# F3-G: Distributed Video Analytics and Anomaly Detection

**Abstract— This project investigates the development of automated tools for video monitoring using multiple cameras in support of security applications. The project has both experimental and theoretical components. In Year 5, we focused on two problems of critical operational interest to surveillance and security in mass transit environments: the automatic, real-time detection of counterflow (e.g., passengers moving the wrong direction through a one-way exit lane) and the re-identification of humans in a wide-area network of non-overlapping surveillance cameras.**

## I. PARTICIPANTS

| Faculty/Staff | | | | |
|---|---|---|---|---|
| **Name** | **Title** | **Institution** | **Email** | **Phone** |
| Richard Radke | Associate Professor | RPI | rjradke@ecse.rpi.edu | 518.276.6483 |
| Venkatesh Saligrama | Professor | BU | srv@bu.edu | 617.353.1040 |
| Mohamed Elgharib | PostDoc | BU | gharib@bu.edu | |
| **Students** | | | | |
| **Name** | **Degree Pursued** | **Institution** | **Email** | **Intended Year of Graduation** |
| Yang Li | PhD | RPI | liy21@rpi.edu | 2015 |
| Ziyan Wu | PhD | RPI | wuz5@rpi.edu | 2014 |
| Yuting Chen | PhD | BU | yutingch@bu.edu | 2016 |
| Jing Qian | PhD | BU | jingq@bu.edu | 2014 |

## II. PROJECT OVERVIEW AND SIGNIFICANCE

The goal of this ALERT project is to provide an automated capability for analyzing crowded videos and detecting anomalous motions that can provide indication of activities of interest to human operators in both indoor and outdoor environments. Real-world tracking in environments of homeland security interest is characterized by many lower-resolution cameras viewing crowds of people, configured so that the cameras do not "work together".

We first continued our Year 4 work on the problem of detecting counterflow motion in videos of highly dense crowds. In Year 5, we focused on improving the detection performance by identifying scene features --- that is, features on motionless background surfaces. We proposed a three-way classifier to differentiate counterflow from normal flow, simultaneously identifying scene features based on statistics of low-level feature point tracks. By monitoring scene features, we can reduce the likelihood that moving features' point tracks mix with scene feature point tracks, as well as detect and discard frames with periodic jitter. We also constructed a Scene Feature Heat Map, which reflects the space-varying probability that object trajectories might mix with scene features. When an object trajectory nears a high-probability region of this map, we switch to a more time-consuming and robust joint Lucas-Kanade tracking algorithm to improve performance. We evalu-

ated the algorithms with extensive experiments on several datasets, including almost three weeks of data from an airport surveillance camera network. The experiments demonstrated the feasibility of the proposed algorithms and their significant improvements for counterflow detection.

We next addressed the problem of human re-identification across cameras with non-overlapping fields of view, one of the most important and difficult problems in video surveillance and analysis. Current algorithms are likely to fail in real-world scenarios for several reasons. For example, surveillance cameras are typically mounted high above the ground plane, causing serious perspective changes. Also, most algorithms approach matching across images using the same descriptors, regardless of camera viewpoint or human pose. In Year 5, we introduced a re-identification algorithm that addresses both problems. First, we build a model for human appearance as a function of pose, using training data gathered from a calibrated camera. We then apply this "pose prior" in online re-identification to make matching and identification more robust to viewpoint. Finally, we integrate person-specific features learned over the course of tracking to improve the algorithm's performance. We evaluated the performance of the proposed algorithm and compare it to several state-of-the-art algorithms, demonstrating superior performance on standard benchmarking datasets as well as a new, very challenging dataset taken at an airport.

## III. RESEARCH AND EDUCATION ACTIVITY

### *A. State-of-the-Art and Technical Approach*

One of the main goals of this ALERT project is identifying anomalous activity in space and time. Currently, this type of video analysis requires significant human supervision. However, since human supervision is often not scalable, most of the video data from surveillance cameras is, at best, stored and rarely processed. Clearly of interest are autonomous analysis algorithms for networked camera systems operating in unstructured and highly-cluttered indoor and outdoor environments. Furthermore, as computing power available in cameras increases, the computational intelligence is expected to move to the network edge, i.e., individual camera nodes, thus calling for distributed solutions. There are fundamental technical challenges to realizing this goal. First, urban scenarios provide a deluge of dynamic data. Identifying relevant information, such as meaningful change detection, in urban clutter is not easy. Second, doing so reliably, i.e., with small false alarm and miss rates, is difficult and perhaps impossible in harsh sensing environments (e.g., camera jitter). Third, combining views from multiple cameras for dynamic scene characterization and to improve detection, localization and tracking performance is challenging.

