Contour matching for a fish recognition and migration-monitoring system

Lee, Dah-Jye, Schoenberger, Robert, Shiozawa, Dennis, Xu, Xiaoqian, Zhan, Pengcheng

Dah-Jye Lee, Robert B. Schoenberger, Dennis Shiozawa, Xiaoqian Xu, Pengcheng Zhan, "Contour matching for a fish recognition and migration-monitoring system," Proc. SPIE 5606, Two- and Three-Dimensional Vision Systems for Inspection, Control, and Metrology II, (16 December 2004); doi: 10.1117/12.571789

SPIE.

Event: Optics East, 2004, Philadelphia, Pennsylvania, United States

Contour matching for a fish recognition and migration monitoring system

Dah-Jye Lee[a], Robert Schoenberger[c], Dennis Shiozawa[b], Xiaoqian Xu[a], and Pengcheng Zhan[a]
[a] Department of Electrical and Computer Engineering,
[b] Department of Integrative Biology,
Brigham Young University, 459 CB, Provo, Utah 84602
[c] Agris-Schoen Vision Systems, Inc., 3320 Mill Springs Drive, Fairfax, VA 22031

ABSTRACT

Fish migration is being monitored year round to provide valuable information for the study of behavioral responses of fish to environmental variations. However, currently all monitoring is done by human observers. An automatic fish recognition and migration monitoring system is more efficient and can provide more accurate data. Such a system includes automatic fish image acquisition, contour extraction, fish categorization, and data storage. Shape is a very important characteristic and shape analysis and shape matching are studied for fish recognition. Previous work focused on finding critical landmark points on fish shape using curvature function analysis. Fish recognition based on landmark points has shown satisfying results. However, the main difficulty of this approach is that landmark points sometimes cannot be located very accurately. Whole shape matching is used for fish recognition in this paper. Several shape descriptors, such as Fourier descriptors, polygon approximation and line segments, are tested. A power cepstrum technique has been developed in order to improve the categorization speed using contours represented in tangent space with normalized length. Design and integration including image acquisition, contour extraction and fish categorization are discussed in this paper. Fish categorization results based on shape analysis and shape matching are also included.

Keywords:
Shape Analysis, Fish Recognition, Polygon Representation, Tangent Space, Power Cepstrum, Fourier Descriptors

1. INTRODUCTION

The so called "Northwest Passage" or Columbia River basin, covers 259,000 square miles and includes the Snake, Deschutes, Okanogan, Wenatchee, Spokane, and Yakima rivers. It is also home to more than 400 dams used for hydroelectric power, irrigation, transportation, flood control, and recreation. Quantifying fish abundance and distribution is critical to viable fishery management and studies of the environmental impacts of these dams. Currently, fish are counted and monitored at 50 percent or more of the US Bureau of Reclamation (USBR) and US Army Corps of Engineers facilities. All monitoring is done manually with outdated equipment (video recorders) and is subject to human error and labor constraints. A large volume of video recordings must be reviewed manually in order to collect this critical information. Other research projects often require the capturing of fish to determine densities, size classes, movements, and species composition. A portable automated fish recognition system will greatly reduce data collection efforts and the impact on fish. According to biologists at the USBR, state governments, universities, and Native American Tribes, an automated fish recognition system is urgently needed. The USDA is supporting this effort.

Monitoring information includes estimates of species composition, densities, fish condition, fish size, timing of migrations, etc. From these, projections of the strengths of runs can be made, long term trends in populations can be compared, and even the relative success of various mitigation measures can be appraised. Condition factors can be used to determine the influence of abiotic factors on fish growth. The timing of migrations can be related to regional sea surface temperatures, and changes in relative species composition can be investigated relative to upstream habitat changes. Even mortalities as fish move progressively upstream can be estimated [1]. Without monitoring, management is not possible. However, monitoring requires labor, either on-site observers to identify and count the species passing through the station, or some mechanism for automated recording of the fish. The mechanisms currently in use are video

Two- and Three-Dimensional Vision Systems for Inspection, Control,
and Metrology II, edited by Kevin G. Harding, Proceedings of SPIE Vol. 5606
(SPIE, Bellingham, WA, 2004) 0277-786X/04/$15 · doi: 10.1117/12.571789

recordings which must still be manually reviewed to obtain the required monitoring information and in either case precision beyond species identification and a general size estimate is not possible.

Research work has been done in monitoring and measuring fish for various applications. Chan et al. [2] used a stereo imaging system to relate salmon morphology to mass. It allows salmon farmers to make informed decisions on feeding, grading and harvesting strategies. Fuzzy C-Mean (FCM) was used for fish recognition in a fishery. This approach requires priori knowledge on the analyzed data [3]. Naiberg and Little [4] built a stereo vision system to measure fish size to monitor the growth of the fish, to determine the feed and medication, and to decide when to harvest. An imaging system was built to analyze fish behavior and movement patterns and to use this information to monitor the presence of acute toxicants in water [5]. An image processing system was built by Gamage and de Silva [6] to measure fish orientation for optimal cutting for fish processing. The system was improved in 1993 using statistical pattern recognition for finding cutting position. The above work has been focused on monitoring size of same species in fish farm or controlling cutting machine for fish processing. However, none of the work has addressed the fish recognition and monitoring needs for biology and environmental research.

We have learned that one existing computer based system that has been utilized to monitor fish migration has major problems with its ability to distinguish between bubbles, debris, and fish. This problem was so significant that the field biologists doing the monitoring work no longer utilize that system. Instead, the existing standard, video tapes are taken and viewed manually. There is a commercial tape reviewing software for fish counting. The software reads images from video tapes and records the fish size. It does not have recognition capability. Also, due to poor image quality, the size measurement is not accurate. Another commercially available technology has been developed by fisheries biologists. The equipment consists of arrays of photo diodes and detectors assembled into a submersed site-specific dimension 'tunnel' that counts and determines size of fish migrating though the array. This technology has the potential to provide data with stand alone, automated equipment which can quantify fish passing a point in a canal, fish ladder, or dam. Similar to the tape reviewing software, it does not have fish classification capability. Size measurement accuracy is affected by the fish moving speed. Other research work has been done in analyzing fish spectral and spatial signature [7]. However, due to lighting and processing speed constraints, the result, although promising, has not been implemented for a fish monitoring system.

