IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Recovering Human Motion Tracking System using Gaussian Process

Prof. T.Venkat Narayana Rao[*1] , Raj Kiran T[2] , D.Saidulu[3]

[*1,2,3]Department of C.S.E, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Hyderabad, India

tvnrbobby@yahoo.com

Abstract

Estimation of motions and tracking the motions with the help of a motion tracking system has been considered as a major
dilemma due to the availability of various frameworks. In the process of tracking, we need to form a mapping from observation space to the state space. We have reviewed Gaussian Process with both discriminative and generative frameworks for the purpose of motion estimation. Both the input and output are shown at the same place. But there are certain limitations due to the dilemma of multimodality. To overcome these limitations we have used a mix of Gaussian process experts. In this paper, we have combined both these information's into a unified system, which acts as a fusion of experts. We have used Gaussian process, dynamic model for learning the movements in latent state space. We had not only estimated human motions but also tracked motions with respect to other objects. The area of study discussed in this paper would be useful to trace out physical and behavioral patterns of living and non living object and this would help in studying the forecast patterns such as natural blooming, industry tool design or earthquakes forecast , monitoring and in control applications along with their impact so as to offer corrective measures before hand.

Keywords: Motion Tracking, process dynamic model, spatial expert, temporal.

Introduction

Analysis of human motions is one of the liveliest areas of research. Motion tracking can be used in surveillance applications such as automatic processes. However, there has always been a problem with these real world applications due to fake inference and highly articulated objects from images. Due to the huge amount of studies, there have been many techniques for mapping, these techniques range from adjoining neighbor retrieval to probabilistic mixture of predictors. It has always been a problem to model small size data in high dimensional space. The spatial information cannot always handle the multimodality issue so the addition of temporal data helps in effectively handling the problem. Both the spatial local expert and the temporal experts are defined online in between small neighborhood for effective learning. The introduction of input sensitive gate function in the conditional mixture of Bayesian experts has helped in modeling visual to pose mapping. But Bayesian mixture expert's model may degrade the data because the performance depends on the data distribution in ambiguous regions. The generative process used for recovering unseen states is achieved with the help of analysis by fusion loop but the approach has been forbidden due to high computational cost to infer the distribution of hidden states. The discriminative methods have much faster test speed compared to that of the generative method. But most of these models are not efficient in dealing with high dimensional data. In our work we have used Gaussian Process with both discriminative and generative framework in the process of human pose/motion estimation along with some natural picture samples. Human motion estimation is build on the basis of Gaussian process regression within the discriminative framework, but the regression model has limitation of expensive computational cost and insufficient capability for handling the problem of multimodality [11] .The thin approximation of Gaussian process can reduce the computational cost since it use only a part/subset of training data. In order to overcome the limitations we have used an alternative approach of using a mixture of Gaussian process experts. The input is divided into different regions with the help of a gating network where a particular Gaussian process expert dominates each input. The computation cost is reduced to only a part of the data rather than the complete dataset [5]. Modeling the mixture of Gaussian process experts is always linked with the gating network where in training the gating network is considered as a vital problem. The blend of Gaussian process experts contain both temporal and spatial information this is very important because using only the spatial information we can't effectively handle multimodality and the human motion estimation itself is ill-posed. Introduction of temporal information is necessary and earlier discriminative models do not have the temporal evaluation except the parametric model, which has the temporal smoothness, constrains, which are added into the Bayesian mixture experts. The Gaussian process dynamic model is used to model the dynamics of human motion. But the model is designed for low dimensional latent space, so it is used to capture the motion priors in latent state space. Both the present method and Gaussian process dynamic model use temporal information but the working of these models is different [7]. The poster estimation comes under the generative method by optimizing the likelihood function. However proposed model is local GPDM is global.