The problem of detecting dominant motions in crowded video and classifying outlying motions has been widely studied [1-4]. Tu and Rittscher [5] introduced a crowd segmentation algorithm by clustering interest points into groups by determining maximal cliques in a graph. However, both the algorithm and experiments are based on videos from overhead views only, which is the easiest case in counterflow detection. Andrade et al. [6] proposed an algorithm for detecting abnormal movements in crowds by applying principal component analysis to optical flow maps and spectral clustering to hidden Markov models, but did not perform any real-world experiments. Also, the performance of this algorithm is highly depending on the training set: a pattern of normal motion may be considered as abnormal if it is not covered by the training set. Brostow and Cipolla [7] used an unsupervised Bayesian detection algorithm to segment low-level feature tracklets based on a spatial prior and a likelihood model of coherent motion. Ali and Shah [8] modeled a highly dense crowd as an aperiodic dynamical system that can be studied with Lagrangian particle dynamics techniques. Antonini and Thiran [4] introduced a trajectory clustering method based on independent component analysis. Junejo et al. [2] applied graph cuts to segmenting tracklets. Cheriyadat and Radke [9] proposed a trajectory clustering algorithm based on non-negative matrix factorization, as well as an automatic dominant motion detection method by clustering trajectories based on longest common subsequences [10].

We next briefly describe a solution to detecting counterflow in a fixed camera that leverages the detection of scene features to improve performance. We first identify low-level features in the initial frame, which are tracked over time using a pyramidal Kanade-Lucas-Tomasi (KLT) optical flow algorithm, which can track large pixel motions while keeping the size of the integration window relatively small. This low-level feature point tracking is often inaccurate, due to both the low resolution and quality of the input videos and periodic jitter. Consequently, it is common for features on foreground objects (corresponding to the allowable/counter flow) to mix or merge with stationary scene features.

Our solution to this problem is to build a three-way classifier over two features extracted from each point track to classify normal flow, counterflow, and scene features. The point tracks are classified at a specified interval (e.g., every 300 frames). The recognized scene features can also be used to compensate for location drift caused by jitter. After the scene features are classified, they can be used to deal with two issues. First, point tracks that were classified as scene points in the previous decision are matched and tracked only after all of the other (flow) features are matched and tracked for each frame. This step significantly reduces the probability that scene features mix with moving features and confuse the tracker/classifier. Second, the statistics of scene features provide an easy way to detect frames with jitter. These frames are then ignored for the purposes of tracking and classification, which substantially improves robustness. Using only the feature tracks classified as scene features, we generate a Scene Feature Heat Map (SFHM) to further reduce false alarms. The SFHM is basically a visualization of the probability that a feature track at the given pixel contains a scene feature. When a tracked feature moves close to a "high-heat" region on the SFHM, it is more likely to mix with scene features. Hence, in this case, we use a Pyramidal Joint Lucas-Kanade algorithm for feature tracking, which combines the Kanade-Lucas and Horn-Schunck optical flow algorithms. Some results are shown in Figure 1.


