CHAPTER 6

6. ANALYSIS OF DYNAMIC DETECTION TECHNIQUES USED FOR IDENTIFYING A FISH IN A VIDEO SYSTEM

6.1. DYNAMIC DETECTION TECHNIQUES USED FOR IDENTIFYING A FISH IN A VIDEO SYSTEM

Existing mean Shift is a potent and resourceful non parametric iterative algorithm that can be used for a lot of purposes like finding modes, clustering etc. Mean Shift was introduced in Fukunaga and Hostetler and has been extended to be applicable in other fields like Computer Vision. The main weakness is that they usually neglect the fact that the background images consist of unlike objects whose circumstances may change frequently. The Mean Shift segmentation is a local homogenization method that is of use for damping shading or tonality differences in local objects. An example is better than many words:

Figure 6.1: Action replaces each pixel with the mean of the pixels in a range-r neighbourhood and whose value is within a distance d

The Mean Shift takes usually 3 inputs:

1. A distance function for measuring distances between pixels. Usually the Euclidean distance, but any other well defined distance function could be used. The Manhattan Distance is another useful choice sometimes.
2. A radius. All pixels within this radius (measured according to the above distance) will be accounted for the calculation.

3. A worth difference. From all pixels inside radius r, we will take only those whose values are within this difference for calculating the mean.

Please note that the algorithm is not well defined at the borders, so dissimilar implementations will give you different results. In this extraction background subtraction is a process to detect a movement or significant differences inside of the video frame, when compared to a reference, and to remove all the non-significant components. A Hierarchical copy is developed from the segmented regions of background using Mean-shift algorithm. This Hierarchical model consists of two models, region model and pixel model. The region model is mainly similar to that of the mixture of Gaussian and is chiefly extracted from the histogram of regions which are specific. The pixel model is mainly made up of images that can take place at the same time and are related to each other. The method proposed in this work involves two processing levels.

These are the steps taking place in existing system.

Figure 6.2: Shark Fish Image with Background

Figure 6.3: Shark Fish Image Subtracted from Background
1. The frames of the video are segmented into regions by mean shift and are taken as the input.

2. Next, according to their position to form uniform segments for a scene region different frames are merged. When this procedure takes place, a dynamic strategy of representing region borders is also developed, which leads to a more robust performance for dynamic background.

3. Then the gray value histograms of these regions are computed to build the region models, and pixel models are computed by the pixel co occurrence within each region.

4. The region models are always built as Gaussian mixture models describing the number of components that are determined by a cluster algorithm.

5. For detecting foreground objects, we first usually segment an input frame according to the uniform segments determined.

6. Next, each region is detected whether it contains foreground objects by a corresponding region model.

7. If the detected result shows that a region contains foreground objects, first we will detect the pixel belonging to the foreground with the help of pixel models. Secondly after detecting each frame, parameters of region models and pixel models will be updated.

Advantages

1. It is not necessary that each model must be set to constant parameters because assigning unlike, parameters according to the region also leads to a more accurate description.

2. The weighted pixels in each region make both the descriptor of region and pixels more precise and clear.

3. The hierarchical models reduce the time cost by just deciding which region contains the foreground and be avoiding other regions that do not contain. Because in some dynamic scenes, the locations of background objects are not fixed, each pixel of the segmented regions is assigned a weight to denote the probability that this pixel belongs to one region.
Disadvantages

1. The main disadvantage is Noise. Mean-shift algorithm does not remove the complete noise in the background subtraction.

2. The second is the shadows. Even the shadows are detected as moving object in the existing system.

6.1.1. Proposed Solution

To rise above the problem usage in existing system, we use an advanced technique called Support-Vector Machine for segmenting the background images. The segmented images are meant to form a hierarchical model. This hierarchical model contains region model and pixel model.

The support vector machine is mainly based on statistic learning. They are defined as supervised learning models with associated learning algorithms that can analyze the data and recognize the patterns, and also used for classification and regression analysis. It is a new mechanism learning theory. The support vector machine has been widely applied to many applications like pattern gratitude, function approximation and system identification because support vector machine is able to deal with both the classification and clustering

6.1.2. Classification and Regression by SVM

Generally the model can be divided into Support Vector Regression and Support Vector Classification. The Training data set is given as \{(ci,di)\}_i=1^N \text{ where } ci \in \mathbb{R}^n \text{ and corresponding binary class label } bi \text{ any where } ai \text{ is the } i^{th} \text{ input vector with known binary target } bi \text{, Let it be a non-linear mapping from the data which are original to a high dimensional feature space, and it is mainly used to replace sample points } ci, cj \text{ and they have their mapping images as } ci, cj\text{respectively.}

The weight and bias of hyper plane is defined as w and b, respectively. We define the hyper plane which may be ready to act as a decision surface in feature space, as such,
First we need to separate the data linearly in the feature space and so the decision function must meet a constraint conditions. The optimization problems are

\[
\sum_{j=1}^{l} w_j \phi_j + b = 0
\]

.......... 6.1

\[
\begin{align*}
\text{Minimize} & : c \sum_{i=1}^{l} e_i \\
\text{Subject to} & : d_i [K(x_i, x_j) - b] \geq 1 - e_i \\
\end{align*}
\]

Where \( e_i \) is defined as a stack variable mainly used to relax the margin constraints which are hard. The regular constant \( C > 0 \) is mainly used to implement the trade-off between the maximal margin of separation and the classification error.

