ISSN: 1231-4005
e-ISSN: 2354-0133
DOI: 10.5604/01.3001.0010.3152

PROJECT ZEUS: VIDEO BASED BEHAVIOURAL MODELLING OF NON-LINEAR TRANSPORTATION SYSTEM FOR IMPROVED PLANNING &URBAN CONSTRUCTION PROJECTS

Jagadeesh Kannan R., Ankush Rai

VIT University
School of Computing Science & Engineering, India
e-mail: jagadeeshkannan.r@vit.ac.in; ankushressci@gmail.com

Janusz Szpytko, Yashesh C. Pandya

AGH University of Science and Technology, Poland
e-mail: szpytko@agh.edu.p, ypandya@student.agh.edu.pl

Abstract

The ability to analyse the traffic and urban mobility pattern with the help of video analytics, which occur in massive volumes of surveillance video will lead us to provide a knowledge based for the urban planners and policy makers to come up with better construction planning. This will soothe the needs of urban commuters and thereby saving unnecessary spillage of money on the construction projects. In this research project, we present an artificial intelligence framework for transport video analytic; which autonomously models behavioural patterns of commuters and flow of traffic, wherein it taxonomically classifies essential patterns based on geometrical feature points of interest to facilitate reality mining. This behavioural pattern of commuter and traffic flow can later be queried and fetched through the newly mathematically programmed ontological data warehousing module, where such reality mined contextual data could be used for sharing essential data.

Keywords: *Algorithms, computer vision; traffic activity recognition; event detection; activity analysis; behaviour modelling*

1. Introduction

Human behaviour recognition is a dynamic point in the field of computer vision. This is because of the quickly expanding measure of video records and the huge number of potential applications taking into account programmed video examination, for example, visual observation, human-machine interfaces, sports video investigation, and video recovery. Among these applications, a standout amongst the most fascinating is human behaviour recognition particularly abnormal state behaviour recognition. A behaviour is a succession of human body developments, and may include a few body parts simultaneously. From the perspective of computer vision, the recognition of behaviour is to coordinate the perception (e.g. video) with beforehand characterized patterns and after that relegate it a label, i.e. behaviour type. Contingent upon multifaceted nature, human activities can be arranged into four levels: gestures, actions, interactions and group activities [1], and much research takes after a base up development of human movement recognition. Significant segments of such frameworks incorporate feature extraction, behaviour learning, classification, behaviour recognition and segmentation [2, 3]. A straightforward procedure comprises of three stages, in particular discovery of human and/or its body parts, following, and after that recognition utilizing the following results. Case in point, to perceive "shaking hands" activities, two man's arms and hands are initially recognized and followed to produce a spatial-temporal description of their development [4, 5]. This description is contrasted and existing examples in the training data to decide the behaviour sort.

2. Methodology

2.1. Parameter free hierarchical approach For Behaviour Recognition

Attempt to perceive intriguing occasions (abnormal state activities) in view of more straightforward or low-level sub-activities. As it were an abnormal state behaviour can be deteriorated into a succession of a few sub-activities, for example, “hand shaking” might be integrated as an arrangement of two hands being expanded, converging into one item, and two hands being pulled back. Sub-activities can be further considered as abnormal state activities until deteriorated into nuclear ones.

The upside of hierarchical methodologies is the ability to display the perplexing structure of human activities and its adaptability for either individual activities, interaction in the middle of humans and/or protests or group activities. In addition, hierarchical models give a natural and helpful interface for incorporating former information and understanding of structure of activities. Hierarchical ways to deal with some degree have a cosy relationship with single layer methodologies. For instance non-hierarchical single layer methodologies can be effortlessly used for low-level or nuclear behaviour recognition, for example, motion location. Some non-hierarchical single layer methodologies can likewise be stretched out to hierarchical models, for example, expanded multi-layered HMMs. Utilizing the scientific classification, hierarchical methodologies are ordered into three groups: measurable methodologies, syntactic methodologies, and description-based methodologies. Here, the contours of the given image of an object are derived using the following algorithm.

Algorithm: Clustered Associative Network for Nerve Feature Mapping (CANNFM)

- Input: $m \times n$ RGB standard Image,

where, m & n are the row & column of the given image.

- Output: *CC,* clustered Associative Networked Pixels of retinopathy images / for breaking down pixels into normalized illumination reflectance field.

