Available online at http://www.journalcra.com

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

International Journal of Current Research
Vol. 6, Issue, 04, pp.6184-6187, April, 2014

ISSN: 0975-833X

RESEARCH ARTICLE

DETECTION AND TRACKING OF HUMAN FROM VIDEO USING LIKELIHOOD ALGORITHM

***Keerthana, K., Loganayaki, A., Meenakshi, S., Nandhini, A. and Dr. Mary Praveena, S.**

Department of ECE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India

ARTICLE INFO

Article History:
Received 26th January, 2013
Received in revised form
24th February, 2014
Accepted 10th March, 2014
Published online 23rd April, 2014

Key words:
Subject moves,
A 3-D appearance description using texture and color is learned progressively.

ABSTRACT

A new approach to tracking human subjects in video sequences. A parametric ellipsoid model in both detection and visual tracking is introduced. For detection, this is projected at static grid positions to find intersections between potential subject positions and foreground image data, as determined by mixture of Gaussian segmentation. For tracking, the ellipsoid is parameterized by position, velocity and height as part of the state vectors of a particle filter. As the subject moves, a 3-D appearance description using texture and color is learned progressively. This allows us to integrate observations from multiple cameras into the likelihood function. The texture and color signature can be used for effective tracking of subjects, with multiple object tracking accuracies on PETS benchmarks of greater than 90%. Further, we have combined this signature with spatial data association to achieve F-measures (combining recall and precision) rates of between 60% and 85% when handing off between cameras with non overlapping views, depending on the nature of the data sets. This measure is further improved by use of the bisection property of the Hungarian method for assignment of identities to subjects.

INTRODUCTION

Image Processing is a technique to enhance raw images received from cameras/sensors placed on satellites, space probes and aircrafts or pictures taken in normal day-to-day life for various applications. Various techniques have been developed in Image Processing during the last four to five decades. Most of the techniques are developed for enhancing images obtained from unmanned spacecraft's, space probes and military reconnaissance flights. Image Processing systems are becoming popular due to easy availability of powerful personnel computers, large size memory devices, graphics softwares etc. Image Processing is used in various applications such as Remote Sensing, Medical Imaging, Film Industry, Forensic Studies, Military, Non-destructive Evaluation, etc.,

Methods of Image Processing

There are two methods available in Image Processing

Analog Image Processing

Analog Image Processing refers to the alteration of image through electrical means. The most common example is the television image. The television signal is a voltage level which varies in amplitude to represent brightness through the image. By electrically varying the signal, the displayed image appearance is altered. The brightness and contrast controls on a TV set serve to adjust the amplitude and reference of the video signal, resulting in the brightening, darkening and alteration of the brightness range of the displayed image.

Digital Image Processing

In this case, digital computers are used to process the image. The image will be converted to digital form using a scanner – digitizer (Intille *et al.*, 1997) (as shown in Figure 1) and then process it. It is defined as the subjecting numerical representations of objects to a series of operations in order to obtain a desired result. It starts with one image and produces a modified version of the same. It is therefore a process that takes an image into another. The term digital image processing generally refers to processing of a two-dimensional picture by a digital computer (Jed Lengyel *et al.*, 1990; Isard and Blake 1998). In a broader context, it implies digital processing of any two-dimensional data. A digital image is an array of real numbers represented by a finite number of bits. The principle advantage of Digital Image Processing methods is its versatility, repeatability and the preservation of original data precision.

The various Image Processing techniques are:

- Image representation
- Image preprocessing
- Image enhancement

****Corresponding author: Keerthana, K.***
Department of ECE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India.

- Image restoration
- Image analysis
- Image reconstruction
- Image data compression

VIDEO BASED HUMAN TRACKING

Vision-based human pose tracking promises to be a key enabling technology for myriad applications, including the analysis of human activities for perceptive environments and novel man-machine interfaces. While progress toward that goal has been exciting, and limited applications have been demonstrated, the recovery of human pose from video in unconstrained settings remains challenging. One of the key challenges stems from the complexity of the human kinematic structure itself. The sheer number and variety of joints in the human body (the nature of which is an active area of biomechanics research) entails the estimation of many parameters. The estimation problem is also challenging because muscles and other body tissues obscure the skeletal structure, making it impossible to directly observe the pose of the skeleton. Clothing further obscures the skeleton, and greatly increases the variability of individual appearance, which further exacerbates the problem. Finally, the imaging process itself produces a number of ambiguities, either because of occlusion, limited image resolution, or the inability to easily discriminate the parts of a person from one another or from the background. Some of these issues are inherent, yielding ambiguities that can only be resolved with prior knowledge; others lead to computational burdens that require clever engineering solutions.