| Day | GT | 2-Class | | 3-Class | | +SFHM | | 3C+SFHM+Joint | |
|---|---|---|---|---|---|---|---|---|---|
| | | TP | FA | TP | FA | TP | FA | TP | FA |
| 1 | 10 | 10 | 257 | 10 | 37 | 10 | 26 | 10 | 0 |
| 2 | 10 | 9 | 320 | 10 | 46 | 10 | 32 | 10 | 0 |
| 3 | 10 | 10 | 198 | 10 | 43 | 10 | 27 | 10 | 1 |
| 4 | 10 | 10 | 212 | 10 | 53 | 10 | 22 | 10 | 0 |
| 5 | 10 | 9 | 223 | 10 | 52 | 10 | 18 | 10 | 0 |
| 6 | 10 | 10 | 315 | 10 | 43 | 10 | 25 | 10 | 0 |
| 7 | 10 | 9 | 275 | 10 | 49 | 10 | 27 | 10 | 0 |
| 8 | 10 | 9 | 231 | 10 | 58 | 10 | 37 | 10 | 0 |
| 9 | 13 | 12 | 198 | 13 | 72 | 13 | 33 | 13 | 0 |
| 10 | 14 | 14 | 152 | 14 | 47 | 14 | 15 | 14 | 0 |
| 11 | 12 | 12 | 202 | 12 | 69 | 12 | 26 | 12 | 0 |
| 12 | 13 | 12 | 171 | 13 | 52 | 13 | 22 | 13 | 0 |
| 13 | 12 | 12 | 239 | 12 | 58 | 12 | 34 | 12 | 0 |
| 14 | 14 | 12 | 277 | 14 | 71 | 14 | 37 | 14 | 0 |
| 15 | 13 | 13 | 208 | 13 | 77 | 13 | 28 | 13 | 0 |
| 16 | 12 | 12 | 132 | 12 | 32 | 12 | 13 | 12 | 0 |
| 17 | 13 | 13 | 157 | 13 | 55 | 13 | 27 | 13 | 0 |
| 18 | 11 | 10 | 189 | 11 | 49 | 11 | 23 | 11 | 0 |
| 19 | 13 | 13 | 288 | 13 | 63 | 13 | 25 | 13 | 0 |
| 20 | 13 | 10 | 215 | 13 | 75 | 13 | 31 | 13 | 0 |
| 21 | 14 | 12 | 375 | 14 | 83 | 14 | 39 | 14 | 1 |
| 22 | 12 | 11 | 358 | 12 | 89 | 12 | 41 | 12 | 1 |
| Total | 249 | 234 | 5218 | 249 | 1273 | 249 | 608 | 249 | 3 |

**Figure 1: (a) Sample flow classification results. Top: normal flow. Bottom: counterflow is detected and the target is located. (b) Results of the counterflow experiment on 3 weeks of airport video. GT denotes the number of ground truth counterflows, TP the number of true positives and FA the number of false alarms.**

### Implementation Details for Counter Flow Detection

Our counter-flow detection metric is an aggregation of detection scores over several frames. Here only non-scene feature trajectories are used. First we estimate the level of counter-flow detection activity at each frame using the following equation:

$$d_n(\boldsymbol{i}) = \left|\left|\left(F_n(\boldsymbol{i}) - F_o(\boldsymbol{i})\right)C\right|\right| \rangle W \qquad (1)$$

Here $F_n(\boldsymbol{i})$ is the i[th] feature point in the nth examined frame while $F_o(\boldsymbol{i})$ is the corresponding feature point in the first frame this feature is available (in time). $d_n(i)$ is a binary value set to 1 if counter-flow exists and 0 otherwise. $(F_n(\boldsymbol{i}) - F_o(\boldsymbol{i}))C$ projects the vector connecting the matching points on to the counter-flow direction. This direction is defined by the unit vector *C*. The || || is the vector magnitude operator and in Eq.2 it estimates how many pixels did the examined point travel in the counter-flow direction. If this displacement is more than *W* pixels then the examined feature point gets flagged as moving in counter-flow. We fix *W* to 50 throughout all experiments.

Equation 1 is applied for every feature point in every examined frame. This assigns a value of either 1 or 0 to each feature point. A detection score is then assigned to the whole frame by summing up all the feature points detection scores. This process is applied every 10 frames. A score is then generated every 50 frames by summing up all the frame detection scores. For a window of 50 frames there are 5 frame detections. This summation will be referred as the Global Detection Score in the rest of this report. The Global Detection Score favors temporally consistent detections and hence it is more robust to noise than single frame detection. The choices of 10 and 50 frames used in the above procedure were made only once at the start of experiments and were fixed afterwards. Counter-Flow is detected at some instant if the most recent Global Detection Score is more than some value. This value is set to 800 in our experiments.

