2010 Fourth Pacific-Rim Symposium on Image and Video Technology

Automated Detection of Fish Bones in Salmon Fillets using X-ray Testing

Domingo Mery, Iván Lillo, Hans Loebel, Vladimir Riffo, Alvaro Soto, Aldo Cipriano, José Miguel Aguilera
School of Engineering
Pontificia Universidad Católica de Chile (PUC)
Santiago de Chile – Chile
Email: dmery@ing.puc.cl

***Abstract*—X-ray testing is playing an increasingly important role in food quality assurance. In the production of fish fillets, however, fish bone detection is performed by human operators using their sense of touch and vision which can lead to misclassification. In countries where fish is often consumed, fish bones are some of the most frequently ingested foreign bodies encountered in foods. Effective detection of fish bones in the quality control process would help avoid this problem. For this reason, we developed an X-ray machine vision approach to automatically detect fish bones in fish fillets. This paper describes our approach and the corresponding validation experiments with salmon fillets. The approach consists of six steps: 1) A digital X-ray image is taken of the fish fillet being tested. 2) The X-ray image is filtered and enhanced to facilitate the detection of fish bones. 3) Potential fish bones in the image are segmented using band pass filtering, thresholding and morphological techniques. 4) Intensity features of the enhanced X-ray image are extracted from small detection windows that are defined in those regions where potential fish bones were segmented. 5) A classifier is used to discriminate between 'bones' and 'no-bones' classes in the detection windows. 6) Finally, fish bones in the X-ray image are isolated using morphological operations applied on the corresponding segments classified as 'bones'. In the experiments we used a high resolution flat panel detector with the capacity to capture up to a 6 million pixel digital X-ray image. In the training phase, we analyzed 20 representative salmon fillets, 7700 detection windows (10×10 pixels) and 279 intensity features. Cross validation yielded a detection performance of 95% using a support vector machine classifier with only 24 selected features. We believe that the proposed approach opens new possibilities in the field of automated visual inspection of salmon and other similar fish.**

***Keywords*-X-ray imaging; automated visual inspection; quality control; fish inspection.**

I. Introduction

In order to ensure food safety, several applications using X-ray testing have been developed for the industry [1]. The inherent difficulties in detecting defects, foreign objects and contaminants in food products have limited the use of X-ray to the packaged foods sector [2]. However, the necessity for NDT has motivated a considerable research effort in this field spanning many decades [3]. Important advances include: bone detection in poultry production [4], identification of insect infestation in citrus [5], detection of codling moth larvae in apples [3], fruit quality inspection like split-pits, water content distribution and internal structure [6], and detection of the granary weevil's larval stages in wheat kernels [7].

In the automated detection of fish bones, often called pin bones, there are few published papers: Andersen mentioned the basic components of a processing line to remove fish bones using a *pin bone detection unit* [8], however, there is no documentation for how the detection works. Han and Shi developed an approach with 85% effectiveness in fish bone detection based on *particle swarm clustering* in regions of interest obtained by thresholding and morphological operations [9]. Lorenc *et al.* elaborated on a general algorithm based on statistical features that can be used to detect fish bones, in which small vertical fragments are segmented in gray value images containing a varying undefined background, however, validation experiments are not reported [10]. Thielemann *et al.* presented an interesting method based on texture analysis of the surface image (not an X-ray image) to predict the positions where fish bones could be present in the fillet [11]. In the literature review, we observed the lack of an approach that could automatically detect fish bones effectively. That is most likely due to the fact that X-ray images of fish fillets with fish bones are very similar to those images where the texture of the surrounding fish fillet is present [8]. As we can see, X-ray testing is playing an increasingly important role in food quality assurance. However, fish bone detection in fish fillets is mainly performed by human operators using touch and vision senses which can certainly lead to misclassification (Fig. 1). In countries where fish is often consumed, fish bones are one of the most frequently ingested foreign bodies

Figure 1. Typical manual operation in a processing line using pliers to remove fish bones in salmon fillets.

978-0-7695-4285-0/10 $26.00 © 2010 IEEE
DOI 10.1109/PSIVT.2010.15

IEEE computer society

encountered in foods [12], so effective fish bone detection in quality control would assist in avoiding this problem. For this reason, we developed an X-ray machine vision approach to detect fish bones in fish fillets automatically with high performance. This paper describes our approach and the corresponding validation experiments with salmon fillets.