This paper discusses research work we have done to design and build an automated Fish Recognition and Monitoring **(FIRM)** system for biologists and researchers. An automated system, such as FIRM, utilizing high resolution digital images, and a computerized image recognition system could significantly reduce manual labor costs and at the same time automatically generate data on species and numbers of fish passing the monitoring station. Fig. 1 shows a typical diversion dam (Fig. 1(a)) and its fish ladder (Fig. 1 (b)) in Prosser, WA. Fish swim upstream through a narrow passage built in the fish ladder. On many western rivers, there can be many diversion dams built along each migration route. Monitoring fish migration at these diversion dams or at narrow passages along the rivers can provide extensive information about the numbers, seasonal timing, periodicity of movement, and passage survival of fish. The FIRM system can be installed at the dams which have a viewing window or at facilities where fish can be guided through a narrow passage so that images can be taken for analysis. The FIRM system can also be used for monitoring fish migration patterns near the inflows and outflows of lakes, as long as fish can be guided to pass through a narrow passage. In the cases without a viewing window, an underwater camera is utilized.


(a)

(b)

Figure 1. Prosser dam on the Yakima River in Prosser, Washington and its fish ladder.

In this paper, we present a fast and accurate shape matching method that compares a fish contour with contours stored in the database. We focused on developing an efficient method for representing the contour extracted from the fish and a fast and accurate method to make real-time shape matching. Previous work focused on finding critical

landmark points on fish shape using a curvature function analysis [8-10]. Fish recognition based on landmark points has shown satisfying results. However, the main difficulty of this approach is that landmark points sometimes cannot be located very accurately. Whole shape matching is used for fish recognition in this paper. Several shape descriptors, such as Fourier descriptors, polygon approximation and line segments, are tested for fish recognition. A power cepstrum technique has been developed in order to improve the speed of categorization and is used for the contours represented in tangent space (normalized length).

In Section 2 we discuss system requirements including a processing flowchart and camera, optics, and lighting selections. In Section 3 we discuss shape extraction methods and present a modified curve evolution method for shape representation. In Section 4 we discuss and compare a few shape matching methods and introduce a fast and accurate algorithm to effectively measure shape similarity for matching or recognition. In Section 5 we present the results of this new method and two other selected methods. We present our conclusions in Section 6.

2. SYSTEM

2.1 Processing Flowchart

A system flowchart shown in Fig. 2 illustrates the overall approach of the automatic fish recognition algorithm. It outlines the need for each of the five processing steps. The first and most important step is image acquisition. How to acquire a high quality image for analysis and lighting and camera selections have been studied. Object detection algorithm detects the presence of an object. Object contour extraction and identification help to identify if the object is indeed a fish. Tracking of the fish object determines the location of the fish and triggers the recognition process when an image of the whole fish can be acquired. Whole fish features including fish contour and significant landmark points can then be extracted for analysis. Shape-based recognition can be performed to obtain size and species information. In addition to fish species, time and fish image (if necessary) can be stored for further analysis by the users. Details of each of these steps and the results from our study are discussed in this paper.

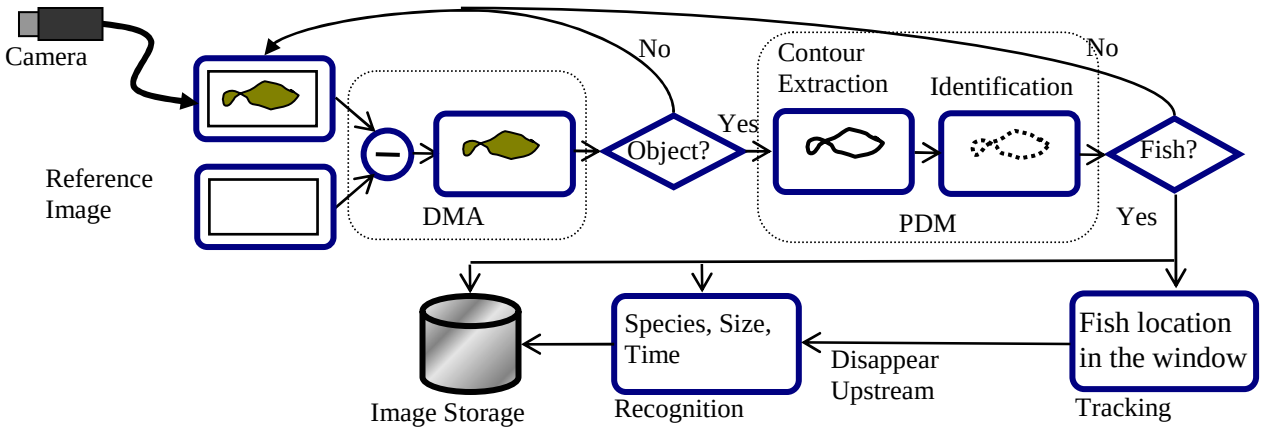


Figure 2. Processing algorithm flowchart

2.2 Camera Selection

A vision system test bed was built for software development, field testing, and data collection. A 10 gallon aquarium was used to simulate the viewing window for testing identification and tracking algorithms before field testing. Fig. 3 shows the vision system and the aquarium. Testing was done by capturing fish movement in the aquarium and comparing live images of fish movement with the stored reference image. The difference between the reference image (background with no fish) and live images indicates the possible objects of interest (fish or debris) in the field of view. Identification and tracking software was first developed in the lab using this setup and then the vision system was brought to the field for testing and data collection.

A Near-infrared (NIR) camera manufactured by Hitachi was used to acquire images at different light wavelengths, visible (< 700 nm) as well as invisible (700 nm ~ 1100 nm). This is not a Thermal Camera. This camera

(Hitachi KP-F2A) was chosen because of its unique features such as near-infrared sensitivity, high-speed electronic shutter for acquiring images of fast moving objects, progressive scanning for full frame resolution and avoiding motion induced interlacing artifacts, and flexible gain and offset adjustments. The extended NIR sensitivity makes it more sensitive to and more suitable for low light applications than comparable visible light cameras.