Figure 2 shows an overview of our detection algorithm. Here locations of scene feature points are shown as in black in the binary mask (first row, second column). Scene features are shown as green circles and their trajectories are shown in red. Finally sites containing counter-flow activity are shown in blue (second row, second column).


**Figure 2: In clockwise direction: Original frame, binary mask of non-scene features; feature points trajectories with green heads and red tails and the resulting counter-flow detection in blue. In the binary mask non-scene features are denoted with white regions.**

## Multi-Camera Fusion

A drawback of using only one camera is the generation of a counter-flow detection score that is view variant. That is an event can be interpreted differently depending on the viewpoint of the examined camera. Figure 3 shows an example of this case. Here there is a passenger lifting his bag while standing in his place. This is not a counter-flow event. However the bag lifting action is seen as a counter-flow activity from the examined camera viewpoint and is misclassified as counter-flow. To reduce false alarms arising from such scenarios we use fuse information from multiple cameras observing the same scene. Figure 4 shows three different cameras observing the same examined corridor of Fig. 3 from different sites. A person moving in the normal-flow direction will be first seen by Camera 3, then by Camera 2 and finally by Camera 1. Directions of counter-flow are shown in red arrows. Hence our detection algorithm is robustified against false alarms by forcing consensus among the detection scores of three cameras. That is, an event is flagged as counter flow if it is detected first in Camera 1 followed by detection in at least one of the remaining two cameras.

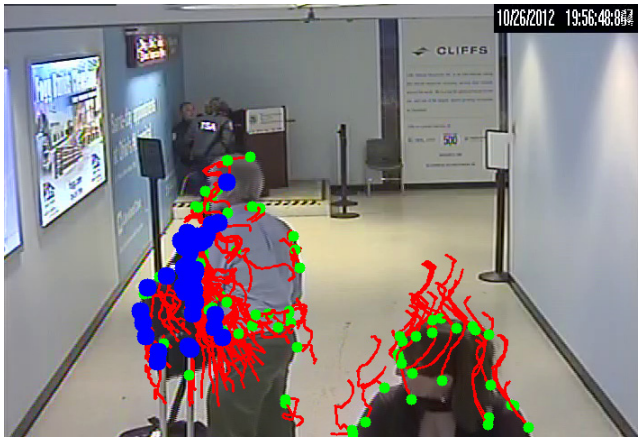


**Figure 3: An activity misclassified as counter-flow from the observed camera.**


**Figure 4: From left: Camera 1, 2 and 3 observing the same exit corridor. Counter-flow direction is shown in red arrows. A counter-flow direction is first seen in Camera 1, followed by 2 and finally 3.**

We evaluated our algorithms on several datasets, both standard ones (e.g., CAVIAR, clips from a shopping center in Portugal) and custom-collected video at a large US airport (Cleveland-Hopkins International). In all cases, we showed that the 3-class classifier combined with the SFHM outperformed both a baseline 2-class algorithm and two competing state-of-the-art algorithms ([8] and [10]). The proposed classifier was able to detect all counterflow events with no misses, and generated only 3 false alarms over 22 days of continuous operation at CLE. Figure 1 and Table 1 on the following page reports example results; please refer to [11,20] for the full evaluation.

## Human Re-identification

We now address the second problem of human re-identification (re-id). Traditional biometric methods such as face or gait recognition are difficult to apply to the re-id problem since most surveillance cameras' resolution is too poor to get a clear image of the target. Instead, most recently proposed re-identification algorithms focus on feature representation and metric learning. Features used in re-id problems include color and texture histograms [12, 13. 14], quantized local feature histograms [15], and spatiotemporal appearance descriptors [16]. Many of these descriptors are high-dimensional and contain some unreliable features; hence metric model learning and feature selection are also critical problems for re-id. Many approaches have been proposed including online relative distance comparison (RDC) [17], RankSVM [14], and set-based methods [13]. Conte et al. [18] proposed the Multiview Appearance Model, which continuously updates an appearance model of the target for each quantized orientation.

