

# Prediction of the Weight of Alaskan Pollock Using Image Analysis

Murat O. Balaban, Melanie Chombeau, Dilşat Cırbın, and Bahar Gümüş

**Abstract:** Determining the size and quality attributes of fish by machine vision is gaining acceptance and increasing use in the seafood industry. Objectivity, speed, and record keeping are advantages in using this method. The objective of this work was to develop the mathematical correlations to predict the weight of whole Alaskan Pollock (*Theragra chalcogramma*) based on its view area from a camera. One hundred and sixty whole Pollock were obtained fresh, within 2 d after catch from a Kodiak, Alaska, processing plant. The fish were first weighed, then placed in a light box equipped with a Nikon D200 digital camera. A reference square of known surface area was placed by the fish. The obtained image was analyzed to calculate the view area of each fish. The following equations were used to fit the view area (X) compared with weight (Y) data: linear, power, and 2nd-order polynomial. The power fit ( $Y = A \cdot X^B$ ) gave the highest  $R^2$  for the fit (0.99). The effect of fins and tail on the accuracy of the weight prediction using view area were evaluated. Removing fins and tails did not improve prediction accuracy. Machine vision can accurately predict the weight of whole Pollock.

**Keywords:** Alaskan Pollock, image processing, view area, weight

**Practical Application:** The weight of Alaskan Pollock can be predicted automatically by taking the image of the fish and using it in one of the correlations developed in this study. The removal of the fins or the fins and the tail did not increase the prediction accuracy of the method. Therefore, intact fish images should be used.

## Introduction

In 2008, Pollock (*Theragra chalcogramma*) represented 28% of U.S. domestic landings by weight (1.04 million tons) as the largest single species captured in the United States (NMFS 2008). Since 2001, the average Pollock catch has been above 1.33 million tons (\$329 million), despite its slight decrease in 2008 (–5% 2007 to 2008). The majority of U.S. Pollock is caught in the Eastern Bering Sea, Alaska (Lanelli and others 2009), with some fishery in the Gulf of Alaska (Dorn and others 2009). It is sold and exported in different forms: whole fish, fillets, surimi, and other products.

Automation of processing operations including sorting brings standardization of quality and objectivity to the process. Increased demands for objectivity, consistency, and efficiency have necessitated the introduction of computer-based image processing techniques. Machine vision (MV) can be used in a number of these operations with a high level of repeatability and flexibility (Gunasekaran 1996). Image-processing techniques have been applied increasingly for food quality evaluation. In the food industry, some quality evaluation is still performed manually by trained inspectors. Computer vision employing image-processing techniques can quantitatively characterize complex size, shape, color,

and texture properties of foods (Du and Sun 2004). MV systems are being used increasingly in food and agricultural industries for quality assurance purposes.

The correlation of view area to weight using MV was tested and demonstrated for different foods such as produce (Eifert and others 2006) and seafood products, for example, shrimp (Balaban and others 1994), Alaska salmon (Balaban and others 2010), oysters (Damar and others 2006), Atlantic Salmon (*Salmo salar*) fillet shape and size (Misimi and others 2008a, 2008b), and the sorting of 3 species of fish (Zion and others 1999). Fish species have also been sorted according to shape, length, and orientation in a processing line (Strachan and others 1990; Strachan 1993). Prawns have been automatically graded by size and packaged in one layer with uniform orientation by combining computer vision and robotics (Kassler and others 1993). Herring roe has been graded using “intelligent machines” (Croft and others 1996). In addition, MV can also evaluate color, shape, species, and visual defects. The combination of these evaluation steps with size sorting would integrate many operations into one efficient step. It is possible that size-sorting Alaska Pollock using MV would make the process more objective and efficient. Using view area to predict weight overcomes many of the potential difficulties of using other criteria such as length, since view area is rotation-, curvature-, and side-independent. A simple count of the number of pixels in the object is sufficient. Development and evaluation of equations to predict weight using view area is a 1st step toward automation of this process. Balaban and others (2010) mentioned the possibility of improving the accuracy of weight predictions using view area, if the fins and/or the tail of the fish were excluded from the analysis, since they bring variability and uncertainty to the process. This type of analysis has not been performed before.