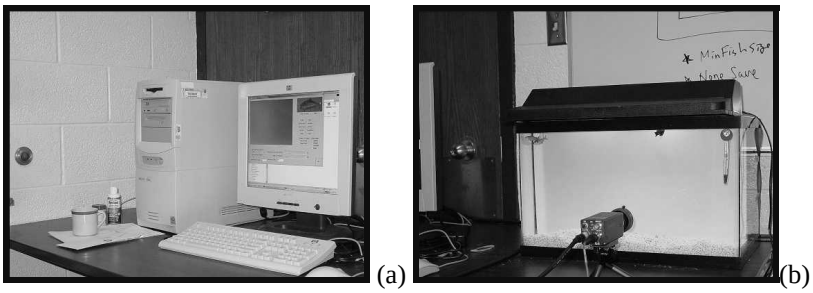
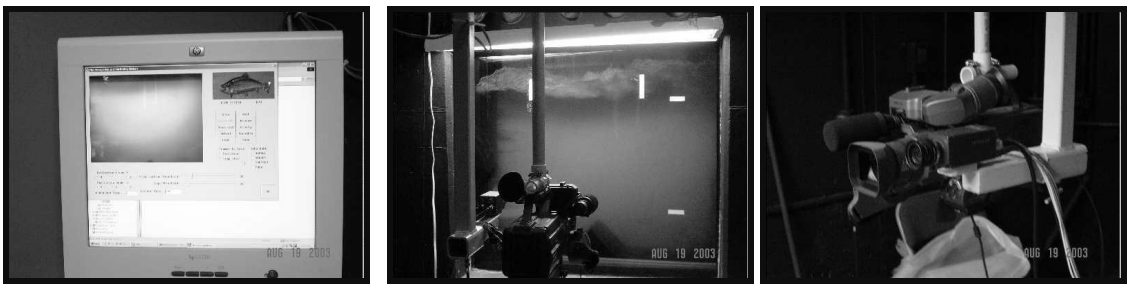

(a) (b)

Figure 3. (a) A vision system and (b) the aquarium.

In order to minimize the lighting effects on fish behavior, the system must be a non-intrusive and passive surveillance system. An extensive study on the impact of illumination on fish behavior was conducted by the US Bureau of Reclamation (USBR) in 2000. The conclusion of Hiebert's study shows that illumination does not have noticeable effect on fish behavior when they pass through the fish passageway [11]. The study also tested Near IR versus visible light and concluded that the fish showed no noticeable change in behavior under these two different types of lighting. These two conclusions indicate that special invisible Near-Infrared light is not needed and regular fluorescent light sources do not affect fish behavior. We used interference filters in different wavelengths and different bandwidths with the NIR camera and found no significant advantage using any specific wavelengths. However, the NIR camera was still selected because of its high sensitivity under low light conditions.

2.3 Field Setup

Pictures in Fig. 4 show (a) the functional graphical user interface, (b) fish viewing window (4'×4') at the test site, (c) camera setup. The viewing window is built in a vault approximately 30 feet below the top surface of the dam. There was a camcorder set up for taping video for off-site manual counting. The video taping system is the current method for monitoring fish migration. It is very labor intensive, medium resolution, and prone to human error. We mounted the NIR camera along side of the camcorder as shown in Fig. 4 (c). The software was developed with friendly graphical user interface for testing. Calibration and parameter adjustment can be done with simple mouse clicks. The viewing window was 4'×4' and the distance allowed for mounting the camera was also 4 feet. Fig. 4 (b) shows the viewing window and the lighting arrangement. Two 40-watt fluorescent lights were mounted on top of the window and two were mounted on the bottom (not shown in (b)). This lighting arrangement was able to provide adequate illumination for image acquisition.



(a) (b) (c)

Figure 4. Functional graphical user interface and pictures from the test site at Prosser Dam in WA.

During field testing we tested and determined the optimal camera shutter speed (1/250 of a second) and the lens focal length (6 mm C-mount lens). Short exposure time ensured good image sharpness but requires more light. We found 1/125 of a second was acceptable considering the lighting available and fish swimming speed. 6 mm lens allowed us to position the camera 4 feet away from the viewing window to cover a 4'×4' area. Using these settings and optics, we were able to acquire sharp and good quality images for analysis.

3. SHAPE EXTRACTION AND REPRESNETATION

3.1 Fish Contour Extraction

Subtraction of images acquired at different times can detect the motion of an object [12]. It is also a simple way to detect the presence of an object assuming a stationary camera position and constant illumination. Fig. 5 (a) shows an image taken without any objects. The only minor variation between frames of the same background is the water turbulence. Averaging of a few frames without objects provides a smooth background image as shown in Fig. 5 (a). Differences between the background image and the image taken at different time shown in Fig. 5 (b) can detect objects distinct from the background as shown in Fig. 5 (c). The difference image contains small pixel clusters (blobs) from water turbulence or image noise. They can be removed with a morphological opening operator. Size and the location of this binary blob can be used for the edge detection process. For small sized fish, edge detection processes can be initiated immediately after the entire fish is in the viewing window. For large fish, that are longer than the viewing window width, subsequent image processing tasks will have to be performed in two steps, one for the head portion of the fish and the other one for the tail. For field testing and data collection, we performed contour extraction for every image that has an object size larger than the set size threshold.

Under normal operation conditions, the difference image should have very good contrast between fish and background as shown in Fig. 5 (c). We used the Canny edge operator to extract the contour because of its contour following (edge tracking) feature. A non-maximum suppression technique based on gradient magnitude is used to *thin* the wide ridges around the local maxima to produce one-pixel wide edges. Once the gradient magnitudes are thinned, extended contour segments can be produced by following high gradient magnitudes from one neighborhood to another [13, 14]. Contour following is initiated only on edge pixels which have high gradient magnitude. However, once it starts, contours are tracked through lower gradient magnitude pixels. A closed fish contour can be detected with this method.

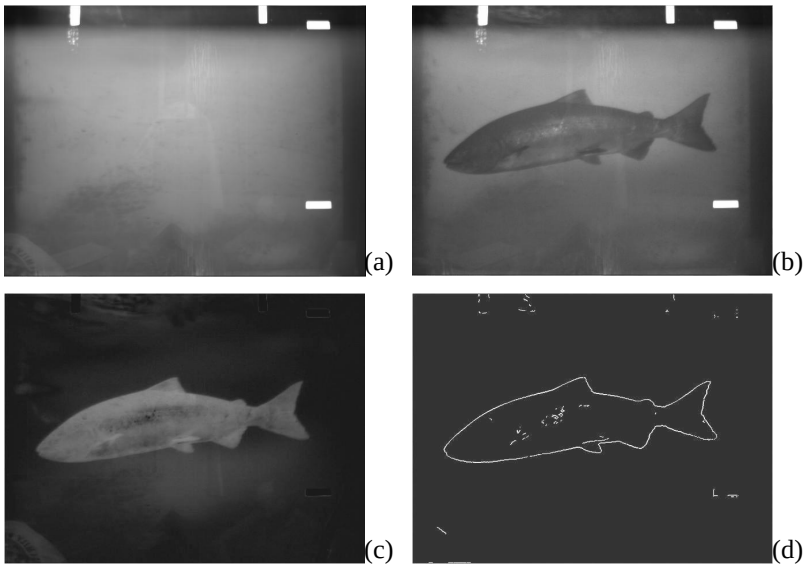


Figure 5. (a) Background image, (b) live image with fish, (c) difference image, and (d) fish contour.

