

# US Patent & Trademark Office

## Patent Public Search | Text View

---

United States Patent Application Publication

20250265860

Kind Code

A1

Publication Date

August 21, 2025

Inventor(s)

Young; Grace Calvert et al.

---

### FISH BIOMASS, SHAPE, SIZE, OR HEALTH DETERMINATION

---

#### Abstract

Methods, systems, and apparatuses, including computer programs encoded on a computer-readable storage medium for estimating the shape, size, mass, and health of fish are described. A pair of stereo cameras may be utilized to obtain off-axis images of fish in a defined area. The images may be processed, enhanced, and combined. Object detection may be used to detect and track a fish in images. A pose estimator may be used to determine key points and features of the detected fish. Based on the key points, a model of the fish is generated that provides an estimate of the size and shape of the fish. A regression model or neural network model can be applied to the fish model to determine characteristics of the fish.

---

**Inventors:** Young; Grace Calvert (Mountain View, CA), James; Barnaby John (Los Gatos, CA), Kimball; Peter (Mountain View, CA), Messina; Matthew (Sunnyvale, CA), Legros; Ferdinand (Mountain View, CA)

**Applicant:** TidalX AI Inc. (Mountain View, CA)

**Family ID:** 1000008578013

**Appl. No.:** 19/070369

**Filed:** March 04, 2025

#### Related U.S. Application Data

parent US continuation 18352971 20230714 parent-grant-document US 12272169 child US 19070369

parent US continuation 17886571 20220812 parent-grant-document US 11756324 child US 18352971

parent US continuation 16734661 20200106 parent-grant-document US 11475689 child US 17886571

---

## Publication Classification

**Int. Cl.:** G06V40/10 (20220101); G06T7/00 (20170101); G06T7/62 (20170101); G06T7/70 (20170101)

**U.S. Cl.:**

**CPC** G06V40/10 (20220101); G06T7/0012 (20130101); G06T7/62 (20170101); G06T7/70 (20170101); G06T2207/10028 (20130101); G06T2207/20081 (20130101); G06T2207/20084 (20130101)

---

## Background/Summary

CROSS REFERENCE TO RELATED APPLICATION [0001] This application is a continuation of U.S. application Ser. No. 18,352,971, filed Jul. 14, 2023, which is a continuation of U.S. application Ser. No. 17/886,571, filed Aug. 12, 2022, now U.S. Pat. No. 11,756,324, issued Sep. 12, 2023, which is a continuation of U.S. application Ser. No. 16/734,661, filed Jan. 6, 2020, now U.S. Pat. No. 11,475,689, issued Oct. 18, 2022, the contents of which are incorporated by reference herein.

### FIELD

[0002] This disclosure generally relates to the marine monitoring systems.

### BACKGROUND

[0003] A manual process of catching and weighing a sample set of fish is often used to estimate the size and weight of a larger number of fish in an area. However, such a process is time-consuming, inaccurate, and requires substantial financial, logistical, and human resources.

### SUMMARY

[0004] According to one example implementation, images of fish are captured using a camera, and are processed and analyzed to construct two-dimensional (2-D) or three-dimensional (3-D) models. Based on processing certain aspects of these models, certain particular features of the fish, including biomass, shape, size, or health, can be determined.

[0005] In general, other innovative aspects of the subject matter described in this specification relate to determining those particular features. In some implementations, a pair of stereo cameras may be utilized to obtain off-axis image views of fish in a defined area. The images may be processed, enhanced, and combined. Object detection may be used to detect and track a fish in images. A pose estimator may be used to determine key points and features of the detected fish. Based on the collected data, algorithms can discern aspects of the fish including biomass, shape, size, or health.

[0006] In some implementations, a 2-D truss network model of the fish can be generated where each truss is a distance between key points detected via image analysis. A regression model or neural network may be applied to the 2-D model to determine specific fish parameters. A regression model or neural network may also be applied to form aspects of the model.

[0007] In some implementations, a 3-D model of the fish can be generated. A regression model or neural network model may be applied to the 3-D model to determine specific fish parameters including biomass, shape, size, or health of the fish. A regression model or neural network may also be applied to form aspects of the model.

[0008] In some implementations, the operations noted above may be repeated and multiple models and weights may be determined for a fish. The multiple models and weights may be averaged to provide results with a greater degree of reliability. In some implementations, each model of a fish

may be evaluated and a score generated to rank the accuracy of a model. Only models that have a score greater than a threshold score may be utilized to determine the key aspects of the fish like weight and health.

[0009] Aspects of the subject matter described in this specification can be embodied in a computer-implemented method. The computer-implemented method can involve one or more processors. These multiple processors can work on one or more images of a fish. Using images of a fish, the method can determine one or more key points associated with one or more features of the fish. These key points can be used to create a model of the fish. The model of the fish can be used to determine a characteristic including biomass, shape, size, and health. The method can then output results to the user or store the data.

[0010] In some implementations, determining characteristics can include determining any of the following health conditions: shortened abdomen, shortened tail, scoliosis, lordosis, kyphosis, deformed upper jaw, deformed lower jaw, shortened operculum, runting or cardiomyopathy syndrome (CMS).

[0011] In some implementations, a single image can be generated from one or more images. From any image, a depth map can be created. With images or depth maps, a convolutional neural network can be used to identify one or more regions of interest on the fish.