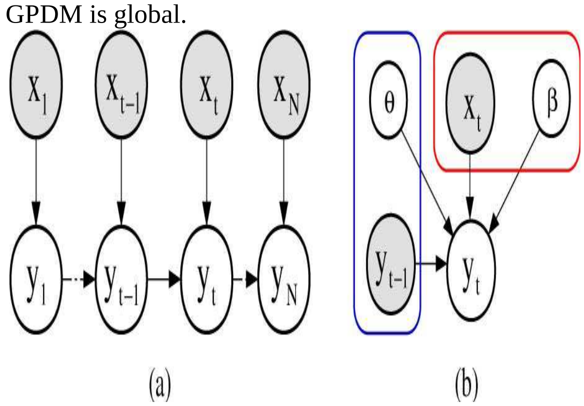

Fig 1. (a)Graphical model showing the temporal chain where x **represents input and y output. Shaded nodes designate the observation and un-shaded nodes indicate modeled variables. Fig 1.(b) Graphical Model of One time node in the temporal chain. The rectangle enclosed blocks in the left and right side represents temporal and spatial experts.**

Based on the methods on human pose inference by sparse GP regression the proposed system is a discriminative approach and it is a non parametric model which has an integration of temporal and spatial information. Previously Gaussian process experts are trained offline where as the local regressors are defined online for each sample [12]. The proposed system overcomes this problem since the process is conducted in the local neighborhood and this model helps in computing the gating network.

Review of Literature

Gaussian process and its variants have been used in the estimation of human poses and motions where in we use can use both discriminative framework and generative frameworks [8]. Gaussian process regression model served as a powerful method. A discriminate method helps in modeling the states directly by learning image to pose mapping. The process of estimation will be fast when the process of training is completed and also recover 3D human pose. Learning from visual observations to that of the articulated body movements and mapping in between them has laid the foundation for the estimation process. There has always been a problem with the pose estimation via monocular video because of various parameters [9], the use of Scaled Gaussian Process Latent Variable Models (SGPLVM) will be effective when small training data sets are available. The full Gaussian process is not that effective in handling the problem of multimodality [10].The sparse approximation of full Gaussian process helps in reducing the computational difficulties. For the approximation of covariance matrix we can use a part of training data or inducing variable [13].Therefore the inputs for every test and the training sample will be involved in the process of inference. There is not always possibility to avoid averaging effect using the effective mechanisms. Gaussian process is a powerful tool for regression but despite of its power and flexibility they also suffer from limitations. A mixture of Gaussian process experts the input space is separated by gating network wherein the specific Gaussian process expert dominates its relative input space. By the use of this mixture of experts the cubic computational cost is reduced to only a part of data. Whereas the mixture of experts are always combined with the determination of gating network [14].

The introduction of temporal information very important because spatial information has not yielded accurate results [16] .The previous methods do not contain temporal information except for the parametric model which has temporal smoothness constrains added to Bayesian mixture experts [13]. Tracking 3D movements from monocular videos has been inadequately constrained [2]. The Gaussian process dynamic model is used modeling the dynamics of human movements. The original GPDM is used to for low dimensional dormant space and can be used for capturing motion priors in latent state space [15]. When the motion priors are available in the state space then the complete process can be used as a generative model based on likelihood function.

Experimental Setup

In this section we have attempted to explain the experimental support with the help of figure 2. The video input signal is given and the initial frames are extracted and every two consecutive frames are compared to find out the changed pixel values and are stored and also the type of Gaussian process expert . Later motion identifiers are drawn based on the changed pixel values and the information provided by the Gaussian process experts. In this experiment the "project AForge" library files are utilized for handling the Image processing routines and filters. We have initially studied whether there is minimum motion in the input signal or not. If there are motions in the video provided then the amount of pixel values changed are identified. Then we create an initial background image and get the video dimensions. Later we lock the image filters for performing operations on the locked data. The calculations on the total amount of changed pixels is carried out. The Disposition of the old background image is created. Then we extract the red channel from the original image. Later we merge red channel with the moving object and replace red channel in the original image. Based on the type of mixture used we can utilize different kinds of edge filters. From the filter sequence, we shift background towards the current frame. The images are locked for representing the identifying marks. In the experiment, we have used graphic objects such as human motion and flowers blooming for study and estimation of motion patterning.