MS 20100558 Submitted 5/21/2010, Accepted 7/22/2010. Author Balaban is with Fishery Industrial Technology Center, Univ. of Alaska Fairbanks, School of Fisheries & Ocean Sciences, Kodiak, Alaska, U.S.A. Author Chombeau is with Agricultural Engineering School, ESITPA, 76 134 Mont Saint Aignan, France. Author Cırbın is with İzmir Inst. of Technology, Güllübağ Köyü-Urla, İzmir, Turkey. Author Gümüş is with School of Fisheries, Akdeniz Univ., Dumluşınar Bulvarı, Antalya 07058, Turkey. Direct inquiries to author Balaban (E-mail: mob@sfos.uaf.edu).

The objective of this study was to develop weight compared with view area correlations for Alaska Pollock as a 1st step toward automation of sorting by size, and to evaluate if exclusion of fins and tail has an effect on the accuracy of weight prediction.

## Material and Methods

### Fish samples

One hundred and sixty whole Alaska Pollock (*T. chalcogramma*) were used for this study. They were obtained fresh, within 2 d after catch, from a Kodiak processing plant in February 2010. They were selected randomly by the plant staff from the process line, in order to have a wide spectrum of different size fish. They were analyzed immediately, using the procedures described below.

### Weighing and imaging

A scale (Mettler Toledo PB 3002-S, Columbus, Ohio, U.S.A.) with a maximum capacity of 3100 g, and an accuracy of 0.1 g was used to weigh the fish. Each fish was placed on a tared plastic tray, weighed and then its picture was taken. A light box, designed and built by the researchers (Luzuriaga and others 1997), with the dimensions 124.5-cm high, 66-cm deep, and 86-cm wide was placed in the pilot plant (Figure 1). Fluorescent bulbs providing D65 illumination were used behind a translucent Polycast acrylic (No. 2447, Polycast Technology Corp., Stamford, Conn., U.S.A.) 6.35-mm thick with 51% light transmission. The door of the light box was closed during picture taking to eliminate interference from



Figure 1—Light box and digital camera used to capture images. The fish and the reference square are also shown.

ambient light. Each fish was placed in the light box, under a Nikon D200 SLR digital camera (Nikon Corp., Tokyo, Japan) with a Nikon AF-S 18-200 mm F/3.5-5.6G ED-IF AF DX VR lens. All adjustments of the camera were set to “manual” for repeatability of pictures (Table 1). The Nikon WT-3 wireless accessory was used to control the camera remotely from a laptop computer, and to transfer the images to the computer immediately after acquisition using the Nikon Camera Control Pro software (Nikon Corp.). The camera was mounted on a Bogen-Manfrotto Maxi Repro Stand Lite assembly (Bogen Imaging, Inc., Ramsey, N.J., U.S.A.) with the lens looking downward. This setup was placed inside the light box. A reference square of 5 × 5 cm<sup>2</sup> was cut from a dark matte carton. This square was used as a reference for size, to calibrate image properties during analysis, and was included in every picture by placing it inside the light box (Figure 2).

### Image analysis

The JPEG images of the side views of the fish from the Nikon camera were entered into the Corel PhotoPaint (Corel Corp., Ottawa, Ontario, Canada) to reduce the size of the images from 1936 × 1296 to 800 × 536 pixels. From previous experience, it was known that images of higher resolution do not improve accuracy, and significantly increase image-processing time. Then, the format of the pictures was changed from JPEG to bmp (bitmap), since image analysis software required bmp files for input. Since the background was gray, and the belly of the fish was mostly grayish, automatic segmentation based on color was not possible.

Table 1—Settings of Nikon D200 camera.

Setting	Specification
Device	Nikon D 200
Lens	VR-18-200 mm F 3.5-5.6 G
Sensitivity	ISO 200
Optimized images	Custom
High ISO NR	Off
Exposure mode	manual
Metering mode	Multipattern
Shutter speed and aperture	1.3 s and f/10
Exposure compensation (in camera)	0 EV
Focus mode	AF-S
Long exposure NR	Off
Exposure compensation (by capture NX)	0 EV
Sharpening	None
Tone compensation	Normal
Color mode	Mode I
Saturation	Normal
Hue adjustment	0
White balance	Direct sunlight



Figure 2—Alaskan Pollock picture taken in the light box, example of side view.

