**Feature Extraction System for C. Elegans Locomotory Behavioral Analysis**

**Part I – Introduction**

1. **Objective**

The purpose of this project is to design a feature extraction and analysis system to extract information for videos of c-elegan locomotion, and generate table data that can be used for analysis of behaviors. The system will be designed using Matlab.

1. **Behaviors of Interest**

The objective of this project is to identify, characterize, and quantify nematode behaviors. This can be accomplished in two primary ways, one discrete, and the other continuous: 1) analysis of signals consisting of time-series feature values, or 2) analysis of a sequence of labels assigned to frames. Table I describes the baseline set of behaviors of interest. It is important to understand that behaviors may consist of other behaviors. For example, a pirouette behavior may consist of a reversal behavior followed by an gamma turn behavior.

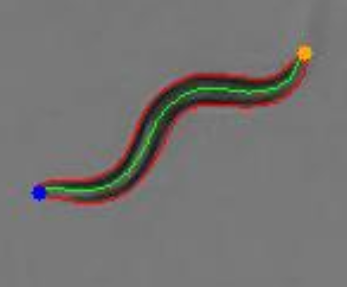
Table I – Baseline Behaviors of Interest

|  |  |
| --- | --- |
| Behavior | Description |
| Running | Sequences during which the worm is moving forward and not executing turns. |
| Resting | Sequences during which the worm is not moving. |
| Reversing | Sequences during which the worm is mving in a straight line in a reverse direction. |
| Pirouetting | A sequence consisting of a reversing behavior followed by a looping behavior. |
| Gamma Looping | A loop behavior where the nematode executes an approximate 180 degree turn by crosses over itself in the manner of a Greek letter gamma. |
| Omega Looping | A loop behavior in which the nematode executes an approximate 180 degree turn in which the head slides past the body without crossing over in the manner of the Greek letter omega. |
| Shallow Turning | A turn of between 0 and 90 degrees without a looping behavior. |

1. **Steps for Image Feature Extraction**

Image feature extraction consists of following steps: 1) Segmenting the individual frames in the video – i.e., identifying the nematode contour, 2) determining the medial axis of the nematode, 3) distinguishing the head and the tail of the nematode, and 4) extracting relevant features for locomotion from these data. Figure 1 shows a typical annotation of a nematode video frame automatically produced by the Feature Extraction System. Single-value image features appropriate for a comma-delimited file are then calculated using information about the foreground pixels (black), the contour (red), the medial axis (green), the head (blue), and the tail (yellow),,

Figure 1 – Typical Annotation of a Nematode Video frame automatically produced by the Feature Extraction System. Single-value image features are calculated using information about the foreground pixels (black), the contour (red), the medial axis (green), the head (blue), and the tail (yellow),,



**Part II – Feature Descriptions**

1. **Overview**

This section describes image features that are extracted. Image features include trajectory, trail, posture, body shape, and experimental features. For illustration, selected graphs of image features are provided based on the extraction of features from two test video sequences:

1. Test Video A - A sequence of 294 frames with 6 sequences of looping behaviors, but with many shot changes, making long-running trajectory information difficult to analyze.
2. Test Video B –
3. Description: Sequence of a shallow turn
4. Resolution: 1080 x 1080
5. Format: 24 bit RGB
6. Frame Rate: 25 frames per second
7. Number of F rames: 535
8. **Trajectory Image Features**

Trajectory features measure the gross movement of the centroid of the nematode. Therefore, the only image information required to be extracted from the video sequence is the location of the nematode centroid in each frame and a time reference for the frame (Table I). All other information can be derived from these data.

