

# Fish recognition and classification system architecture



**ASSIGN  
BUSTER**

## **1. 1 Introduction**

In the previous chapter, the architecture and the approaches of object recognition and classification system were shown in details. Moreover, the features of shape characters of fish that will be used for classification stage are provided. Therefore, this chapter focuses on some background of literature approaches on related work and concepts in the field of object recognition and classification. In particular, a main component to design fish recognition and classification system architecture is used; it will show these experiments history of development in several cases.

The following literature review is divided into four main sections. The fish recognition and classification, first aspect is covered. The second aspect relates with image segmentation techniques to segment underwater images are presented. Investigates most of the feature extraction and selection by shape representation and description, the third aspect is applied. Finally, the classifiers technique for object recognition and classification in aspect of support vector machine is reported.

## **1. 2 Fish Recognition and Classification**

Recently, there were many researchers who attempted to design and apply the interaction between an underwater environment and learning techniques to develop the recognition and classification system in order to classify the fish. Therefore, Castignolles et al., (1994) used off-line method detection with static thresholds to segment the images that recorded by S-video tapes and enhance image contrast by using background lighting. Furthermore, to recognize the species a Bayes classifier was tested after extract twelve geometrical features from fish images. However, this method needs control <https://assignbuster.com/fish-recognition-and-classification-system-architecture/>

on the light of background, determine the value of threshold and multiple imaging. Moreover, where fish are lined up close to each other, the applications tend to be impractical for the real-time.

The moment-invariant for features extraction is fast and very easy to implement. Therefore, Zion et al. (1999) stated the features extraction from dead fish tails by using moment-invariants in order to identify of species. Moreover, the image area is used to estimate fish mass. Furthermore, the accuracy of 99%, 93% and 93%, respectively, for grey mullet, St. Peter's fish and carp is got for identification of fish species. Therefore, later Zion et al., (2000) tested this method with live fish swimming in clean water. The accuracies were 100%, 91% and 91%, respectively for fish species identification. However, the features of the tail in the image which were extracted by the moment-invariant are strongly affects by the water opaqueness and fish motion. This method needs clear environments and all features appear clearly.

An automatic system to select the desirable features for recognition and classification object is needed. Therefore, Chan et al., (1999) developed a 3D point distribution model (PDM) in order to extract the lateral length measurement automatically from the images without an extensive user interaction to locate individual landmark points on the fish by using an n-tuple classifier as a tool to initiate the model. Moreover, the WISARD architecture is used as a Look-Up Table (LUT) which holds information about the pattern that the classifier tries to recognize, in order to assess the performance and usefulness of the n-tuple classifier in the application of fish

recognition. However, this method needs to fix the pre-defined threshold value, amount of prior knowledge for the fish and the bigger training set.

Determine the landmarks as tips of snout or fins for fish are very important to recognize the fish. Therefore, Cardin & Friedland (1999) stated the morphometric analysis by biometric interpretation of fish homologous landmarks as tips of snout or fins for fish stock discrimination. However, they did not refer to algorithms for determining landmarks and the external points are not satisfactory because their locations are subjective.

From other aspect, Cardin (2000) reviewed the landmarks of shape by using morph metric methods for stock identification of fish. Moreover, Winans (1987) used the fins points, extremities point and arbitrarily landmarks in order to identify the fish from those points. Therefore, the attachment of fin membranes were found to be more effective for finfish group discrimination than landmarks located on extremities. Furthermore, Bookstein, (1990) stated the homologous landmarks were found to be more effective in describing shape than other arbitrarily located landmarks. However, these methods should be considered fish sample size, life history, stage of development and the features discriminating power.

Fourier descriptor for geometric features description is very famous algorithm. Therefore, Cadieux et al., (2000) stated the Fourier descriptors of silhouette contours, the geometric features described by seven of moment-invariants stated by Hu (1962) are developed in order to count fish by species from fish ways mounted next to river. Therefore, the 78% of accuracy is achieved by using a majority vote of three classification

<https://assignbuster.com/fish-recognition-and-classification-system-architecture/>

methods. However, this method needs sensors that generate silhouette contours as the fish swim between them and the hardware based on a commercial biomass counter.