Therefore, the pictures were cleaned to eliminate the background, then to erase the fins, and finally to erase both the fins and the tails by hand (Figure 3). Fins and tail have much less weight per unit area compared to, for example, the middle of the fish. The area of the fish would increase disproportionately if the fins were erect, compared to when the fins were not erect. This would also apply to the tail, whether it is fanned out, or compressed. The LensEye software (Engineering & Cyber Solutions, Gainesville, Fla., U.S.A.) was used to analyze the images. The number of pixels of the fish, and that of the reference square were counted by the software, and the view area of the fish was calculated by converting the number of pixels to cm<sup>2</sup>. Since the surface area of the reference square was 25 cm<sup>2</sup>:

$$\text{Fish view area (cm}^2\text{)} = \text{Fish pixels/reference square pixels} \times 25 \quad (1)$$

### Regression analysis

The following correlations were tried, based on previous experience with other fish: linear, power, and 2nd-order polynomial:

$$\text{Linear : } Y = A + BX \quad (2)$$

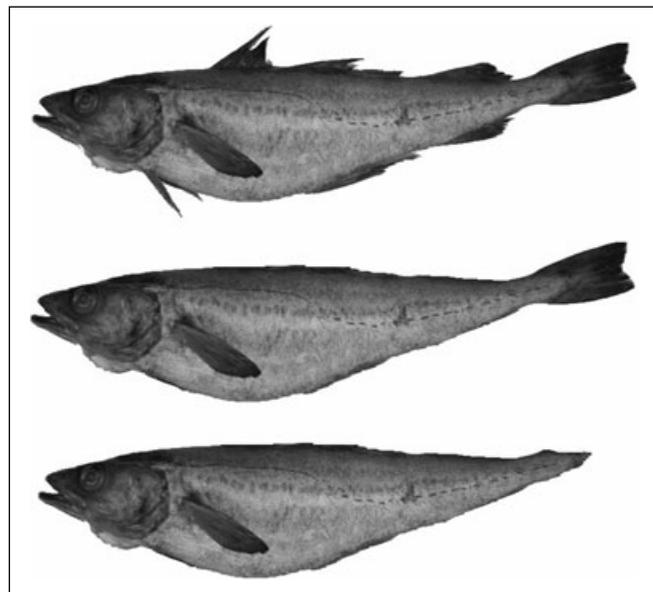


Figure 3—Example of intact Pollock, Pollock without fins, and Pollock without fins and tail.

$$\text{Power : } Y = AX^B \quad (3)$$

$$\text{Polynomial : } Y = C_0 + C_1X + C_2X^2 \quad (4)$$

In the equations above,  $Y$  = weight (kg),  $X$  = view area (cm<sup>2</sup>),  $A$ ,  $B$ ,  $C_i$  are coefficients. These equations were chosen, since they have shown good results during the analysis of other seafood (Balaban and others 1994; Zion and others 1999; Balaban and others 2010). The  $R^2$  for the fit was calculated for each fit.

### Results and Discussion

Table 2 summarizes the results of equation fits to the weight compared with view area data for whole, no fins, and no fins no tail images. In general, the power (Eq. 3) fit had the highest  $R^2$  value compared with the linear (Eq. 2) and the polynomial (Eq. 4) fits. This is expected, since the weight (which is related to volume, assuming the density of the fish is constant) and the view area (a surface) are bound with a dimensional relation such as (Balaban and others 2010):

$$\text{Volume dimension} = (\text{Area dimension})^{3/2} \quad (5)$$

### Whole fish

For whole fish, it was observed that the power (Eq. 3) fit had the highest  $R^2$  value (0.993). However, when all equation fits were compared, it was seen that all  $R^2$  values were very close, 0.987 for the linear (Eq. 2) and 0.985 for the 2nd-order polynomial (Eq. 4).

The plot of linear, power, and 2nd-order polynomial curves fits for intact (whole) fish are shown in Figure 4, with the experimental weight data displayed to compare with the predicted weights. The  $R^2$  values for power (0.993) and 2nd-order polynomial fits (0.985) were close, and their curves were almost superposed, at least for the lower weight region. The experimental weight, represented by markers on the plot, was distributed closely around the power and 2nd-order polynomial curves. Some fish deviated much from the predicted weight. An example is given in Figure 5 for fish (A) that has a difference of predicted weight from experimental data as -13%, and fish (B), with a 0.2% difference. It is apparent that fish (A) has a “fuller” belly, and all fins erect, while fish (B) is “slimmer” and with fins leveled. This difference among the fish can be attributed both to the normal variation between the fish, and to the “state” of the fins and tail when the picture was taken. It should be noticed that for smaller fish the prediction of weight is better. As the fish get bigger, there is more scatter, possibly due to the females carrying roes at this time of the season.