Loop: *for i: m,*
Loop: *for j: n,*

$$r_{i,j} = \frac{R}{(R+G+B)}, g_{i,j} = \frac{G}{(R+G+B)}.$$

End /for masking the pixels representing nerve region:

$$\text{If } (\gamma_M \leq \textstyle\sum_M),$$

$$\{P_{nerve} = P(nerve|rg, N \quad \text{Else } P_{background} = (background|rg, \gamma_M)\},$$

where M is the model of nerve Colour also embarked as low intensity pixels, $\gamma_M \& \sum_M$ are mean & covariance of the pixel distribution based on intensities in *rg* Colour scheme after pre-processing.

Now, (x_1, y_1), (x_2, y_2), …, $(x_i, y_j) \in P_{nerve}$, while $l \leq \min\ (\sum_{i=1}^{N} D(v_i, n_i))$ //D is the displacement vector / evaluates the feasible value of target nodes (pixels)T_{target}:

$$T_{target} = \textstyle\sum_{i=1}^{N} D(x_i, y_i).$$

Create target vectors for feasible neighbouring nodes:

Loop: *for* 1 to n_i:

$$\textstyle\sum_{i=1}^{N} T(x_i, y_i) \leq R_{target}.$$

To create an associative logical network of P_{nerve},

$$CC(i,j) = \textstyle\sum_{i=1}^{m} \sum_{j=1}^{n} sgn(D(x_i, y_i) + \lambda T(x_i, y_i)),$$

where λ is the Lagrange multiplier.

End for loop / End while loop / End Process.

Thus, the nodes emerged out from the above CANNFM algorithm represents the positioning of the pixels involved in the feature sets. The variable R_{target} is determined through the following sets of equations, as follows:

$$f = arg \sum_{i=1}^{m} \sum_{j=1}^{n} \delta_{i,j}^{C} . \sigma \left\| \mathrm{a_i} - \mathrm{b_j} \right\|_{n}^{m},$$

where, $\delta_{i,j}^{C}$ degree of membership of $\mathrm{a_i}$ in cluster set$\mathrm{b_j}$; $\mathrm{a_i}$ is the multidimensional data block; $\mathrm{b_j}$ is the center of the cluster of the cluster of multi dimensional data blocks for the congregated pixel position in respect to the nerve Colour model M; σ is a scalar weighted by the regularization term derivable by the disparity of homogeneity within the cluster C for the timed interval of pixels m and n. Therefore, the threshold value of neighbouring pixel nodes in the spatial locality of the nerve colour model M is derivable from the pixels sets corresponding to that of nerve only.

$$R_{target}(i,j) = \sum_{m=1}^{n} \frac{f_{i,j+1}^{2}}{f_{I+1,j}^{2}}.$$

This fulfils the requirement to develop an agent to get feature sequences with its neighbouring pixel module. Forming a logical chain of module, this is susceptive enough to accommodate pervasive operations. Now, in order to classify the logically encoded micro-expression from the features of the facial image during the self-deterministic sampling time we employ the scheme discussed in next section, using the UCPC algorithm. Here, the algorithm is sued for mining the feature for ranging associative parametric evaluation at every instant in the due process. This is where our proposed algorithm comes in to play which remodel the Cascaded Neural Network (CNN) for increasing its effectiveness against multivariable dependencies of time dependent sampling transactions with several attributes to data mine in a logical sequencing. The modelled algorithm using a hybrid of cascaded neural network is formalized below:

Algorithm: Unsupervised Cascaded Profile Classifier (UCPC)

- Input: List of t transactions made & $CC(i,j)$ i.e., template classes of the trained vectors for the t-1 nerve model,
- Output: $CC'(i,j)$, final state of the feature vectors (1 or 0 for validation).

For 1: $\mathrm{b_{jq}}$//for each instances / run the conventional water-filling algorithm for set CC (i, j), get water level W:

$$W = P(CC(i,j) = P_{t+1} | P_{t+n} = S_i), 1 \leq i < N, N \leq j \leq 1,$$

where P is the probability of states and S the number of states S= $\{\mathrm{S_1}, \mathrm{S_2}, \ldots \mathrm{S_n}\}$ & N is the number of water channels.

for t = 1 to X:

$$H_n = \tanh(w_{HX}X_N + w_{HH}H_{N-1} + B_N),$$

$$k_i = w_{HH}.X_N + B_N + \sum_{i+1}^{N} \tanh(\delta_H \ .(1 - \mathrm{y_{N-1}}).$$

Move sub channels to temporal set *CC'' (i, j) =* $\{CC \Big| \sum_{1}^{N} H_N{}^{-1} \leq w_{HH} - \frac{\Delta W}{N}\}$ //for each hidden layer.