[0012] In some implementations, pose estimation can be used to find key points within regions of interest. Images from multiple acquisition devices can be used.

[0013] In some implementations, determining a characteristic of the fish using the model of the fish includes applying a linear regression model to the model of the fish.

[0014] In some implementations, additional images can be captured of the fish. The additional images can be used to help determine characteristics or key points of the fish.

[0015] In some implementations, a neural network classifier can be trained using a pose estimation model to predict likely key points of the fish.

[0016] Other aspects include corresponding methods, systems, apparatus, computer-readable storage media, and computer programs configured to implement the operations of the above-noted methods.

[0017] The above-noted aspects and implementations further described in this specification may offer several advantages. For example, because the determination of fish biomass, shape, size, or health relies on images obtained from cameras, the system may be selective in which images are chosen for determining characteristics because a large number of images can be obtained at a relatively low cost. This selectively provides a greater degree of accuracy in the determination. Furthermore, fewer images may need to be processed to obtain accurate results thereby also providing the benefit of reducing required processing time and resources. Another advantage is that skeletal deformities including spinal deformities like scoliosis, lordosis, and kyphosis, can be screened for and monitored. Health conditions not only affect the quality and rate of reproduction but can also directly affect the economic viability of hatcheries and fish owners. For example, visible skeletal deformities can impact the marketability and price of a fish.

[0018] In some implementations, a single fish or a population of a large number of fish may be continuously monitored and a profile may be created for each fish or for the population as a whole. Thus, for users who are interested in verifying the age, history, or quality of a single fish or the population as a whole, this system provides comprehensive information about physical changes and qualities. The fish, or fish population, profiles may be provided to any interested party.

[0019] In addition, neural networks and machine learning techniques used to implement pose estimation and feature identification provide an adaptive system, which can continuously be trained using previous results and training data to improve the accuracy of feature identification and estimation of the biomass, shape, size, or health of a fish. Since cameras and image processing techniques are used to identify fish and estimate their characteristics, the reliance on humans to catch fish to determine fish characteristics is reduced. The implementations described herein also

provide greater accuracy and consistency in determining fish characteristics including biomass, shape, size, or health. Additionally, the systems can be scaled to cover larger areas of fish relatively easily by adding more cameras in additional locations.

[0020] The details of one or more aspects described in this specification are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages of the subject matter will become apparent from the description, the drawings, and the claims.

---

## Description

### BRIEF DESCRIPTION OF THE DRAWINGS

[0021] FIG. 1A depicts a lateral view of an exemplary fish tank with a fish monitoring system.

[0022] FIG. 1B depicts an aerial view of an exemplary fish tank with a fish monitoring system.

[0023] FIG. 2 depicts a top view of an image acquisition system that captures images of fish in a fish tank.

[0024] FIGS. 3A and 3B depict an exemplary flow diagram of a method for determining a biomass, shape, weight, and health of a fish.

[0025] FIG. 4 depicts an exemplary system for determining a size, shape, and weight of a fish.

[0026] FIG. 5A depicts an image of an example fish with labels corresponding to features of the fish.

[0027] FIG. 5B depicts an image of an example fish with labels corresponding to key points and lines corresponding to truss lengths on the fish.

[0028] Like reference numbers and designations in the various drawings indicate like elements.

### DETAILED DESCRIPTION

[0029] Aspects of the disclosed subject matter are described in further detail with respect to the figures.

[0030] FIGS. 1A and 1B depict a lateral view and an aerial view, respectively, of a fish tank **180** with a fish monitoring system. The tank **180** may have various shapes and sizes. FIGS. 1A and 1B depict an exemplary implementation in which the tank includes a cone base with a circular or cylindrical top. The tank **180** may include different types of water, e.g., fresh water, salt water, water at different salinity levels, and may include one or more species of fish. The tank **180** may be made of any suitable material, such as glass, concrete, acrylic, plastic, or combinations thereof. Additional devices such as air pumps, water pumps, lighting systems, heating and cooling systems, and filtering systems may be used to regulate conditions in the tank **180**.

[0031] The fish monitoring system may include multiple devices and sensors such as light sensors, thermometers, filters, salinity sensors, and image acquisition systems. The image acquisition systems may include a camera system **185** with one or more cameras configured to obtain images and videos of fish in the tank **180**.

[0032] As described in further detail below, the camera system **185** may be connected to a computer system located outside the tank **180**. The computer system may control multiple parameters of the cameras such as position, lens focal length, or zoom, and may control the camera system **185** to obtain still or moving images of fish. The camera system **185** may include a motor configured to maneuver cameras in particular directions based on instructions received from the computer system. The computer system may receive the images from the camera system **185** for further processing.

[0033] The camera system **185** may be deployed in different locations within a tank **180**. In general, the camera system **185** may be located at a position in the tank **180** that enables images of good quality, e.g., clear images of fish without blurriness and optimal capture angle and view, to be captured by the camera system **185**. For example, as illustrated in FIGS. 1A and 1B, the camera system **185** may be located in a relatively central location of the tank **180**. The camera system **185**

may be held in position in the tank **180** in various ways such as by using a surface buoy, sub-surface buoys, fixed structures, or cable lines.