The rest of the paper is organized as follows: In Section II, the proposed X-ray machine vision approach is explained. In Section III, the results obtained in several experiments on salmon fillets are shown. Finally, in Section IV some concluding remarks are given.

II. Detection of Fish Bones

The key idea of our work is to use a machine vision methodology (Fig. 2), to automatically detect fish bones in fish fillets. The steps involved in this methodology are [13]:
1. Image acquisition: An X-ray digital image of the fish fillet being tested is taken and stored in the computer.
2. Pre-processing: The digital image is improved in order to enhance the details.
3. Segmentation: Potential fish bones are found and isolated from the scene's background.
4. Feature extraction/selection: Significant features of the potential fish bones and their surroundings are quantified.
5. Classification: The extracted features are interpreted automatically using a priori knowledge of the fish bones in order to separate potential fish bones into detected fish bones or false alarms.
6. Post-processing: Using morphological approaches the fish bones are detected and isolated in the X-ray image.

A. Image Acquisition

The X-ray source generates X-ray photons which irradiate the fish fillet being inspected. The fish fillet absorbs energy according to the principle of *differential absorption* [14]. Thus, internal elements of the fillet (such as fish bones, regular structures of the muscles, discontinuities, or foreign objects) modify the expected radiation received by the X-ray detector [15]. We used a flat panel detector that forms the digital X-ray image in three steps [16]: *i)* A scintillator sheet absorbs the X-ray photons and converts them into light photons, *ii)* an amorphous silicon photodiode array converts the light photons into electronic charges, and *iii*) each photodiode is read by an analog-to-digital converter (ADC) forming each pixel of the digital X-ray image. In our experiments, as shown in Fig. 3, we used:

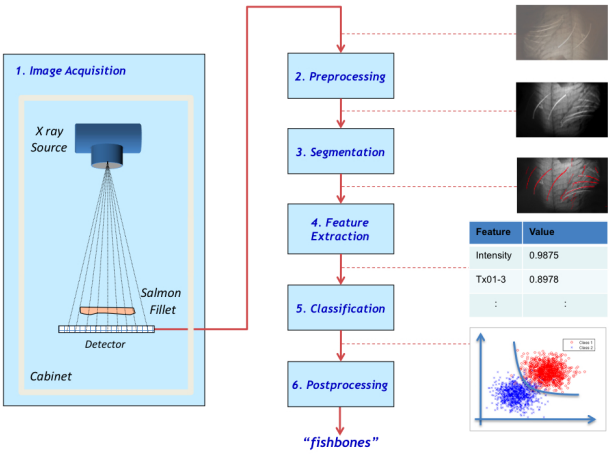


Figure 2. Machine vision schema used to automatically detect fish bones in fish fillets.

1. X-ray Source: A battery powered X-ray system Poskom XM-20BT (tube focal spot: 1.2mm, max. output: 100kV, 20mA).

2. Flat Panel Detector: A digital radiography system Cannon CXDI-50G (detection size: $35\times43\text{cm}^2$, image size: 2.208×2.688 pixels (5.93 million pixels), pixel size: 160 microns, grayscale: 4.096 (12-bits) gray value).

The X-ray source, the fish fillet exposed to X-rays, and the flat panel detector are enclosed in a lead cabinet that provides enough radiation attenuation and prevents access to the X-ray beam. The voltage and composite factor of the X-ray source were set to 40kV and 21mAs respectively by maximizing the contrast and minimizing the noise of more than twenty X-ray images of salmon fillets.

B. Pre-processing and Segmentation

The fish bones are only present in certain space frequencies of the spectrum: they are not too thin (minimal 0.5mm) nor too thick (maximal 2mm). The segmentation of potential fish bones is based on a band pass filter to enhance the fish bones with respect to their surroundings as shown in 4. Our approach to detect potential fish bones has four steps:

1. Enhancement: The original X-ray image $\mathbf{X}$ (Fig. (4b) is enhanced linearly by modifying the original histogram in order to increase contrast [13]: We obtain an enhanced image $\mathbf{Y} = a\mathbf{X} + b$.

3. Band pass filtering: The enhanced image $\mathbf{Y}$ is filtered using a radial symmetric 17×17 pixels mask $\mathbf{H}$ (Fig. 4a). Mask $\mathbf{H}$ was estimated from twenty X-ray images by minimizing the error rate proposed by Canny [17] applied to fish bones (all fish bones should be found and there should be no false alarms). The filtered image $\mathbf{Z} = \mathbf{Y} * \mathbf{H}$ is obtained (see Fig. 4c).