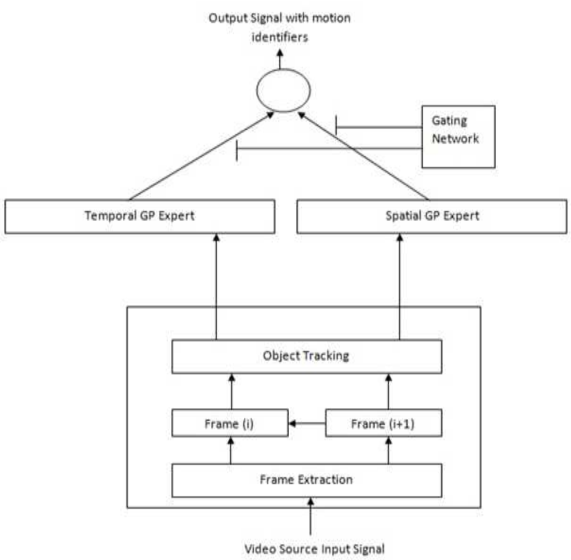

Figure 2. Experimental setup

Implementation

The Gaussian process experts are defined in the integrated input output space because of this the samples are composed in both input and output space, and this method is different from the preceding models where in the neighborhood is defined separately in input and output spaces. Because of this cause, there is a possibility of failure when dealing with many to one mapping because the relationship between the neighborhoods from the output space may change in the input space [6]. Therefore, the proposed model has the capability to deal two way mapping. In the Gaussian process experts model we get the temporal experts though the temporal chain. In the unified space, both the temporal experts and spatial experts are integrated into one, in order to make the predictions and to effectively handle the problem of multimodality [16]. In the proposed model, the temporal information is trained offline so if we lose the information we can switch back to the spatial information to complete the prediction. The proposed process of motion tracking involves five steps:

Step 1: Select a monocular video signal from an offline data or a online data source.

Step 2: Derive all the inner classifications and enable support for tracking.

Step 3: Perform extraction of frames from the monocular video if the data from an offline source but if data is taken from online source then utilize the available frames.

Step 4: Track all the variations in the pixels using mapping techniques.

Step 5: Reconstruct the frames with motion identifying marks.

Steps involved in implementation of MTS:

Step 1: Video Source Input signal
Step 2: Monocular Videos
Step 3: Video Classification and Enable Support
Step 4: Frame Extraction
Step 5: Object Tracking
Step 6: Frame reconstruction with motion identifiers
Step 7: Video Output signal

Observations & Findings

Initial experiments were implemented on human motions. Where in, we concentrated on the overall motion of the human body. From figure 3, we can clearly observe the motion identifying marks around the human body based on his body movements. We used monocular videos for tracking the motions. The below are the results of normal tracking.

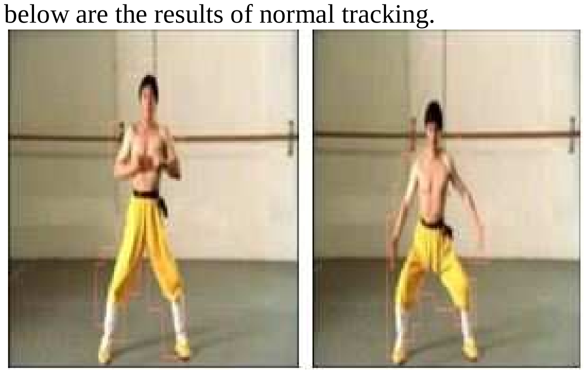

Figure 3: Showing the normal tracking of human motions based on the body movements

In the subsequent process, we have focused on tracking based on the temporal information in which different tracking results were obtained as shown in Figure 4. From the figure, it is evident how the object is tracked based on the temporal information by the means of differentiating frames.

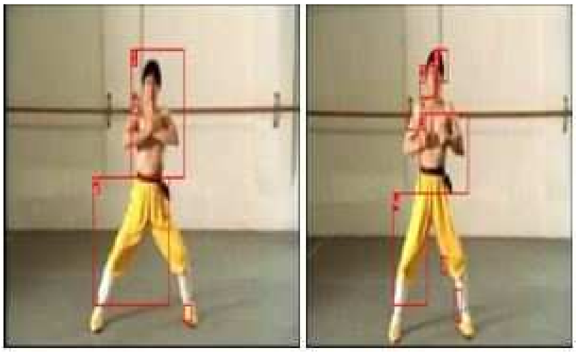

Figure4: Showing the tracking of human motions with temporal information

The further implementation is based on tracking with reference to the spatial information. As shown in figure 5, the outcome of the tracking is based on the spatial information is showcased. There is a chance of switching between each expert system as we indented to mix the experts processes.

