TOWARDS BENCHMARKING OF VIDEO MOTION TRACKING ALGORITHMS

Mukesh Motwani[#1], Nishith Tirpankar[*2], Rakhi Motwani [#1], Monica Nicolescu[#1] and Frederick C. Harris, Jr.[#1]

[#1]*Univerity of Nevada, Reno USA*

[1]mukesh@cse.unr.edu

[3]fredh@cse.unr.edu

[*2]*KPIT Cummins Infosystems Ltd., Pune India*

[2]nishitht1@kpitcummins.com

ABSTRACT

The environment in which video motion needs to be tracked, places several constraints on the design of the tracking system. Current datasets which are used to evaluate and compare video motion tracking algorithms use a cumulative performance measure without thoroughly analyzing the effect of these different constraints imposed by the environment. There is need to build a heuristic framework which analyses these constraints as parameters of the framework and their effect on selection or design of tracking algorithm. The emphasis in this paper is to identify these parameters which will lay a foundation for defining subjective measures for the comparison of performance evaluation of tracking algorithms.

Index Terms— video tracking, metric, design.

1. INTRODUCTION

Video tracking is very challenging problem since the tracking environment is unique for every situation. Top down design approach of a tracking system consists of analysis of the coverage area leading to positioning of the camera, specification of pan and tilt unit, specification of the camera, number of cameras used and design of the algorithm. There are several video tracking algorithms to choose from as highlighted in survey [1]. The publicly available datasets such as PETS [2], i-LIDS [3],[4] which are currently used for testing the performance of motion tracking algorithms present very limited scenarios which may not be relevant to the current scenario. As shown in Fig.1 a sequence image from CAVIAR [5] dataset, considers only tracking humans in an indoor environment.

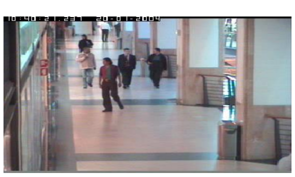

Fig. 1: Image sequence from CAVIAR

The current evaluation systems compare algorithms in very specific environments against a single metric [3] related to deviation of the tracked path from the actual ground truth. This metric is cumulative and does not identify the cause of deviation or measure the deviation due to specific outlier. This can result in the failure of the evaluated algorithm in an environment with certain unaccounted conditions such as occlusion since there is no object in the scene and thus absence of ground truth. The current evaluation metrics based on track (path of object of ground truth is compared with path of object of tracker) or frame based evaluations [6], [7] are not sufficient to cover all the operating scenarios for tracking algorithms. Thus, no generic evaluation metrics exist which can be used to test the performance of tracking algorithms in the presence of outliers. This lack of analysis results in a flawed method of comparison which leads to poor selection of tracking algorithms for a system. It is a challenge to arrive at a true comparison metric for tracking systems. The metric of deviation from ground truth path can still be used and is a valid measure provided there are image sequences which account for these variations in isolation. Thus, there is a need to create a cumulative metric which is derived from these subjective metrics corresponding to different constraints in the environment. It is imperative to identify these constraints in the environment for video tracking algorithms to be benchmarked. There is also a need to create datasets that enable to test the performance of tracking algorithms which have scenarios with these specific constraints in isolation to compute these subjective metrics corresponding to the scenarios.

The scope of this paper is limited to proposing heuristic parameters based on which motion tracking systems can be designed and compared. The purpose of this paper is to thoroughly analyze these constraints to lay foundation for computation of a metric or benchmarking system which compares algorithms in the presence of these constraints. This framework would enable to meaningfully compare tracking algorithms in spite of wide variation in testing environments.

Following sections list heuristics which should be considered in the process of design or selection of an

978-0-7695-3960-7/10 $26.00 © 2010 IEEE
DOI 10.1109/ICSAP.2010.84

algorithm for tracking in any environmental condition and also for comparison of algorithms in these environments.