The estimation of 3D human pose is currently possible in constrained situations, for example with multiple cameras, with little occlusion or confounding background clutter, or with restricted types of movement. Nevertheless, despite a decade of active research, monocular 3D pose tracking remains largely unsolved. From a single view it is hard to escape ambiguities in depth and scale, reflection ambiguities where different 3D poses produce similar images, and missing observations of certain parts of the body because of self-occlusions.

PERSON DETECTION PROCESS

Although our feature sets are use full in general, our experiments will focus on the problem of finding people in images and videos. Person detection is a challenging task, with many applications that has attracted lot of attention in recent years. Consider the case of personal digital content analysis, where typical content is images taken during a vacation, at a party or at some family occasion. Statistics show that even digital camera owners who use their cameras only occasionally can take as many as 10,000 photos in just 2-3 years, at which point it becomes tedious to manually search and locate these photos. Intelligent digital content management software that automatically adds tags to images to facilitate search is thus an important research goal. Most of the images taken are of people, so person detection will form an integral part of such tools. For commercial film and video contents, person detection will form an integral part of applications for video on demand and automatic content management. In conjunction with face and activity recognition, this may facilitate search for relevant contents or searches for few relevant sub-sequences. Person detectors are also being explored for the detection of pedestrians by smart cars. Typically, information from multiple sensors such as stereo and infra-red cameras is fused and domain specific knowledge such as the fact that pedestrians often traverse cross walks is exploited, but performance is still far below than needed for such systems to be used in the real world. More robust person detectors would certainly help to improve the overall system performance Another application is in video surveillance and security where real-time systems are needed to analyze and process video sequences for intrusion detection.

CHALLENGES OF PERSON TRACKING IN VIDEO

Firstly, the image formation process suppresses 3-D depth information and creates dependencies on viewpoint such that even a small change in the object's position or orientation w.r.t. the camera centre may change its appearance considerably. A related issue is the large variation in scales under which an object can be viewed. An object detector must handle the issues of viewpoint and scale changes and provide invariance to them. Secondly, most natural object classes have large within-class variations. For example, for humans both appearance and pose change considerably between images and differences in clothing create further changes. A robust detector must try to achieve independence of these variations. Thirdly, background clutter is common and varies from image to image. Examples are images taken in natural settings, outdoor scenes in cities and indoor environments. The detector must be capable of distinguishing object class from complex background regions. The previous two difficulties present conflicting challenges, which must be tackled simultaneously. A detector that is very specific to a particular object instance will give less false detections on background regions, but will also miss many other object instances while an overly general detector may handle large intra-class variations but will generate a lot of false detections on background regions. Fourthly, object colour and general illumination varies considerably, for example direct sunlight and shadows during the day to artificial or dim lighting at night. Although models of colour and illumination invariance have made significant advances, they still are far from being effective solutions when compared to human and mammalian visual systems which are extremely well adapted to such changes, c.f. (Land 1959, Daw 1984, Ingle 1985). Thus a robust object detector must handle colour changes and provide invariance to a broad range of illumination and lighting change.

TRACKING AND DETECTION

Video surveillance can be an effective tool for today's businesses—large and small—in security surveillance, production monitoring, and deterring predatory and purloining behaviors. Since the introduction of analog video surveillance systems back in the 1970s, tremendous strides have been made in sensing, storage, networking, and communication technologies. The consequence is that, instead of employing video surveillance mainly as an "after-effect" forensic tool, it is now feasible to deploy digital, network-based surveillance systems to provide interactive, real-time monitoring and surveillance. This research proposes a software framework for video analysis to enable robust and real-time human activity detection and recognition. Our research makes the following contributions:

For activity detection and tracking, we propose a realtime, robust algorithm that is well suited for analyzing outdoor, far-field activities. In particular, we improve the robustness in activity analysis by providing intelligent control and fail-over mechanisms, built on top of low-level motiondetection algorithms such as frame differencing and featurecorrelation. These mechanisms improve overall robustness and accuracy by maintaining tracking and recovering from both non-catastrophic errors (such as occasional, short periods of occlusion and silhouette merging) and catastrophic errors (such as long periods of disappearance of activities from a camera's field of view). These fail-over mechanisms include efficient multi-hypothesis tracking to disambiguate figures from cluttered background, and a two-level, hierarchical activity representation scheme for bottom-up data fusion and top-down information guidance.