While *j<X:*

$$k_j = w_{HO}.X_N + \textstyle\sum_{i+1}^{N} \tanh(\delta_O \ . \ \mathrm{y_{N-1}}), \quad O_N = w_{OH}H_N + B_O.$$

Update CC' (i, j) (attribute classes) with the channel *N &* Δw_{XX} as weighted sub channels to buffer the sate sequences:

$$\Delta w_{XX} = \eta.\, \delta_O.\, k_i.\, O_N + \frac{\Delta W}{N},$$

$$CC''(i,j) = \sum_{i=1}^{m}\sum_{j=1}^{n} sgn\left(\Delta w_{XX}\,.\; O(k_i) + \Delta W\,.\; Y(t', k_j)\right) + \sum_{i,j}^{m,n} P(i,j)\; // \; attribute\ classes,$$

where, m, n belongs to runtime feature associations made in real time.

End loop / End while loop.

$$If\, P(i,j) = \frac{\|CC''(i,j) - CC(i,j)\|}{|W|} < N,$$

$$CC'(i,j) = \prod_{i=1}^{N} \Delta w_{XX}\, t_{ft}(R_i, L_i) \;,$$

else:

$$min\left\{\int_{\Omega} \sum_{i,j}^{m,n} P(i,j)\, |\nabla\delta_O| + \lambda|\delta(w_{XX}(x))| dx\right\}.$$

End //if.

End loop.

Given a sequence of input vectors $(X_1, X_2, X_3, ..., X_n)$ from CC (i, j), a sequence of hidden states $(H_1, H_2, H_3, ..., H_n)$, and a sequence of outputs $(O_1, O_2, O_3, ..., O_n)$ are generated in due process. Notations in above sets of equations are namely, w_{HX} is the input-to-hidden weight matrix, w_{HH} is the hidden-to-hidden (or recurrent) weight matrix, w_{OH} is the hidden-to-output weight matrix, and the vectors B_N and B_O are the biases. The expression replaces the inputs received form feedback loops with a special initial bias vector checked for nonlinearity while ensuring that the training is done coordinate wise. η is the learning rate, t' is the time of the next frame and k_i is the local induced field of activation potential for the i*th* neuron, k_j is the co-activation neuron field for the next sequence of activation units, δ_H & δ_O are the pointer variable for the field & sub-field trace of an aneurism, respectively.

The experimental setup for the application and using the data from the previous algorithm is used for Parameter free Geometric Modelling and Ellipse Fitting as. Whereupon, the algorithm is executed to run over the semantic contour network in order to learn the data slip and latter recover the missed data through the proposed PFGMEF algorithm. Unlike other algorithm, we take a different route, and focuses on creating an unsteady checkpoint based system, which reacts to the sequences in a constructive and destructive manner. While the past studies are focused on discarding instances to keep the bounded state of the support set, instead we employ a sequence based fusion techniques that reacts to the pattern in a way which can let it build itself the past instances by keeping the biasing and computing through the sequences to dynamically create a support set and thereafter adjoining the essential part of it which latter shall help develop the same checkpoints by the application of the similar sequence. The variation in hypothesis is encoded & decode through a sequence, which is based on trading equilibrium of sparseness. The algorithm described below as follows: Our work is based on the idea of the finding the topology of the features and fitting in the ellipse & then coding it in logical sequence with the cascaded neural network (Fig. 1).

We use the following algorithm for the breakdown of pixels in the described logical framework as follows:

Algorithm: Parameter free Geometric Modelling and Ellipse Fitting (PFGMEF)

- Input: s_i, s_j are the time instances of the geometrical pairs, S_t is the support set, H_t is the geometrical fitting hypothesis at time t & p_t is the pair equilibrium sequence,

- Output: H_o is the final geometrical fitting hypothesis & updated p'_t.
- Step 1: For t = 1, 2, 3 …;
- Step 2: Receive New Instance S_t;
- Step 3: Attach time instances of the pairs:

$$p_t = \frac{[s_i * s_j * q]}{[s_i * s_j]} \; paired\ equilibrium\ sequence, \quad S_t = \sum_i \; p_t\,(t - t_i) \;\; //Binding\ Process;$$