The manual measurement for the landmarks points is more accurate to identify the object. Therefore, Tillett et al., (2000) proposed the modification of point distribution model (PDM) in order to segmented fish images by means is proposed. Moreover, the edge and its proximity in order to attract landmarks are considered. Furthermore, the average accuracy of 95% by estimating fish length to manual measurement is compared. However, this method required manual placement of the PDM in an initial position close to the centre of the fish, thereby affecting the accuracy of the final fitting. Also, neighboring fish images forced the PDM away from the correct edges and fish whose orientation was very different from the initial PDM or were smaller than the initial values could not be correctly fitted.

The combining of more than one classifier is important to get more accuracy to classify the objects. Therefore, Cedieux et al., (2000) proposed intelligent system by combining the result of three classifiers in order to recognize the fish. Therefore, Byes maximum quantification classifier, a learning vector classifier and One-class-One-Network of neural network classifier are used by analysis algorithm of an infrared silhouette sensor to acquire the fish and the majority vote. Moreover, the results depended on at least two from three classifiers should be show the same result. However, this method needs other approach for feature selection in order to improve the recognition performance and to optimize the selection of relevant characteristics for fish

classification. Moreover, it needs more computational to identify and classify the object.

Detection, representation the features of object and then the classification are the main steps for any recognition and classification system. Therefore, Tidd & Wilder (2001) stated a machine vision system to detect and classify fish in an estuary by using a video sync signal to drive and direct the strobe lighting through a fiber bundle into a 30 cm $\times$ 30 cm $\times$ 30 cm field of view in a water tank. Moreover, to segment fish images and eliminate partial fish segments, the window-based segmentation algorithm and an aspect ratio are used by means of the segment aspect ratio and a length test.

Furthermore, Bayes classifier is used to classify three fish species from extracted fish image area and aspect ratio. However, this method is tested on only 10 images of each of the species, and needs more computation.

Moreover, they concluded that the system and method have the potential to operate in situ.

The monitoring objects in underwater is difficult problem. Therefore, Rife & Rock (2001) proposed Remotely Operated Vehicles (ROV) in order to follow the marine animal in underwater. However, this method needs continuous hours of the pieces movements.

Locating the critical points of object is very important to determine the length, weight and the area of the objects. Therefore, Martinez et al., (2003) stated an underwater stereo vision system is used to calculate the weight of fish from their length by using a prior knowledge of the species in order to find points of the fish image and linking them with real-world coordinates.

<https://assignbuster.com/fish-recognition-and-classification-system-architecture/>

Moreover, in order to find caudal fin points and the tip of the head, the template matching with several templates is used. Furthermore, accuracy of 95% and 96% for estimated fish weight is reported. However, this method needs a prior knowledge of the species, critical points to calculate the length and only used to find the weight.

The shape of object is very important feature to recognize and identify the objects. Therefore, Lee et al., (2003) developed automated Fish Recognition and Monitoring (FIRM) system, as shape analysis algorithm in order to locate critical landmark points by using a curvature function analysis. Moreover, the contour is extracted based on these landmark points. Furthermore, from this information species classification, species composition, densities, fish condition, size, and timing of migrations can be estimated. However, this method utilizes high-resolution images and determines the location for the critical points of fish shape.

In a conventional n-tuple classifier, the n-tuple is formed by selecting multiple sets of n distinct locations from the pattern space. Therefore, Tillett & Lines (2004) proposed an n-tuple binary pattern classifier with the difference between two successive frames in order to locate the initial fish image for detecting the fish head. Moreover, the dead fish hanging in a tank are used to estimate the mean mass. However, the estimation accuracy was low for live fish images due to poorer imaging conditions and larger fish population density.