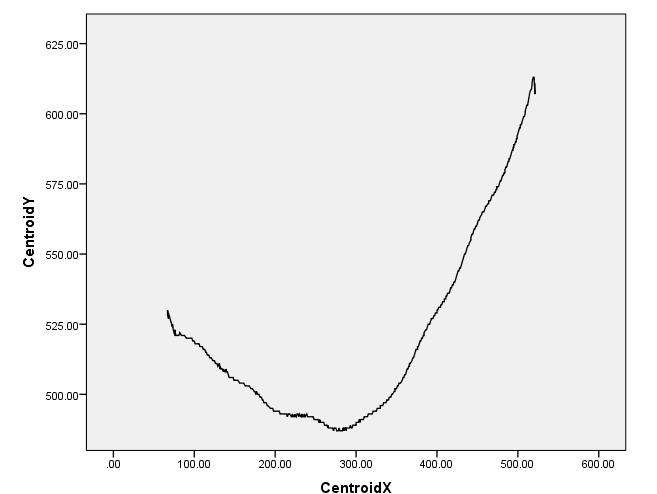
Several other trajectory parameters can be derived from the centroid location. The parameters described in Table 2 currently implemented in the software are based on frame to frame values, but more useful information for analysis will require taking into consideration average values over different time scales. The mean physical wavelength or head to tail distance of the nematode may be an important consideration in determining what these scales are. Figure 2, which is a plot of CentroidX versus CentroidY demonstrates the possible usefulness of trajectory information in understanding nematode behavior (in this case, showing a shallow turn).

Table 1 – Trajectory Low Level Image Features

|  |  |  |
| --- | --- | --- |
| **Feature** | **Description** | **Calculation** |
| ElapsedTime | The relative time of the frame shot | Elapsed time since start |
| CentroidRow | Centroid of the nematode in the vertical direction. | Mean value of the row subscript of all pixels in the segmented nematode. |
| CentroidCol | Centroid of the nematode in the horizontal direction | Mean value of the col subscripts of all pixels in the segmented nematode. |

Table 2 – Derived Trajectory Features

|  |  |  |
| --- | --- | --- |
| **Feature** | **Short Description** | **Calculation** |
| CentroidX | X coordinate of the centroid in the Cartesian coordinate system | Equal to CentroidCol |
| CentroidY | Y coordinate of the centroid in the Cartesian coordinate system | Equal to NumRows – CentroidRow. Where NumRows is the number of pixel rows in the frame |
| Velocity Vector Angle | Centroid | Direction of each centroid movement vector between successive frames |
| Instantaneous Velocity | Centroid | The distance travelled per unit time between successive frames. |
| Instantaneous Acceleration | Centroid | The change in velocity per unit time between successive frames. |
| Average Velocity | Centroid | The distance travelled per unit time over a user-defined time interval. |
| CurrentRange | Current Range | Euclidean distance in pixels from the starting centroid location to the current centroid location . |

Figure 2 – Trajectory Plot For Test Video B, showing the nematode executing a shallow turn

1. **Trail Image Features**

While the trajectory describes the gross movement of the worm, the trail describes the path the worm traces out as it moves. Theoretically a wild-type worm would trace out a sinusoidal wave when it is not turning, and each point along the medial axis would follow this same path. Reality deviates from this theoretical ideal in three ways. First, nematodes slip as they move due to lack of traction, secondly, the head moves rapidly from side to side independent of locomotion (so called foraging behavior), and third, some mutant varieties of nematodes have nonstandard movement characteristics (e.g., lazy tails). In order to account for these factors, we define the trail as the path traced by the seventh pixel from the head end, in order to discount local movement of the head. Non-standard movement may be captured by comparing the path of this point to the actual points of locations along the nematode body. Another important feature of the trail is the skewer angle, which defines the angle of vector connecting the tail to the head. Because the skewer angle varies smoothly and continuously from frame to frame it is used in this system to associate ends of the worm across frames and to distinguish the head from the tail. Table 3 shows the trail image features. Figure 3 shows the trail plot for Test Video B, which shows both the shallow turn and the sinusoid path traced by the nematode in executing the turn. Figure 4 shows the skewer angle (the angle of the line connecting the head to the tail with respect to the horizontal), plotted as a function of frame sequence number for Test Video B during the execution of the turn. The skewer angle should modulate in a sinusoidal fashion during a run, with a change of mean value when the nematode executes a turn.