- Step 4: Evaluate risk of pairing & remaining time instance of memory bound:

$$while\ k_i \neq k_n,$$

$$\{R[k_i] = \frac{1}{p_t} \sum_{l=1}^{t} (S_t - S_{t-1}) \int_t \; k_i\, dt,$$

$$M_t = \sum_k k\,(t - t_i), \;\; where, k_i = \frac{(p_t\,\Delta t)^k}{k!} \exp(-p_t \,.\, \Delta t);$$

- Step 5: Update the hypothesis with chain sequence by checking for risks involved:

$$If\, R[k_i] \leq R[S_t],$$

$$\{\sum_{k=l}^{\infty} \sum_{k_i}^{m} \frac{(p_t)^{k-k_i}}{(M_t - t_i)!} . \exp(-p_t \,.\, \Delta t) . \frac{(p_t)^{k_i}}{(k_i)!}\},$$

else:

{ *Print* "Failed to Update",

$$k_i = Null\ //Assign,$$

$$k_i = k_{i+1}\};$$

- Step 6: Add instances to the support set:

$$If\, p_t t_i \leq k_i,$$

$$\{\sum_{k=l}^{\infty} \frac{(p_t t_i + p_{t+1} t_{i+1})^k}{k!} \exp(-p_t \,.\, \Delta t)\},$$

else:

$$\{S_t = S_{t-1}\},$$

} //end while;

- Step 7: End process.

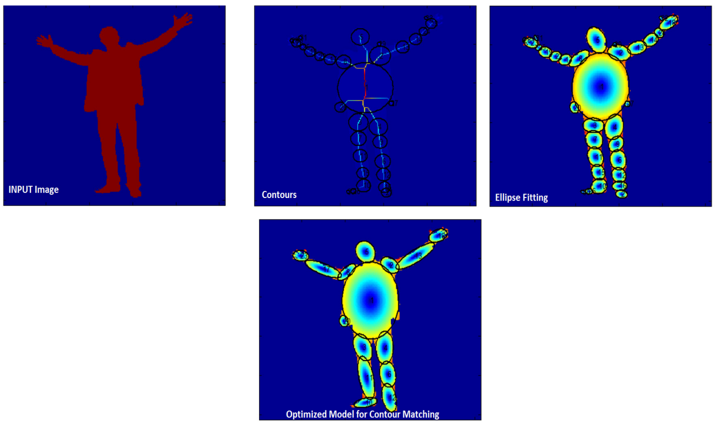

Fig. 1. Output of the proposed algorithm for parameter free behaviour recognition

Here, for matching the hierarchical geometry of the parts; let us suppose that for the same geographical area $I_b^a \& I_{b+1}^a$ are the two distinct geometrically fitted images (template and test image) of the same object with consecutive frames; where a is the multiple ellipse number and b is the frame number. The size of this images can be given as m x n. The images are pre-processed with the employment of ABLATA (Attribute Based Level Adaptive Thresholding Algorithm)thresholding algorithm; converting the images into binary images and extracting the object region and its multiscale object region extracted using Attribute based Level Adaptive Thresholding Algorithm (ABLATA) algorithm [5, 6].

It is required to reduce the noises of the images with the additional requirement of image normalization, which is achieved, with the deployment of DPCA algorithm (Decomposable Pixel Component Analysis Algorithm) [7, 8]. Since, the values derived from standard deviation or STD issued to trace the changes with respect to the time in the consequent frame of multiple view SAR images which is given by:

$$\delta_i(x,y) = \frac{\sum(I_{xy}^i(t) - \overline{I_{xy}^i(t)})^2}{T},$$

where, $I_{xy}^i(t)$ is the consequent frame with the location of each pixel value in the form of (x, y) for the frame at time t, $I_{xy}^i(t)$ averaged over information of all $I_{xy}^i(t)$ value for time t. Hence, STD for the frame comprising of object (δ_i^0) and its multiscale object region (δ_i^s) can be mapped as:

$$\delta_i^0(x,y) = \begin{pmatrix} \delta_i^s(1,1) & .. & .\delta_i^s(1,y) \\ \vdots & \ddots & \vdots \\ \delta_i^s(x,1) & .. & \delta_i^s(x,y) \end{pmatrix} \& \ \delta_i^s(x,y) = \begin{pmatrix} \delta_i^s(1,1) & .. & .\delta_i^s(1,y) \\ \vdots & \ddots & \vdots \\ \delta_i^s(x,1) & .. & \delta_i^s(x,y) \end{pmatrix}.$$