The different features can used together to classify the object. Therefore, Chambah et al., (2004) proposed Automatic Color Equalization (ACE) in order

to recognize the fish spaces. Furthermore, the segmentation by using background subtraction was presented. The geometric features, color features, texture features and motion features are used. Then, Bayes classifier is used to classify the selected fishes to one of the learned species. However, this method depends on the color features that need lightness constancy and color constancy to extract visual information from the environment efficaciously.

The semi-local invariants recognition is based on the idea that a direct search for visual correspondence is the key to successful recognition. Therefore, Lin et al., (2005) proposed neighbor pattern classifier by semi-local invariants recognition to recognize the fish. Moreover, when they compare it with integral invariants, they found it less mismatching. Furthermore, they compare wavelet-based invariants with summation invariants and found it has more strong immunity to noise. However, this method needs some critical point of the fish shape.

The Bayesian filter was originally intended for statistical recognition techniques, and is known to be a very effective approach. Therefore, Erikson et al. (2005) proposed fish tracking by using Bayesian filtering technique. Moreover, this method models fish as an ellipse having eight parameters. However, this method considers only counting the fish without looking into its type. Furthermore, the fish may be having varying in number of the parameters.

From other aspect, Lee et al., (2008) stated several shape descriptors, such as Fourier descriptors, polygon approximation and line segments in order to



categorize the fish by using contour representation that extracted from their critical landmark points. However, the main difficulty of this method is that landmark points sometimes cannot be located very accurately. Moreover, it needs a high quality image for analysis.

**Table 1. 1: Critical Analysis of Relevant Approaches**

Author	Algorithm	Remarks
Castignolles et al. 1994	Off-line method	This method needs control on the light of background, determine the value of threshold. Moreover, where fish lined up close to each other, the applications tend to be impractical for the real-time.
Zion et al., 1999	Moment-invariants	The features of the tail in the image which were extracted the moment-invariant are strongly affects by the water opaqueness and fish motion. Therefore, this method not clear environments and all features appear clearly.
Chan et al. 1999	PDM	this method needs to fix the pre-defined threshold value amount of prior knowledge for the fish and the bigger set.
Cardin and Friedland 1999	Morphometric analysis	They did not refer to algorithms for determining landmarks and the external points are not satisfactory because the locations are subjective.
Cardin 2000	Develop Morphometric	These methods should be considered fish sample size, history, stage of development and the features' discrimination

	analysis	power.
Cadieux et al. 2000	Fourier descriptor	This method needs sensors that generate silhouette contours as the fish swim between them and the hardware based commercial biomass counter.
Tillett et al. 2000	Modify PDM	This method required manual placement of the PDM in its initial position close to the centre of the fish, thereby affecting the accuracy of the final fitting. Also, neighboring fish forced the PDM away from the correct edges and fish orientation was very different from the initial PDM or values smaller than the initial values could not be correctly fitted.
Cedieux et al. 2000	Intelligent System	This method needs other approach for feature selection in order to improve the recognition performance and to overcome the selection of relevant characteristics for fish classification. Moreover, it needs more computational resources to identify and track the object.
Tidd and Wilder 2001	Machine Vision System	This method is tested on only 10 images of each of the species, and needs more computation. Moreover, they concluded that the system and method have the potential to operate in situ.
Rife and Rock 2001	ROV	This method needs continuous hours of the pieces movements.
Martinez et al.	Template	This method needs a prior knowledge of the species, contours

al., 2003	Matching	points to calculate the length and only used to find the
Lee et al. 2003	FIRM	This method utilizes high-resolution images and determines the location for the critical points of fish shape.
Tillett and Lines 2004	n-tuple	The estimation accuracy was low for live fish images due to poorer imaging conditions and larger fish population density.
Chambah et. al. 2004	ACE	This method depends on the color features that need color constancy and color constancy to extract visual information from the environment efficaciously.
Lin et al., 2005	Neighbor Pattern Classifier	This method needs some critical point of the fish shape.
Erikson et al. 2005	Bayesian Filtering Technique	This method considers only counting the fish without identifying into its type. Furthermore, the fish may be having varying number of the parameters.
Lee et al. 2008	Several Shape Descriptors	The main difficulty of this method is that landmark points sometimes cannot be located very accurately. Moreover, it needs a high quality image for analysis.

