

Multi Target Tracking Access with Data Association in Distributed Camera Networks

Azmira Krishna, CMAK Zeelan Basha, Pradeep Raj Savarapu, Soumya Ranjan Nayak,
S. Sivakumar

Abstract--- *Data Association in Distributed Camera Network is a new method to analyse the large volume of video information in camera networking. It is an important step in multi camera multi target tracking. Distributed processing is a new paradigm to analyse the videos in camera network and each camera acts on its own and all cameras cooperatively work together to achieve a common goal, In this paper, we have addresses the problem of Distributed Data Association(DDA) to obtain the feet position of the object. These positions are shared with its immediate neighbours and find local matches using homography. By propagating these local matches across the network in order to obtain the global associations. In this proposed method DDA is less complex and improves the high accuracy compared to the centralized methods (STSPIE, EMTIC, JPDAEKCF, CSPIF, and CEIF).*

Keywords--- *DDA, Networks, STSPIE, EMTIC, JPDAEKCF, CSPIF, CEIF.*

I. INTRODUCTION

Owing to availability of low cost cameras and problems in single camera system, there is an increase in usage of camera networks. The applications of camera networks are visual surveillance, traffic control, sports analysis, and activity recognition. The problems in single camera system are occlusions and limited area coverage. Camera network is a collection of cameras connected together either wire or wireless. Camera networks can provide wide area coverage and different views of the same seen. These facilitate to solve the problem of occlusion. In our project we are assuming cameras are smart i.e., cameras can not only capture images but also they have process and communication capabilities.

Figure-(1) shows typical example of camera network. There are five cameras C_1, C_2, \dots, C_5 observing an area (black rectangle) consist of seven targets T_1, T_2, \dots, T_7 . The solid blue lines show the communication channels between different nodes.

Target Tracking is the process of locating the targets in an image or world plane in consecutive frames. Tracking the target is an important task used for behavioural understanding, autonomous vehicle navigation and activity recognition.

Manuscript received September 16, 2019.

Azmira Krishna, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

CMAK Zeelan Basha, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

Pradeep Raj Savarapu, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

Soumya Ranjan Nayak*, Chitkara University Institute of Engineering and Technology Chitkara University, Punjab, India. (e-mail: soumya.nayak@chitkara.edu.in)

S. Sivakumar, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

In camera networks, multi target tracking has an additional challenge of inter camera data association. Given the object measurement across views object association deals with finding the detections corresponding to same object.

To effectively develop a multi target tracking algorithm first we need to address the object association problem. Cameras are directional sensors and have obliged field of view (FoV). In a general camera mastermind a goal is secured by couple of cameras. This makes camera frameworks confront another test known as naivety [1]. The camera which does not have data about the objective from neighbours is insinuated as credulous camera. These cameras influence the last outcome in the appropriated preparing.

Authors in [3] composed Kalman consensus filter (EKCF) with Joint Probabilistic Data Association filter (JPDA), created JPDA-EKCF.

In any case, EKCF does not work admirably [1] on camera systems. In [2], the author developed Extended Multi Target Information Consensus filter (EMTIC) by extending Information Consensus Filter (ICF) to represent nonlinear estimation display and to coordinate with joint probabilistic data association filter (JPDA) to represent estimation vulnerability.

Another expansion of ICF was created in [4], to represent nonlinear process and estimation display. However in this paper the authors have not thought about false positive along with missed detections.

Both JPDA-EKCF and EMTIC depends upon expanded Kalman filter principles. Processing in camera networks can be done in two methods, centralized and distributed, see Figure (2). In centralized method, videos from all the cameras are sent to a central place where single camera techniques can be used.

But centralized method has the following disadvantages: Centralized method needs high communication bandwidth, high memory and high computational requirement at the central server. If the central server fails then whole network will fail. To circumvent these issues a new paradigm is developed known as distributed processing. The distributed process has the listed attributes: i) there would be no central server. Every camera follows up on its own. ii) Cameras complete nearby handling and communicate with prompt neighbours to repeatedly and enhance their outcomes. iii) The outcomes at every camera ought to be equivalent to that of centralized technique.