[0034] Various factors may determine the position of the camera system **185** in the fish tank **180**. For instance, in some cases, if fish in the tank **180** are the type of fish that swim against the current, the camera system **185** may be positioned substantially parallel to the current, as depicted in FIG. 2. Other fish may swim with the current or may not have swimming patterns that are dependent upon a current. In some cases, the particular species of fish in the fish tank **180/280** may swim at particular depths or areas that have particular temperatures or amounts of light, and the camera system **185/285** may be positioned within the tank **180/280** to enable cameras **285A** and **285B** to focus on fish in these particular depths or areas.

[0035] As shown in FIG. 2, a camera system **285** in a fish tank **280** may include a plurality of cameras, such as a left stereo camera **285A** and a right stereo camera **285B** or an upper camera and a lower camera. In general, various camera configurations may be used. Each of the cameras **285A** and **285B** may be positioned to obtain images and videos of fish in the fish tank **280**. In some implementations, the cameras **285A** and **285B** may be positioned to obtain images of fish at approximately perpendicular angles relative to the fish so that a lateral view of one or more fish may be obtained. The multiple cameras **285A** and **285B** may provide more than one image for a particular fish from slightly different angles. The multiple images may be used to improve characterization of the fish as described below with respect to FIGS. 3A and 3B.

[0036] In some implementations, the cameras **285A** and **285B** in the camera system **285** are calibrated before obtaining fish images. To calibrate the cameras **285A** and **285B**, the cameras **285A** and **285B** may capture images of reference patterns at different angles and distances relative to the camera lens, and a root mean square (RMS) error may be calculated by determining the difference between the captured images of the patterns and the reference patterns. If the RMS error satisfies an error threshold, settings of the cameras **285A** and **285B** may be adjusted to recalibrate the cameras **285A** and **285B**. Adjusting the settings of the cameras **285A** and **285B** may include any operation that modifies a captured reference image. The operations may include, but are not limited to, one or more of adjusting a position of a camera, adjusting a lens position of the cameras **285A** and **285B**, and adjusting an amount of zoom of the cameras **285A** and **285B**.

[0037] After adjusting the settings of the cameras **285A** and **285B**, another set of images may be captured and a second RMS error may be calculated. The calibration process may be repeated until the RMS error no longer satisfies the error threshold.

[0038] FIGS. 3A and 3B depict an exemplary flow diagram of a method for determining biomass, shape, size, or health of a fish. The method may be implemented by the system described further with reference to FIG. 4. The system may include cameras that are calibrated as described above and configured to obtain one or more images of fish in a fish tank (**S305**). The images may include a left stereo image **305A** and a right stereo image **305B** obtained from a left stereo camera and a right stereo camera, respectively. Other implementations may use different camera angles including top and bottom stereo cameras as additions to, or replacements for, left and right stereo cameras.

[0039] The captured images **305A** or **305B** may be preprocessed (**S310**). The preprocessing may include image enhancement and rectification. For example, images **305A** or **305B** may be enhanced by performing one or more of histogram equalization, filtering, dehazing, deblurring, or denoising to improve image quality. In some cases, light levels may be boosted, for example, by merging multiple images obtained in a burst mode. In some cases, color in an image may be enhanced by performing adaptive histogram equalization.

[0040] In some cases, in response to capturing images **305A** or **305B** with poor image quality, the cameras may be recalibrated as described above. For example, a captured image **305A** or **305B** may be evaluated to determine a quality of the image or the depiction of a fish in the image. If the image **305A** or **305B** is significantly blurred, has occlusions, or the fish is at an undesired angle relative to the camera (e.g., a longitudinal axis of the fish is not perpendicular to the camera), the

cameras may be recalibrated or another image may be captured.

[0041] In some implementations, as part of the preprocessing, an identification of a fish in an obtained image may be determined. For example, a fish that has been tagged or marked using methods such as, e.g., morphological marks, genetic marks, microtags, passive integrated transponder tags, wire tags, radio tags, may be identified by its tag or marker. In some implementations, obtained images may be examined to identify a unique spot pattern of a fish. This unique dot pattern may correspond to a signature of the fish and may be used to identify the fish in subsequent and previous images. Standard object tracking techniques including but not limited to optical flow, visual similarity, Kalman filtering, may be tuned for fish movements and can also be used to track and identify the fish in subsequent and previous images.

[0042] In some implementations, as part of the preprocessing, the left and right stereo images **305A** and **305B** may be combined to form a single image using any suitable image combination or merging technique such as stereo correspondence techniques. Object detection may be performed to detect fish in multiple, preprocessed images or the single, preprocessed image **310A** (**S310**). In some implementations, faster recurrent convolutional neural network (RCNN) may be utilized to perform the object detection.

[0043] In some implementations, semantic segmentation may be performed to segment a fish in an image from the background in the image. Semantic segmentation may make it easier to analyze detailed features of a fish. In general, various suitable object detection techniques may be used to detect fish in a single, preprocessed image **310A**.

[0044] As shown in FIG. **3A**, bounding boxes may be used to identify detected objects in an image **310A**. The bounding boxes may include measured dimensions based on depth measurement and an indication of a margin of error in the measured dimensions. The bounding boxes or detected objects may correspond to regions of interest in the image **310A** such as the images of fish in the image **310A**. If multiple frames are being processed, a nearest neighbor algorithm may be used to find the most likely match of objects between frames.

[0045] In some implementations, a depth map may be generated to determine a distance of a fish from the camera. The depth map may be generated using any suitable technique. For example, Rayleigh scattering or image array depth reconstruction may be used to create a depth map. In addition, one or more of stereoscopic cameras, sonars, acoustic cameras, or lasers may be utilized to determine the distance of a fish from the camera lens.