3. Thresholding: Those pixels in $\mathbf{Z}$ that have gray values greater than a certain threshold θ are marked in a binary image $\mathbf{B}$. The threshold is defined to ensure that all fish bones are detected, *i.e.*, false alarms are allowed in this step.

4. Removal of small objects: All connected pixels in $\mathbf{Z}$ containing fewer than A pixels are removed as shown in Fig. 4d.

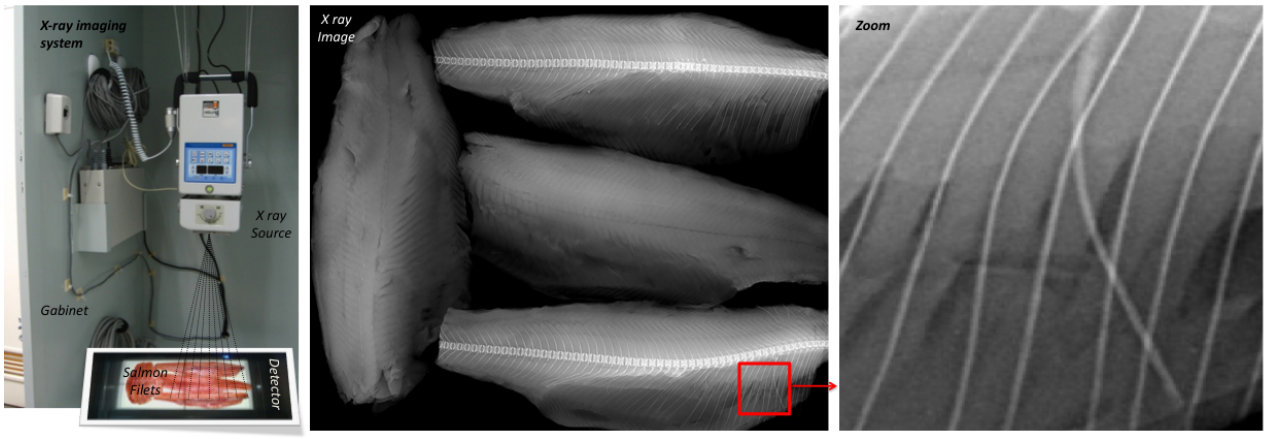


Figure 3. X-ray imaging system: fish fillets are irradiated by X-rays and the flat panel detector captures a high resolution X-ray image in which the fish bones are detectable.

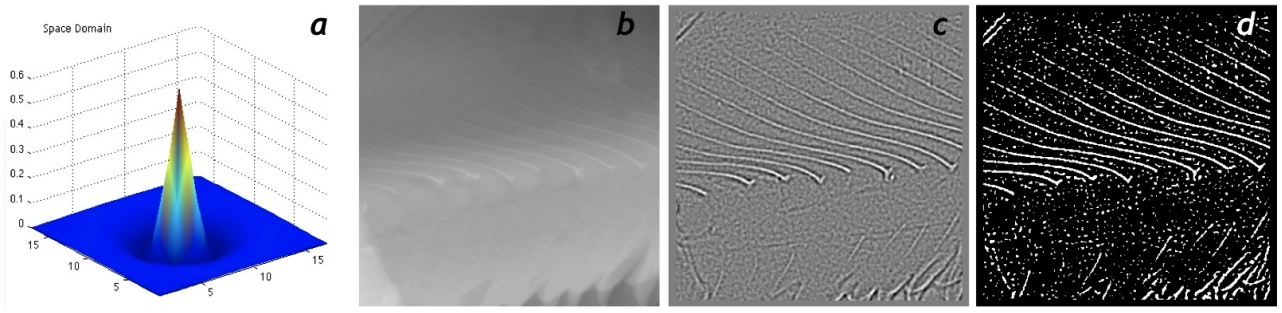


Figure 4. Segmentation of potential fish bones: a) Convolution mask $\mathbf{H}$ in space domain, b) original X-ray image $\mathbf{X}$ of a salmon fillet, c) filtered image $\mathbf{Z}$, d) potential fish bones after thresholding and removing objects deemed too small.