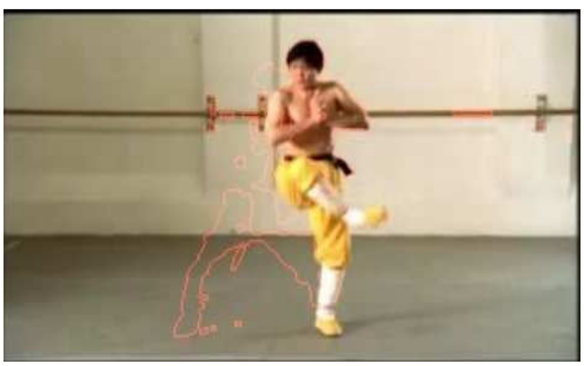

Figure 5: Showing the tracking of human motions with spatial information

From all the previously mentioned figures, we have tracked down the motions based on the temporal information and spatial information with a particular emphasis on human motions. In the consequent experiments, we have performed the same experts i.e., the temporal information and spatial information on other objects such as natural objects like flower blooming.

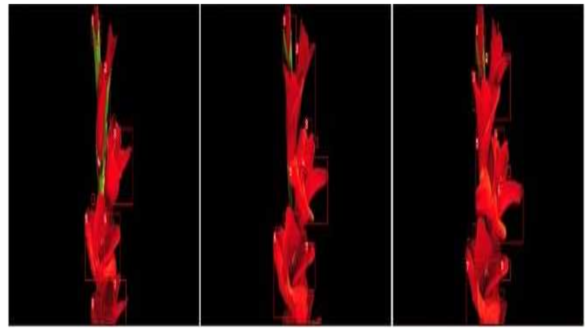

Figure 6: tracking based on temporal information

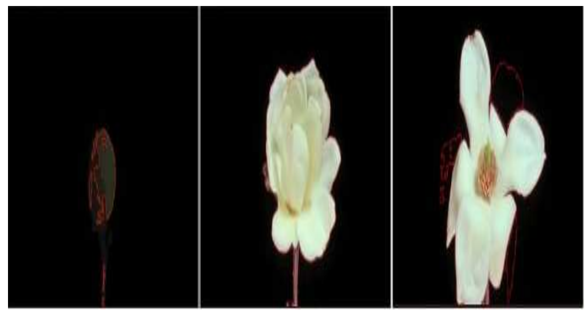

Figure 7: tracking based on spatial information

From the Figure 6 we have tracked the details based on the temporal information. In Figure 7 we used the spatial expert system for tracking down the motion and witnessed the overall tracking of the flower blooming.

The Applications and Merits

The analysis of motion is one the most active areas of interest for many research studies due to the number of revolutionary applications. Our system can be applied to the application areas such as surveillance, control etc. For instance, by using this system, we can monitor the places where huge numbers of people pass through daily in busy areas like airports or railway stations and computations can be proposed to calculate the crowd size in the direction flow and make analysis on congestion. If we consider the advanced applications based on this system, it is also possible to analyze the actions, activities and behavior of the crowd and individuals.

The security issues can also be handled in shopping malls and any confidential areas for detection of abnormal activities with the help of capture devices.

In control applications, this system can be modeled for estimating motions in order to control or investigate scheduled and unscheduled events. It may be an interface for a game software program such as EyeToy which is similar to a webcam but uses the computer vision and motion recognition in order to make the player interact with the game to be more precise and help in Human Computer Interactions. It can be used in the entertainment industry also where in the generation of graphic models make the products more attractive and natural. The system can be used for analysis-based applications for example for investigation and produce optimized results in the athlete's performances. It is also helpful in manufacturing process to attain accuracy and ruggedness. Using high-speed video cameras and motion analysis, we can actually detect inefficiencies or malfunctions. in the process of manufacturing industries and also high speed video clips analysis to study the impact of projectiles. Due to the intensified security issues, our system is designed for automatic tracking of motions. The object tracking can also be used in the process of counting the particles, such as bacteria, viruses, ionic polymer-metal composites, etc. These are some of the possible range of applications, which can use this motion tracking system, and there are many more applications which can be made functional and useful to the mankind.

Merits

1. Mapping of state space and observation space due for the multimodal conditional distribution
2. Use of monocular videos as input signal and tracking the 2D and 3D motions
3. Incorporation of unified space for both the input and output information.
4. The system is made to function automatically so that the manual hardship is reduced.
5. This system can handle linear human motions as well as non linear human motions and also in manufacturing industries, surveillance supervising and tracking movement in objects otherwise not possible by naked eye.

