

In an embodiment a video action detection system according to claim 12 is provided.

~~Brief description of the drawing~~

- 5 ~~—————In view of the above, according to a first aspect of the present invention, there is provided a method of automated support vector machine detection of presence of an action of a plurality of predetermined types of action from video data, the method comprising the steps of~~
- 10 ~~—detecting spatiotemporal interest points (STIPs) in the video data;
—extracting descriptors from the video data in spatiotemporal areas at the detected spatiotemporal interest points;
—assigning the detected spatiotemporal interest points to bins in a feature vector based on the extracted descriptors;
—computing bin values of the feature vector for each respective bin, the bin~~
- 15 ~~value of each bin comprising sum of contributions associated with descriptors that have been assigned to the bin, with a bin dependent adjustment of a size of the sum and/or the contributions of at least one of the bins dependent on the predetermined type of action;
—computing a support vector machine (SVM) match score from the feature~~
- 20 ~~vector and reference vectors for the predetermined type of action.~~
- detecting spatiotemporal interest points (STIPs) in the video data;
- extracting descriptors from the video data in spatiotemporal areas at the detected spatiotemporal interest points;
- assigning the detected spatiotemporal interest points to bins in a feature
- 25 histogram based on the extracted descriptors;
- computing bin values of the feature histogram for each of the bins, the bin value of each bin comprising sum of contributions associated with
- descriptors that have been assigned to the bin, a size of the sum and/or the contributions of respective bins being adjusted by multiplication with a
- 30 weighting factor that depends on the respective bin and the predetermined type of action;

- using a decision function of a support vector machine (SVM) to compute a match score from the feature histogram and reference vectors of the support vector machine for the predetermined type of action.

According to a second aspect of the present invention, there is provided a method of automated detection of actions from a plurality of predetermined types of action, comprising executing the method in accordance with the first aspect for each of the plurality of predetermined types of action respectively, using, for each respective different predetermined type of action, a respective different bin dependent adjustment of a size of the sum and/or the contributions of ~~at least one of the~~ the respective bins dependent on the predetermined type of action and respective different sets of reference vectors.

According to a third aspect of the present invention, there is provided a computer readable medium, comprising a program of instructions for a programmable computer system that, when executed by the programmable computer system, will cause the programmable computer system to execute the method according to any one of the preceding aspects.

According to a fourth aspect of the present invention, there is provided a video action detection system for detecting presence of an action of a predetermined type of action from video data, the system comprising

- a source of video data or video data input;
- a spatiotemporal interest point (STIP) detector coupled to the source of video data or video data input;
- a feature data extractor coupled to the source of video data or video data input and the STIP detector, configured to extract feature data from video data from the source of video data or video data input at STIPs detected by the STIP detector;
- a descriptor quantizer coupled to the descriptor extractor and configured to assign the extracted descriptor values to bins of a feature ~~vector~~ histogram;
- a computing circuit configured to compute bin values of the feature ~~vector~~ histogram for each ~~respective of the~~ bins, the bin value of ~~each the~~ bin

comprising a sum of contributions associated with STIPs that have been assigned to the bin, with a bin dependent adjustment of a size of the sum and/or the contributions of ~~at least one of the~~respective bins dependent on the predetermined type of action; and to compute a support vector machine (SVM) match score from the feature ~~vector~~histogram and reference vectors for the predetermined type of action.

Brief description of the drawing

These and other objects and advantageous aspects will be come apparent from a description of exemplary embodiments, using the following figures

- | | |
|---------------|--|
| Figure 1 | shows an action type detection system |
| Figure 2 | shows a flow chart of video processing |
| Figure 3 | shows a flow-chart of support vectors |
| determination | |
| Figure 4 | shows a flow chart of factor selection |

Detailed description of exemplary embodiments

Figure 1 shows a video processing system for detecting action types, the system comprising an input video data memory 10, a feature processor 12 with an input coupled to input video data memory 10, a weight memory 14, a multiplier 15 with inputs coupled to outputs of feature processor 12 and weight memory 14, a detector 16 with an input coupled to an output of multiplier 15, a control computer 19, a support vector memory 17 with an output coupled to detector 16, a detection parameter memory 18 with an output coupled to detector 16. A camera 11 is shown coupled to input video data memory 10. Feature processor 12 comprises a feature detector 120 with an input coupled to input video data memory 10, a feature value extractor 122 with inputs coupled to feature detector 120 and input video data memory 10, a histogram binning (sometimes also called

quantization) unit 124 with an input coupled to histogram binning unit 124 and a binning parameter memory 126 coupled to histogram binning unit