C. Feature extraction and selection

The segmented potential fish bones are divided into small 10×10 pixels windows called *detection windows*. In a training phase, using a priori knowledge of the fish bones, the detection windows are manually labeled as one of two classes: 'bones' and 'no-bones'. The first class corresponds to those regions where the potential fish bones are indeed fish bones. Alternatively, the second class corresponds to false alarms. Intensity features of the enhanced X-ray image $\mathbf{Y}$ are extracted for both classes. Features extracted from each area of an X-ray image region are divided into four groups as shown in Table I:

1. Standard: Simple intensity information related to the mean, standard deviation of the intensity in the region, mean first derivative in the boundary, and second derivative in the region [18].

2. Statistical textures: Texture information extracted from the distribution of the intensity values based on the Haralick approach [19]. They are computed utilizing *co-occurrence matrices* that represent second order texture information (the joint probability distribution of intensity pairs of neighboring pixels in the image), where mean and range of the following variables were measured: Angular Second Moment, Contrast, Correlation, Sum of squares, Inverse Difference Moment, Sum Average, Sum Entropy, Sum Variance, Entropy, Difference Variance, Difference Entropy, Information Measures of Correlation, and Maximal Correlation Coefficient.

3. Filter banks: Texture information extracted from image transformations like Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) [13], and Gabor features based on 2D Gabor functions, *i.e.*, Gaussian-shaped bandpass filters, with dyadic treatment of the radial spatial frequency range and multiple orientations, which represent an appropriate choice for tasks requiring simultaneous measurement in both space and frequency domains (usually 8 scale and 8 orientations) [20].

4. Local binary patterns: Texture information extracted from occurrence histogram of *local binary patterns* (LBP) computed from the relationship between each pixel intensity value with its eight neighbors. The features are the frequencies of each one of the histogram's 59 bins. LBP is very robust in terms of gray-scale and rotation variations [21].

Table I
EXTRACTED FEATURES

Group	Name and references
1. Standard	Mean Intensity, Standard deviation Intensity, Mean Laplacian, Mean Gradient, etc. [18].
2. Statistical textures	Tx(k, p) (mean/range) for k=1. Angular Second Moment 2. Contrast, Correlation, 4. Sum of squares, 5. Inverse Difference Moment, 6. Sum Average, 7. Sum Entropy, 8. Sum Variance, 9. Entropy, 10. Difference Variance, 11. Difference Entropy, 12.,13. Information Measures of Correlation, 14. Maximal Correlation Coefficient, and p=1,...5 pixels [19].
3. Filter Banks	DFT (1,2;1,2) and DCT (1,2;1,2) [13]. Gabor (1,...,8;1,...8), max(Gabor), min(Gabor), Gabor-J [20].
4. Local Binary Patterns	LBP (1,...,59) [21].

In our experiments, $n = 279$ features are extracted from each detection window. Afterwards, the features must be selected in order to decide on the relevant features for the two defined classes.

The n extracted features for sample i are arranged in an n-vector: $\mathbf{f}_i = [f_{i1}...f_{in}]$ that corresponds to a point in the n-dimensional measurement feature space. The features are normalized yielding a $N \times n$ matrix $\mathbf{W}$ which elements are defined as:

$$w_{ij} = \frac{f_{ij} - \mu_j}{\sigma_j} \tag{1}$$

for $i = 1, ..., N$ and $j = 1, ..., n$, where f_{ij} denotes the j-th feature of the i-th feature vector, N is the number of samples and μ_j and σ_j are the mean and standard deviation of the j-th feature. Thus, the normalized features have zero mean and a standard deviation equal to one. Those high correlated features can be eliminated because they do not provide relevant information about the food evaluation quality.

In feature selection, a subset of m features ($m < n$) that leads to the smallest classification error is selected. The selected m features are arranged in a new m-vector $\mathbf{s}_i = [s_{i1}...s_{im}]$. This can be understood as a matrix $\mathbf{S}$ with $N \times m$ elements obtained from m selected columns of the large set of normalized features $\mathbf{W}$.

The features can be selected using several state-of-art algorithms documented in literature like Forward Orthogonal Search [22], Least Square Estimation [23], Ranking by Class Separability Criteria [24] and Combination with Principal Components [25] among others. However, in our experiments the best performance was achieved using the well-known Sequential Forward Selection (SFS) algorithm [26]. This method selects the best single feature and then adds one feature at a time that, in combination with the selected features, maximizes classification performance. The iteration is halted once no considerable improvement in the performance is achieved by adding a new feature. By evaluating selection performance we ensure: *i)* a small intraclass variation and *ii)* a large interclass variation in the space of the selected features. For the first and second conditions the intraclass-covariance $\mathbf{C}_b$ and interclass-covariance $\mathbf{C}_w$ of the selected features $\mathbf{S}$ are used respectively. Selection performance can be evaluated using:

$$J(\mathbf{S}) = \text{trace}\left(\mathbf{C}_w^{-1}\mathbf{C}_b\right), \tag{2}$$

where 'trace' means the the sum of the diagonal elements. The larger the objective function J, the higher the selection performance.