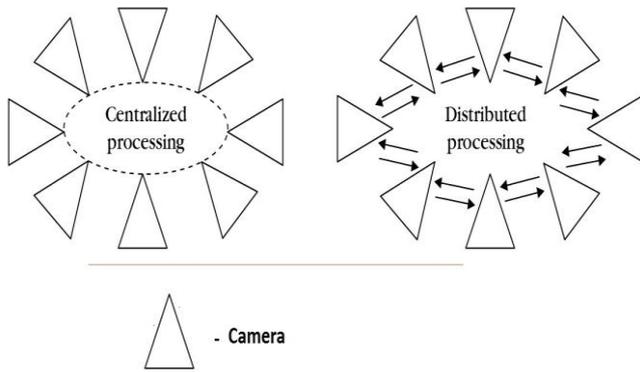


Figure 1: Different processing methods in camera networks

II. PRIOR RELATED WORK

Assignment algorithms like the Hungarian Algorithm [6] linked to the approach reported in [5] can be used as data association for single camera [7]. Greedy matching algorithms performance is same as Hungarian algorithm but with reduce complexity. Some approaches utilize the evaluation of a probabilistic state space [8]. Matching is another approach widely used for data association.

Berclaz et.al [9] used the output of object detectors to create Probabilistic Occupancy Maps (POM) to estimate the maximum number of objects and their locations, which are further used for association by graph matching. Another approach that uses matching was presented by Rudakova et al [10] where the matching operates on bipartite graphs. Matching can be used directly to relate the features of an object with the most likely ones over time as presented by Andersen et al [11]. In [12], the object was detected using standard Gaussian mixture model (GMM) and learning of ground plane homo-graphy imperative over various views. In [13], authors proposed a technique based on principal axes of people, to arrange individuals crosswise over different cameras. The likelihood correspondence is developed by the connection between "ground-points" of people recognized in every camera view and the convergence of fundamental axes detected in various camera views then it changes to the ground plane. This technique requires camera alignment and exact movement recognition presented in Figure 1.

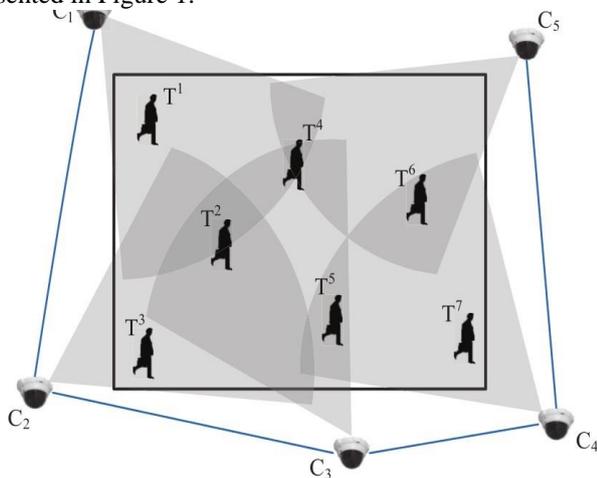


Figure 2: Typical example of a camera network

In this paper, we utilized a DDA algorithm in perspective of cooperative particle channels. The objective is followed in every camera view and ground plane. In any case, we use a circulated normal calculation which combines in constrained time, with the objective that we can get the correct intermingling to the consequences of brought together strategy. It in like manner diminishes the correspondence and computation cost. Here we report results on recreation information. Finally our proposed DDA method with less complex and high accuracy can be compared to other methods.

III. PROBLEM STATEMENT

Consider a set of cameras $C = \{ C_1, C_2, \dots, C_{N_c} \}$ with overlapping field of views (FoV) observing an area of interest and suppose that the targets are moving in this area. We assume targets are moving on the ground plane. We also assume that the homo-graphies between camera image plane and ground plane are provided. We consider the possibility of missed detections and false positives. Let us denote the object detections of camera i on image plane as $z_i^t, t = 1, 2, \dots, l_i$, where l_i indicates the number of detections.

Our problem is to get global object associations in distributed manner, i.e., every camera should know the correspondences between its estimations and the estimations of every single other camera, even though they don't communicate directly. Initially we considered centralized method; here our goal is to match the object detection across cameras assuming detections from all cameras are available.

