International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

Advances in Motion Detection and Tracking of Human as a Target

J. Biju, D.Shanthi

Abstract*— Motion estimation of a target is the major area with higher computational complexity in video processing. It is the progression of discovery the motion patterns that describe the transformation from one frame to another in a sequence of video. Therefore, it is reasonable to carry out motion estimation only where movement is present. Image data in an image series remains mostly the same between frames along the target motion. To make use of the image statistics redundancy in image sequences, there is a need to guess motion. Motion estimation is valid for video compression improvement, stereo correspondence, object tracking and finding optical flow. Many precise methods have been proposed in the framework of one or more of these applications. Most motion estimation algorithms either operate directly in the image domain or finding the similar metric that measures how alike two pixels or two patches of pixels. In this paper, a review of a variety of motion estimation technique is presented.

***Key words*— optical flow, target tracking, segmentation, motion detection, representation of targets**

I. INTRODUCTION

In many visual processing algorithms, finding point-wise relationship between set of images play a vital role such as motion estimation and tracking for video compression and object tracking. Motion estimation is the technique of decisive motion vectors that portray the conversion from one two dimensional image to another image from neighboring frames in a video frame sequence. In sequence of video frames, motion is a main source of information. Image data in the video sequence remains mostly the analogous between frames along the motion. To discover the data redundancy in image sequence, there is a need to estimate the pattern of motion. Motion arises due to moving objects in the three dimensional scene or motion in camera. The use of the motion estimation is to recover back the information by interpreting the image content. Motion estimation evaluates the movement of objects in a video sequence to obtain motion vectors representing the defined motion. To initiate a system the data should be initialized, the movement of the target should be tracked, the exact location of the object should be estimated and finally the parameters should be evaluated to recognize the actions.

Manuscript published on 30 September 2019.
***Correspondence Author(s)**
J. Biju, Asst. Professor, Department of Computer Science and Engineering, Theni Kammavar Sangam College of Technology Theni, Tamilnadu.
Dr. D.Shanthi, Professor and Head, Department of Computer Science and Engineering, PSNA College of Engineering & Technology, Dindigul, Tamilnadu.

In this survey, we analyze the various literature exists in research based on the four steps in motion estimation and tracking.

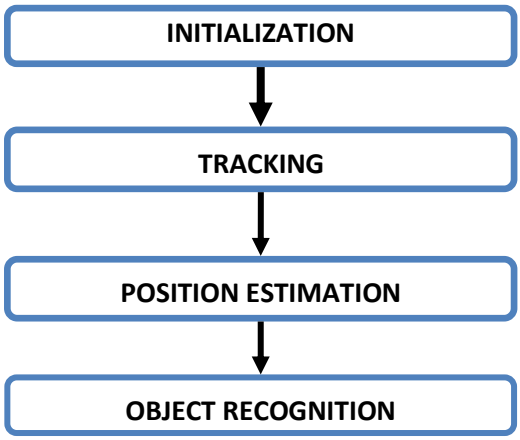

Fig 1: Steps in motion estimation

II. INITIALIZATION

Initialization involves camera lens calibration, adapting to visual characteristics, and model initialization which are the actions used to preprocess the data[1] , [2]. Few initialization steps are done before starting any operation and also as first stage in most of the apporaches.

Model initialization depends on the initial position of a target and the model which is used to represent the target.

Model-based algorithm is proposed [3] to determine the location of a target describes the sequence as assuming the initial position of target and incrementing it in each frame. The same idea is used in [4] by having a special operator to specify the position in every frame.

In some systems, [5] the same technique is employed to track the target and determine the position. The outcome of this approach is dependent on the target and how it fits the preliminary model in designing the system. Some systems use mean data of and others measure the presen state of the target and create offline data in offline [6]

III. TRACKING

Tracking is a process of enabling a continuous relationship of the target and their connecting links between video frames. It is a separate process, as a tool to prepare data for object recognition [7]. It involves detecting the moving objects by identifying the moving objects and tracking them with respect to time.