D. Classification and Validation

A classifier decides whether the detection windows are 'bones' or 'no-bones'. We tested several classifiers, such as statistical or those based on neural networks [25], however, the best performance was achieved using support vector machines (SVM) [27]. SVM transforms a two-class feature space, where the classes overlap, into a new enlarged feature space where the classification boundary is linear. Thus, a simple linear classification can be designed in the transformed feature space in order to separate both classes . The original feature space is transformed using a function $h(\mathbf{s})$, however, for the classification only the kernel function $K(\mathbf{s}, \mathbf{s}') = \langle h(\mathbf{s}), h(\mathbf{s}')\rangle$ that computes inner products in the transformed space is required. In our case, the best classification was obtained using a *Gaussian Radial Basis* (RBF) function kernel defined by [28]:

$$K(\mathbf{s}, \mathbf{s}') = e^{-||\mathbf{s}-\mathbf{s}||^2} \tag{3}$$

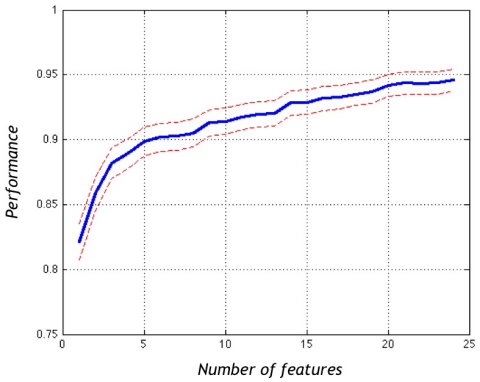
where the linear boundary, *i.e.*, the separating hyperplane in the transformed space, is computed using the Least–Squares approach [24].

The performance of the classifier was defined as the ratio of the detection windows that were correctly classified to the total number of detection windows. The performance was validated using cross-validation, a technique widely implemented in machine learning problems [29]. In cross-validation, the samples are divided into F folds randomly. $F - 1$ folds are used as training data and the remaining fold is used as testing data to evaluate the performance of the classifiers. We repeated this experiment F times rotating train and test data. The F individual performances from the folds are averaged to estimate the final performance of the classifiers.

III. EXPERIMENTAL RESULTS

We experimented with 20 representative salmon fillets obtained at a local fish market. The average size of these fillets was 15×10 cm^2. According to pre-processing and segmentation techniques explained in Section II-B we obtained several regions of interest where fish bones could be located. The area occupied by these regions of interest corresponds to aprox. 12% of the salmon fillets as shown in Fig. 4.

From the mentioned regions of interest we obtained 7697 detection windows of 10×10 pixels. Each window was



1.	LBP_{48}	2.	LBP_{11}	3.	$Gabor_{7,5}$	4.	$DFT_{2,2}$	5.	$Tx_{12,3}$(mean)	6.	LBP_{59}
7.	$Tx_{14,5}$(mean)	8.	$DFT_{2,1}$	9.	$Tx_{10,3}$(mean)	10.	$Tx_{2,5}$(range)	11.	$Tx_{1,5}$(range)	12.	$Tx_{12,1}$(mean)
13.	StdIntensity	14.	LBP_{20}	15.	$Tx_{2,5}$(mean)	16.	$DFT_{1,2}$	17.	$Tx_{8,5}$(range)	18.	$Gabor_{2,6}$
19.	$Tx_{2,1}$(range)	20.	$Gabor_{8,8}$	21.	$Tx_{12,2}$(mean)	22.	$Gabor_{3,5}$	23.	$Tx_{6,4}$(range)	24.	$Tx_{5,3}$(range)

Figure 5. Classification performance with a 95% confidence interval using the first m features (refer to Table I to see a description of the features).

labeled with '1' for class *bones* and '0' for *no-bones*. From each window 279 features were extracted according to Section II-C. After the feature extraction, 75% of the samples from each class were randomly chosen to perform the feature selection. The best performance was achieved using Sequential Forward Selection. The best 24 features are shown in Fig. 5 in ascending order.