BLOCK DIAGRAM

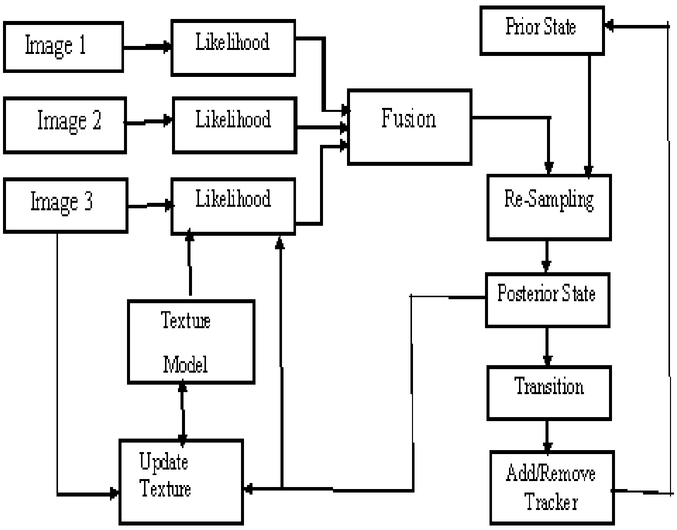

Fig. 1. Block diagram of Tracking

In general, there are two necessary activities, detection and tracking. It is necessary to detect a new presence, caused for example by entry into the field of view or reemergence from occlusion or blind areas. Detection must search the entire observable space, and thus can have considerable complexity. Tracking is an estimation process, which includes prediction and evaluation. Prediction includes a motion model that limits the search space for evaluation of subject presence in succeeding video frames. Evaluation requires a comparison between the prediction of position, velocity and appearance, and the observed data. Tracking implies the recovery of trajectories. In tracking humans, there are different levels at which tracking can be performed. At the highest level, the whole body is tracked without paying attention to the details of the posture and limbs. At a lower level, the posture and limbs are tracked. At an even lower level, one or two parts of the body (such as hands) are tracked. The finest level would be tracking the fingers of a hand or facial features. The type of application determines the tracking level. Table 1 summarizes the applications mentioned earlier and the required tracking level. The different levels are denoted as (1) whole, (2) limb, (3) head and hand, and (4) face and finger. Action denotes that action recognition is required for the corresponding application. Most available techniques for detecting moving objects have been designed for scenes acquired by a stationary camera. These methods allow to segment each image into a set of regions representing the moving objects by using a background differencing algorithm (Haritaoglu *et al.*, 1998 and Intille *et al.*, 1997). More recently, (Grimson *et al.*, 1998) have proposed a local modeling of the background using a mixture of K-Gaussian allowing to process video streams with time varying background. These methods give satisfactory results and can be implemented for real time processing without dedicated hardware.

Table 1. Error and number of particles

Dataset	η max	f(%)	m(%)	s(%)
PETs09	128	1.178	4.367	0.02
∑gt=4923	256	1.300	3.555	0.05

SIMULATION RESULTS

Fig. 2. PETs09 Dataset

The images consisting of people are given as input

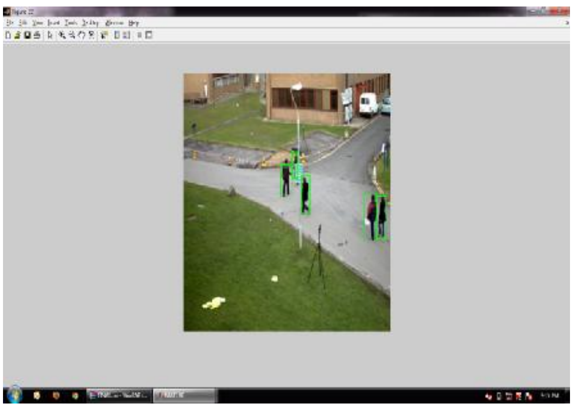

Fig. 3. Tracking of people

The people in each of the images are tracked and enclosed in a grid (rectangular box).