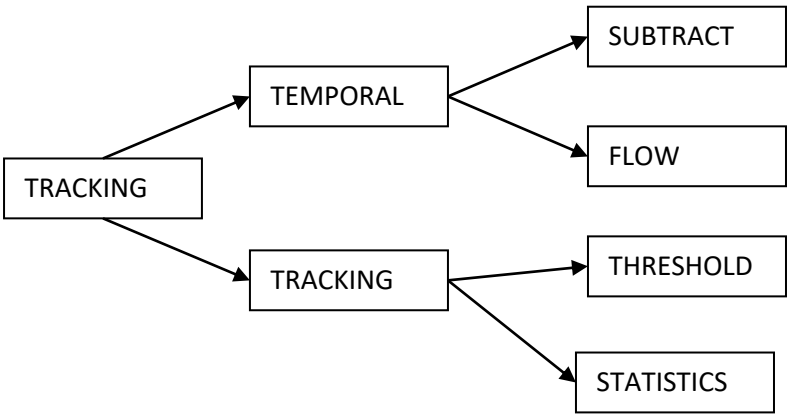

Fig 2 : Tracking classification

The features like edges, shapes and color are extracted to determine the position estimation [8]. Here, edges and other shapes are extracted and correlated with the edges of a target model to finalize the spot of the model.

3.1. Temporal images

Fixed camera and stable background forms the basic for temporal data based extraction. Here, the differences between images in a video series starts when there is a movement in the frame due to motion of the object. Subtraction is widely employed which involves simple subtraction of previous frame and current frame [9] in every pixel with respective intensity value[9]. For static videos, background without any target and noise eliminated value is used as reference frame [10]. Flow is denoted as the property describing the movement pattern in frame[11]. They calculate the pattern of movement of a target with optical flow with respect to a reference model.

3.2. Spatial images

Spatial data depends on statistical and thresholding approach. Thresholding is a basic technique dependent on exceptional background conditions and statistical technique is advanced than basic subtraction of frames. Basically thresholding is easy if the color information of the image data varies in each frame. Each and every pixel has individual property in background which can be extracted in statistical approaches. This majorly involves mean, variance, mode and other numerical quantitative property of the image for analysis.

IV. POSITION ESTIMATION

Position estimation is a post processing step developed for recognizing the target and its location in an image. The basic idea is to estimate position is to develop a target model and comparing it with the incoming models. To exactly compare the models prior knowledge and position information is required. In model free based approaches no previous information is required. They depict a model based on simple structures and stick diagrams. The location is dependent on the position of cluster of points.

Incase of Indirect Model a reference model based on previous use knowledge of position of target in past frame. In direct mode the model is periodically updated keeping the reference model as dynamic unlike direct mode.

V. RESULTS & DISCUSSIONS

The final step in target detection is recognizing the target and naming it depending on previous knowledge. The work is done by classifying the target based on its motion patterns and actions performed. Static recognition deals with the spatial information and it processes one frame per time. It uses information of previously processed frames. Dynamic pattern recognition uses temporal nature of the video sequence and gets updated frame to frame possessing the dynamicity.

Table 1 – comparative study of existing articles

S.No	Existing algorithm	Concept/Technique	Advantages	Disadvantages
1.	extraction of articulated objects	• Involves extraction of articulated objects. • contours parts of human body are extracted	Exact definition of human structure	Need for 3D data
2.	Tracking & Counting Moving People	• Counts people passing a counting line.	Efficient preprocessing	Complexity in determining reverse entry

3.	Human Movement with Explicit Motion Models	• Modeling of cloth as texture	general camera position is enough	cause complex illumination
4.	Model based technique to extract and generate human motion	• model based motion of a target	save the cost in 3D sensing	generating motion for multimedia
5.	Estimating 2 $^{1}_{2}$ D human model	• Outcome model based approach	Advanced than 2D structure	Issues in developing a virtual structure
6.	Contour tracking with active shape models	• dynamic filtering with a stretchy shape model • Kalman filter is worn to control spatial feature	incorporates variability in shape	Uses less system parameters for analysis
7.	Human body tracking with multi-ocular vision	• Modeling of limbs and associated articulations • Kalman filter is employed to predict model pose	cycling sequences	Numerous state variables in kalman filter
8.	Pfinder	• real-time tracking system for interpreting target actions	Used in wireless interfaces, low-bandwidth coding	Needs 10Hz and a standard SGI computer
9.	Low rank recognition of target motion	• Discusses on physical movement like running, walking	Usable for predefined speed and motion	Issues in unpredicted motion estimation
10.	Motion estimation by 3-successive video frames	• Motion estimation in continuous 3 frames	Easy in predicting the next state	Problem arises when the target is stable for more than 3 frames
11.	Human tracking by distributed video systems	• Background subtraction based tracking	Noise free	Used for static camera only
12.	Human motion analysis by robot arm model	• robot arm model	Uses gradient scheme of modeling	Issues in unpredicted motion estimation
13.	Vision and behavior – study	• Active vision techniques are used to view the target	Multiple targets can be used	Similar targets creates confusion
14.	Posture estimation by structural motion models	• fast posture estimation algorithm by location information targets	The unknown parameters are predicted and adjusted by a minimization method.	Usable for human models only
15.	Lower limb kinematics by medial axis transformation	• simple model for free-speed human walking	3D kinematic data.	Usable for standard database only
16.	Multiple cues for human motion capture	• A "touch-free" computer vision solution • analysis-by-synthesis approach is used	Tested on 3D movements and successfully estimate the position	Unclear processing flow
17.	Maximum likelihood estimation by EM algorithm	• Technique for monotone behaviour of the likelihood and divergence is derived.	missing value situations, applications to grouped, censored or truncated data	Less statistical evaluation