Hence, the above two matrix can be utilized as the feature matrix across frames suppressing the cluttered or static parts and reading the moving parts corresponding to each pixel. This is used to extract the path of the targeted and its path of its multiscale object region. Thus, the quantitative information about its trajectory in continuous frame sequence can be derived from:

$$s_i = \sum(I_{xy}(t) - I_{xy}(t-1))^2.$$

Now, we need to derive the symmetry of the object and its multiscale object region to optimize the classification process. The symmetry equation of the process of two symmetrical parts a and b and angle θ for orientation is given as:

$$S_s(s_i, \delta_i, \theta) = (f_a(s_i, \delta_i, \theta), f_b(s_i, \delta_i, \theta))^T.$$

This symmetry breakdown allows us to ease the identification problem of the object by looking for the inter-correlation between symmetry of the object and the symmetry of its multiscale object region. Hence, the relationship between it can be learned by using the one shot learning algorithm for object classification, which is given as:

Algorithm: One Shot Learning Algorithm for Object Classification

k←O(k) // Number of Sample Object in the frame.

For s=1…k do //Mapping Multiscale object regions of objects in the frame:

$$n_s \leftarrow P(\delta_i^s | k, S_s).$$

For j=1…n_s do // Mapping objects in the frame:

$$n_o \leftarrow P(\delta_i^o | S_s).$$

End for

Compute symmetry for object and its multiscale object region:

$$S_o(s_i, \delta_i, \theta) = (f_a(s_i, \delta_i, \theta), f_b(s_i, \delta_i, \theta))^T,$$

$$S_s(s_i,\delta_i,\theta) = (f_a(s_i,\delta_i,\theta), f_b(s_i,\delta_i,\theta))^T.$$

If cross relation STD between symmetric parts:

$$\delta_{ab}(t) \approx \frac{cov\,(I_a(t), I_b(t))}{\sigma_{xy_a(t)} \cdot \sigma_{xy_b(t)}} < S_o.$$

Update:

$H_i \leftarrow P(H_H|n_s, n_o)$ //updating classification history of cross relation.

Else exit.

End for.

The main advantage posed by the above algorithm is the setup it put forth for scheduling training and processing based on memory bound to the support set on timely basis and thus forming a decomposable or expanding sequence when the other instance pairs are added in relational to the previous trained hypothesis, such that the trained hypothesis is always bounded and deducible from the other previous pairs of instances. The training is achievable in small number roof instances with high accuracy. Here, the Fig. 1 represents the onset of the data slip rate due to the bounded memory units to update the online hypothesis. However, by the end of the recovered image the data slip rate have not shown any fluctuation in loss of data, thereby converging the results at the saturation level with respect to the time. The advantage that our algorithm put forth is its adaptability in re-organizing the scale invariant geometrically hypothesis so generated during the process run. This lead to a curvelet transformation for the process to end at the point where the front data and the half pattern length saturates. The process is cyclic in nature and does not require unnecessary updation in due process.

2.2. Self-modelling of behavioural patterns based on population coding of neural network

The first step is to define the synchronous machine model of distributed systems with its weighted sensed data x_{ij} and control action sets y_{ki}. Here, we are using membrane computing based model of neural system to define the correlation between the sensed data and the control action such that the final sets derivable would be optimized and weighted to achieve optimization for high dimensional decision space, then in the next step it shall be forwarded to semantically filter out optimal policy (i.e., correlated state action pair) with higher reward through the help of distributed reinforcement learning.

Now, for the first step let us suppose that we have a total of Excitatory neurons (E_N) and Inhibitory neurons (I_N) where they are in the ratio of $I_N = 0.2 \times E_N$. Now, that we need to find an evolutionary Hodgkin-Huxley equation for self-managing neurons. Therefore, to model the phenomenon of building the learning model for biological neurons we require merging the properties of artificial neural network with the biological neurons. Where, the sequence of inputs of the firing neurons is affects the other subsequent sequence and consequently synapse formation before giving a unitary idea of a stimuli. Thus, W_{ij} be the weightage for the connection strength from neuron i to neuron j, similarly W^{IE}, W^{EE} and W^{EI} represents weightage for inhibitory to excitatory connections, excitatory to excitatory and excitatory to inhibitory connections respectively. The W^{EE} and W^{EI} are initialized as sparse random matrices with the range of connection probabilities between the value of 0.1 and 0.2. Initially, the W_{IE} connections are meant to freeze at their random initial values which are depicted from uniform distribution, latter followed by normalization [9, 10]. Altogether, the sum of connections entering a neuron is in a sequence of 1 & 0; thereby the binary vectors is given by x (E_N) ∈ {0, 1} and y (I_N) ∈ {0, 1} for the excitatory and inhibitory neural activity at time step t, respectively. Hence, the sequencization of the network states at time step t + 1 is equivalent to:

$$x_{ij}(t+1) = \Theta\left(\sum_{j=1}^{E_N} W_{ij}^{EE}(t)x_{ji}(t) - \sum_{k=1}^{I_N} W_{ik}^{EI}(t)y_{ki}(t) - T_{ij}^{E}(t) + \xi E_i(t)\right),$$

$$y_{ki}(t+1) = \Theta\left(\sum_{j=1}^{N^E} W_{ij}^{IE} x_j(t) - T_i^I(t) + \xi I_i(t)\right).$$

As the network, equation continues to evolve x_{ij} & y_{ki} , the synaptic weights are given as:

$$\Delta W_{ij}^{EE}(t) = \eta\left(x_{ij}(t)x_{ji}(t-1) - x_{ij}(t-1)x_{ji}(t)\right),$$

$$\Delta W_{ij}^{EI}(t) = -(1-\eta)y_{ji}(t-1)\left(1 - x_{ji}(t)\left(1 + \frac{1}{\mu_{ji}}\right)\right).$$

Θ is the Heaviside step function; T^E and T^I are the threshold values for excitatory and inhibitory neurons, where it is initially drawn from the uniform distribution within the interval [0, T_{max}^E] and [0, T_{max}^I]. $\xi E_i(t)$ and $\xi I_i(t)$are white Gaussian noise processes with $\mu_\xi \in$ [0.01, 0.05]. Here one time step corresponds roughly to the duration of window of the spike time dependent plasticity and η is the learning rate. Now, that the threshold value of the excitatory neurons in response for a sequence of activated neurons is made pass through the previously generated targeted sequence code blocks of firing neuron states $S_{ij}^{code\ block}$, which is determined by the adaptation rate η_{AD} as:

$$T_{ij}^E(t+1) = T_{ij}^E(t) + \eta_{AD}\left(x_{ji}(t) - S_{ij}^{code\ block}\right).$$

The inhibitory spike-timing dependent plasticity ji rule regulates the weights backward from inhibitory to excitatory neurons, which stabilize the amount of excitatory and inhibitory to drive sensory information through the excitatory control neurons. Therefore, the evolutionary dynamics of the neuronal membrane potential that mediates the excitatory & inhibitory sequences through the network of membranes is governed by the above-deduced equation.

2.3. Modelling of scalable ontological database and parallel querying method

A content element can be represented as r items for which there is a sequence of m number of fish schools given by $X \in \{X_{x_0}{}^{(1)}, X_{x_0}{}^{(2)}, \dots, X_{x_0}{}^{(n)}\}$. Here X is the set of possible values of a fish school. A fish school could be composed of a short sequence of related joint labels, or a cluster of it. Content fish schools may overlap spatially, temporally, or both. Here, overlapping time windows is 2 sec long and starts every 185 ms; with overlap of 15/16 are used as fish schools. Let us suppose that C, B, and Y be matrix of filtered output, Y be the matrix of filters for stimulant variable and response variable for each X, such that C = XBY. Then C is a super frame of B. The length of a fish school S is equivalent to the total number of fish schools in it and is denoted by |S|.

Now:

$$X_{x_0} = \begin{bmatrix} X_{x_0}{}^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_{x_0}{}^{(r)} \end{bmatrix}, \quad X_{x_0}{}^{(j)} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, j = 1, 2, \dots r\ ,$$

where X_{x_0} is the covariance map from X, which asserts to the association, formed between the fish schools $S(t)$with that of stimulant and response variable for the query, X_i is the position of input record. C_l is the cluster value, which contains various values from 1 to l.