3.2 Fish Detection and Tracking

The flowchart of the fish detection and tracking algorithms is shown in Fig. 6. We first calibrated the system to acquire a reference image (without fish) by averaging 10 or more frames of image to smooth out the image background noise. After the reference image was established, it was then subtracted from every frame of live image acquired from the camera to obtain a difference image. The acquired live image is sent to an edge detector and an edge threshold was applied to the edge image to determine the fish location. A difference threshold was applied to the difference image to obtain a binary image of the fish or object.

The binarized difference image and the binarized edge image were then sent to the contour detection subroutine to detect the fish/object contour. For the binarized edge image, we treated it as a mask, and "AND" it with the binarized difference image. Combining the contour and the mask from the edge image, we were able to extract the fish/object bounding box. This bounding box was checked to see if it falls in the area of interest. If this bounding box falls outside the area of interest, then the acquired image will be ignored and a new image will be acquired to repeat the same process. If the bounding box falls within the area of interest, then the image will then be fully processed.

We learned from field testing at the Prosser Dam that sunlight variation, turbulence, and debris could cause false triggering because they are significantly different from the background reference image. To solve this problem, we implemented a tracking mechanism that monitors the bounding box of the fish candidate. If the bounding box is detected several times (programmable) in a row at the same location, the system will automatically recalibrate and create a new reference image that includes the debris or turbulence. This new auto-calibrate feature was tested and showed very good results. As a result of this added feature, we were able to leave the system at the Prosser Dam to collect image data for almost two months. This detection and tracking software was developed so that we were able to perform field testing to experiment with different tasks and to learn and improve the image acquisition system. We also used this software and the vision system to acquire real fish images for the development of our fish recognition algorithms.

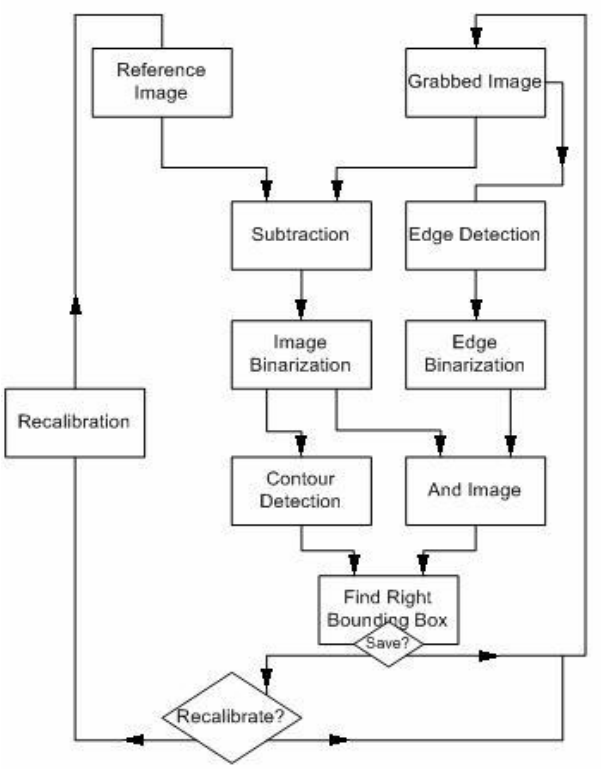


Figure 6. Flowchart of the fish detection algorithms.

3.3 Shape Representation

Shape descriptors and classification algorithms must be invariant to translation, rotation, and scaling because objects can be viewed from different angles, locations, and with varying size. In order to accurately classify fish species, it is first necessary to reduce the number of data points on the contour to a reasonable number that can be evaluated using shape similarity measurement. Many of the data points obtained in the contour extraction algorithm are redundant. Additionally, we want to filter out the data points that contain edge noise. We found that a reduced data set of 40 points was sufficient to retain the important shape features for comparison. Another reason for reducing data points is to be able to represent contours in tangent space for l_2 norm calculation or using bend angle function for calculating Fourier descriptors.

Data reduction can be accomplished through a curve evolution technique that iteratively compares all the relevance measures of the vertices on the contour [15-17]. A higher relevance measure means that the vertex makes a larger contribution to the overall shape of the contour, and thus is more important to be retained. For each iteration, the vertex with the lowest relevance was removed and a new polygon was created by connecting the remaining vertices with a straight line. We modified the relevance measure, K, for the curve evolution method in [15-17] to remove redundant points while maintaining the significance of the contours. The new relevance measure is shown in (1) where β is the turn angle on the vertex between line segments s_1 and s_2 and $l(s_1)$ and $l(s_2)$ are the normalized length from the vertex to the two adjacent vertices [18-19].

$$K(s_1, s_2) = \frac{|(\beta(s_1, s_2) - 180)|\, l(s_1) l(s_2)}{l(s_1) + l(s_2)} \tag{1}$$

This modified curve evolution method reduces short, straight line segments that provide little information about the overall shape of the object. This method preserves a fixed number of data points (40) for each contour, which makes it easier to measure shape similarity. Unlike equal space sampling that may lose data points containing significant shape information, this modified curve evolution preserves detail shape information. Fig. 7 (a) shows the original data set of a dog that was obtained from the contour extraction algorithm. Fig. 7 (b) shows the reduced data set obtained using Equation (1). Although there is a slight distortion in the shape of Fig. 7 (b), the basic shape and the detail of the object was retained. These 40 data points, although not equally spaced, make the most significant contribution to the shape and can be easily used for shape representation. Fig. 7 (c) shows the 40 data points obtained by equal space sampling. It contains redundant data points that form a straight line, which should be removed. It also loses the detail of the contour especially on the round corners or curves.

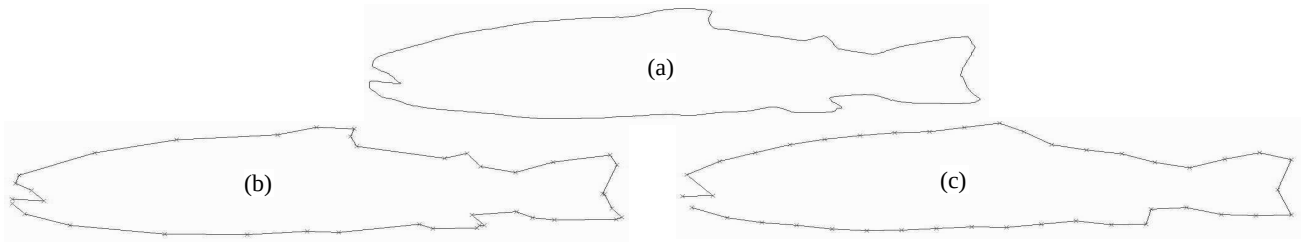


Figure 7. (a) Original data set, (b) reduced to 40 points using Equation 1, and (c) equal space sampling.