18.	Human posture recognition using single synthetic images	• attempting to get better out of the 2D structure of a multi-component target	Can be extended for 3D	static posture descriptions such as only
19.	Image sequence analysis	• Three primary problems of image sequence analysis are investigated - modeling, connection of feature points and analysis of occluding regions	Use key frame features and window code matching	Complex steps in evaluation of performance
20.	Image analysis of human motion with constraint propagation	• forecast at the semantic level and verifications are sought at the image level.	High speed	Used for small sized image data
21.	Understanding purposeful human motion	• Pixel-level, probabilistic processes including models of dynamic restraint on human motion.	Improves the perception of target motion.	Issues in adding new feature space
22.	Visual motion perception	• Detecting the displacement of retinal image skin texture using spatial relationships	computationally straightforward	Eye is alone considered as feature, not applicable for long vision images

VI. CONCLUSION

Thus, a review of a variety of motion estimation technique is presented. The details on Motion estimation of a target which possess higher computational complexity in video processing were discussed. It is reasonable to carry out motion estimation only where movement is present so that the complexity can be eliminated. Here, Motion estimation is evaluated for video compression improvement, object tracking and finding optical flow. Many precise methods were proposed in the framework of one or more of these applications. Most motion estimation algorithms either operate directly in the image domain or finding the similar metric that measures how alike two pixels or two patches of pixels. In this paper, The applications of target detection and tracking are enormous and it is expected that there will be continuous growth in the resources in future.

REFERENCES

1. M. Rossi and A. Bozzoli, *Tracking and Counting Moving People*, Technical Report 9404-03, IRST, Trento, Italy, April 1994
2. D. Meyer, J. Denzler, and H. Niemann, Model based extraction of articulated objects in image sequences, Proceedings of *Fourth International Conference on Image Processing, 1997*.
3. Rohr K et al, *Human Movement Analysis with Explicit Motion Models*, chap. 8, pp. 171–198, Dordrecht/Boston, 1997
4. Perales and Torres, A system for human motion matching between synthetic and real images based on a biomechanic graphical model, in proceedings of *Motion of Non-Rigid and Articulated Objects workshop, 1994*, pp. 83–88.
5. Njastad, S. Grinaker, and Storhaug, Estimating parameters in a 2 $\frac{1}{2}$ D human model, in proceedings of *11th Scandi-navian Conference on Image Analysis 1999*.
6. C. R. Wren, Azarbayejani, Darrell, and Pentland, Pfinder: real-time tracking of the human body, *Transactions of Pattern Analysis and . Machine Intelligence* **19**(7), 1997, 780–785
7. Tsukiyama and Shirai, Detection of the movements of persons from a sparse sequence of TV images,in proceedings of *Pattern Recognition* **18**, 1985, 207–213.
8. Hogg, *Interpreting Images of a Known Moving Object*, thesis for PhD, University of Sussex, UK, 1984
9. Haritaoglu, Harwood, and Davis, Who? When? Where? What? - A real time system for detecting and tracking people, in proceedings of *International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998.*
10. Kameda and Minoh, A human motion estimation method using 3-successive video frames, in proceedings of *Virtual Systems and Multimedia conference, 1996*.
11. Bregler, Learning and recognizing human dynamics in video sequences, in proceedings of *Computer Vision and Pattern Recognition conference, 1997*