amp the amplitude for a distinct frequency, $F$.

Modifications of the model (Models 1–5) are summarized in Table 2. Some DFAs included fish length (Models 4–5) and some did not (Models 1–3). All DFAs included acoustic incidence angle and all amplitude variables. All models were basically...
additive models, but included different first-order interactions between specific amplitude variables and acoustic incidence. The different interactions in the models are shown in Table 2. To validate the results obtained by the linear DFA, a cross-validation was performed comparing the different models. Here, half the observations were arbitrarily sampled and used to establish a prediction surface from linear discrimination. This surface was used to test how well one can predict in the other 50% of the data. This experiment is repeated a number of times through simulation, to establish the probability of success of correct classification. Figure 9 shows the results of this, along with the probability (or proportion) of correct classification with 95% confidence limits by the tested DFA model.

The results (Table 2, Figure 9) indicated that DFAs can discriminate between different species. The correlation ranged from 0.69 for the pure additive model without the length variable included (Model 1) to 0.92 for the model including length and four interaction effects between the amplitude variables and the acoustic incidence (Model 5). The apparent power of this analysis may be an artefact of discriminating a few species groups using many variables extracted from many observations of a few fish. Consequently, the DFA method must be assessed further and tested on more fish, and in situ. However, the cross-validation test indicates some robustness in the statistical results.

### Discussion

In this exercise, we studied the spectral characteristics of the backscattering from single free-swimming fish with the purpose of extracting possible useful features for direct recognition of species. The beam angles of the measurement transducers were frequency-dependent. To measure the frequency-dependence of the backscattering directivity pattern, it was necessary to monitor continuously the positions of the fish in the beam. This was done both acoustically and optically. The results strongly indicate narrow main lobes, suggesting that the pitch or tilt angle of the fish is the dominant determinant of variations in target strength. This is even true for relatively small, single fish. The frequency response is relatively flat for small pitch angles, but variations increase with increasing pitch angle. These findings agree with the modelling results of Clay and Horne (1994) and Jech and Horne (2001).

The formulae above are based on the assumption that for a relatively limited target-angle range of $\pm 3.5^\circ$, it is sufficient only to consider the acoustic tilt angle and to ignore roll. The shape of the scattering pattern from fish (swimbladder) is relatively circularly symmetrical around the roll axis, as shown clearly by the modelling results of Jech and Horne (2001) and Towler et al. (2003).

The cage appeared to work well during the measurements. There were some distinct, but relatively weak echoes, somewhat variable in amplitude, corresponding to the top and bottom of the cage, but no noticeable interference in the range of distances at which the fish echoes were accepted.

The inherent advantages of the standard-target acoustic calibration method are that it easily gives corrections of the initial calibration constants, takes both the electrical and acoustic properties of the signal path from transmitter to receiver, and is usable for ship-mounted transducers. Moreover, the whole frequency range is measured simultaneously, and only a few sphere positions are needed for the daily routine calibrations. However, owing to the valleys in the sphere spectrum (Figure 5), the disadvantage of the method is that the signal-to-noise ratio was low at some frequencies. It was, therefore, necessary with interpolation to obtain corrections for all frequencies. Acoustic calibration should be improved by using several calibration spheres with different diameters to get more evenly distributed high-level calibration data, even though this would increase the required calibration time.

The acoustic-calibration procedure included the assumption that one can interpolate the correction constants linearly between the values at the measured frequencies (Figure 5). The difference lines shown in Figure 5 indicate that this assumption is reasonable, and that uncertainty is only in the relatively low-amplitude values.