| | | Single Camera | Multiple Cameras |
|---|---|---|---|
| Week 1 | Processed Data (hours) | 104 | 104 |
| | Correct Detections | 50/50 | 50/50 |
| | Missed Detections | 0 | 0 |
| | False Detections | 9 | 1 |
| Week2 | Processed Data (hours) | 168 | 168 |
| | Correct Detections | 70/70 | 70/70 |
| | Missed Detections | 0 | 0 |
| | False Detections | 14 | 3 |
| Week 3 | Processed Data (hours) | 168 | 168 |
| | Correct Detections | 63/70 | 70/70 |
| | Missed Detections | 0 | 0 |
| | False Detections | 9 | 2 |
| Archived Data | Processed Data (hours) | 440 | 440 |
| | Correct Detections | 190/190 | 190/190 |
| | Missed Detections | 0 | 0 |
| | False Detections | 32 | 6 |

**Table 1: Counter-flow detection results for three weeks testing at Cleveland Hopkins International Airport. Here we examined single and multiple camera settings. As shown our approach was able to achieve 100% detection in all cases. In addition, the use of multiple cameras was able to reduce false alarms by more than 80%.**

The critical issue for our problem of interest is that these previous algorithms are not generally viewpoint invariant or suitable to low-quality, low-resolution cameras. Also, the standard datasets used to evaluate the re-id algorithms are all images taken from cameras whose optical axis has a small angle with (or is even parallel to) the ground plane, which is generally not the case in real-world surveillance applications, where the angle between the camera optical axis and the floor is usually large.

In Year 5, we proposed a novel viewpoint-invariant approach to re-identify target humans in cameras that don't share overlapping fields of view. The approach is designed to be directly applicable to typical real-world surveillance camera networks. First, we introduced a sub-image rectification method to cope with perspective distortion. Second, we proposed a viewpoint-invariant descriptor that takes into account a pose prior learned from training data. Finally, we showed how discriminative features for a particular person can be learned online for adaptive re-identification. We tested our algorithms on both standard benchmarking datasets and a new, very challenging dataset acquired at a US airport, and demonstrate that the proposed algorithm yields better re-id performance compared to many previous approaches.

The key step in our new algorithm is the generation and estimation of a "pose prior" that takes into account the viewpoint dependence in the image of a person. We divide the image of a human into horizontal strips or sub-regions. In each strip, histograms of color and texture information are extracted that form a feature vector. Let $X_a$ and $X_b$ be two descriptors extracted from the images of targets A and B, and W be a classifier trained to distinguish between A and B. The pose prior is used to make the descriptor distance invariant to viewpoint changes, as represented by a new distance function: $f(I_a, I_b)=W^T|P(I_a, \theta_a)-P(I_b, \theta_b)|$, in which $I_a$ and $I_b$ are the images of targets A and B. $\theta_a$ and $\theta_b$ are the estimated viewpoint angles corresponding to targets A and B. $P(I, \theta)$ is the converted descriptor of I with respect to the pose prior at $\theta$. Instead of directly extracting descriptors for each strip of the target, $P(I, \theta)$ weights the contribution at each pixel of a strip based on the estimated pose of the target. Figure 5 on the following page illustrates the idea.