Table 3 – Trail Image Features

|  |  |  |
| --- | --- | --- |
| **Feature** | **Short Description** | **Calculation** |
| TrailPointRow | Row subscript of a single pixel on the skeleton | Row subscript of the 7th pixel form the head pixel |
| TrailPointCol | Column subscript of a single pixel on the skeleton | Column subscript of the 7th pixel from the head pixel. |
| TrailPointX | X coordinate of a single pixel on the skeleton | TrailPointCol |
| TrailPointY | Y coordinate of a single pixel on the skeleton | NumRows - TrailPointrow |
| Skewer Angle | Orientation of the sinusoid with respect to the horizontal axis | The angle with respect to the horizontal of the line connecting the head to the tail of a worm. |

Figure 3 – Trail Plot for Test Video B showing the path the nematode traces with its sinusoidal movement while executing a shallow turn

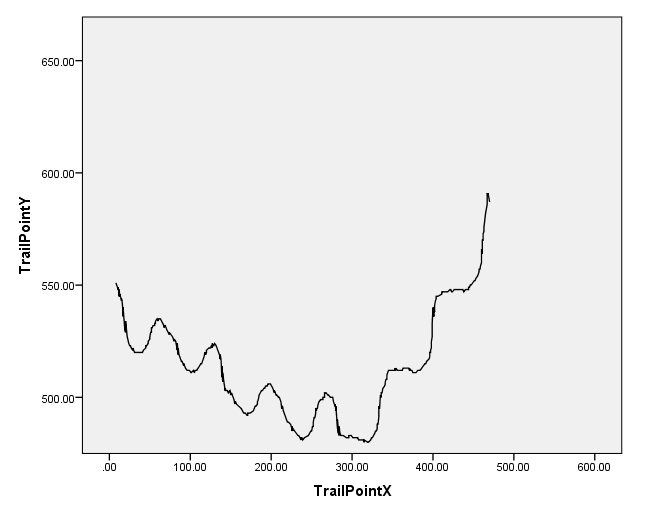
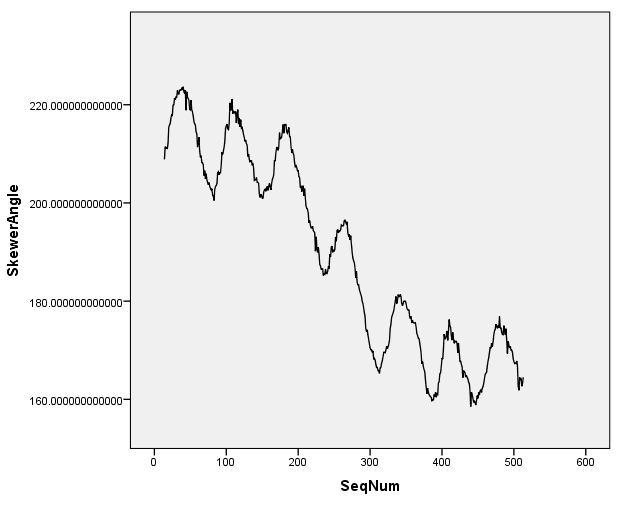


Figure 4 - Skewer Angle Plot as a function of Frame Sequence Number during the execution of a turn. The skewer angle should modulate in a sinusoidal fashion during a run, with a change of mean value when the nematode executes a turn.



1. **Posture Image Features**

Posture image features provide information about the configuration of the nematode in a particular frame. Though sequences of posture features may represent behaviors, these features themselves are not behavior labels. By definition, behavior labels require multi-frame analysis. For example, a gamma turn behavior may consist of a change in non-loop posture (perhaps, wavelength shortening), followed by a sequence of delta loop postures, followed by a sequence of alpha loop postures. Furthermore, some behaviors are independent of posture, such as reversal behavior, which may be indistinguishable from running behavior based on posture. Posture image features are shown in Table 4, and a description of some of these features is given graphically in Figure 5. As an example of how posture features can be used to indicate behaviors, Figure 6 shows skeleton elongation factor (SktElgFactor) as a function of time for Test Video A. Note that looping behavior may create characteristic peaks in the skeleton elongation factor. Also, Figure 7 shows SktElgFactor as a function of sequence number of Test Video B. A large peak in SktElgFactor amplitude may accompany a shallow turn. Finally, Figure 8 shows the categorical variable, Posture, as a function of frame number for Test Video A. The feature identifies 6 series of looping behaviors in the video sequence – four omega loops, one gamma loop, and a delta posture (type of loop unknown) at the end of the sequence .