2. BACKGROUND

The background of the image can be either static or moving. In case the background is static, the complexity of the background needs to be considered such as whether the background is simple or if it contains cluttered components. If the background is not static and contains multiple motions, then the tracking system needs to decide if the motions are of objects of interest or are they simply motions which should be ignored. If the background motions are not of importance, then the algorithm needs to see if the motions occur close to or away from the object of interest. Associated motions in the background due to the object motion such as due to shadows or due to object debris such as left by the trail of water of a boat need to be accounted. Based upon the above scenarios of the background, the following parameters are proposed.

2.1. Multiple motion traceability

Consider scenario where there are multiple ships which are part of a fleet as shown in Fig. 2. The tracking algorithm may require tracking all of them. In any frame, there can be multiple objects moving at the same time. As per the requirements based on the scenario, there could be a requirement to track multiple objects. Special metrics are required to handle this parameter as conventionally used video sequences do not test this parameter [8].

Fig. 2: Tracking multiple objects

2.2. Immunity to changes in background (distracters)

Consider the image shown in Fig 3. If the car on the left is the object being tracked, the motion of the people or vehicles should not affect the tracking algorithm.

Fig. 3: Background objects – distracters

If there is motion in the frame other than that of the desired objects, the tracking algorithm should not deviate from its current targets.

2.3. Ability to distinguish other close by motions

In Fig. 4, if the target to be tracked is the 2 wheeler at the centre of the frame, the motion of the vehicles around it should not cause the algorithm to loose track of the target.

Fig. 4: Distinguish motion of the object of interest

This is the parameter which decides if motions in the vicinity of the object being tracked will affect the tracker. Morphological based algorithms may not perform well in this case.

2.4. Shadows

Shadows can cause problems in tracking especially when they become larger than the object to be tracked. Shadows are longer in the evenings than in the noon in an outdoor environment, as shown in Fig. 5.

Fig. 5.1 Fig. 5.2

Fig. 5: Presence of shadows

2.5. Presence of complex or cluttered static background

In the Fig. 6 below, the lady is walking in a cluttered background with clothes of different colors.

Fig. 6: Cluttered background objects

This can upset algorithms based on tracking color. Thus, the tracking algorithm needs to take into account the complexity of the background of the environment.

2.6. Merge and split

Consider a scenario in which two people come together, shake hands and move away. At the point when the bodies merge, the tracking algorithm can detect the combined bodies as one. Also, when they split, the tracking algorithm may follow the wrong body upon splitting up.

3. VELOCITY OF OBJECTS

For any tracking system, velocity of the object or variation in rate of change of velocity is an important factor to be considered in the design. These factors are described below.

3.1. Variation in velocity of object

In Fig 7, it can be seen that the car is decelerating as it covers lesser distances in the time period of 0.2 seconds.

Fig. 7.1: t = 0.666666 s

Fig. 7.2: t = 0.866666 s

Fig. 7.3: t = 1.066666 s

Fig. 7.4: t = 1.266665 s

Fig. 7: Change in velocity of object

Algorithms which use motion history can rely on the constancy of velocity or motion of the object.

3.2. Objects moving too fast or too slow

There is always a problem of tracking objects which move too fast or too slowly. For example, consider the tracking of cars in a street. It is possible that a car can move fast enough for its relative displacement in consecutive frames to be large. Fast motion of objects causes poor motion continuity and can be overcome by using larger search space and feature based tracking as is proposed in the paper [9]. In case of objects which move too slowly, some algorithms which rely upon object motion tend to classify the slow objects as background and thus eliminate them as prospective tracking targets. If the object is small and its motion is slow, then it may be identified as noise.

3.3. Directionality

Consider an application of tracking the players in a soccer match as shown in Fig. 8. The objects tend to change their direction of motion suddenly. Motion in any direction in the plane of the image should be traceable irrespective of sudden changes in the direction of motion. Algorithms which use motion history, like the Kalman filter cannot work in this case.

Fig. 8.1

Fig. 8.2

Fig. 8.3

Fig. 8.4

Fig.8: Abrupt changes in direction of motion of player

4. OBJECT DISAPPEARANCE

If the object is not visible in the scene, there are two possibilities. Either it has left the field of view, or it has been hidden behind an occluding object in the field of view. Based on this consideration, following parameters are proposed.