[0046] After detecting a fish in one or more images, e.g., a combined single image, a stereo image pair, or a sequence of images, and using bounding boxes to identify regions of interest, a pose estimator may be used to identify key points in each region of interest (**S315**). The pose estimator is described further with reference to FIG. **4**. In some implementations, the pose estimator may execute DeepPose operations, multi-fish pose estimation operations, or convolutional neural network operations. As shown in image **315A** (an enlarged version is shown in FIG. **5**), the key points may be associated with features of the fish such as an eye, nostril, gill plate, operculum, auxiliary bone, pectoral fin, lateral line, dorsal fin, adipose fin, pelvic fin, anal fin, and caudal fin. Key points may be labeled by numerical coordinates reflecting pixel positions in an image and may be associated with a particular feature of the fish.

[0047] In some implementations, when the key points and associated features may be partially occluded or non-viewable in an image, the pose estimator can still identify likely key points and associated features based on the probability of a key point and associated feature being present at a particular location. The probability of a key point location may be based on one or more of a likely shape, size, or type of the fish in the image or the location of other features of the fish. For example, using FIG. **5A** as a reference, even though the adipose fin may not be shown in an image, the location of the adipose fin may be estimated using a probability model based on the position of the caudal fin and the dorsal fin in the image. Further processing through neural networks and models can infer, based on models of the particular fish species, where the occluded key points are

located.

[0048] In FIG. 3B, a model **322A** of the fish may be generated using the identified key points from pose estimator **S315** in FIG. 3A. The key points can be stored in item **S320** and used as input for the model process shown in FIG. 3B. The model can be 2-D or 3-D depending on implementation. [0049] For 2-D model implementation, key points stored in **S320** can be used in the construction of a truss network. The truss network is a set of values that indicate distances which separate specific key points. A graphical example of a truss network is shown in FIG. 3B item **322A** or in greater detail in FIG. 5B. An individual truss is the distance between two specific key points. The model format used within the system can be a vector, or group of vectors, where each truss length is an element in a vector.

[0050] For 3-D implementation, various 2-D to 3-D conversion techniques may be used. For example, in some implementations, key points in the 2-D images may be mapped to a 3-D model of the fish using the depth map. The depth map may be determined using various techniques such as a block matching algorithm, depth from motion, or stereo processing by semiglobal matching and mutual information. Objects, i.e., fish, in the stereo images (e.g., left and right images) may be detected, the depths from the cameras determined, and disparities between the images and detected objects may be used to generate the 3-D model.

[0051] In some implementations, the generated model **322A** may be scored and ranked. The score and rank reflect a quality factor of a generated model and the captured image of a fish. The scoring of the model **322A** may be determined based on a number of parameters including one or more of an elevation angle of a fish relative to the camera, a flatness level of the fish relative to the camera, a pose or perpendicularity of the fish relative to the camera, a distance of the fish relative to the camera, or neural network models for scoring particular poses. Values for the elevation angle, flatness level and perpendicularity of the fish and the distance of the fish from the camera may be determined in the previous operations such as when determining a depth map and determining the locations of key points. In some cases, the various parameters may be assigned different weights.

[0052] For example, in some cases, fish having higher elevation angles or fish at greater distances from the camera may have a lower score. In some cases, images of fish in which the fish does not appear relatively perpendicular or flat to the camera may be scored lower. In some cases, the number of determined key points may be used to calculate a score. For example, a higher score may be given to images for which a greater number of key points were determined from the image or fewer key points were determined using a probability model due to a lack of one or more key points being visible in the image. In general, the higher the score, the better the quality of the image and model.

[0053] The score of the model **322A** may be ranked alongside other scores of other models for the same fish, if available (**S325**). For example, as shown in item **325A** in FIG. 3B, the model **322A** of a fish assigned an identification, such as A312, has a score of 86 and is ranked 23. In general, various types of scoring systems and ranking systems using the criteria described above may be utilized.

[0054] If the score or ranking satisfies a threshold, the model **322A** may be utilized to determine biomass, shape, size, or health of the fish (**S330**). For example, if the threshold is a score of 85 or higher or a rank of 25 or higher, the model **322a** may satisfy the threshold based on the score and rank shown in item **325A**. The threshold may be set differently for different fish, environments, or fish tanks.

[0055] To determine key characteristics including weight and health of the fish, various trusses, meaning lengths between key points, of the fish can be used. FIG. 5B shows a number of possible truss lengths including upper lip **500** to eye **501**, upper lip **500** to leading edge dorsal fin **503**, upper lip **500** to leading edge pectoral fin **502**, leading edge dorsal fin **503** to leading edge anal fin **505**, leading edge anal fin **505** to trailing low caudal peduncle **507**, trailing lower caudal peduncle **507** to trailing upper caudal peduncle **508**. Other key points and other separations, including permutations

of key points mentioned, can be used. For different fish, different key points may be generated. For any set of key points, a truss network may be generated as a model.