Claims

1. A method of automated support vector machine detection of
 5 | presence of an action of a plurality of predetermined types of action from
 video data, the method comprising the steps of
 - detecting spatiotemporal interest points (STIPs) in the video data;
 - extracting descriptors from the video data in spatiotemporal areas at the
 detected spatiotemporal interest points;
 - 10 | - assigning the detected spatiotemporal interest points to bins in a feature
~~vector-histogram~~ based on the extracted descriptors;
 - computing bin values of the feature ~~vector-histogram~~ for each ~~respective of~~
~~the bins~~, the bin value of each bin comprising sum of contributions
 associated with descriptors that have been assigned to the bin, ~~with a bin-~~
 15 | ~~dependent adjustment of a size of the sum and/or the contributions of at~~
~~least one of the respective bins being adjusted by multiplication with a~~
~~weighting factor that depends on the respective bin and dependent on the~~
 predetermined type of action;
 - using a decision function of computing a support vector machine (SVM) to
 20 | compute a match score from the feature histogram~~vector~~ and reference
vectors of the support vector machine for the predetermined type of action.
2. A method according to claim 1, comprising using a parameter to
 control said adjustment of the size of the sum and/or the contributions, by
 making the weighting factors~~size~~ dependent on a value of the control
 25 | parameter, the method comprising selecting the reference vectors and the
 value of the parameter, the method comprising
 - determining an initial set of reference vectors by means of an automatic
 training process using training bin values obtained without said
 adjustment;
 - 30 | - computing bin scores for individual ones of the respective bins of the
feature histogram, the bin score for each of the respective bins of the feature

histogram being computed from a scores for a matches between bin values of bins-vector components for the respective bin of in the set of reference vectors and a counts of STIPs in the respective bin from a set of video segments that have been designated to show the predetermined type of action;

- determining the value of the parameter based on the bin scores for the individual ones of the respective bins, the parameter being set to increase or decrease the bin value of the at least one of the respective bins relative to the bin values of other ones of the bins according to whether the bin score for the at least one of the bins respective bin is larger or smaller than the bin score for the other ones of the bins respectively;

- determining the reference vectors for use in the support vector machine by means of an automatic training process using bin values obtained with said adjustment controlled by the parameter value.

3. A method according to claim 2, wherein the parameter is a the weighting factor, the size of the sums of the respective bins and/or the contributions to the respective bins being adjusted by multiplying a count of the STIPs that have been assigned to the respective bins, and/or contributions to said sum of contributions count by from individual STIPs that have been assigned to the respective bins, by the weighting factor.

4. A method according to claim 1, wherein the weighting factor for each of the respective bin is selected dependent on a count value of STIPs in the respective bin based on a plurality of support vectors for the predetermined type of action, the weighting factor being increased with increasing count value when at least one of the support vectors for the predetermined type of action defines a component value for the respective bin that is higher than the count value of STIPs in the respective bin.

5. A method according to claim 2, 3 or 4, wherein the bin score for each individual one of the respective bins is computed as a sum of a bin dependent bias value and the score for the match between the bin value and the count of STIPs assigned to the respective bin, the method comprising

- determining an overall bias value for the predetermined type of action by means of said automatic training process using training bin values obtained without said adjustment;

- assigning respective fractions of the overall bias value, for use as the bin dependent bias value, to the respective ~~ones of the~~ bins, dependent on said counts for the respective ~~one of~~ bins.

6. A method according to claim 5, wherein the fractions of the overall bias value are assigned to each of the ~~respective ones of the~~ bins in proportion to an average of said counts for the respective ~~one of~~ bins, the average being taken over said set of video segments, or the fractions of the overall bias value are assigned to the ~~respective ones of the~~ bins in proportion to a value of said counts for the respective ~~one of~~ bins for one of said set of video segments.

7. A method according to claim 5, the method comprising
- computing values of the parameter for a plurality of different ways of assigning fractions of the bias value to each of the ~~respective ones of the~~ bins, including said assignment dependent on said counts for the respective ~~one of~~ bins,

- comparing reliability of the detection of the presence of actions of the predetermined type obtained by detection using respective ones of the parameter values;

- selecting one of the parameter values that yields a highest reliability;

- using the reference vectors by means of an automatic training process using bin values obtained with said adjustment with said adjustment controlled by selected one of the parameter values.

~~8. A method according to claim 1, wherein the bin dependent adjustment comprises multiplying the sum and/or the contributions of at least one of the bins by a weighing factor that depends on the at least one of the bins and the predetermined type of action.~~

98. A method according to claim 1, wherein the SVM match score is computed by computing mutual scores between the feature vector histogram

and each the reference vectors for the predetermined type of action, and summing products of the mutual scores with factors for the reference vectors

109. A method according to claim 98, wherein the mutual scores
 5 between the feature ~~vector~~-histogram and the reference vectors are computed for each of the reference vector from a sum over all of the bins of the smallest of the bin value of the feature ~~vector~~-histogram for the bin and the bin value the reference vector for the bin.