4.1. Disappearance of object from scene

If the object leaves the frame, the algorithm should then identify the absence of the object and indicate the same based on timeout conditions set. The algorithm should be able to decide if the object is merely occluded or has left the scene.

4.2. Occlusion

Consider the airplane as the tracked object as shown in Fig. 9. The aircraft is occluded partially by the clouds.

Fig. 9.1

Fig. 9.2

Fig. 9.3

Fig. 9.4

Fig. 9: Partial Occlusion

The tracker should be able to recover from partial occlusion or even complete occlusions under certain circumstances. Thus, objects which either partially or completely obscure the objects of interest from view should not result in the losing the track of the object. Occlusions are handled using appearance models [10] and the temporal information [11] or using multiple camera feeds [12].

5. LIGHTING CONDITIONS

In order to find out the effect of the lighting conditions on the design or for comparison of tracking algorithms, following parameters are proposed.

5.1. Outdoor Day/Night operation

Consider the case of tracking people in an outdoor environment both during the night and the day. Typically, tracking systems using infrared cameras can operate under such conditions. The algorithm should be evaluated based whether it can track effectively during night as well.

5.2. Immunity to variation of intensity of light

Changes in the ambient or local light intensity could cause a change in the appearance of the object or the background. The light source intensity could cause change in pixel values as shown in Fig 10.

Fig. 10.1

Fig. 10.2

Fig. 10: Variation in light intensity

5.3. Change due to reflectance

The reflectance of the object can change if its orientation with respect to the camera changes. Also, if the light source is directional and not diffused, object appearance can change due to the relative motion between light source and object. Additionally, the object may have its own light source as shown in Fig. 11. Thus, the variability of surface reflectance over time can cause a change in the perception of the color of the object [13]. This can upset trackers which rely upon color information. The tracker should be able to follow objects irrespective of color changes.

Fig. 11: Pixel density change – object with own light source

6. CAMERA POSITIONING

The position of the camera with respect to the object dictates the area of coverage of the scene. Following is a description of how these parameters can affect the tracking algorithm.

6.1. Area of coverage of tracking-Scale immunity

As seen in Fig.12, the size of the airplane approaching the camera keeps increasing. The tracking algorithm must be able to track objects undergoing any scale change.

Fig. 12.1

Fig. 12.2

Fig. 12.3

Fig. 12.4

Fig. 12: Change of scale of object

The motion of the object toward or away from the fixed camera changes the appearance of the size of the object without causing any motion in the plane of the image. Algorithms like optical flow face the “aperture” or “correspondence” problem due to this. Algorithms should be able to adapt the size of their search windows to accommodate this change in scale. This is especially important for automatic zoom calculation for the camera. The zoom of the camera is calculated based upon the perspective size and adjusted to preserve it [14], [15], [16].

6.2. Objects too small in the field of view

The camera position determines the field of view, which in turn determines the size of the object. If the field of view is too large, the objects may appear too small. Objects that are too small may not be detected by the algorithm. The tracking algorithm to be evaluated should be able to detect the object of interest irrespective of the size of the object in the field of view. If the selected tracking algorithm cannot track distant objects less than specific size, then a camera with automatic zoom should be selected.

7. CHANGE IN OBJECT APPEARANCE

Objects in the scene can change their appearance in the course of tracking. This can be due to the object being non-rigid or object pose change.

7.1. Non-rigid bodies

A perfect example of a non-rigid object is the human body or a bird as shown in Fig 13. The motion of the wings from the body gives an impression that the entire shape of the bird has changed. This can be a cause of concern for algorithms which use templates which rely on matching shape of the object or optical flow which computes velocity vectors for each point separately.

Fig. 13.1

Fig. 13.2

Fig. 13: Non-rigid bodies

7.2. Change in shape of body due to orientation

The shape of a non rigid body may change due to its change in orientation or pose. Feature based algorithms are not suitable in such scenarios.