Table 2—Equation fit results for the weight compared with view area.  $Y$  = weight (g),  $X$  = view area (cm<sup>2</sup>),  $A$ ,  $B$ ,  $C_i$  are coefficients

Equation	Intact		No fins		No fins, no tail		
	Value	CI	Value	CI	Value	CI	
$Y = A + BX$	A	-248.80	22.72	-273.08	23.98	-268.40	23.33
	B	3.58	0.08	3.91	0.09	4.24	0.10
	$R^2$	0.987		0.978		0.980	
$Y = A \cdot X^B$	Value	CI	Value	CI	Value	CI	
	A	0.18	1.11	0.16	1.12	0.18	1.12
	B	1.47	0.02	1.51	0.02	1.51	0.02
$Y = C_0 + C_1X + C_2X^2$	$R^2$	0.993		0.993		0.993	
	$C_0$	-94.08		-69.30		-81.19	
	$C_1$	2.16		1.92		2.25	
	$C_2$	$2.83 \times 10^{-3}$		$4.27 \times 10^{-3}$		$4.62 \times 10^{-3}$	
	$R^2$	0.985		0.987		0.986	

CI = 95% confidence interval.

**Impact of fish fins and tail on the predicted weight**

Since there was a possibility that fins and tail could cause deviations in the predicted weight, the fins were manually erased from the images, regression analyses were performed, and then the tail was also erased before the regressions were repeated (Figure 3).

For the power equation in Table 2,  $R^2$  values for the intact, no fins, and no fins no tail images are the same: 0.993. For the other 2 equations,  $R^2$  values were also close: the 2nd-order polynomial fit had  $R^2$  values of 0.985, 0.987, and 0.986 for intact, no fins, and no fins no tail forms, respectively.

The 95% confidence intervals for the A values for the linear fits for all 3 forms, shown in Table 2, overlapped. This suggests that there is no significant difference in the A value (intercept) for the linear fits. However, the 95% confidence intervals of the B values of the linear fit (slope) do not overlap. There is increasing slope values going from the intact form, to the no fins, and to no fins no tail forms. This is expected, since for the same weight of the fish (Y), the intact image has the largest view area (X), and requires the smallest slope (3.58), and the no fins no tails form has the smallest view area (X), and requires the largest slope (4.24).

Figure 6 shows the plot of power fits for intact, no fins and no tail fish images, since this equation best predicts the weight from view area. It was observed that curves followed similar trajectories, and that  $R^2$  values were the same (0.993) for all forms of images. Therefore, at least for the Pollock pictures used in this study, the removal of the fins and tail did not improve the weight prediction accuracy of the correlations used for the view area.

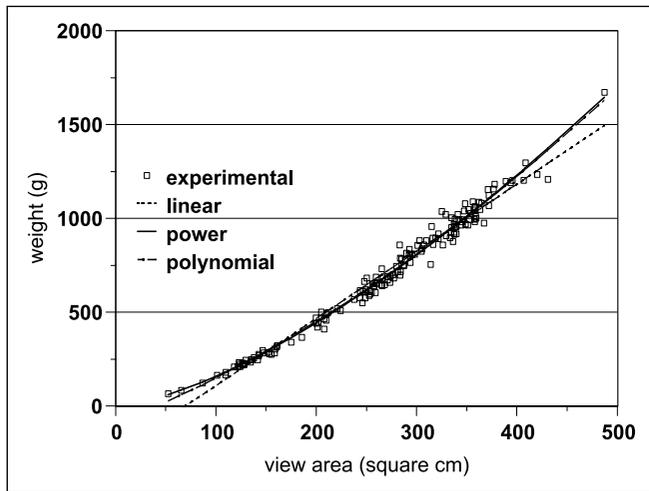


Figure 4—The weight prediction for intact fish (all equations).