In this work we use the ground plane homo-graphy limitation to relate protests in a multi camera framework. A 2D point (x, y) in a picture can be spoken to as a 3D vector $y=(y1, y2, y3)$ where $y = y1=y3$ and $z = y2=33$. This is known as the homogeneous depiction of a point and it lies on the projective plane P^2 .

A homo-graphy is an invertible mapping of focuses and lines on the projective plane P^2 . Distinctive expressions for this change consolidate collineation, projectivity, and planar projective transformation. The principle used in our work is explained in Figure 3. For simplicity, the following explanation is given for two cameras. The object feet positions are obtained using parts based HOG model [14-15]. Let us denote detections as z_i^p and z_j^q in the perspective of camera i and j , where p and q are real labels in the two perspectives before implementing object association z_j^p and z_j^q are considered as a set of corresponding points in the event that they can project as a similar point on the world ground plane. Given the homo-graphy network between the ground planes in perspective of camera i , and the world ground plane as H_i . We can locate the corresponding p point of z_i^p on the world ground plane G_i^p ,

$$\begin{bmatrix} G_i^p; 1 \end{bmatrix} = C_{ip} H_i \begin{bmatrix} z_i^p; 1 \end{bmatrix} \tag{3.1}$$

Where c_{ip} is a scaling factor to ensure above equation is fulfilled when homogeneous coordinates are used,

Similarly G_j^q can be find as

$$\begin{bmatrix} G_i^q \\ 1 \end{bmatrix} = C_{iq} H_j \begin{bmatrix} z_j^q \\ 1 \end{bmatrix} \quad (3.2)$$

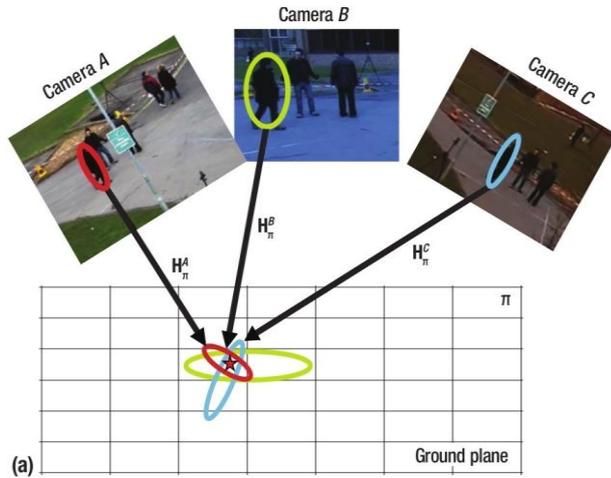


Figure 3: Figure shows object association among cameras that projecting each image plane locations to ground plane using the homography [14]

Where c_{jq} is a scaling factor to ensure above equation is fulfilled when homogeneous coordinates are used, H_j is the homography framework in middle of the ground plane in the perspective of camera j and the world ground plane.

Theoretically G_i^p and G_j^q ought to represent a similar point on the world ground plane.

But, due to noise in object detection and homography estimation, these two points may not concur with each other. In this way, we calculate the Euclidean separation between these two points as

$$D_g = \sqrt{(G_i^p(1) - G_j^q(1))^2 + (G_i^p(2) - G_j^q(2))^2} \quad (3.3)$$

Threshold T_g is set according to the tolerance of error. When

$$D_g < T_g \quad (3.4)$$

G_i^p and G_j^q are thought to be a similar point on the world ground plane, consequently objects that are localized at z_i^p and z_j^q ought to be allotted same labels. Else, they represent to various protests, and ought to have different labels. Following steps give the synopsis of the implemented method:

1. Video sequence is converted into frames. We set the number of targets to 2. Hence we considered only those frames in which two targets are visible in both cameras.
2. For each frame we run object detection algorithm. We used parts based on HOG model. The output of detection algorithm may contain false positives and missed detections.

Table 1: Dataset Results on applying centralized algorithm

Data set	No of Cam	1 target I/P-O/P errors	2 target I/P-O/P errors	3 target I/P-O/P errors	4 target I/P-O/P errors	5 target I/P-O/P errors
EPFL Lab6	2	10 - 0	49 - 0	108 - 10	180 - 18	160 - 20

3. These detections are mapped from image plane to ground plane using homography matrix.
4. We calculate Euclidean distance between each detection and each ground truth and ground plane location.
5. Using a threshold we matched detections of one camera to other.
6. We also calculated the number input errors and output errors.

