

first image data set within each background data set, and generating the second set of one or more background data sets may include maintaining multiple samples of the second image data set within each background data set.

[0014] Generating each first background data set may include selecting from the multiple samples one value that is representative of the background for each element within the first image data set, and generating each second background data set may include selecting from the multiple samples one value that is representative of the background for each element within the second image data set. Selecting may include selecting the median value from all sample values in each of the background data sets.

[0015] In other implementations, generating may include comparing the first image data set to a subset of the background data set, and comparing the second image data set to a subset of the background data set.

[0016] In other implementations generating a first difference map further may include representing each element in the first image data set as one of two states, and generating a second difference map further may include representing each element in the second image data set as one of two states, where the two states represent whether the value is consistent with the background.

[0017] In still other implementations, detecting may include identifying a cluster in each of the first and second difference maps, where each cluster has elements whose state within its associated difference map indicates that the elements are inconsistent with the background.

[0018] Identifying the cluster further may include reducing the difference map to one row by counting the elements within a column that are inconsistent with the background. Identifying the cluster further may include identifying the column as being within the cluster and classifying nearby columns as being within the cluster. Identifying the column as being within the cluster also may include identifying the median column.

[0019] Identifying the cluster further may include identifying a position associated with the cluster. Identifying the position associated with the cluster may include calculating the weighted mean of elements within the cluster.

[0020] Detecting further may include classifying the cluster as the object of interest. Classifying the cluster further may include counting the elements within the cluster and classifying the cluster as the object of interest only if that count exceeds a predefined threshold. Classifying the cluster further may include counting the elements within the cluster and counting a total number of elements classified as inconsistent within the background within the difference map, and classifying the cluster as the object of interest only if the ratio of the count of elements within the cluster over the total number of elements exceeds a predefined threshold.

[0021] The step of detecting further may include identifying a sub-cluster within the cluster that represents a pointing end of the object of interest and identifying a position of the sub-cluster.

[0022] In the above implementations, the object of interest may be a user's hand, and the method may include controlling an application program using the absolute position of the object of interest.

[0023] The above implementations further may include acquiring a third image and a fourth image representing different viewpoints of the object of interest, processing the third image into a third image data set and the fourth image into a fourth image data set, and processing the third image data set and the fourth image data set to generate the background data set associated with the background. The method also may include generating a third difference map by determining differences between the third image data set and the background data set, and a fourth difference map by determining differences between the fourth image data set and the background data set, and detecting a third relative position of the object of interest in the third difference map and a fourth relative position of the object of interest in the fourth difference map. The absolute position of the object of interest may be produced from the first, second, third and fourth relative positions of the object of interest.

[0024] As part of this implementation, the object of interest may be a user's hand, and also may include controlling an application program using the absolute position of the object of interest.

[0025] In another aspect, a method of tracking an object of interest controlled by a user to interface with a computer is disclosed. The method includes acquiring images from at least two viewpoints, processing the acquired images to produce an image data set for each acquired image, and comparing each image data set to one or more background data sets to produce a difference map for each acquired image. The method also includes detecting a relative position of an object of interest within each difference map, producing an absolute position of the object of interest from the relative positions of the object of interest, and using the absolute position to allow the user to interact with a computer application.

[0026] Additionally, this method may include mapping the absolute position of the object of interest to screen coordinates associated with the computer application, and using the mapped position to interface with the computer application. This method also may include recognizing a gesture associated with the object of interest by analyzing changes in the absolute position of the object of interest, and combining the absolute position and the gesture to interface with the computer application.

[0027] In another aspect, a multiple camera tracking system for interfacing with an application program running on a computer is disclosed. The multiple camera tracking system includes two or more video cameras arranged to provide different viewpoints of a region of interest and are operable to produce a series of video images. A processor is operable to receive the series of video images and detect objects appearing in the region of interest. The processor executes a process to generate a background data set from the video images, generate an image data set for each received video image and compare each image data set to the background data set to produce a difference map for each image data set, detect a relative position of an object of interest within each difference map, and produce an absolute position of the object of interest from the relative positions of the object of interest and map the absolute position to a position indicator associated with the application program.

[0028] In the above implementation, the object of interest may be a human hand. Additionally, the region of interest