Now, at each step we calculate:

$$S(t) = \sum_{i=-\infty}^{+\infty} \sum_{k=1}^{N_{SC}} C_{li} S_l(t - iF_s)\,, \qquad S_l(t) = \prod(t) e^{j2\pi f_l t}\,, \qquad \prod(t) = \begin{cases} 1, 0 < t \le T_s \\ 0, t \le 0, t > T_s \end{cases},$$

where C_{li} is the i^{th} vertex label at the 1^{th} sub node sequence (when output of one iteration is propagated to the input of the other), T_s is the time of query execution, S_l is the position indices of sub nodes for the l^{th} query labels, N_{sc} is the number of labels (number of matching iterations), f_l is the frequency of the labels found in association and π(t) is the pulse shaping function. Following this process to complete the dataset in all records. Thus, the dynamics of the equation for a computational querying job from the above mentioned ontological tree model working on cost flow could be computed as shown below:

At time t, when transition index of each fish schools to neighbouring school p_t is requested:

Then:

Fetch x *(p_t, r (p_t, t))* $\leftarrow [forwardshare(i,j), backwarshare(i,j)]$.

it will increase in times t' > t.

where:

$$forwardshare(i,j) = \sum_{\substack{k\le i\\ l>j}} \left\langle \begin{matrix} forward\ weight\ of\ reference \\ from\ process\ task\ k\ to\ memeory\ instance\ l \end{matrix} \right\rangle,$$

$$backwarshare(i,j) = \sum_{\substack{k> i\\ l\le j}} \left\langle \begin{matrix} backward\ weight\ of\ reference \\ from\ prcoess\ task\ k\ to\ memeory\ instance\ l \end{matrix} \right\rangle,$$

$$z(i) = number\ of\ memeory\ insertion\ points\ among\ F_1, F_2, \dots F_j, and$$

$$z(j) = index\ of\ last\ bottom\ inserting\ points\ among\ F_1, F_2, \dots F_j.$$

Compute weighted relationship between shared vertices by fish schools – Boolean sequence function to determine the feasible sequence of to form for a tree branch F in 2D relationship can be given as:

$$C_{i+1} = \left(\int_I |F_i(x-1, y-1) - F_i(x+1, y+1)|\, dx \right)^{1/2}.$$

Check if $C_{i+1} < C_{Max}$ // check for fitness variation / based on weighed ordering in history indices of tree structure the collective volatile component of the movement is given by:

$$H(i,p) \leftarrow \left(\sum_{t=t(i,j)+1}^{t(i,j+1)-1} C_j(t) \right) - z(i,j) - w(i).$$

Here, *x(i, p)* be an indicator to the event that the solution is in state *i* during the p_{th} phase of memory instance and n_i be the number of phases of state *i*. Thus, forming a dynamic sequence. The converging fish school network is topologically complex in several ways. The nodal degree distribution is fat-tailed with high-degree hub nodes to be located using sequence of information to excite the necessary displacement of fish school and asses the information in an associative form. This enables the machine not only to learn but it enables it to embark the cross relationship between various data for prediction of indices based on logical querying. Here, the cost flow constraints to maintain equality and balance in the ontological tree at the converged edges can be expressed as equilibrium cost constraint given by:

$$\min_{\mathrm{P}} \left[\mu(C_I) + (1-\mu)\left(\sum_{i\in C_n}^{N} C_i + \sum_{i\in C_n}^{N} C_i \right) \right].$$

3. Results & Discussion

HMMs can be considered as a straightforward instance of dynamic Bayesian networks. A HMM represents the condition of the world utilizing a solitary discrete random variable however DBN represents the condition of the world utilizing an arrangement of random variables. Multiple levels of hidden states shape a representation of hierarchical human activities. Past research endeavours on factual methodologies for the most part harp on utilizations of augmented HMMs and dynamic Bayesian networks: 2-layered hierarchical hidden Markov models (HMMs) and dynamic probabilistic networks (DPNs). The proposed methods add a measurable approach to deteriorate the body into a hierarchical structure. A hierarchical complex space is learnt to portray the motion designs. Geometric condition is tested using UCPC & CANNFM and is utilized to anticipate these motion designs. PFGMEF and one shot semantic extraction of matching points is utilized to order last human actions in view of the motion designs. This hierarchical representation of human behaviour as opposed to straightforward non-hierarchical pack of-words representation. In hierarchical K-implies tree is additionally used to represent the feature signs. The issue of inadequate integrating so as to train data is handled in with area information. To begin with, request rationale based area information is abused for dynamic Bayesian system learning, both the structure and the parameters. In comparison with the presented method, syntactic methodologies incorporate actions as a series of images. An image in this context is really the nuclear sub-activities said in the past area. Nuclear sub-activities can be perceived utilizing any of the past hierarchical or non-hierarchical systems. However, actions represented as a series of images results in a constraint for simultaneous behaviour recognition.