### 1.3 Image Segmentation Techniques

Basically, there are different techniques that would help to solve the image segmentation problems. Therefore, Jeon et al., (2006) categorized these techniques into, thresholding approaches, contour approaches, region approaches, clustering approaches and other optimization approaches using

a Bayesian framework, neural networks. Moreover, the clustering techniques can be categorized into two general groups: partitional and hierarchical clustering algorithms. Furthermore, partitional clustering algorithms such as K-means and EM clustering are widely used in many applications such as data mining, compression, image segmentation and machine learning (Coleman & Andrews 1979; Carpineto & Romano 1996; Jain et al., 1999; Zhang 2002a; Omran et al., 2006). Therefore, this research will focus on the literature review relates with image segmentation techniques to segment fish of underwater images by using k-means algorithm and background subtraction approaches.

### **1. 3. 1 K-Means Algorithm for Image Segmentation**

In general, the standard K-means clustering algorithm is employed in order to cluster a given dataset into k groups. Therefore, the standard K-means algorithm consists of four steps: Initialization, Classification, Centroid computation and Convergence condition. Moreover, several methods attempt to improve the standard K-means algorithm related to several aspects associated to each of the algorithm steps. Furthermore, regarding the computational of the algorithm the steps that need more improvements are initialization and convergence condition (Amir 2007 & Joaquín et al., 2007). Therefore, the following sections will be focused on this step in order to represent and address the review for this step.

#### **1. 3. 1. 1 The Initialization Step of K-Means Algorithm**

Basically, the earliest reference to initialize the K-means algorithm was by Forgy in 1965 that choose points randomly and used as the seeds. Therefore, MacQueen, introduced to determine a set of cluster seeds by using an online <https://assignbuster.com/fish-recognition-and-classification-system-architecture/>

learning strategy (MacQueen 1967 & Stephen 2007). However, this method can be choosing the point near a cluster centre or outlying point. Moreover, repeating the runs is the increased time taken to obtain a solution.

The approach in order to divide the dataset to classes without prior knowledge of classes is required. Therefore, Tou & Gonzales (1974) suggested the Simple Cluster Seeking (SCS) method by Calculating the distance between the first instance in the database and the next point in the database, if it is greater than some threshold then select it as the second seed, otherwise move to the next instance in the database and repeat until K seeds are chosen. However, this method depends on the value of threshold, the order of pattern vectors to be processed and repeating the runs is the increased time taken to reach the seeds chosen.

For optimal partition of dataset which can achieve better variation equalization than standard. Therefore, Linde et al., (1980) proposed a Binary Splitting (BS) method, based on the first run for  $K = 1$ , Then split into two clusters until convergence is reached and the cycle of split and converge is repeated until a fixed number of clusters is reached, or until each cluster contains only one point. However, this method increased the computational complexity by split and the algorithm must be run again.

Good initial seeds for clustering algorithm are significant in order to rapidly converge to the global optimal structure. Therefore, Kaufman & Rousseeuw (1990) suggested selecting the first seed as the most centrally located instance, then the next seed selected based on the greatest reduction in the

distortion and continue until K seeds are chosen. However, this method needs more computation in choosing each seed.

In order to select the optimal seed artificial intelligence (AI) is used.

Therefore, Babu & Murty (1993) and Jain et al., (1996) proposed a method by using genetic algorithms based on the various seed selections as population, and then the fitness of each seed selection is assessed by running the K-means algorithm until convergence and then calculating the Distortion value, in order to select of near optimal seed. However, this method should be run K-means for each solution of each generation. Moreover, a genetic algorithms result depends on the choice of population size, and crossover and mutation probabilities.