Normally there is a problem of which data cannot be linearly separated in sorting. So to avoid this problem, the Support Vector Machine can map the data which are given as input into a feature space which are high dimensional. Usually the Support Vector Machine constructs an optimal hyper level surface in the high dimensional gap which is transferred into a non-linear decision boundary by first converting it into original space. The non-linear expression for the classification function is given as equation.

\[
t(x) = \sum_{i=1}^{l} a_i d_i [K(x_i, x_j) + b]
\]

..........6.3

The routine of SVM also includes the choice of this nonlinear charting function. The SVM applied uses the basis function to perform the operation of mapping. This meaning is expressed in (4).

\[
K(x, d) = \exp(3 - s(x - d)^2)
\]

.......... 6.4

The s parameter in the above equation shows the reflection of the degree of generalization that is made to apply to the data used. Less generalization can be achieved in Support Vector Machine by attaining more data. When there is little s, it may reflect more generalization and a big one reflects less generalization. When the
input data is not regulated, this parameter can perform a normalization task. The
classification scheme may be also defined with the case of the regression. In this
case, the main idea for training the SVM by using d values different from +1 and -1.
Then, an approximation function is derived that fits approximately the known values
only.

Figure 6.4: SVM Basis Function to Perform Mapping
6.1.3. Silhouette Detection Algorithm

In this work it also applies Image-space silhouette algorithms. A very simple and useful way to create the silhouettes of a 3D scene is identifying the silhouette edges in the image space. Image space algorithms utilize discontinuities in the image barriers and extract silhouette edges using image giving out methods. The most thorough way to find silhouettes would be to detect edges in the colour buffer, detecting color discontinuous in the image buffer. A big limitation is the image resolution and complicated texture. Highly textured plane will engender many edges that are unrelated to the object shape; likewise, no edge is sensed between two overlapping objects with the equal colour. Review all the image space silhouette detecting algorithms, the relatively effective way is detecting silhouettes in specially rendered images by developing the geometry information of the 3D scene.

The popular knowledge can be categorized into the two classes: The first type algorithms perceive silhouettes in strength map and normal map individually. By compositing the detecting results of depth map and regular map, we can be satisfied with the last image. Depth map is a system detecting geometry alternating by depth multiplicity. Silhouette detection Algorithm is mainly used to detect the angle point of the image that is moving. This Algorithm is largely concerned to do the geometry inspection mainly on the basis of image’s gray-scale. Then it segregates the pixels into 3 main points. They are angle point, edge point and flat area. To suit a different value when measured in different orders we need to apply about template to figure.

The middle pixel of the pattern is always labelled as nucleus. While sensing the edge, we want to shift the template which is in and around to the image, and then compare every gray piece of the pixel in the pattern along with the nucleus. If the D-value is slighter than the entrance value, mark that this point has like gray-scale to the nucleus.

\[
q(n_0) = \begin{cases} 
1, & |d(r) - d(r_0)| \leq d \\
0, & \text{else}
\end{cases} \quad ----6.5
\]
The parameter d is defined as the threshold value. If the d value is smaller than the threshold value, then that point is marked as to that similar gray-scale to the nucleus. \( q(w,w_0) \) is defined as the pixel function. \( O(r_0) \) is the gray-scale value of the nucleus that is in the center of template. \( O(r) \) is defined as the gray-scale value of other pixels in the templates. D is defined as the threshold value.

Therefore, for any image area the pattern goes through, the area which is formed by all the pixels to satisfy the formula (1) called as similar nucleus value area (SUSAN)[6]. The size of SUSAN area is as follows,

\[
\eta(r_0) = \sum q(W,W_0) \ldots \ldots 6.7
\]

There are two main aspects to consider by using Silhouette detection Algorithm for detecting image edge:

1) The template selection
2) To determine the value for d & g’s threshold value.

The two elements used here are used to determine the efficiency of edges that are detected.

![Image](https://via.placeholder.com/150)

**Figure 6.5:** To Obtain Initial Background Model

(Left) Training Sequence, (200-600 frames), (Right) Estimation of The Times Past Map During Training Run for each Pixel.(D=Starting cell,1a.1b=Sub Cells)

### 6.2. RESULT AND DISCUSSIONS

As the image is digitalized, the template found cannot be the real surrounding. Hence to conquer this we use rectangle template \((2m+1) \times (2m+1)\) instead.

**Algorithm:** The background Selection algorithm

- Step 1. Background Image Set
Step 2. Filter with the current classifier. True goto step 3. False

Step 3. Correctly classified background images.


Step 5. Select Randomly N images from the new Background Image Set

**Performance Analysis Related to Accuracy & Prediction**

![Performance Analysis Graph]

Figure 6.6: Performance Analysis Related to Accuracy and Prediction

To determine the value for d & g’s threshold value

Threshold value D is used to determine the contrast ratio of the object and background that are recognizable. Area with slighter contrast ratio, D should be minor.
Figure 6.7: Working Environment of Image Processing
http://matlabserver.cs.rug.nl/

Figure 6.8: Edge Detection of Shark Fish

Implementation and Experimental Results

This solution was applied using MATLAB. It is a high-level language and interactive environment for numerical computation, visualization, and programming. Using MATLAB, you can analyze data, develop algorithms, and create models and function. The language, tools, and built in math functions enable you to explore multiple approaches and reach a solution faster than with spreadsheets or traditional programming languages [10]. It is important that after the application by the proposed model, it has good results over existing techniques. The proposed method’s performance is compared with many several available methods.
The performance has been observed by calculating noise in the background subtraction.

6.3. SUMMARY

The objective is to develop a motion detection Technique just to extract foreground objects in a wide range of environmental conditions. The new technique was designed to handle the problems typically associated with the background subtraction done by hierarchical model. To overcome the problem of noise and shadows an advanced technique called support vector machine for segmenting the background images can be used. The segmented imagery is meant to form a hierarchical model. It contains region and pixel model. The support vector machine is based on statistic learning. The background subtraction using support vector mechanism can be used to overcome the problem of mistiness and noise and can provide complete feature data.