3.4 Tangent Space and Bend Angle

Polygon approximation of the reduced data points can be expressed as turn angle vs. normalized length (tangent space) or bend angle vs. normalized length. The turn angle is calculated for each segment by referencing to the reference horizontal line. The bend angle is calculated so that the clockwise turn gives a negative angle whereas a counter clockwise turn gives a positive angle. The representations of these two functions are shown in Fig. 8.

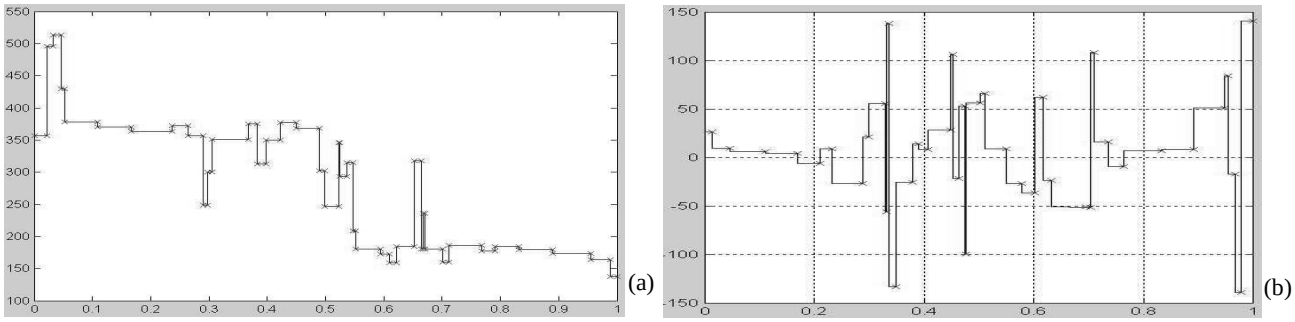


Figure 8. (a) Turn angle and (b) bend angle function vs. normalized length

They are both represented as a function of normalized length to meet the scaling invariant requirement. They are translation invariant because the turn angle or bend angle and length do not contain information about the shape location. The object rotation angle becomes irrelevant when calculating the turn angle and bend angle functions because they are both based on the relative angles. In other words, both functions will stay the same even if the object is rotated, translated, or resized. Depending on the data point selected as the first point in the data sequence, these two functions may shift along the x axis. Moreover, rotation will cause the turn angle function to shift along the y axis.

4. SHAPE MATCHING

We reviewed different shape matching methods that are capable of reporting similarity between shapes. Methods reviewed include global shape characteristics such as perimeter, convex perimeter, major axis length and angle, minor axis length and angle, compactness, and roughness, invariant moments, matching in tangent space, and Fourier descriptors using bend angle function. From our past research work [8, 9, 18, 19], we learned that the first two methods do not yield accurate matching results especially for minor shape variations. In this paper we used turn angle function to represent shapes in tangent space and computed shape similarity using l_2*-norm.* Alignment of two turn functions for similarity measurement is a time consuming process. We developed a much faster matching algorithm that uses power cepstrum to align two similar signals before the similarity can be measured. We also used bend angle function for shape representation and computed Fourier descriptors for performance evaluation.

4.1 Matching Bend Angle Functions Using Fourier Descriptors

For the bend angle function approach, the Fourier descriptors can be used to measure the similarity between two shapes. The Fourier expansion of $\Theta(l)$ is expressed as

$$\Theta(l) = \mu_0 + \sum_{n=1}^{\infty} (a_n \cos nl + b_n \sin nl) \tag{2}$$

, where a_n and b_n are coefficients for each frequency component. Since $\Theta(l)$ is a step function, μ_0, a_n, *and* b_n *can be derived as*

$$\mu_0 = -\pi - \frac{1}{L}\sum_{k=1}^{m}\lambda_k\theta_k$$

$$a_n = -\frac{1}{n\pi}\sum_{k=1}^{m}\theta_k \sin\frac{2\pi n\lambda_k}{L} \qquad b_n = \frac{1}{n\pi}\sum_{k=1}^{m}\theta_k \cos\frac{2\pi n\lambda_k}{L} \tag{3}$$

$$where\ \ \lambda_k = \sum_{i=1}^{k} l_i \quad and \quad L = \sum_{i=1}^{m} l_i = the\ total\ length$$

Power spectrum of the bend angle function is invariant to the shift in length (l in this case). Because of this property, Fourier descriptors of a bend angle function (function of normalized length) meet all invariant requirements for shape description for shape matching using the bend angle function. Power spectrum (A_n) and phase angle information (F_{jk}) can be calculated as follow [20, 21]:

$$A_n = \sqrt{a_n^2 + b_n^2} \quad and \ \ \alpha_n = \tan^{-1}(b_n / a_n)$$
$$F_{jk} = j^*\alpha_k - k^*\alpha_j \quad where\ j^* = j/\gcd(j,k) \tag{4}$$
$$\gcd : greatest\ common\ divisor$$

4.2 Tangent Space Searching

An important requirement for shape similarity measure is that the shift on the starting point of the polygon should not have any effect on similarity measurement calculations. For rotation or starting point shift, turn function remains the same except shifting vertically when there is a rotation and moving horizontally when there is a shift in starting point. This means turn functions of two shapes must be aligned to compensate for the shifts caused by rotation and starting point shift before they can be compared. The distance between two turn functions, Θ_A and Θ_B , can be measured as

$$\delta_2(A,B) = \|\Theta_A - \Theta_B\|_2 = \sqrt{(\int_0^1 |\Theta_A - \Theta_B|^2 ds)} = \sqrt{\min_{\theta\in R, t\in[0,1]}(\int_0^1 |\Theta_A(s+t) - \Theta_B(s) + \theta|^2 ds)} \tag{5}$$

In most cases, the two turn functions are not identical because of the differences in shape. The alignment can only be achieved through minimizing the distance while shifting one turn function. In other words, the distance between two turn functions is obtained by performing a two-dimensional search to find the minimum distance. Another approach is to reduce the search to one dimension by calculating the best value of the turn angle θ [15-17]. The best value of θ is a function of length shift t in the x axis to minimize

$$h(t,\theta) = \int_0^1 |\Theta_A(s+t) - \Theta_B(s) + \theta|^2 ds \quad \text{when} \tag{6}$$

$$\theta'(t) = \int_0^1 (\Theta_B(s) - \Theta_A(s+t))\, ds = \alpha - 2\pi t \ \ \text{,where}\ \ \alpha = \int_0^1 \Theta_B(s)\, ds - \int_0^1 \Theta_A(s)\, ds\,. \tag{7}$$

For each searching step in x (length) direction, the best value of θ was calculated according to Equation 7. The distance δ was calculated and recorded. After shifting the turn function through the searching range, the minimum δ is the distance between the two turn functions. Fig. 9 shows the one dimensional searching result for comparing two shapes expressed as turn functions. The difference between the two shapes (dissimilarity) can be calculated as the total distance between two aligned signals.