Table 4 – Posture Image Features

|  |  |  |
| --- | --- | --- |
| **Feature** | **Short Description** | **Calculation** |
| IsLoop | Flag to indicate a looping posture | 1 if Worm head or tail is in contacts with another part of the body, determined by the presence of a hole in the segmentation of the worm. |
| Posture | Categorical variable specifying whether the posture is indicative of a gamma, omega, delta loop, or nonloop posture | 1 – Non-Loop Posture  2 – Delta Loop Posture  3 – Gamma Loop Posture  4 – Omega Loop Posture |
| Major Axis Length | Estimate of the length of the minimum rectangle that encloses the body | The major axis of the ellipse as defined by the first eigenvalue of the covariance matrix of the bw image. |
| Minor Axis Length | Estimate of the width of the minimum rectangle that encloses the body. | The minor axis of the ellipse as defined by the second eigenvalue of the covariance matrix of the bw image. |
| Elongation | Estimate of elongation of the minimum rectangle that encloses the body. | The minor axis divided by the major axis. |
| TrackAmplitude | Amplitude of the skewer-fitted skeleton | Height of the bounding rectangle of the skewer fit |
| SkewerLength | Length of the skewer line | Width of the bounding rectangle of the skewer fit |
| SktAmpRatio | Meauure of the balance of the posture amplitude | The ratio of the maximum distance the skeleton reaches above the skewer fit line to the maximum distance below the skewer fit. |
| SktElgFactor | Ratio of the height to width of the bounding rectangle of the skewer-fitted skeleton |  |
| SktCmptFactor | SkewerFit Skeleton | Skeleton length divided by the area of the bounding rectangle angle of the skewer fitted skeleton. |
| TrackPeriod | The period of the best fit sine wave to the skewer-fitted skeleton |  |

Figure 5 – Graphical Depiction of Selected Non-Loop Posture Features

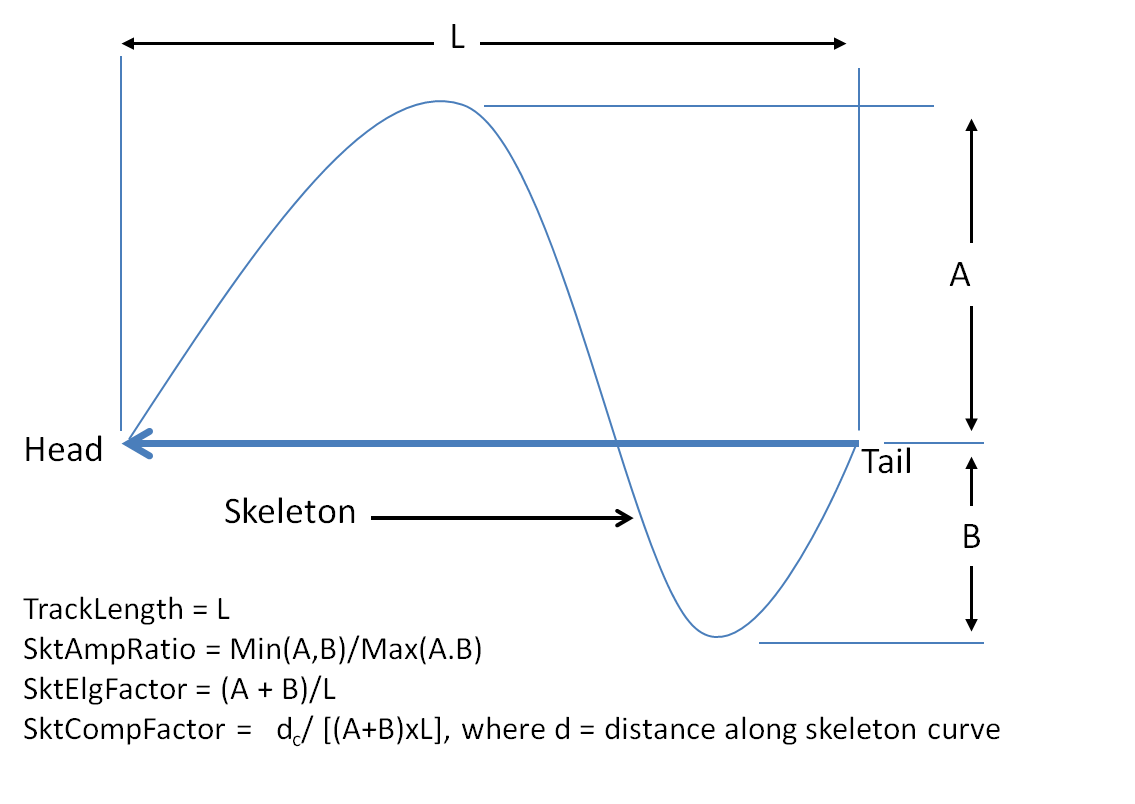


Figure 6 – SktElgFactor as a Function of Frame Sequence Number in Test Video A. Peaks represent the execution of looping behaviors.