[0056] Models like **322A** which have passed a threshold, may be compared with existing models. In the 2-D case, both new models and existing models can exist as vectors with value elements corresponding to specific truss lengths. An existing vector may be associated with relevant data including biomass, shape, size, or health of a particular fish. The process of determining characteristics **S330** of a new fish can relate to using a linear regression model, specific clustering algorithms (e.g. K-means or Mixture of Gaussians), or other computational system to find existing vectors in high dimensional proximity to a new vector representing a part, or all of model **322A**. Separate or multiple vectors can be used to compare specific fish characteristics.

[0057] Various health conditions can be monitored for including but not limited to: shortened abdomen, shortened tail, scoliosis, lordosis, kyphosis, deformed upper jaw, deformed lower jaw, shortened operculum, runting and cardiomyopathy syndrome (CMS). Some health conditions, like the shortened operculum condition, can directly increase susceptibility to parasites, reduce health functions like respiration, and increase the rate of mortality within affected specimens. Skeletal deformities including spinal deformities like scoliosis, lordosis, and kyphosis, can not only affect the quality and rate of reproduction but can also directly affect the economic viability of hatcheries and fish owners. Visible skeletal deformities can affect the market price of a fish.

[0058] Determining specific health conditions, including average health and others, can involve focusing on truss length models and using a neural network to compare with a larger data set. In some conditions, involving more or less truss lengths within a model can be advantageous. For diseases like deformed upper jaw, shortened operculum, or others, specific truss lengths can become more heavily weighted within the neural network. Linear regression models as well as clustering algorithms can be used.

[0059] Determined characteristics can be displayed for the user or stored as data. A possible display option is shown in item **330A** of FIG. **3B**. The weight can be displayed in any given weight metric. The health status can be displayed via alphanumeric IDs corresponding to particular health conditions. The imaged fish having an ID of A312 may have an estimated weight of 23 lbs. Fish ID A312 may also have a health condition label, **134c**. Health condition may represent a number of specifiers including health concerns, average health, above average health, specific abnormalities or characteristics. Individual identifiers may be subsets of larger distinctions. Health condition IDs may be used to retrieve analytics on specific fish health as well as the trends and specific values of the health of a fish population. Other forms of representing characteristics of a fish can be used.

[0060] The estimated biomass, shape, size, health, and model of a fish captured in an image may then be output as results (**S335**). The results may be output in several manner. For example, in some cases, the model **322a** along with specific characteristics may be displayed on the display **335A** of a computer device. In some cases, the results may be stored in a fish profile for the fish in a database. The results may be added or aggregated to previous results associated with the fish. New average values for the weight and size dimensions may be determined periodically or each time new results are generated.

[0061] In some implementations, the stored fish data could provide a track record of the fish. For example, a fish could be tracked through its lifetime in a fish tank. A fish may be tracked from birth and through its growth to a fully developed adult fish. As such, details of the timing and type of changes a fish underwent may be recorded. If a party, such as a researcher or fish purchaser, is interested to learn more about a fish's history, the fish database may be queried to retrieve information about the fish's history.

[0062] In some implementations, the results may be provided to train the pose estimator. For example, an image of a fish and its determined model, estimated weight, shape, and size may be provided as a reference to train the pose estimator as training data or to use as a weighted average for the overall fish weight computation. If feedback for the results is available, the feedback may



also be provided as training data. For example, if a reviewer after viewing the results indicates that the results are poor estimates, the reviewer's feedback may be provided as training data to the pose estimator.

[0063] In some implementations, the results may be provided to train the health monitor. For example, an image of a fish and its determined 3-D or 2-D model, estimated weight, shape, size and health condition may be provided as a reference to train the health monitor as training data or to use as a weighted average for the overall fish health computation. If feedback for the results is available, the feedback may also be provided as training data. For example, if a reviewer after viewing the results indicates that the results do not accurately reflect the health condition of the fish, the reviewer's feedback may be provided as training data to the health computation.

[0064] In general, fish may be tracked over long periods of time and over short periods of time. For short-term tracking, a continuous video of the fish may be obtained by controlling a camera system so that cameras in the camera system may continuously capture images of the fish as it moves. In some cases, the camera system may be programmed to automatically track fish movement. In some cases, the camera system may be controlled manually by a user, e.g., systems administrator, to track fish movement.

[0065] For long-term tracking, periodic images of a fish may be obtained, for example, every few days, weeks, or months. Methods to identify the fish may be used to confirm the identity of a fish in an image, and update the identified fish's profile. For example, in some cases, the method to identify a fish may include extracting features from a fish image through representation learning that uses a metric loss to learn a feature extractor based on positive image samples, multiple images of the same fish, and negative image samples, at least one image of a fish and at least one other image of a different fish. In some cases, hand engineering may be used to extract features from a fish image.

[0066] The result of the feature extraction is a function mapping images of the fish to a vector in a high dimensional vector space. Each detection of a fish in an image is either a new observation (first sight) or is close to a cluster of other examples (repeat visit). Clustering algorithms (e.g. K-means or Mixture of Gaussians) may be used to compute clusters. Over time as the fish mature, the cluster may drift or expand and this evolution can be tracked.

[0067] Referring back to FIG. 3B, in some implementations, after determining characteristics of an imaged fish, the system may determine if more data for the fish is requested by a user or required. If more data for the fish is requested or required, the system will repeat the operations in FIGS. 3A and 3B beginning from operation S305. If no more data for the fish is requested or required, the method for determining a fish's characteristics may be terminated.