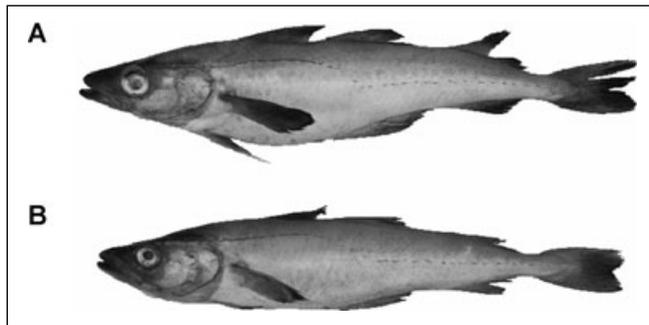


Figure 5—Example of two fish: (A) fish with -13% difference between predicted (power curve) and experimental weight, (B) fish with 0.2% difference between predicted (power curve) and experimental weight.

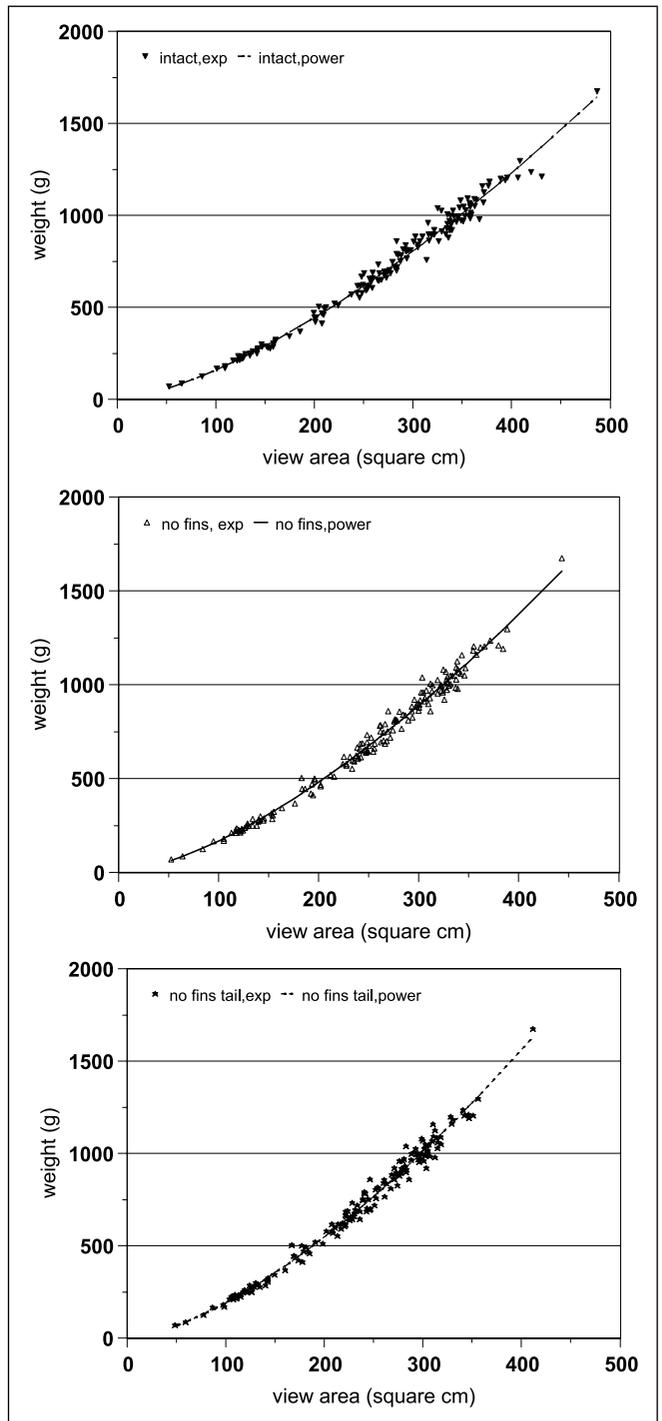


Figure 6—Comparison between intact (A), no fins (B), no fins no tail (C) Pollock pictures for the power fit ( $Y = A \cdot X^B$ ).

**Conclusion**

This study demonstrated that it is possible to predict the weight of whole Pollock from its view area, using analysis of its image from MV. The best equation to use was the power equation.

When fish images are analyzed without fins and without fins and tail, compared to whole fish images, there was no significant improvement in the  $R^2$  for the fits. Therefore, it would be possible to use all image types to predict the weight of fish, without loss

of accuracy. Since the removal of fins and tail would require extra work and computing time, the use of whole images is recommended.