Tab. 1. Comparison of hierarchical approaches

APPROACH	CATEGORY	KTH	WZMN	OTHER
Yin'10	Statistical	82%		
Zeng'10	Statistical	92.1%	100%	
Han'10	Statistical			CMU:98.27%
Wang'11	Statistical	92.5%		HOHA:37.6% UCF:68.3%
Ijsselmuiden'10	Description-based			GroupActivities:74.4%
Morariu'09	Description-based			Basketball:72%
Proposed Framework	Hierarchy based Parameter free Geometry Modelling	98.5%	99.2%	CMU:98.19%

Fig. 2. The sample snapshot of the preliminary results of the work under development (a) For Mugging Scenario: The system identifies event, where the perpetrator is marked by the system in thick red bounding box whereas the victim in thick green bounding box and the snatching action in thin red bounding box. (b) Hit & Run Scenario: The system identifies event, where the perpetrator is marked by the system in thick red bounding box whereas the victim in thick green bounding box.

A part of this algorithm is in combination with description-based methodologies and varies from measurable and syntactic methodologies through a capacity unequivocally to express human activities' spatiotemporal structures. In this manner, such routines can perceive both sequential and simultaneous actions rather being constrained to sequential actions. Essentially, description-based methodologies model human activities as an event of implanted sub-activities (Fig. 2). Such events must fulfil determined temporal, spatial and legitimate relationships that are signatory of an abnormal state behaviour. Subsequent to the presentation of Allen's temporal interim predicates, they have been embraced for description-based human movement recognition for both sequential and simultaneous relationships. Context free grammars have additionally been used for description-based methodologies. Performance comparison of the presented method with that of past techniques is shown below in Tab. 1.

4. Conclusion

In this study, we give a survey and a new framework for automated human behaviour recognition. A substantial gathering of techniques is recognized. Among them, 50 particular and powerful recommendations of the most recent three years are accounted for. The examination utilizes the same scientific classification as a past survey taking into account whether the behaviour is perceived straightforwardly from the images or low-level sub-actions. Our objective was to cover the best in class improvements in every classification, together with the datasets utilized as a part of approval. The writing surveyed demonstrates that much research has been committed to recognition of events and human actions in transport settings specifically from the recordings or images in a solitary layered way. The experiment of presented approach on this practical datasets, for example, Hollywood films and YouTube recordings has averaged recognition rate of around 99.45%. The precision of other methods reported is low in the writing surveyed here. In view of the aftereffects of low-level actions, we trust more research will be done in the zone of abnormal state behaviour recognition in datasets and genuine scenes.

References

[1] Aggarwal, J., Ryoo, M., *Human activity analysis: A survey* - ACM Computing Surveys 43, 1-43, 2011.

[2] Aggarwal, M. S. R. J. K., *Semantic representation and recognition of continued and recursive human activities*, International Journal of Computer, Vision 82, 1-24, 2009.

[3] Blank, M., Gorelick, L., Shechtman, E., Irani, M., Basri, R., *Actions as space-time shapes*, IEEE International Conference on Computer Vision (ICCV), pp. 1395-1402, 2005.

[4] Blei, D., Lafferty, J., *Correlated Topic Models*, Advances in neural information processing systems, 18, 147, 2006.

[5] Blei, D. M., Ng, A. Y., Jordan, M. I., *Latent Dirichlet allocation*, Journal of Machine Learning Research, 3, 993-1022, 2003.

[6] Rai, A., *Artificial Intelligence for Emotion Recognition*, Journal of Artificial Intelligence Research & Advances, 1(2), 24-30, 2014.

[7] Rai, A., Sakkaravarthi Ramanathan, Kannan, R. J., *Quasi Opportunistic Supercomputing for Geospatial Socially Networked Mobile Devices*, Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), IEEE 25th International Conference, 2016.

[8] Rai, A., *Multispectral Image Denoising using Bi-Directional Recurrent Neural Network with DPCA Algorithm*, Journal of Image Processing & Pattern Recognition Progress, 2(1), 25-30, 2015.

[9] Rai, A., *Attribute Based Level Adaptive Thresholding Algorithm for Object Extraction*, Journal of Advancements in Robotics, 1(1), 29-33, 2015.

[10] Rai, A., *Artificial Intelligence for Emotion Recognition*, Journal of Artificial Intelligence Research & Advances, 1(2), 24-30, 2014.