[0068] The request for additional data for the fish may be explicit or implicit. For example, in some cases, the system may be programmed to obtain multiple sets of data to determine characteristics of a fish, and the measurements may be repeated until the requisite number of data sets has been obtained. In some cases, the system may receive a request from a user to obtain additional data for a particular fish.

[0069] FIG. 4 depicts an exemplary system **400** for determining a size, shape, and weight of a fish. The system **400** may include a fish tank **480**, a computing device **490**, and a server **475**. The fish tank **480** corresponds to the fish tank described with reference to FIGS. 1A, 1B, and 2. The fish tank **480** may include a fish monitoring system that includes multiple sensors, such as light sensors, thermometers, filters, salinity sensors, and a camera system **401**.

[0070] The camera system **401** may include one or more video/photographic cameras, stereo cameras, or optical sensing devices configured to capture images. For instance, the camera system **401** may be configured to capture images of one or more fish at various depths and lighting conditions in tank **480**. The camera system **401** may be configured to capture single, static images of fish and also video images of fish in which multiple images of fish may be periodically captured. The camera system **401** may be controlled based on commands received from the computing

device **490**.

[0071] The camera system **401** may be triggered by several different types of techniques. For instance, motion sensors may be built into the camera system **401** and used to trigger the camera system **401** to capture one or more images when motion is detected. In some implementations, the camera system **401** is configured to receive a command to capture an image from the computing device **490** or a sensor.

[0072] In some examples, the camera system **401** may trigger integrated or external illuminators (e.g., Infrared, Z-wave controlled “white” lights, lights controlled by the computing device **490**) to improve image quality. An integrated or separate light sensor may be used to determine if illumination is desired. Activating the illuminators may result in increased image quality.

[0073] The camera system **401** may be programmed with any combination of time/day schedules, system activation commands, or other variables to determine when images should be captured. The camera system **401** may enter a low-power mode when not capturing images. In some cases, the camera system **401** may be powered by internal, replaceable batteries. In some cases, the camera system **401** may employ a small solar cell to recharge the battery when light is available.

[0074] The camera system **401** may be connected to computing device **490** through cables, and data, such as image **405A**, may be communicated to the computing device **490** through the cables. The computing device **490** may include a pose estimator **492**, memory **494**, processor **496**, and input/output devices **498**. The pose estimator **492** may include a neural network and may be trained using training data and various machine learning methods. The training data may include numerous images of fish with variations in the positions and type of features. Based on the training, the pose estimator **492** may determine probable locations of the fish features and variations in the properties of the features, such as a shape, size, and color of the feature. The pose estimator **492** may also be trained to determine how the variations in the shape and size of a fish and locations of features in the fish affect the weight of a fish.

[0075] Memory **494** may be implemented as one or more mass storage devices, for example, magnetic, magneto optical disks, optical disks, EPROM, EEPROM, flash memory devices, and may be implemented as internal hard disks, removable disks, magneto optical disks, CD ROM, or DVD-ROM disks for storing data. In some implementations, the memory **494** may store fish profile data, which may include size, shape, weight, health, score, and ranking data associated with each profiled fish. The fish profile data may also include one or more images and models of the fish. In some implementations, memory **494** may store training data for training the pose estimator **492** and data for training the RCNN.

[0076] Input/output devices **498** may include input devices such as a keyboard, a pointing device, a mouse, a stylus, and/or a touch sensitive panel, e.g., a touch pad or a touch screen. Output devices may include displays, screens, speakers, and, in general, any device that can output digital data. Input/output devices **498** may also include a transceiver that includes a transmitter and a receiver and may be utilized to communicate with server **475**. The transceiver may include amplifiers, modulators, demodulators, antennas, and various other components. The transceiver may transfer or route data between devices connected to the server **475**. The transceiver may route data communicated between tank **480** and server **475** and between computing device **490** and server **475**. For example, after determining a fish's weight, shape, size, health, and model, the computing device **490** may transmit, via transceiver, fish profile information **430A** such as fish identification, score, rank, weight, and health information to a server **475**.

[0077] Processor **496** may be coupled to the pose estimator **492**, memory **494**, and input/output device **498** for executing instructions to implement the methods described in this specification. In some implementations, executable instructions may be stored in the memory device **110**. The processor **496** may be programmed by encoding an operation as one or more executable instructions and providing the executable instructions in the memory device **110**. The processor **496** may include one or more processing units, e.g., without limitation, in a multi-core configuration.

The term processing unit, as used herein, refers to microprocessors, microcontrollers, reduced instruction set circuits (RISC), application specific integrated circuits (ASIC), logic circuits, and any other circuit or device capable of executing instructions to perform operations described herein. The pose estimator **492** may be implemented as part of the processor **496** or electrically connected to the processor **496**.

[0078] In some implementations, the server **475** may be implemented as a plurality of servers and various components of the server **475** may be distributed across the plurality of servers. Server **475** may be connected to computing device **490** through one or more networks. One or more operations of the method depicted in FIGS. 3A and 3B may be implemented in the computing device **490** or server **475** such that portions of the method may be executed by computing device **490** and other portions by server **475**.

[0079] Server **475** may include any suitable computing device coupled to the one or more networks, including but not limited to a personal computer, a server computer, a series of server computers, a mini computer, and a mainframe computer, or combinations thereof. For example, server **475** may include a web server (or a series of servers) running a network operating system. In some implementations, the server **475** may be connected to or may be integrated with one or more databases, such as fish profile database.