Enhancement approach in order to improve the clustering quality and overcome computational complexity is required. Therefore, Huang & Harris (1993) stated the Direct Search Binary Splitting (DSBS) method, based on Principle Component Analysis (PCA), in order to enhance splitting step in Binary Splitting algorithm. However, this method also required more computational to reach k seeds chosen.

Calculating the distance between all points of dataset in order to select the seed is used. Therefore, Katsavounidis et al. (1994) proposed the algorithm as the KKZ algorithm based on preferably one on the edge of the data as the first seed. Then, chosen the second seed based on the point which is furthest from the first seed. Moreover, choosing the furthest point from its nearest seed is repeated until K seeds are chosen. However, this method obvious pitfall from any noise in the data as preferably seed.

In order to increase the speed of the algorithm based on divide the whole input domain into subspaces is required. Therefore, Daoud & Roberts (1996) proposed approach to divide the whole input domain into two disjoint volumes, and then this subspace is assumed that the points are randomly distributed and that the seeds will be placed on a regular grid. However, this methods refers at the end into randomly choose.

The mean of the any dataset is important value in order to estimate the seed depends on it. Therefore, Thiesson et al., (1997) suggested approach to calculate the mean of the entire dataset based on randomly running time of the algorithm to produce the K seeds. However, this method uses the random way to repeat the steps until reach the desirable clusters.

In order to find better clustering initialization of k-means algorithm, Forgy's method is used. Therefore, Bradley & Fayyad (1998) presented a technique that begins by randomly breaking the data into 10, or so, subsets. Then it performs a K-means clustering on each of the ten subsets, all starting at the same set of initial seeds, which are chosen using Forgy's method. However, this method needs to determine the size of the subset and used the same initial seed for each subset.

A way of reducing the time complexity of initialization for k-means algorithm calculation is to use structures like k-d trees. Therefore, Likas et al., (2003) stated a global K-means algorithm which aims to gradually increase the number of seeds until K is found, by using the kd-tree to create K buckets and use the centroids of each bucket as seeds. However, this method needs to test the results to reach the best number of clusters.

The performance of iterative clustering algorithms depends highly on initial cluster centers. Therefore, Mitra et al., (2002) and Khan& Ahmad (2004) proposed a Cluster Centre Initialization Method (CCIA) based on the Density-based Multi Scale Data Condensation (DBMSDC) by estimating the density of the dataset at a point, and then sorting the points according to their density and examining each of the attributes individually to extract a list of possible seed locations. The process is repeated until a desired number of points remain. However, this method depends on other approach to reach the desired seeds, which lead to more computation complexity.

On the other read, in order to reduce the time complexity of initialization for k-means algorithm calculation is to use structures like k-d trees. Therefore, Stephen & Conor (2007) presented a technique for initializing the K-means algorithm based on incorporate kd-trees in order to obtain density estimates of the dataset. And then by using the distance and the density information sequentially to select K seeds. However, this method occasionally failed to provide the lowest value of distortion.

**Table 1. 2: Critical Analysis of Relevant Approaches**

Author	Algorithm	Remarks
Forgy 1965 and MacQueen 1967	Random initial K-means	This method can be choosing the point near a cluster or outlying point. Moreover, repeating the runs is the increased time taken to obtain a solution.