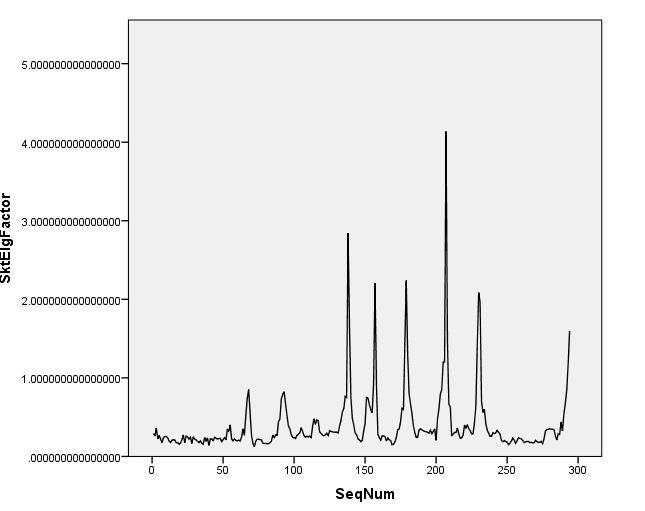


Figure 7 – Plot of SktElgFactor as a function of frame sequence number for Test Video B, indicating possible change in amplitude during the shallow turn.

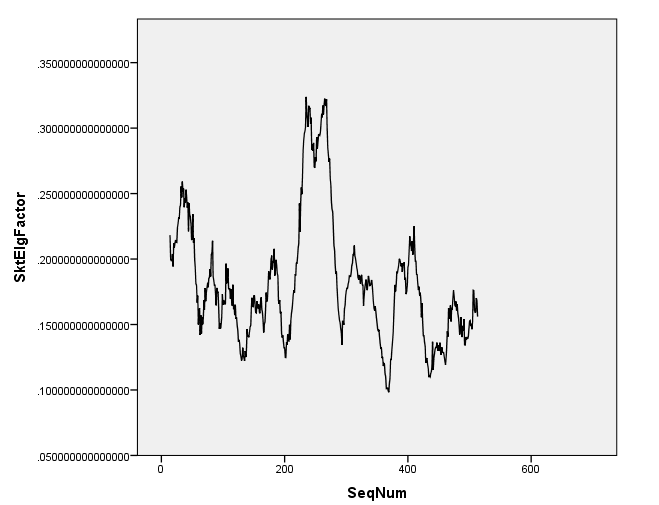
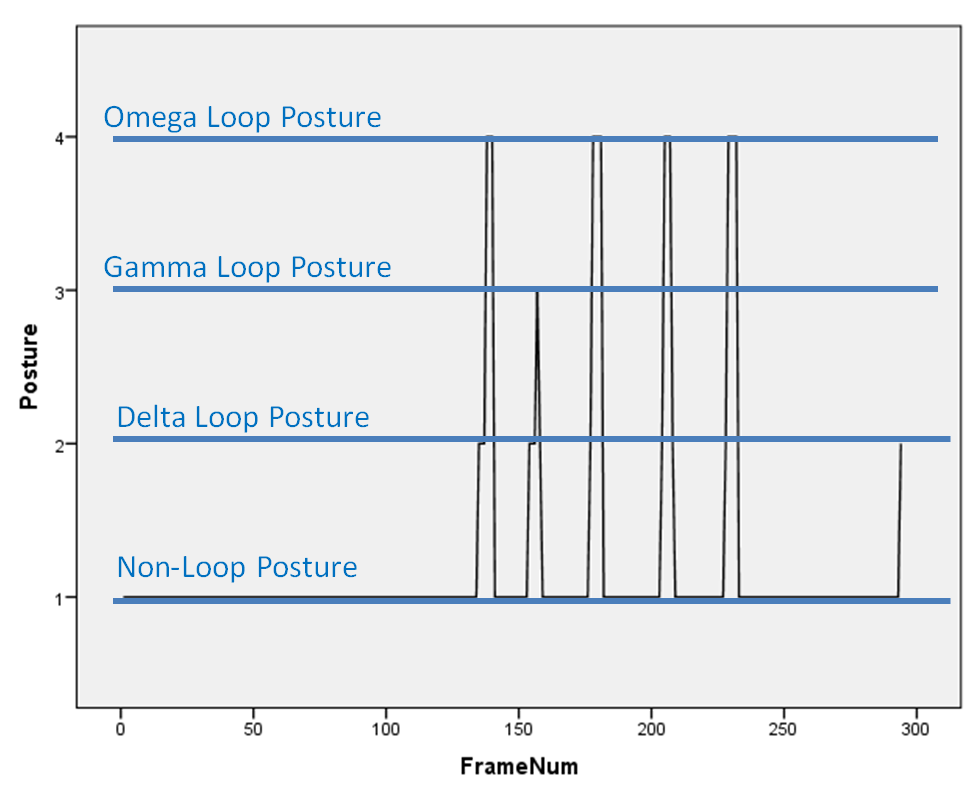