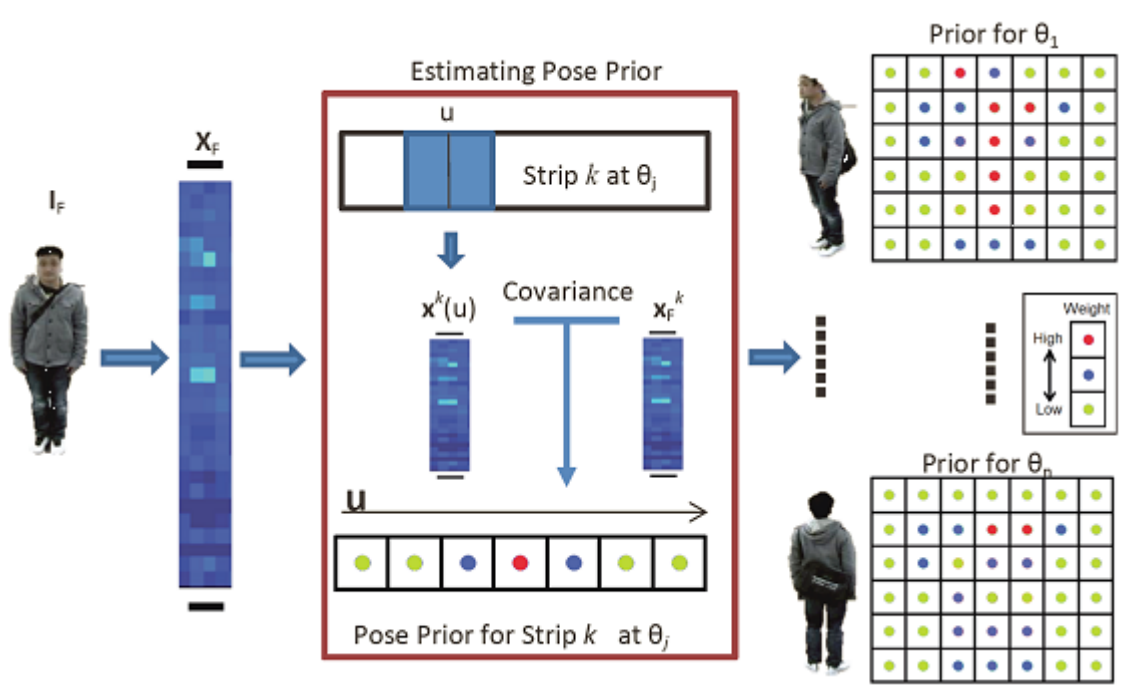


**Figure 5: Learning the pose prior.**

We conducted several experiments to evaluate the performance of the proposed algorithm, comparing our algorithm against state-of-the-art competitors on standard benchmarking datasets, as well as a new custom dataset created from videos from a surveillance camera network at a medium-sized US airport. We analyzed four synchronized video streams from this camera network, indicating a target in one view, and then automatically extracting descriptors of the target and detecting its reappearance in the other cameras.

| Method | 1 | 5 | 10 | 20 |
|---|---|---|---|---|
| Combined | **29.7** | **68.4** | **84.3** | **95.1** |
| SVM+PP | 21.2 | 56.7 | 76.8 | 88.7 |
| SVM+DF | 25.0 | 59.6 | 67.6 | 84.8 |
| SVM+R | 23.1 | 64.7 | 78.0 | 89.9 |
| SVM[19] | 17.2 | 50.8 | 66.0 | 83.6 |
| RDC[24] | 17.5 | 49.4 | 64.8 | 83.1 |

**Figure 6: Re-identification results comparing the proposed algorithm, SVM and RDC on the difficult airport camera network experiment.**

We tagged 86 targets in one of the four cameras. After the target leaves the camera in which he/she was tagged, 50 candidates are selected from the other views, only one of which is a correct match. For each probe image, we compute the result of the classifier for all the other images in the gallery set and record the rank of the true match. We compared our proposed algorithm to the high-performing RDC algorithm [17] and SVM algorithm [14]. The results of this realistic experiment are summarized in Figure 6. It can be seen that the proposed algorithms outperform RDC and SVM at every rank. Please refer to [19] for the full evaluation.

**Matte Propagation**

Another approach for solving the re-identification problem is by using ideas from the Matte Propagation problem [21-24]. Matte propagation is the task of tracking an object throughout an entire image sequence. First a rough outline of the object of interest is drawn, called a matte, and then this matte is propagated throughout the entire sequence. Matte propagation is commonly used in cinema post-production to mainly track actors. The estimated propagated mattes are then used to pull the actor from the original background

image and placing him over a different background using Matte Extraction [25-26].

We use ideas from Ring et al. [21,27] to solve the problem in hand of human re-identification. Once a matte is drawn around the object of interest, we extract KLT feature point trajectories around the examined matte. Those feature points are usually sparse in their spatial coverage. Hence we generate spare-to-dense spatial coverage using the image geodesic distance [27]. The idea here is that pixels not separated by a strong edge are treated as lying on the same object. Hence they are assigned the motion of the nearest feature point trajectory. A 2D affine model is then fit to the dense feature point trajectories and the matte is propagated according to this model.