Conclusion

In this paper a new approach to tracking human subjects in video sequences is discussed. A parametric ellipsoid model in both detection and visual tracking is introduced. For detection, this is projected at static grid positions to find intersections between potential subject positions and foreground image data, as determined by mixture of Gaussian segmentation. For tracking, the ellipsoid is parameterized by position, velocity and height as part of the state vectors of a particle filter.

REFERENCES

Andriluka M., S. Roth, and B. Schiele. People-tracking-by-detection and people-detection-by-tracking. In IEEE Comp. Vision and Pattern Rec., 2008.

Berclaz J., F. Fleuret, and P. Fua. Robust people tracking with global trajectory optimization. In IEEE Comp. Vision and Pattern Rec., 2006.

Daw, N.W. The psychology and physiology of colour vision. Trends in Neurosciences, 1:330–335, 1984.

Everingham M., J. Sivic, and A. Zisserman, “Taking the bite out of automated naming of characters in TV video,” IVC, vol. 27, pp. 545–559, 2009.

Grimson, W.E.L. L. Lee, R. Romano, and C. Stauffer. Using adaptive tracking to classify and monitor activities in a site. In CVPR98, pages 22–31, 1998.

Haritaoglu I., D. Harwood, and L.S. Davis. W4S: A real-time system for detecting and tracking people in 2 1/2-d. In ECCV98, 1998.

Ingle, D. The goldfish as a retinex animal. Science, 227:651–654, 1985.

Intille S.S., J.W. Davis, and A.F. Bobick. Real time closed world tracking. In CVPR97, pages 697–703, 1997.

Isard M. and A. Blake, “Condensation - conditional density propagation for visual tracking,” IJCV, vol. 29, no. 1, pp. 5–28, 1998.

Jed Lengyel, Mark Reichert, Bruce R. Donald, and Donald P. Greenberg. Real-time robot motion planning using rasterizing computer graphics hardware. SIGGRAPH Computer Graphics, 24(4):327-335, 1990.

Kass A., M. Witkin and D. Terzopoulos. Snakes: Active contour models. International Journal of Computer Vision, pages 321–331, 1988.

Kim M., S. Kumar, V. Pavlovic, and H. Rowley, “Face tracking and recognition with visual constraints in real-world videos,” CVPR, 2008.

Land E. H. Color vision and the natural image. part ii. Proceedings of the National Academy of Sciences of the USA, 45:651–654, 1959b.

Li Y., C. Huang, and R. Nevatia. Learning to associate: Hybridboosted multi-target tracker for crowded scene. In IEEE Comp. Vision and Pattern Rec., 2009.

Li Y., H. Ai, T. Yamashita, S. Lao, and M. Kawade, “Tracking in low frame rate video: A cascade particle filter with discriminative observers of different lifespans,” CVPR, 2007.

Lim J., D. Ross, R.S. Lin, and M.H. Yang, “Incremental learning for visual tracking,” NIPS, 2005.

Mark Harris. Real-Time Cloud Simulation and Rendering. Ph.D. thesis, University of North Carolina at Chapel Hill, 2003.

Michael D. Breitenstein, Fabian Reichlin, Bastian Leibe, Esther Koller-Meier and Luc Van Gool, “Online Multi-Person Tracking-by-Detection from a Single, Uncalibrated Camera”, IEEE Transactions On Pattern Analysis And Machine Intelligence, Submitted January 2010, Revised October 2010.

Schulz D., W. Burgard, D. Fox, and A. Cremers, “People tracking with a mobile robot using sample-based joint probabilistic data association filters,” International Journal of Robotics Research (IJRR), vol. 22, no. 2, pp. 99–116, 2003.

Topp E. and H. Christensen, “Tracking for following and passing persons,” in Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), Alberta, Canada, 2005.

Wikipedia: Graphics Processing Unit. Retrieved November 2010.

Wikipedia: Video Card. Retrieved November 2010.

****