Figure 8 – Plot of Posture Categorical Variable as a function of frame number showing identification of 6 series of looping behaviors in the video sequence – four omega loops, one alpha loop, and a delta posture (type of loop unknown) at the end of the sequence .



1. **Body Shape Features**

Body shape features measure the shape and size of the segmented worm body. Table 5 lists body shape features and how they are calculated. Body shape features should in theory be invariant to nematode posture and behavior.

Body shape features are useful for calibrating the image extraction system and quantifying segmentation and pixilation errors. We would expect the features to be normally distributed based on random Gaussian noise. For example, Figure 9 shows the histogram of nematode area in pixels for Test Video B.

In addition, since many of these features compare perimeter to area of the nematode based on its segmentation, looping postures create outliers. This is because overlapping worm bodies create large mismatches between perimeter and area of the segmented image. Figure 10 shows four body shape features plotted as a function of frame sequence number for Test Video A. The Y axis is a log scale to separate the graphs of each feature. Outliers identify six cases of looping behaviors in the video sequence.

Table 5 – Body Shape Features

|  |  |  |
| --- | --- | --- |
| **Feature** | **Short Desciption** | **Calculation** |
| Width | Width (in pixels) of the nematode at its midpoint. | Measured as the value of the distance transform at the midpoint of the skeleton. |
| Fatness | Area to length Comparison | Area/Length |
| Heywood Circularity Factor | Perimeter to Area Comparison | Ratio of the region perimeter to the perimeter of the circle with the same perimeter. |
| Hydraulic Radius | Perimeter to Area Comparison | Ratio of the region area to its perimeter |
| Rectangle Big Side | Perimeter to Area Comparison | The length of the longest side of the rectangle that has the same area and perimeter as the region |
| Area | Area | Number of pixels representing the animal in the binary image. |
| Perimeter | Perimeter | The Quasi Euclidean perimeter of the segmented region |
| Thickness | Width a midpoint of skeleton | Measured as the length of the shortest line across a worm passing through the midpoint of the worm skeleton, as determined by the minimum of three Euclidean measurements from the pixel to the perimeter made at 90 degrees and +/- 5 degrees |
| Length | Length of the nematode | Quasi-Euclidean Length in pixels of the skeleton curve |

Figure 9 – Histogram of Area for Test Video B, showing Normal Distribution. Gaussian Noise caused by Segmentation Error and Pixelation Effects