Figure 7 outlines our approach in more detail. The first row shows the object of interest with yellow matte drawn around it. Feature points trajectories are extracted and the matte is propagated along the red arrows. The bottom row shows the geodesic distance of the person of interest. As shown the jacket is correctly segmented as one object as it is not separated by any strong edge. The same goes for the pants and the head of the person.


**Figure 7: First row; a yellow matte draw around the examined person and extracted feature points trajectories having green heads and red tails. Here the matte is propagated along the red arrows in the last image. Second row; examined person, image geodesic distance and dense motion segmentation.**

## *B. Major Contributions*

The overall breakdown of the research contributions of this project over the 5 years of ALERT are as follows:

Year 5: Transition of counterflow algorithms into operational evaluation at CLE; development of human re-identification algorithms.

Year 4: Development of counterflow algorithms, design and detailed evaluation of algorithms for keeping a pan-tilt-zoom surveillance camera calibrated.

Year 3: Construction of true-scale airport security screening checkpoint simulation; design and evaluation of real-time human tracking and baggage association algorithms

Year 2: Construction, design, and calibration of overhead camera array for human tracking in large, reconfigurable environment on RPI campus

Year 1: Design and evaluation of multi-object detection and tracking algorithms in single videos of crowded environments; automatic segmentation of multiple motions and detection of anomalies.

## IV. FUTURE PLANS

The counterflow algorithms have been evaluated for the past 9 months at the Cleveland airport in on-site testing as part of the video analytics transition task. Our current algorithm achieves real-time performance on the supplied test dataset with high detection rates and few false alarms. We are currently in the process of evaluating the performance of the algorithm as a function of a sensitivity threshold, which will be used to create a Receiver Operating Characteristic (ROC) curve that can be used to set a desired operating point. We are also integrating the algorithms with a hardware system that uses Zigbee wireless communication to create a visible and audible alarm at the exit lane podium, in preparation for creating a more generalizable, field-ready system. We are also changing our focus to the "tag and track" problem of following an indicated passenger through a large network of non-overlapping cameras at CLE, which will combine human tracking algorithms with our rank-based re-id approach. Hardware procurement and infrastructure for this project is underway.

## V. RELEVANCE AND TRANSITION

The presence of ALERT hardware and software on-site in Cleveland is expected to produce a wealth of new data and research problems of direct DHS/TSA interest for several years. The research projects in this area were directly motivated (and in fact, requested) by TSA officials as critical needs for their surveillance infrastructure. The F3-G team has visited the Cleveland airport on a nearly weekly basis since Fall 2012 to carry out the transition task research collaboration with Michael Young and Edward Hertelendy from DHS/TSA This substantial multi-institutional collaboration (RPI, NU, BU, Siemens) on-site at Cleveland allows the transitioning of algorithms designed in the lab to on-line problems of operational importance.

## VI. LEVERAGING OF RESOURCES

The computer vision algorithms addressed in this project are of direct interest to ALERT industry partners (e.g., Siemens), as well as other vendors of video analytics hardware and software (e.g., Bosch Security Systems, Honeywell, ObjectVideo, GE Security, 3VR, IDSS). Additional funding for this project from DHS/TSA is being pursued independently of ALERT core funding.

## VII. PROJECT DOCUMENTATION AND DELIVERABLES

### *A. Peer Reviewed Archival Publications*

1. Z. Wu and R.J. Radke, Keeping a Pan-Tilt-Zoom Camera Calibrated, Accepted at IEEE Transactions on Pattern Analysis and Machine Intelligence, July 2012.

### *B. Peer Reviewed Conference Proceedings*

1. Z. Wu, Y. Li, and R.J. Radke, Viewpoint Invariant Human Re-Identification Using Pose Priors, in review forIEEE International Conference on Computer Vision, April 2013.
2. G. Castanon, A. Caron, V. Saligrama, P. Jodoin, Exploratory Search in Long Surveillance Videos, ACM