[0080] Server **475** may also implement common and standard protocols and libraries, such as the Secure Sockets Layer (SSL) protected file transfer protocol, the Secure Shell File Transfer Protocol (SFTP)-based key management, and the NaCl encryption library. Server **475** may be used for and/or provide cloud and/or network computing. Although not shown in the figures, the server **475** may have connections to external systems providing messaging functionality such as e-mail, SMS messaging, text messaging, and other functionalities, such as encryption/decryption services, cyber alerts, etc.

[0081] The one or more networks may provide network access, data transport, and other services to the server **475**. The one or more networks may include and implement any commonly defined network architectures including those defined by standards bodies, such as the Global System for Mobile communication (GSM) Association, the Internet Engineering Task Force (IETF), and the Worldwide Interoperability for Microwave Access (WiMAX) forum. For example, the one or more networks may implement one or more of a GSM architecture, a General Packet Radio Service (GPRS) architecture, and a Universal Mobile Telecommunications System (UMTS) architecture. The one or more networks may implement a WiMAX architecture defined by the WiMAX forum or a Wireless Fidelity (WiFi) architecture. The one or more networks may include, for instance, a local area network (LAN), a wide area network (WAN), the Internet, a virtual LAN (VLAN), an enterprise LAN, a layer 3 virtual private network (VPN), an enterprise IP network, corporate network, or any combination thereof. In some implementations, the one or more networks may include a cloud system that provides Internet connectivity and other network-related functions.

[0082] Server **475** may be connected to or may be integrated with one or more databases, such as a fish profile database. The one or more databases may include a cloud database or a database managed by a database management system (DBMS). A DBMS may be implemented as an engine that controls organization, storage, management, and retrieval of data in a database. DBMSs frequently provide the ability to query, backup and replicate, enforce rules, provide security, do computation, perform change and access logging, and automate optimization. Examples of DBMSs include Oracle® database, IBM® DB2, Adaptive Server Enterprise, FileMaker®, Microsoft® Access®, Microsoft® Structured Query Language (SQL) Server, MySQL™, PostgreSQL®, MongoDB, Mondo/ES JavaScript Object Notification (JSON), and a NoSQL implementation. A DBMS typically includes a modeling language, data structure, database query language, and transaction mechanism. The modeling language may be used to define the schema of each database in the DBMS, according to the database model, which may include a hierarchical model, network model, relational model, object model, or some other applicable known or convenient organization.

Data structures can include fields, records, files, objects, and any other applicable known or convenient structures for storing data. A DBMS may also include metadata about the data that is stored.

[0083] Embodiments and all of the functional operations and/or actions described in this specification may be implemented in digital electronic circuitry, or in computer software, firmware, or hardware, including the structures disclosed in this specification and their structural equivalents, or in combinations of one or more of them. Embodiments may be implemented as one or more computer program products, for example, one or more modules of computer program instructions encoded on a computer readable medium for execution by, or to control the operation of, data processing apparatus. The computer-readable medium may be a machine-readable storage device, a machine-readable storage substrate, a memory device, a composition of matter effecting a machine-readable propagated signal, or a combination of one or more of them. The term “data processing apparatus” encompasses all apparatus, devices, and machines for processing data, including by way of example a programmable processor, a computer, or multiple processors or computers. The apparatus may include, in addition to hardware, code that creates an execution environment for the computer program in question, for example, code that constitutes processor firmware, a protocol stack, a database management system, an operating system, or a combination of one or more of them. A propagated signal is an artificially generated signal, for example, a machine-generated electrical, optical, or electromagnetic signal that is generated to encode information for transmission to a suitable receiver apparatus.

[0084] A computer program, also known as a program, software, software application, script, or code, may be written in any form of programming language, including compiled or interpreted languages, and it may be deployed in any form, including as a standalone program or as a module, component, subroutine, or other unit suitable for use in a computing environment. A computer program does not necessarily correspond to a file in a file system. A program may be stored in a portion of a file that holds other programs or data in a single file dedicated to the program in question, or in multiple coordinated files. A computer program may be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

[0085] The processes and logic flows described in this specification may be performed by one or more programmable processors executing one or more computer programs to perform actions by operating on input data and generating output. The processes and logic flows may also be performed by, and apparatus may also be implemented as, special purpose logic circuitry, for example, an FPGA (field programmable gate array) or an ASIC (application specific integrated circuit).

[0086] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processors of any kind of digital computer. Generally, a processor will receive instructions and data from a read only memory or a random access memory or both. A processor may include any suitable combination of hardware and software.

[0087] Elements of a computer may include a processor for performing instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include, or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, for example, magnetic, magneto optical disks, or optical disks. Moreover, a computer may be embedded in another device, for example, a user device. Computer-readable media suitable for storing computer program instructions and data include all forms of non-volatile memory, media and memory devices, including by way of example, semiconductor memory devices, such as EPROM, EEPROM, and flash memory devices; magnetic disks, for example, internal hard disks or removable disks; magneto optical disks; and CD ROM and DVD-ROM disks. The processor and the memory may be supplemented by, or incorporated in, special purpose logic

circuitry.

[0088] While this specification contains many specifics, these should not be construed as limitations on the scope of the disclosure or of what may be claimed, but rather as descriptions of features specific to particular embodiments. Certain features that are described in this specification in the context of separate embodiments may also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single embodiment may also be implemented in multiple embodiments separately or in any suitable sub-combination. Moreover, although features may be described above as acting in certain combinations and may even be claimed as such, one or more features from a claimed combination may in some cases be excised from the combination, and the claimed combination may be directed to a sub-combination or variation of a sub-combination.