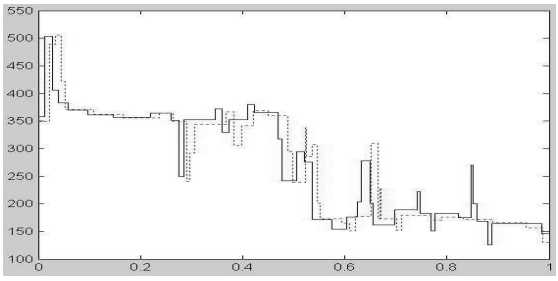


Figure 9. Two aligned normalized turn angle functions in tangent space.

4.3 Matching Turn Functions Using Power Cepstrum

As shown in Section 4.2, matching in tangent space can be simplified to a 1D search. The best estimated value θ can be calculated in Equations 7. However, searching in turn angle is still a time consuming process. In this section we introduced a new searching method that uses power cepstrum technique. The cepstrum technique was first described by Bogert et al. in 1962 [22-26]. It was used to analyze data containing echoes. This technique can be divided into power cepstrum and complex cepstrum. Power cepstrum is usually used to determine the echoes while complex cepstrum can be used to both detect and remove the echoes. This technique has been extended to analyze 2D signals for image registration and 3D vision [27-30]. We chose cepstrum technique over phase correlation method because of better noise tolerance [28, 30], which is equivalent to the shape difference that can be seen in matching two turn functions.

Power Cepstrum is defined as the power spectrum of the logarithm of the power spectrum of a signal [24-26]. This definition can be expressed in Equation 8 as:

$$power\ cepstrum = \left| F(\ln(\left| F(\Theta(l)) \right|^2)) \right|^2 \tag{8}$$

, where $\Theta(l)$ is the turn function and $|F(.)|^2$ is the power spectrum. In our turn angle function matching application, $\Theta(l)$ is the superposition of two turn angle functions, $\Theta_A(l)$ and $\Theta_B(l)$ as shown in Section 4.2. In ideal cases, the only differences between these two turn angle functions are a scale factor α that represents the rotation between two shapes and an index shift l_o between two functions. This is shown in Equation 9 as:

$$\Theta(l) = \Theta_A(l) + \Theta_B(l) = \Theta_A(l) + \alpha\, \Theta_A(l - l_o) \tag{9}$$

The result of applying power cepstrum to $\Theta(l)$ is

$$\begin{aligned} P[\Theta(l)] &= \left| F(\ln(\left| F(\Theta_A(l)) \right|^2)) \right|^2 + A\,\delta(l) + B\,\delta(l \pm l_o) + C\,\delta(l \pm 2l_o) + \cdots \\ &= \text{power cepstrum of the reference turn function } \Theta_A \text{ plus a train of impulses} \\ &\quad \text{occurring at integer multiples of the shift } l_o \end{aligned} \tag{10}$$

By detecting the occurrence of these impulses, shape indexing shift can be determined. The distance (dissimilarity) between the two turn angle functions can be calculated by aligning the two functions according to the detected shape indexing shift.

In turn angle function matching application, the ideal case shown in Equation 10 will not be possible because there will be difference between two contours. Power cepstrum results will have noise around the impulses that makes the detection of the local maxima difficult. However, inaccurate detection of the impulses will yield an incorrect shape indexing shift which will increase the dissimilarity value. This results in higher dissimilarity for shapes that are different and lower dissimilarity for shapes that are similar. Impulses can be detected accurately when comparing similar shapes as shown in Fig. 10. In Fig. 10 (a) two very similar contours with different data point indexing show a small shift between turn functions. Impulses can be detected accurately as shown in Fig. 10 (b). However, when there are noticeable differences between two contours, power cepstrum result will be affected and the detection of impulses will not be as accurate. Fig. 11 (a) shows turn functions of two identical contours but with rotation, scaling, and shift in data indexing. Power cepstrum result is not affected by these variations as shown in Fig. 11(b). Fig. 11 (c) shows turn functions of two contours with noticeable differences. Power cepstrum result becomes noisy and although impulses may not be a good indication of the index shift, the shape similarity measurement still shows that they are different.

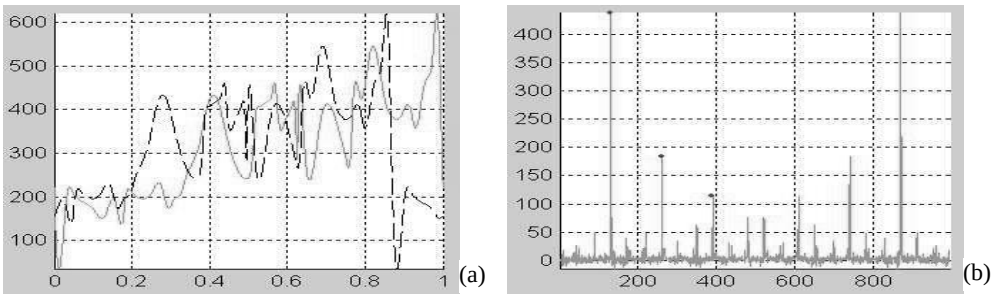


Figure 10. (a) turn functions of two very similar contours, and (b) power cepstrum result.