Conclusion

The work presented in paper is a blend of Gaussian process expert's model in which would track not only the human pose/motions but also the motions of other objects from the monocular video. The most important part of our work is the mixing of Gaussian process experts into a unified system where in we extract the temporal information and the spatial information and efficiently handled the problem of multimodality. The data which is used is present it the local neighborhood where the input and output are unified into a particular space. By providing this unified input-output space we have overcome the problem of one way mapping and effectively handle the issue of two-way multimodality. The adaptation to the scenarios is very straightforward. We have conducted our experiments on human motion tracking and also on object tracking based on the blooming of flowers through which we learned the working of Gaussian process experts with the help of temporal and spatial information can be blended and unified model can be defined. This would be the fast track technological domain development in the days to come for studying the human and non-human elements relating to movements and oscillations which are basically not visible by a naked eye.

References

[1] Xu Zhao, Member, Yun Fu, and Yuncai Liu, Human Motion Tracking by Temporal-Spatial Local Gaussian Process Experts, IEEE Transactions on Image Processing., 2011 vol.20 pp. 1141–1151.

[2] C. Sminchisescu, A. Kanaujia, Z. Li, and D. Metaxas, "Discriminative density propagation for 3-D human motion estimation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2005, vol. 1, pp. 390–398.

[3] A. Agarwal and B. Triggs, "Recovering 3-D human pose from monocular images," IEEE Trans. Pattern Anal. Machine Intell., vol. 28, no. 1, pp. 44–58, Jan. 2006.

[4] A. Elgammal and C. S. Lee, "Inferring 3-D body pose from silhouettes using activity manifold learning," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2004, vol. 2, pp. 681–688.

[5] X. Zhao, H. Ning, Y. Liu, and T. S. Huang, "Discriminative estimation of 3-D human pose using Gaussian processes," in Proc. 19th Int. Conf. Pattern Recognit., 2008, pp. 1–4.

[6] X. Zhao, Y. Fu, H. Ning, Y. Liu, and T. S. Huang, "Human pose regression through multiview visual fusion," IEEE Trans. Circuits Syst. Video Technol., vol. 20, pp. 957–966, 2010.

[7] C. E. Rasmussen and C. K.Williams, Gaussian Processes for Machine Learning. Cambridge, MA: MIT Press, 2006.

[8] R. Urtasun and T. Darrell, "Local probabilistic regression for activityindependent human pose inference," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2008, vol. 2, pp. 1–8.

[9] R. Urtasun, D. J. Fleet, A. Hertzmann, and P. Fua, "Priors for people tracking from small training sets," in Proc. IEEE Int. Conf. Comput. Vis., 2005, vol. 1, pp. 403–410.

[10] X. Zhao, Y. Fu, and Y. Liu, "Temporal-spatial local Gaussian process experts for human pose estimation," in Proc. 9th Asian

Conf. Comput. Vis., 2009, vol. 5994, pp. 364–373.

[11] J. Quiñonero-Candela and C. E. Rasmussen, “A unifying viewof sparse approximate Gaussian process regression,” J. Machine Learning Res., vol. 6, pp. 1939–1959, 2005.

[12] N. D. Lawrence, M. Seeger, and R. Herbrich, “Fast sparse Gaussian process methods: The informative vector machine,” in Proc. Advances in Neural Information Processing Systems, 2003, pp. 625–632.

[13] E. Snelson and Z. Ghahramani, “Sparse Gaussian processes using pseudo-inputs,” in Proc. Advances in Neural Information Processing Systems, 2006, pp. 1257–1264.

[14] C. E. Rasmussen and Z. Ghahramani, “Infinite mixtures of Gaussian process experts,” in Proc. Advances in Neural Information ProcessingSystems, 2002, pp. 881–888.

[15] R. Urtasun, S. Epfl, D. J. Fleet, and P. Fua, “3D people tracking with Gaussian process dynamical models,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2006, vol. 1, pp. 238–245.

[16] J. M.Wang, D. J. Fleet, and A. Hertzmann, “Gaussian process dynamical models,” in Proc. Advances in Neural Information Processing Systems, 2005, pp. 1441–1448.