The prediction of weight of Alaska Pollock with MV is possible. It would be desirable to investigate if the weight and/or view area correlations can be used to determine the sex of fish. Separation of the females carrying roe from the males would be desirable, since roe is a valuable product. This would increase the efficiency of the operations.

## Acknowledgments

The authors would like to thank Westwards Seafoods, Inc., especially Mr. Hiroyuki Kozuma, Vice President of Operations, Mr. Mitch Kilborn, General Manager, and Mr. Chris Sannito, QC Manager for allowing the use of their plant, and their fish for the study. Partial funding was provided by the Univ. of Alaska Fairbanks.

## References

- Balaban MO, Bergmann Y, Yearalan S, Otwell WS. 1994. Determination of count and uniformity ratio of shrimp by machine vision. *J Aquat Food Prod Technol* 3(3):43–58.
- Balaban MO, Ünal-Şengör GF, Gil-Soriano M, Guillén-Ruiz E. 2010. Using image analysis to predict the weight of Alaskan salmon of different species. *J Food Sci* 75(3):E157–162.
- Croft EA, de Silva CW, Kurnianto S. 1996. Sensor Technology integration in an intelligent machine for Herring Roe Grading. *IEEE/ASME TransMechatron* 1(3):204–15.
- Damar S, Yagiz Y, Balaban MO, Ural S, Oliveira ACM, Crapo CA. 2006. Prediction of oyster volume and weight using machine vision. *J Aquat Food Prod Technol* 15(4):5–17.
- Dorn M, Aydin K, Barbeaux S, Guttormsen M, Megrey B, Spalinger K, Wilkins M. 2009. Assessment of the Walleye Pollock Stock in the Gulf of Alaska. North Pacific Groundfish Stock Assessment and Fishery Evaluation Reports for 2010. 2009. 61–164 p. Available from: <http://www.afsc.noaa.gov/refm/docs/2009/GOApollock.pdf>. Accessed May 17, 2010.
- Du C-J, Sun D-W. 2004. Recent developments in the applications of image processing techniques for food quality evaluation. *Trends Food Sci Tech* 15:230–49.
- Eifert JD, Sanglay GC, Lee D-J, Sumner SS, Pierson MD. 2006. Prediction of raw produce surface area from weight measurement. *J Food Eng* 74(4):552–6.
- Gunasekaran S. 1996. Computer vision technology for food quality assurance. *Trends Food Sci Tech* 7(8):245–56.
- Kassler M, Corke P, Wong P. 1993. Automatic grading and packing of prawns. *Comput Electron Agric* 9:319–33.
- Lanelli JN, Barbeaux S, Honkalehto T, Kotwicki S, Aydin K, Williamson N. 2009. Assessment of the walleye pollock stock in the Eastern Bering Sea. North Pacific Groundfish Stock Assessment and Fishery Evaluation Reports, 2010. 2009. 49–148 p. Available from: <http://www.afsc.noaa.gov/refm/docs/2009/EBSpollock.pdf>. Accessed May 17, 2010.
- Luzuriaga D, Balaban MO, Yeralan S. 1997. Analysis of visual quality attributes of white shrimp by machine vision. *J Food Sci* 62(1):1–7.
- Misimi E, Erikson U, Skavhaug A. 2008a. Quality grading of Atlantic salmon (*Salmo salar*) by computer vision. *J Food Sci* 73(5):E211–7.
- Misimi E, Erikson U, Dirge H, Skavhaug A, Mathiassen JR. 2008b. Computer vision based evaluation of pre and postrigor changes in size and shape of Atlantic cod (*Gadus morhua*) and Atlantic salmon (*Salmo salar*) fillets during rigor mortis and ice storage: effects of perimortem handling stress. *J Food Sci* 73(2):E57–68.
- National Marine Fisheries Service (NMFS). 2008. Fisheries of the United States 2008. Silver Spring, Md.: U.S. Dept. of Commerce.
- Strachan NJC, Nesvabda P, Allen AR. 1990. Fish species recognition by shape analysis of images. *Pattern Recog* 25:539–44.
- Strachan NJC. 1993. Length measurements of fish by computer vision. *Comput Electron Agric* 8:93–104.
- Zion B, Shklyar A, Karplus I. 1999. Sorting fish by computer vision. *Comput Electron Agric* 23:175–87.