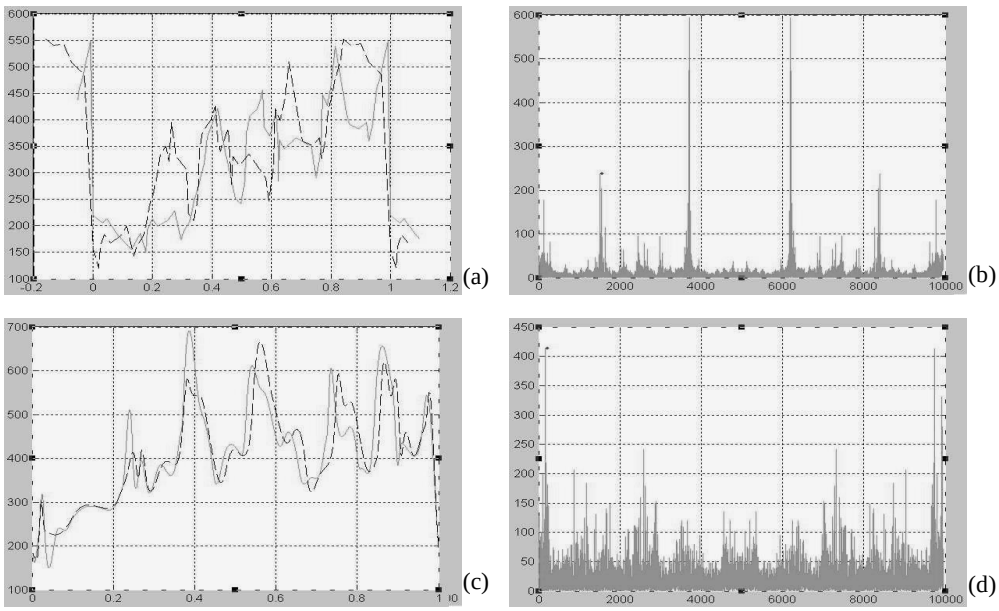


Figure 11. (a) Turn functions of two identical contours but with different data point indexing, scaling, and rotation. (b) Power cepstrum result. (c) Turn functions of two contours with noticeable difference. (d) Power cepstrum result of (c).

5. RESULTS

5.1 Image Collection

Seven fish species that have similar shape characters were chosen for study. They are Chinook Salmon (SA), Winter Coho (WC), brown trout (BT), Bonneville cutthroat (BC), Colorado River cutthroat trout (CRC), Yellowstone Cutthroat (YC) and mountain whitefish (WF). We used a total of 64 images to generate shapes of these seven species and subspecies in our database and they are 10 Chinook salmon, 8 Winter Coho, 10 brown trout, 7 Bonneville cutthroat, 10 Colorado River cutthroat, 10 Yellowstone cutthroat and 9 mountain whitefish. Fig. 12 shows examples (from top to bottom in the order they are mentioned above) of these species.

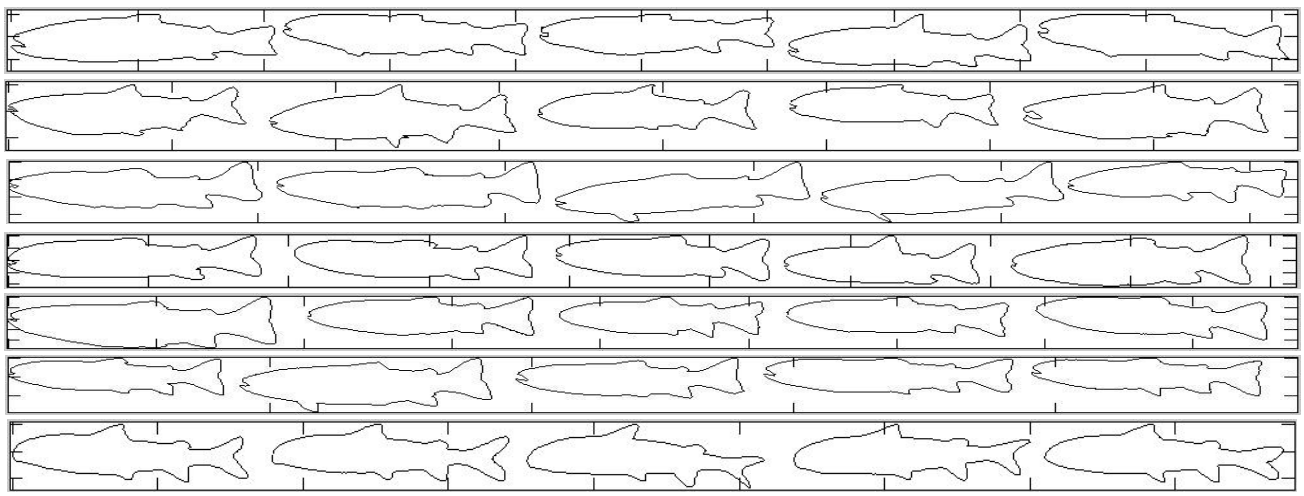


Figure 12. Examples of extracted contours of the seven selected species.

5.2 Analysis

Four matching methods were implemented and tested with this data set. Each contour was used to compare against other contours in the data set. Matching result (dissimilarity) was recorded and ranked to determine the best matched shape category. Performance was evaluated based on the accuracy in correct recognition. Table 1 shows the summary of test result. Fourier descriptors using bend angle function has the highest recognition accuracy of 64%. Power cepstrum matching has slightly lower precision of 60.9% than the Fourier descriptor method. Invariant moments and matching in tangent space have the lowest recognition accuracy.

Table 1. Fish Species Recognition Results

Category	Invariant Moments	Fourier Descriptors	Tangent Space	Power Cepstrum
SA	7/10	10/10	5/10	7/10
WC	5/8	3/8	4/8	4/8
BT	3/10	6/10	4/10	6/10
BC	0/7	3/7	2/7	4/7
CRC	4/10	6/10	8/10	5/10
YC	3/10	6/10	2/10	4/10
WF	8/9	7/9	9/9	9/9
%	46.9%	64%	53.2%	60.9%

6. CONCLUSIONS

This paper presents the design of an automatic fish species recognition system. Component selections and lighting consideration are discussed. Shape recognition methods were implemented for testing the prototype system. We measured shape similarity between the test fish contour and the contours stored in database to determine the fish species. Seven species were selected for experiment. We developed a matching technique using power cepstrum to align two functions in tangent space for comparison. Two other methods were implemented to compare the performance. Although the performance of power cepstrum is slightly lower than the Fourier descriptors, it improves the searching process greatly. Although recognition accuracy is near 60%, they system still has great potential for biological and environment research because images with low recognition confidence can be saved for manual review. It still provides a 60% reduction of labor. Further research work will be conducted to improve recognition accuracy.