### Table 2. Results from statistical linear DFA performed in R.

<table>
<thead>
<tr>
<th>Model</th>
<th>Length variable</th>
<th>Interaction effects</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Not included</td>
<td>None</td>
<td>0.69</td>
</tr>
<tr>
<td>Model 2</td>
<td>Not included</td>
<td>Acoustic Incidence<em>Amp83</em>Amp91<em>Amp99</em>Amp107</td>
<td>0.82</td>
</tr>
<tr>
<td>Model 3</td>
<td>Not included</td>
<td>Acoustic Incidence<em>Amp208; Amp177</em>Amp185<em>Amp192</em>Amp200</td>
<td>0.75</td>
</tr>
<tr>
<td>Model 4</td>
<td>Included</td>
<td>None</td>
<td>0.86</td>
</tr>
<tr>
<td>Model 5</td>
<td>Included</td>
<td>Length<em>Acoustic Incidence</em>Amp83<em>Amp91</em>Amp99</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Ampxx, mean square amplitude in dB averaged over 7.8 kHz; xx, the frequency at the midpoint in kHz (see also the plots in Figure 8).
Two methods of obtaining the position of the fish based on the fix-point positions in the three-dimensional images were considered at the start of the project. In the direct method, the parameters for lines in the object space corresponding to light rays passing through the fix-points in each image and the aperture centre point of the corresponding camera have to be calculated. In theory, the intersection points of these lines in the object space will give the positions of the object fix-points. However, because the image fix-points were marked on a small target with low contrast, there were errors in the estimates that were sometimes large and clearly varied from point to point in the same fish image. This would have made it very difficult to make a proper estimate of the actual position of the fish, particularly in the direction in and out of the picture. Therefore, in the indirect method used here, a set of points outlining a fish line-grid model with fix-points was defined in the object space, and the positions of the corresponding sets of model points in the images were then calculated. The position and angular attitude of the model were then changed by manual or automatic iteration until a suitable fit of the fish and model images had been obtained. This has several advantages over the direct method. As contrast in the images was generally low, the small errors in placing the image fix-points would have caused the calculated ray lines not to intersect, making it difficult to calculate probable positions of the object fix-points. With the indirect method, the relative positions of the calculated object fix-points were always the same. Therefore, an averaging effect with regard to the position of the centre point of the model could be obtained. Also image-position data for fix-points, which were visible in only one of the images in a pair, could be used, and points or contours other than those defined as fix-points could be used visually to aid the position fitting. Finally, obviously erroneous automatic fits could relatively easily be spotted and corrected manually.

More data should be obtained from individual and aggregated fish of more species to obtain the parameters required for accurate species recognition. Future laboratory experiments should also include a broader frequency range (e.g., 20–400 kHz). All data collections should be synchronized centrally and immediately, avoiding uncertainty in the synchronization process. Multiple calibration spheres with different diameters should be used to obtain measurements with high SNR levels across the full bandwidth. Low-light cameras with high pixel densities and colours should be used to improve image contrasts and optical tracking of fish. The dynamic range of the broad-bandwidth system should also be increased.

The corrections for incidence angles did not reduce the variations in echo-amplitude vs. tilt angle, as expected. Therefore, more work may be needed to model the acoustic tilt angle and the influence of swimming movements on the backscatter. The shapes of the backscattering-directivity pattern may be used to identify fish species and their sizes. This could be measured in situ from multiple observations as a fish swims through the broad-bandwidth beam, or more synoptically using a multibeam, broad-bandwidth transducer. Because broad-bandwidth systems have shorter detection ranges than most narrow-bandwidth echosounders, they may need to be deployed on towed bodies or on autonomous underwater vehicles to allow them to get closer to the fish.

To conclude, synchronized, broad-bandwidth, acoustic backscattering and accurate three-dimensional positions and angular orientations of individual free-swimming fish have been made under controlled conditions. This was done through specially developed experimental design and data-analysis software. The characteristics of the broad-bandwidth, backscattering-directivity patterns were investigated for possible acoustic fish-species recognition methodology. The results, while inconclusive, suggest that it is possible to discriminate acoustically between some free-swimming juvenile gadoids, at least under experimentally controlled conditions.

Acknowledgements

This paper has been produced under equal authorship. Assistance was given by the mechanical and electronic workshops at DTU-Aqua, North Sea Centre, Denmark. Our colleague Peter Faber designed the image-data collection software (Image Grab), our colleague Kasper Christensen advised on the type of statistical modelling, and Henning Nielsen of Ålborg University advised on procedures for camera calibration and the processing of the fish images. We thank the RESON Company, Denmark, for providing us with the initial acoustic calibration data of the RESON transducers, and Dave Demer, NOAA/NMFS/SWFS/Fisheries Resources Division, La Jolla, CA, USA, for valuable editorial comments on the draft manuscript.

References

Korneliussen, R. J., and Ona, E. 2002. An operational system for acoustic tagging of individual free-swimming juvenile gadoids, at least under experimentally controlled conditions. This was done through specially developed experimental design and data-analysis software. The characteristics of the broad-bandwidth, backscattering-directivity patterns were investigated for possible acoustic fish-species recognition methodology. The results, while inconclusive, suggest that it is possible to discriminate acoustically between some free-swimming juvenile gadoids, at least under experimentally controlled conditions.

Acknowledgements

This paper has been produced under equal authorship. Assistance was given by the mechanical and electronic workshops at DTU-Aqua, North Sea Centre, Denmark. Our colleague Peter Faber designed the image-data collection software (Image Grab), our colleague Kasper Christensen advised on the type of statistical modelling, and Henning Nielsen of Ålborg University advised on procedures for camera calibration and the processing of the fish images. We thank the RESON Company, Denmark, for providing us with the initial acoustic calibration data of the RESON transducers, and Dave Demer, NOAA/NMFS/SWFS/Fisheries Resources Division, La Jolla, CA, USA, for valuable editorial comments on the draft manuscript.

References

A method for the possible discrimination of juvenile gadoids