The performance of the classification using SVM classifier and the first m selected features was validated using an average of ten cross-validation with $F = 10$ folds as explained in Section II-D. The results are shown in Fig 5. We observe that by using 24 features, the performance was almost 94.7% with a 95% confidence interval between 94.3 and 95.1%.

IV. Conclusion

The need for more information on the quality control of several fish types by means of quantitative methods can be satisfied using X-ray testing, a non-destructive technique that can be used to measure, objectively, intensity and geometric patterns in non-uniform surfaces. In addition the method can also determine other physical features such as image texture, morphological elements, and defects in order to automatically determine the quality of a fish fillet. The promising results outlined in our work show that we achieved a very high classification rate in the quality control of salmon when using a large number of features combined with efficient feature selection and classification. The key idea of the proposed method was to select, from a large universe of features, only those features that were relevant for the separation of the classes. We tested our method on 20 representative salmons yielding a performance of 95% in accuracy using 24 features and support vector machines. Although the method was validated with salmon fillets only, we believe that the proposed approach opens new possibilities not only in the field of automated visual inspection of salmons but also in other similar fish.

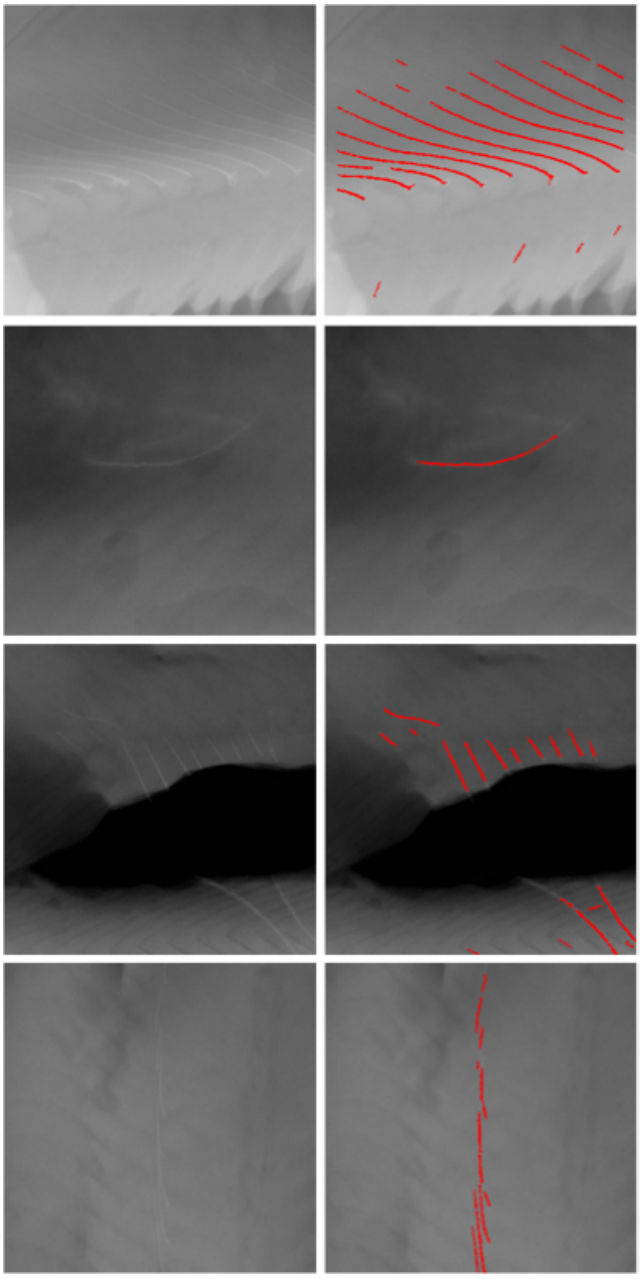


Figure 6. Results obtained in four X-ray images. The columns correspond to enhanced images, classified fish bones and post processed fish bones. The first row corresponds to the example shown in Fig. 4.

Acknowledgment

The authors would like to thank to Grant FONDEF No. D07I1080 - Chile.

REFERENCES

[1] E. Davies, *Image Processing for the Food Industry*. World Scientific, 2000.

[2] J. Kwon, J. Lee, and W. Kim, "Real-time detection of foreign objects using x-ray imaging for dry food manufacturing line," in *Proceedings of IEEE International Symposium on Consumer Electronics (ISCE 2008)*, April 2008, pp. 1–4.