REFERENCES

1. Dauble, D. and R. P. Mueller, “Upstream passage monitoring: difficulties in estimating survival for adult Chinook salmon in the Columbia and Snake Rivers.”, Fisheries 25: pp. 24-34, 2000.
2. Chan, D.; Hockaday, S.; Tillett, R.D.; Ross, L.G., “A trainable n-tuple pattern classifier and its application for monitoring fish underwater”, Seventh International Conference on Image Processing And Its Applications, pp. 255 - 259 vol.1, Manchester, UK, 1999.
3. Menard, M.; Loonis, P.; Shahin, A. , “A priori minimization in pattern recognition. Application to industrial fish sorting and face recognition by computer vision”, Proceedings of the Sixth IEEE International Conference on Fuzzy Systems, pp. 1045 - 1050 vol.2, Barcelona, Spain, 1997.
4. Naiberg, A.; Little, J.J., “A unified recognition and stereo vision system for size assessment of fish”, Proceedings of the Second IEEE Workshop on Applications of Computer Vision, pp. 2 – 9, Sarasota, FL, USA, 1994.
5. Nogita, S.; Baba, K.; Yahagi, H.; Watanabe, S.; Mori, S. , “Acute toxicant warning system based on a fish movement analysis by use of AI concept”, Proceedings of the International Workshop on Artificial Intelligence for Industrial Applications, pp. 273 – 276, Hitachi City, Japan, 1988.
6. Gamage, L.B.; de Silva, C.W. , “Use of image processing for the measurement of orientation with application to automated fish processing”, 16th Annual Conference of IEEE Industrial Electronics Society, pp. 482 - 487 vol.1, Pacific Grove, CA, USA, 1990.
7. Bendall, C., Hiebert, S., and Mueller, G., “Experiments in In Situ Fish Recognition Systems Using Fish Spectral and Spatial Signatures”, US Department of the Interior, Open File Report 99-104.

8. Lee, D. J., Redd S., Schoenberger, R., Xu, X., and Zhan, P., "An Automated Fish Species Classification and Migration Monitoring System", Proceedings of The 29th Annual Conference of the IEEE Industrial Electronics Society, p. 1080-1085, Roanoke, Virginia, November 2-6, 2003.
9. Lee, D. J., Zhan, P., Shiozawa, D.K., and Schoenberger, R., "An Automated Fish Recognition and Migration Monitoring System for Biology Research", The 2004 Annual Meeting of the Western Division of the American Fisheries Society, Salt Lake City, UT, March 1-4, 2004.
10. Strout, ., Shiozawa, D.K., and Lee, D.K., "Computerized Fish Imaging and Population Count Analysis", The 2004 Annual Meeting of the Western Division of the American Fisheries Society, Salt Lake City, UT, March 1-4, 2004.
11. Hiebert S., Helfrich, L., Weigmann, D., and Charles Liston, "Anadromous Salmonid Passage and Video Image Quality under Infrared and Visible Light at Prosser Dam, Yakima River, Washington", North American Journal of Fisheries Management, pp. 827-832, 2000.
12. Sonka, M., Hlavac, V., and Boyle, R., "Image Processing, Analysis, and Machine Vision", pp. 682-685 and 744-750, PWS Publishing, 1999.
13. Shapiro, L. and Stockman, G., "Computer Vision", pp. 256-261 and pp. 297-301, Prentice-Hall Inc., Upper Saddle River, New Jersey , 2001.
14. Trucco, E. and Verri, A., "Introductory techniques for 3-D computer vision", pp. 71-80 and 183-191, Prentice-Hall, 1998.
15. Latecki, L.J. and R. Lakämper, "Application Of Planar Shape Comparison To Object Retrieval In Image Databases". Pattern Recognition, 35(1):15-29, 2002.
16. Latecki, L.J. and R. Lakämper, " Shape Description and Search for Similar Objects in Image Databases", *State-of-the-Art in Content-Based Image and Video Retrieval*, Kluwer Academic Publishers, 20001.
17. Arkin, Esther M., L. Paul Chew, Daniel P. Huttenlocher, Klara Kedem, and Joseph S. B. Mitchell, "An Efficient Computable Metric for Comparing Polygon Shapes", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 3, pp. 209-216, March 1991.
18. D. J. Lee, Daniel Bates, Christopher Dromey, and Xiaoqian Xu, "A Vision System Performing Lip Shape Analysis for Speech Pathology Research", *Proceedings of The 29th Annual Conference of the IEEE Industrial Electronics Society*, Roanoke, Virginia, November 2-6, 2003.
19. D.J. Lee, D.M. Bates, C. Dromey, X. Xu, and S. Antani, "An Imaging System Correlating Lip Shapes and Tongue Contact Patterns for Speech Pathology Research", *Proceedings of The 16th IEEE Symposium on Computer-Based Medical Systems***,** p. 307-313, New York, NY, USA, June 26-27, 2003.
20. Gonzalez, Rafael and Richard Woods, *Digital Image Processing*, pp.655-659, Prentice Hall, 2002.
21. Zahn, Charles and Ralph Roskie, "Fourier Descriptors for Plane Closed Curves", *IEEE Transaction on Computer*, vol. C-21, no. 3, March 1972.
22. B. P. Bogert, M. J. R. Healy, and J. W. Tukey, "The Quefrency Analysis of Time Series of Echoes: Cepstrum and Saphe Cracking," in Proceedings, Symposium held at Brown U., 11-14 July 1962, p. 209, M. Rosenblatt, Ed. (Wiley, New York, 1963).
23. D. E. Dudgeon, "The Computation of Two-Diemnsional Cepatra," IEEE Trans. Acoustic Speech Signal processing, ASSP-25, 476 (1977).
24. D. G. Childers, D. P. Skinner, and R.C. Kemerait, "The Cepstrum: A Guide to Processing,", Proceedings, IEEE 65, 1428 (1977).
25. R.C. Kemerait and D. G. Childers, "Signal Detection and Extraction by Cepstrum Techniques," IEEE Trans. Inf. Theory IT-18, 745 (1972).
26. D.J. Lee, S. Mitra, and T.F. Krile, *"Analysis of Sequential Complex Images Using Feature Extraction and 2-D Cepstrum Techniques"*, Journal of Optical Society of America, vol. 6, p. 863-870, June 1989.
27. D.J. Lee, T.F. Krile, and S. Mitra, *"Power Spectrum and Cepstrum Techniques Applied to Image Registration"*, Applied Optics, vol. 27, p. 1099-1106, March, 1988.
28. S. Mitra, D.J. Lee, and T.F. Krile, *"3-D Representation from Time-Sequenced Biomedical Images Using 2-D Cepstrum"*, IEEE Intl. Conference on Visualization in Biomedical Computing, p. 401-408, Atlanta, GA, May 1990.
29. D.J. Lee, S. Mitra, and T.F. Krile, *"Accuracy of Depth Information from Cepstrum-Disparities of a Sequence of 2-D Projections"*, SPIE in Intelligent Robots and Computer Vision, vol. 1192, p. 778-788, Philadelphia, PA, November 1989.
30. D.J. Lee, S. Mitra, and T.F. Krile, *"Noise Tolerance of Power Cepstra and Phase Correlation in Image Registration"*, Optical Society of America Meeting, Santa Clara, CA, November 1988.