Multi-Media (ACM MM) 2012

3. G. Castanon, A. Caron, V. Saligrama, P. Jodoin, Real-time activity search in Surveillance Videos, AVSS 2012

## VIII. REFERENCES

[1] D. Buzan, S. Sclaro , G. Kollios, Extraction and clustering of motion trajectories in video, in: Proceedings of the 17th International Conference on Pattern Recognition.
[2] I. Junejo, O. Javed, M. Shah, Multi feature path modeling for video surveillance, in: 17th International Conference on Pattern Recognition.
[3] J. Alon, S. Sclaro , G. Kollios, V. Pavlovic, Discovering clusters in motion time-series data, in: CVPR 2003.
[4] G. Antonini, J. Thiran, Counting Pedestrians in Video Sequences Using Trajectory Clustering, IEEE Transactions on Circuits and Systems for Video Technology 16 (2006) 1008–1020.
[5] P.H. Tu, J. Rittscher, Crowd Segmentation Through Emergent Labeling, Lecture Notes in Computer Science 3247 (2004) 187–198.
[6] E. Andrade, S. Blunsden, R. Fisher, Modelling Crowd Scenes for Event Detection, in: 18th International Conference on Pattern Recognition.
[7] G. Brostow, R. Cipolla, Unsupervised Bayesian Detection of Independent Motion in Crowds, in: IEEE Conference on Computer Vision and Pattern Recognition, volume 1.
[8] S. Ali, M. Shah, A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis, in: IEEE Conference on Computer Vision and Pattern Recognition.
[9] A.M. Cheriyadat, R.J. Radke, Non-negative matrix factorization of partial track data for motion segmentation, in: IEEE 12th International Conference on Computer Vision.
[10] A. Cheriyadat, R. Radke, Detecting Dominant Motions in Dense Crowds, IEEE Journal of Selected Topics in Signal Processing 2 (2008) 568–581.
[11] Z. Wu and R.J. Radke, Improving Counterflow Detection in Dense Crowds with Scene Features, in second review at Pattern Recognition Letters, Special Issue on Pattern Recognition and Crowd Analysis, submitted May 2013.
[12] D. Gray. Viewpoint invariant pedestrian recognition with an ensemble of localized features. ECCV , 2008.
[13] W. Zheng, S. Gong, and T. Xiang. Transfer re-identification: From person to set-based verification. CVPR , 2012.
[14] B. Prosser, W.-S. Zheng, S. Gong, and T. Xiang. Person reidentification by support vector ranking. BMVC, 2010.
[15] L. F. Teixeira and L. Corte-Real. Video object matching across multiple independent views using local descriptors and adaptive learning. Pattern Recognition Letters, 30(2):157–167, Jan. 2009.
[16] N. Gheissari and T. Sebastian. Person reidentification using spatiotemporal appearance. CVPR, 2006.
[17] W. Zheng, S. Gong, and T. Xiang. Re-identification by relative distance comparison. IEEE PAMI, 35(3):653–668, June 2012.
[18] D. Conte, P. Foggia, G. Percannella, and M. Vento. A multiview appearance model for people re-identification. AVSS, 2011.
[19] Z. Wu, Y. Li, and R.J. Radke, Viewpoint Invariant Human Re-Identification Using Pose Priors, in review forIEEE International Conference on Computer Vision, April 2013.

[20] M. Elgharib, and V. Saligrama, Counter-Flow Detection, Boston University Report Submitted to DHS, April 2013.
[21] D. Ring and A. Kokaram. Feature-Cut: Video Object Segmentation Through Local Feature Correspondences. In ICCV ’09, Japan.
[22] A. Kokaram, B. Collis and S. Robinson. A Bayesian framework for recursive object removal in movie post-production. In ICIP ’03, 937-940.
[23] G. Baugh and A. Kokaram. Semi-automatic motion based segmentation using long term motion trajectories. In ICIP ’10, 3009-3012.
[24] D. Corrigan, S. Robinson and A. Kokaram. Video Matting Using Motion Extended GrabCut. In CVMP ’08, London, UK.
[25] Y. Chuang, Brian Curless, D. Salesin and R. Szeliski. A Bayesian Approach to Digital Matting. In CVPR ’01, 264-271.
[26] A. Levin, D. Lischinski and Y. Weiss. A Closed Form Solution to Natural Image Matting. In CVPR ’06, New York.
[27] D. Ring and F. Pitie. Feature-Assisted Sparse to Dense Motion Estimation using Geodesic Distance. In IMVIP ’09, Ireland.