8. CONCLUSION AND FUTURE WORK

Heuristic measures for evaluation of tracking systems using subjective measures have been proposed in this paper. This paper attempts to identify in detail the different constraints imposed on a video tracking algorithm. Having identified the important parameters, a subjective metric for each of the parameter proposed here will be defined in future work. Based on these subjective measures, a cumulative metric will be derived and used for benchmarking. These parameters will also help in the design process of a tracking algorithm and comparing existing algorithms. This cumulative metric will also lead to algorithms which would allow the tracking system to detect change in these parameters and automatically select and switch to appropriate tracking algorithms. Thus, a general purpose tracker can be constructed.

9. REFERENCES

[1] Trucco, E.; Plakas, K., "Video Tracking: A Concise Survey," *Oceanic Engineering, IEEE Journal of* , vol.31, no.2, pp.520-529, April 2006

[2] Young, D.P. and Ferryman, J.M., “PETS metrics: on-line performance evaluation service”, Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005. 2nd Joint IEEE International Workshop on 15-16 Oct. 2005 Page(s):317 – 324

[3] i-LIDS (HOSDB, UK) dataset: http://scienceandresearch.homeoffice.gov.uk/hosdb/cctv-imaging-technology/video-based-detection-systems/i-lids/.

[4] Krinidis, M., Stamou, G., Teutsch, H., Spors, S., Nikolaidis, N., Rabenstein, R., Pitas, L., “An Audio-Visual Database For Evaluating Person Tracking Algorithms” IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). Volume 2, March 18-23, 2005 Page(s):237 - 240

[5] CAVIAR, Context Aware Vision using Image-based Active Recognition (projet européen IST): http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/

[6] Tordoff, B., Murray, D., “Reactive control of zoom while fixating using perspective and affine cameras”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 26, Issue 1, Jan. 2004 Page(s):98 – 112

[7] Kasturi, R., Goldgof, D., Soundararajan, P., Manohar, V., Garofolo, J., Bowers, R., Boonstra, M., Korzhova, V. and Jing Zhang “Framework for Performance Evaluation of Face, Text, and Vehicle Detection and Tracking in Video: Data, Metrics, and Protocol”,Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 31, Issue 2, Feb. 2009 Page(s):319 – 336

[8] Perera, A.G.A., Hoogs, A., Srinivas, C., Brooksby, G., Wensheng Hu, “Evaluation of Algorithms for Tracking Multiple Objects in Video”, Applied Imagery and Pattern Recognition Workshop, 2006. AIPR 2006. 35th IEEE, 11-13 Oct. 2006 Page(s):35 – 35

[9] Yuan Li, Haizhou Ai, Yamashita, T., Shihong Lao, Kawade, M., “Tracking in Low Frame Rate Video: A Cascade Particle Filter with Discriminative Observers of Different Lifespans”, IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR '07, 17-22 June 2007 Page(s):1 – 8

[10] Campbell-West, F., Hongbin Wang, Miller, P., “Where is It? Object Reacquisition in Surveillance Video”, International Machine Vision and Image Processing Conference, 2008, 3-5 Sept. 2008 Page(s):182 - 187

[11] Jiyan Pan, Bo Hu, “Robust Occlusion Handling in Object Tracking”, IEEE Conference on Computer Vision and Pattern Recognition, 2007, CVPR '07, 17-22 June 2007 Page(s):1 – 8

[12] Velipasalar, S., Wolf, W.,Image Processing, “Multiple object tracking and occlusion handling by information exchange between uncalibrated cameras” ICIP 2005.Volume 2, 11-14 Sept. 2005 Page(s):II - 418-21

[13] Renno, J.-P., Makris, D., Ellis, T., Jones, G.A “Application and Evaluation of Colour Constancy in Visual Surveillance” Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005, 2nd Joint IEEE International Workshop on 16-16 Oct. 2005 Page(s):301 – 308

[14] Poelman, C.J., Kanade, T., “A paraperspective factorization method for shape and motion recovery” Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 19, Issue 3, March 1997 Page(s):206 – 218

[15] Tordoff, B., Murray, D., “Reactive control of zoom while fixating using perspective and affine cameras”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 26, Issue 1, Jan. 2004 Page(s):98 – 112

[16] Wei, Y., Badawy, W., “A novel zoom invariant video object tracking algorithm (ZIVOTA)”, Canadian Conference on Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Volume 2, 4-7 May 2003 Page(s):1191 - 1194 vol.2