[3] R. Haff and N. Toyofuku, "X-ray detection of defects and contaminants in the food industry," *Sensing and Instrumentation for Food Quality and Safety*, vol. 2, no. 4, pp. 262–273, 2008.

[4] M. Graves, "X-ray bone detection in further processed poultry production," *Machine Vision for the Inspection of Natural Products*, pp. 421–449, 2003.

[5] J. Jiang, H. Chang, K. Wu, C. Ouyang, M. Yang, E. Yang, T. Chen, and T. Lin, "An adaptive image segmentation algorithm for x-ray quarantine inspection of selected fruits," *Computers and Electronics in Agriculture*, vol. 60, pp. 190–200, 2008.

[6] Y. Ogawa, N. Kondo, and S. Shibusawa, "Inside quality evaluation of fruit by X-ray image," *International IEEE/ASME Conference on Advanced Intelligent Mechatronics (AIM2003)*, vol. 2, pp. 1360–1365, July 2003.

[7] R. Haff and D. Slaughter, "Real-time X-ray inspection of wheat for infestation by the granary weevil, sitophilus granarius (l.)," *Transactions of the American Society of Agricultural Engineers*, vol. 47, pp. 531–537, 2004.

[8] K. Andersen, "X-ray techniques for quality assessment," in *Quality of Fish from Catch to Consumers*, J. B. Luten, J. Oehlenschläeger, and G. Olafsdottir, Eds. Wageningen Academic Publ., 2003, pp. 283–286.

[9] Y. Han and P. Shi, "An Efficient Approach for Fish Bone Detection Based on Image Preprocessing and Particle Swarm Clustering," *Communications in Computer and Information Science*, vol. 21, pp. 940–948, 2007.

[10] A. Lorencs, I. Medniek, and J. Sinica-Sinavskis, "Fast object detection in digital grayscale images," *Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences*, vol. 63, no. 3, pp. 116–124, 2009.

[11] J. Thielemann, T. Kirkhus, T. Kavli, Schumann-Olsen, O. Haugland, and H. Westavik, "System for estimation of pin bone positions in pre-rigor salmon," *Lecture Notes in Computer Science*, vol. Advanced Concepts for Intelligent Vision Systems (ACIVS2007), no. 4678, pp. 888–896, 2007.

[12] Y. Akazawa, S. Watanabe, S. Nobukiyo, H. Iwatake, Y. Seki, T. Umehara, K. Tsutsumi, and I. Koizuka, "The management of possible fishbone ingestion," *Auris Nasus Larynx*, vol. 31, no. 4, pp. 413 – 416, 2004.

[13] R. Gonzalez and R. Woods, *Digital Image Processing*, 3rd ed. Pearson, Prentice Hall, 2008.

[14] H. Haken and H. Wolf, *The Physics of Atoms and Quanta: Introduction to Experiments and Theory*, 6th ed. Berlin, Heidelberg: Springer, 2000.

[15] R. Halmshaw, *Non-Destructive-Testing*, 2nd ed. London: Edward Arnold, 1991.

[16] M. Spahn, M. Strotzer, V. Markus, S. B. B. Geiger, G. Hahm, and S. Feuerbach, "Digital radiography with a large-area, amorphous-silicon, flat-panel X-Ray detector system," *Investigative Radiology*, vol. 35, no. 4, pp. 260–266, 2000.

[17] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

[18] M. Nixon and A. Aguado, *Feature Extraction and Image Processing*, 2nd ed. Academic Press, 2008.

[19] R. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[20] A. Kumar and G. Pang, "Defect detection in textured materials using gabor filters," *IEEE Trans. on Industry Applications*, vol. 38, no. 2, pp. 425–440, 2002.

[21] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, Jul 2002.

[22] H.-L. Wei and S. Billings, "Feature subset selection and ranking for data dimensionality reduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 162–166, Jan. 2007.

[23] K. Mao, "Identifying critical variables of principal components for unsupervised feature selection," *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 35, no. 2, pp. 339–344, April 2005.

[24] MathWorks, *Matlab Toolbox of Bioinformatics: User's Guide*. Mathworks Inc., 2007.

[25] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[26] A. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.

[27] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.

[28] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, August 2003.

[29] T. Mitchell, *Machine Learning*. Boston: McGraw-Hill, 1997.