[0089] Similarly, while actions are depicted in the drawings in a particular order, this should not be understood as requiring that such actions be performed in the particular order shown or in sequential order, or that all illustrated actions be performed, to achieve desirable results. Moreover, the separation of various system components in the embodiments described above should not be understood as requiring such separation in all embodiments, and it should be understood that the described program components and systems may generally be integrated together in a single software product or packaged into multiple software products.

[0090] It should be understood that the phrase one or more of and the phrase at least one of include any combination of elements. For example, the phrase one or more of A and B includes A, B, or both A and B. Similarly, the phrase at least one of A and B includes A, B, or both A and B.

[0091] Thus, particular implementations have been described. Other implementations are within the scope of the following claims. For example, the actions recited in the claims may be performed in a different order and still achieve desirable results.

## Claims

1. (canceled)

2. A method comprising: obtaining (i) sensor data from a set of one or more sensors and (ii) one or more images of an object from one or more cameras, wherein the sensor data is different than the one or more images; generating, based on (i) the sensor data and (ii) the one or more images, a biomass estimation of the object; and outputting a representation of the biomass estimation of the object for display or storage.

3. The method of claim 2, wherein obtaining the sensor data from the set of one or more sensors comprises: obtaining motion data from one or more motion sensors, wherein obtaining the one or more images of the object from the one or more cameras occurs in response to obtaining the motion data.

4. The method of claim 3, wherein the one or more motion sensors are coupled to the one or more cameras, and wherein obtaining the one or more images of the object from the one or more cameras occurs in response to the one or more motion sensors transmitting a command, using the coupling between the one or more motion sensors and the one or more cameras, to the one or more cameras to capture the one or more images.

5. The method of claim 3, comprising: transmitting a command to the one or more cameras to capture the one or more images.

6. The method of claim 2, comprising: processing the sensor data; and activating, based on processing the sensor data, a set of one or more illuminators configured to illuminate the object.

7. The method of claim 6, wherein processing the sensor data comprises: determining that an illumination level indicated by the sensor data satisfies an illumination threshold.

8. The method of claim 6, wherein obtaining the one or more images of the object from the one or more cameras occurs after activating, based on processing the sensor data, the set of one or more

illuminators configured to illuminate the object.

**9.** The method of claim 2, wherein obtaining the one or more images of the object from the one or more cameras occurs after obtaining the sensor data from the set of one or more sensors.

**10.** The method of claim 2, wherein obtaining the one or more images of the object from the one or more cameras comprises: obtaining one or more images of a fish within an aquaculture environment.

**11.** The method of claim 2, comprising: transmitting instructions to the one or more cameras indicating a period of time to not capture images.

**12.** The method of claim 2, comprising: identifying one or more key points on the object depicted in the one or more images; and generating a model of the object using the one or more key points, wherein generating the biomass estimation of the object uses the model of the object.

**13.** One or more non-transitory computer storage media encoded with computer program instructions that when executed by one or more computers cause the one or more computers to perform operations comprising: obtaining (i) sensor data from a set of one or more sensors and (ii) one or more images of an object from one or more cameras, wherein the sensor data is different than the one or more images; generating, based on (i) the sensor data and (ii) the one or more images, a biomass estimation of the object; and outputting a representation of the biomass estimation of the object for display or storage.

**14.** The media of claim 13, wherein obtaining the sensor data from the set of one or more sensors comprises: obtaining motion data from one or more motion sensors, wherein obtaining the one or more images of the object from the one or more cameras occurs in response to obtaining the motion data.

**15.** The media of claim 14, wherein the one or more motion sensors are coupled to the one or more cameras, and wherein obtaining the one or more images of the object from the one or more cameras occurs in response to the one or more motion sensors transmitting a command, using the coupling between the one or more motion sensors and the one or more cameras, to the one or more cameras to capture the one or more images.

**16.** The media of claim 14, wherein the operations comprise: transmitting a command to the one or more cameras to capture the one or more images.

**17.** The media of claim 13, wherein the operations comprise: processing the sensor data; and activating, based on processing the sensor data, a set of one or more illuminators configured to illuminate the object.

**18.** The media of claim 17, wherein processing the sensor data comprises: determining that an illumination level indicated by the sensor data satisfies an illumination threshold.

**19.** The media of claim 17, wherein obtaining the one or more images of the object from the one or more cameras occurs after activating, based on processing the sensor data, the set of one or more illuminators configured to illuminate the object.

**20.** A system comprising: one or more computers and one or more storage devices on which are stored instructions that are operable, when executed by the one or more computers, to cause the one or more computers to perform operations comprising: obtaining (i) sensor data from a set of one or more sensors and (ii) one or more images of an object from one or more cameras, wherein the sensor data is different than the one or more images; generating, based on (i) the sensor data and (ii) the one or more images, a biomass estimation of the object; and outputting a representation of the biomass estimation of the object for display or storage.

**21.** The system of claim 20, wherein obtaining the sensor data from the set of one or more sensors comprises: obtaining motion data from one or more motion sensors, wherein obtaining the one or more images of the object from the one or more cameras occurs in response to obtaining the motion data.

---