110. A method according to claim 1, comprising determining whether a
 10 result of said summing the products exceeds a threshold value.

1211. A method of automated detection of actions from a plurality of
 predetermined types of action, comprising executing the method of claim 1 for each of the plurality of predetermined types of action respectively, using, for each respective different predetermined type of action, a respective
 15 different bin dependent adjustment of a size of the sum and/or the contributions of ~~at least one of the~~ the respective bins dependent on the predetermined type of action and respective different sets of reference vectors.

1312. A computer readable medium, comprising a program of
 20 instructions for a programmable computer system that, when executed by the programmable computer system, will cause the programmable computer system to execute the method according to any one of the preceding claims.

1413. A video action detection system for detecting presence of an action
 of a predetermined type of action from video data, the system comprising
 25 - a source of video data or video data input;
 - a spatiotemporal interest point (STIP) detector coupled to the source of video data or video data input;
 - a feature data extractor coupled to the source of video data or video data input and the STIP detector, configured to extract feature data from video
 30 data from the source of video data or video data input at STIPs detected by the STIP detector;

- a descriptor quantizer coupled to the descriptor extractor and configured to assign the extracted descriptor values to bins of a feature ~~vector~~histogram;

- a computing circuit configured to compute bin values of the feature ~~histogramvector~~ for each ~~respective of the~~ bins, the bin value of ~~each the~~ bin comprising a sum of contributions associated with STIPs that have been assigned to the bin, with a bin dependent adjustment of a size of the sum and/or the contributions of ~~at least one of the~~respective bins dependent on the predetermined type of action; and to compute a support vector machine (SVM) match score from the feature ~~histogramvector~~ and reference vectors for the predetermined type of action.

~~15~~14. A video action detection system according to claim 14, comprising a training module, the training module being configured to

- determine a set of reference vectors by means of an automatic training process using training bin values of the feature histogram obtained without said adjustment;

- compute bin scores for individual ones of the respective bins from scores for matches between ~~bin-component~~ values ~~of for the respective~~ bins of the further set of reference vectors and counts, in the respective bins, of spatiotemporal points of interest from a set of video segments that have been designated to show the predetermined type of action;

- determine the value of the parameter based on the bin scores for ~~the each~~ individual ones of the respective bins, the parameter being set to increase or decrease the bin value of the ~~at least one of the~~respective bins relative to the bin values of other ones of the bins according to whether the bin score for the ~~at least one of the~~respective bins is larger or smaller than the bin score for the other ones of the bins respectively;

- determine the reference vectors by means of an automatic training process using bin values obtained with said adjustment controlled by the parameter value.

15. A method according to claim 1, comprising using a parameter to control said adjustment of the size of the sum and/or the contributions, by

making the weighting factor dependent on a value of the control parameter, the method comprising selecting the reference vectors and the value of the parameter, the method comprising:

- determining an initial set of reference vectors by means of an automatic
training process using training bin values obtained without said
adjustment;

- computing bin scores for individual ones of the respective bins of the
feature histogram, the bin score for each of the respective bins (i) of the
feature histogram being computed from values $v = c + \text{Sum } \alpha(j) * \min$
(p_i, z_{ji}) wherein c is a bin independent bias value, z_{ji} are the vector
components for the respective bin (i) in vectors (j) from the set of reference
vectors, the sum is taken over the set of reference vectors,- and p_i is a count
of STIPs in the respective bin (i) obtained from a set of video segments that
have been designated to show the predetermined type of action;

- weighting factors for the respective bins (i) based on the bin scores for the
individual ones of the respective bins (i);

- determining the reference vectors for use in the support vector machine by
means of an automatic training process using bin values obtained with said
adjustment controlled by the parameter value.

16. A method according to claim 15, the method comprising

- determining a bias value c for the predetermined type of action by means
of said automatic training process using training bin values obtained
without said adjustment;

- assigning respective fractions of the bias value, for use as the bin
dependent bias value, to the respective bins, dependent on said counts for
the respective bins.

17 A method according to claim 16, wherein the fractions of the
overall bias value are assigned to each of the respective bins in proportion to
an average of said counts for the respective bin, the average being taken
over said set of video segments, or the fractions of the overall bias value are

assigned to the respective bin in proportion to a value of said counts for the respective bin for one of said set of video segments.

18. A method according to any one of the preceding claims, wherein the bin dependent adjustment comprises multiplying the contributions to the respective bins by a weighing factor that depends on the respective bin and the predetermined type of action and position information of the detected STIP in the video data.

19. A method according to claim 18, wherein at least part of the respective bins is associated with spatiotemporal coordinates (x_i, y_i, t_i) for the respective bin, the step of computing the bin value comprising computing a function of the difference between the detected spatiotemporal coordinates of the detected STIP feature and the spatiotemporal coordinates (x, y, t) associated with the respective bin to which the feature is assigned.