Tou and Gonzales 1974	SCS	This method depends on the value of threshold, the pattern vectors to be processed and repeating the run. Increased time taken to reach the seeds chosen.
Linde et al., 1980	BS	This method increased the computational complexity and the algorithm must be run again.
Kaufman and Rousseeuw 1990	Selecting the first seed.	This method needs more computation in choosing each seed.
Babu and Murty 1993	GA	This method should be run K-means for each solution in each generation. Moreover, a genetic algorithm's result depends on the choice of population size, and crossover and mutation probabilities.
Huang and Harris 1993	DSBS	This method also required more computation to reach the seeds chosen.
Katsavounidis et al. 1994	KKZ	This method has an obvious pitfall from any noise in the data. A preferably seed.
Daoud and Roberts 1996	two disjoint volumes	This method refers at the end into randomly choosing seeds.
Thiesson et al. 1997	the mean of dataset	This method uses the random way to repeat the steps until it reaches the desirable clusters.
Bradley and Fayyad 1998	randomly breaking	This method needs to determine the size of subset and the number of clusters.

	technique	same initial seed for each subset.
Likas et al. 2003	Global K-means	This method needs to test the results to reach the best number of clusters.
Khan and Ahmad 2004	CCIA	This method depends on other approach to reach the best seeds, which lead to more computation complexity.
Stephen and Conor 2007	kd-trees	This method occasionally failed to provide the lowest of distortion.

### 1.3.2 Background Subtraction for Image Segmentation

The basic approach for automatic object detection and segmentation methods is the background subtraction. Moreover, it is a commonly used class of techniques for segmenting out objects of a scene for different applications. Therefore, Wren et al., (1997) proposed running Gaussian Average based on ideally fitting a Gaussian probability density function on the last  $n$  pixel's values in order to model the background independently at each pixel location. Moreover, to increase the speed the standard deviation is computed. Therefore, the advantage of the running average is given by the low memory requirement instead of the buffer with the last  $n$  pixel values are used. However, the empirical weight as a tradeoff between stability and quick update is often chosen.

The detection of objects is usually approached by background subtraction based on multi-valued background. Therefore, Stauffer & Grimson (1999) proposed the multi-valued background model in order to describe the foreground and the background values. Moreover, the probability of

<https://assignbuster.com/fish-recognition-and-classification-system-architecture/>

observing a certain pixel value at specific time by means of a mixture of Gaussians is described. However, this method needs assigning the new observed pixel value to the best matching distribution and estimating the updated model parameters.

Density estimators can be a valuable component in an application like in the use of object tracking. Therefore, Elgammal et al. (2000) proposed a non-parametric model based on Kernel Density Estimation (KDE) by using the last  $n$  background values, in order to model the background distribution.

Moreover, the sum of Gaussian kernels centered as one sample data by the most recent  $n$  background values as background is given. However, complete model estimation also requires the estimation of summation of Gaussian kernels.

The eigen-decomposition methods are computationally demanding by involving the computation of each eigenvector and corresponding eigenvalues. Therefore, Oliver et al., (2000) proposed eigen backgrounds approach based on eigenvalues decomposition by using the whole image instead of blocks of image. Moreover, this method can be improving its efficiency, but depend on the images used for the training set. However, this method not explicitly specified what images should be part of the initial sample, and whether and how such a model should be updated over time.

In order to generate and select of a plurality of temporal pixel samples derived from incoming image, the temporal median filter is used. Therefore, Lo & Velastin (2001) proposed temporal median filter based on the median value of the last  $n$  frames as the background model. Moreover, Cucchiara et

al. (2003) developed the temporal median filter by computing the last  $n$  frames, sub-sampled frames and the time of the last computed median value in order to compute the median on a special set of values. However, the disadvantage of the temporal median filter approach, the computation by a buffer with the recent pixel values is required. Moreover, the median filter does not provide a deviation measure for adapting the subtraction threshold.

The information of the difference frames is accumulated, in order to construct a reliable background image. Therefore, Seki et al., (2003) proposed the background subtraction based on co-occurrence of image variations. Moreover, it works on blocks of  $N \times N$  pixels treated as an  $N^2$  component vector, instead of working at pixel resolution. However, this method offers good accuracy against reasonable time and memory complexity. Furthermore, a certain update rate would be needed to cope with more extended illumination changes.

Background modeling of a moving object requires sequential density estimatio