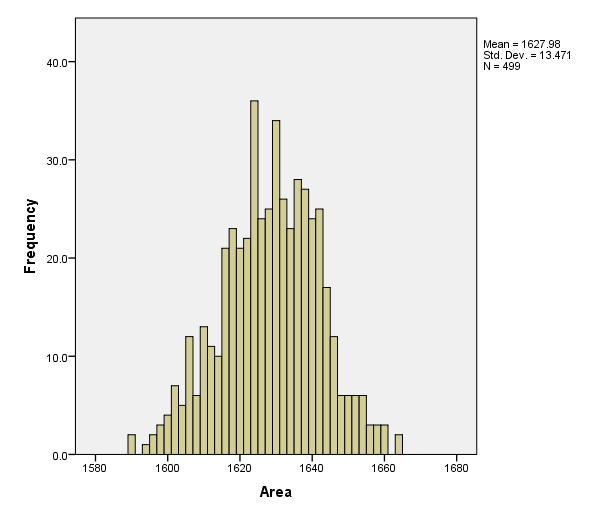
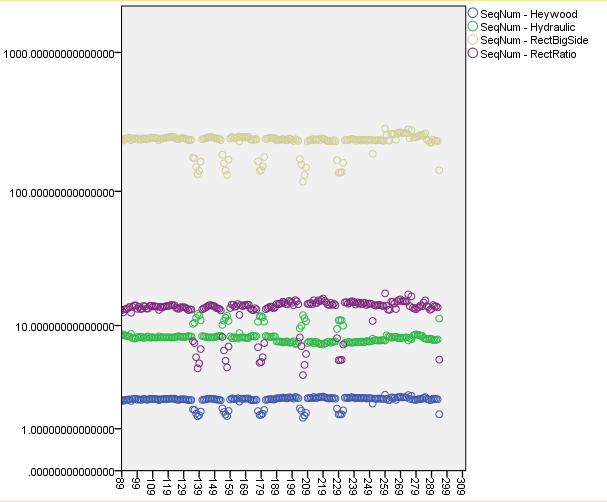


Figure 10 – Time series Plot of Body Shape Features Showing outliers during Looping Behaviors.



1. **Experimental Features**

The following features have been calculated by other researchers in the field, and are currently under evaluation for inclusion in the standard feature set. These features are currently calculated by the prototype software.

Table 6 – Experimental Features

|  |  |  |
| --- | --- | --- |
| **Feature** | **Description** | **Implementation Status** |
| Flex | Maximum skeketon point vector angle difference for a 32 frame window. | No |
| Frequency | Fourier calculated frequency of flexing at each sampled skeleton point. | No |
| Mean Skeleton Angle | Mean value of a sktpVector angle in a sktp set | Yes |
| Mean Point Distance to Length Ratio | Mean distance between adjacent skeleton points divided by the length (Note: should equal the | Yes |
| Max Point Distance to Length Ratio | Maximum distance between adjacent skeleton points divided by the skeleton length | Yes |
| Min Point Distance to Length Ratio | Minimum distance between adjacent skeleton points divided by the skeleton length | Yes |
| Max Skeleton Angle | Maximum SktVector angle in a sktp set | Yes |
| Y axis Moment of Inertia of the Skeleton | Moment of Inertia about the line though the centroid parallel to the vertical axis | Yes |
| X-axis Moment of Inertia of the Skeleton | Moment of Inertia about the line through the centroid parallel to the horizontal axis | Yes |
| Product of Inertia for the Skeleton | Moment of Inertia about the centroid point | Yes |
| Length To Pixel Ratio | Length/ (number of pixels in the skeleton) | Yes |
| Inertia XX | Moment of inertia of the BW image in reference to the x axis | Yes |
| Inertia YY | Moment of inertia of the BW image in reference to the y axis | Yes |
| Inertia XY | Product of the x and y moments of inertia | Yes |
| Number of Pixel in Skeleton | Count of pixels in skeleton. | Yes |
| Flex | Maximum skeketon point vector angle difference for a 32 frame window. | No |
| PushRatio | Ratio of the mean distance all skeleton points travelled between t and t+1 (with superimposed centroids) | No |
| IsReversal | Result is equal to 1 if all the following conditions are met, else 0:   * the skewer angle between successive frames has changed less than 10 degrees * between frame i-1 and i , the head skeleton point moves towards where the tail skeleton point was at time t-1 * between time i-1 and i the tail skeleton point moves away from where the head skeleton point was at time t-1 * the average distances moved by the head and tail during the time interval is larger than 2.5% of the body length | Yes |

**References**

1. Zhaoyang Feng, Chirstopher J. Cronin, John H. Wittig Jr, Paul W. Sternberg, William R Shafer, “An imaging system for standardized quantitative analysis of C. elegans behavior,” BMC Bioinformatics, 26 August 2004.
2. Zhaoyang Feng, Chirstopher J. Cronin, Paul W. Sternberg, William R Shafer, “Quantitative analysis of C. elegans: Algorithms to calculate behavioral and morphologica features,” Supplementary material to Ref 1.
3. IMAQ Vision Users Manual, Chapter 8, Quantitative Analysis,