3.1 Experimental results

We have performed experiments with the real world data. We used a Terrace video sequence [9] provided by EPFL CVLab. The video sequence was shot by 4 unsynchronized digital video cameras and synchronized the video streams by hand afterwards. All cameras were at head level covering the area of interest from different angles. The ground is at with a regular pavement. All cameras are provided with homo-graphies are provided between each camera plane and ground plane (top view). Among these four cameras we have considered two cameras in our experiments. One of the cameras (camera 2) is taken as reference camera i.e., objects detections of other camera are matched to the detections of this camera. Hence, initially for this camera we need to label correctly each target. This can be done using any standard matching algorithm. To avoid this step we have generated ground truth for this camera.

The results are shown in Figure 4. First column shows the frame number, second column shows the detections before object association and third column shows the detections after object association. In frame 15 two targets are assigned with correct labels. The target 1 is shown in red box and target 2 is shown in green box. In frame 51 we observe that the targets are assigned with wrong detection labels. As showed in column 3 of that row, detections are correctly labelled after object association algorithm. In frame 74 occlusion of one target has occurred. We have detected one object only. Correspondingly in output also we have one detected value in that camera. In frame 83, false positive has been detected, as shown in figure the object association algorithm resolves this issue also.

We say an error has occurred if a wrong label is assigned to the target. We calculated the input and output errors. The input errors are 49 and output errors are 0. First column shows the frame number, second column shows the detections before object association and third column shows the detections after object association. In Table 1, we will implement algorithm on each dataset which corresponds to all cameras and finding the accurate resulted output. Here input error is the detection in the frame wrongly assigned count, and output error means after data association, if there is any problems in assigning the correct target. Few datasets give output errors, without detection we can't get the object location. Using the filter is other case. But corresponding to the other camera we got the corrected results after correction we get accurate results.

MULTI TARGET TRACKING ACCESS WITH DATA ASSOCIATION IN DISTRIBUTED CAMERA NETWORKS



Figure 4: Figure shows object association results

1	0	0	1	0	0	0	0
0	1	0	1	0	0	0	0
0	0	1	0	1	0	0	0
1	1	0	1	0	1	0	0
0	0	1	0	1	0	1	0
0	0	0	1	0	1	0	0
0	0	0	0	1	0	1	1

Initial vector: Each camera i initializes n_i vectors each of length $m_{sum} = 7$. If there is a match then the component value is 1 or 0 otherwise.

1	1	0	1	0	1	0
1	1	0	1	0	1	0
0	0	1	0	1	0	1
1	1	0	1	0	1	0
0	0	1	0	1	0	1
1	1	0	1	0	1	0
0	0	1	0	1	0	1

Final vector: Cameras exchange vectors and if two measurements are related, then by the transitivity of the association, they are likewise connected with the various estimations related with them.

IV. PROPOSED METHOD ON DISTRIBUTED DATA ASSOCIATION ALGORITHM

In this segment we clarify distributed data association method.

Here our point is that, each camera should know the correspondences between its estimations and the estimations of each and every other camera, despite the fact that they don't convey straightforwardly.

In association with this we find neighbourhood facilitates between cameras i and its neighbours and communicate these crosswise network we get global associations.

Steps

Require: S_i and A_{ij}

Ensure: All the global matches and inconsistencies are found

Initialize m_i vectors

$$y_r^i(0) = [A_{i1}]r, 1, \dots, [A_{in}]r, m_n$$

Propagation of the matches

repeat

Send $y_r^i(t)$ to the neighbours

Receive $y_s^j(t)$ from the neighbours

for all $r, r' \in S_i$ do

if f_r^i is associated to f_s^j ($[A_{ij}]_{r,s} = 1$) then

$$y_r^i(t+1) = y_r^i(t) \vee y_s^j(t)$$

end if

if f_r^i and f_s^j are both associated to any f_s^j ,

$$([y_r^i]_s(t) = [y_s^j]_r(t) = 1, \text{ for some } s)$$

then

$$y_r^i(t+1) = y_r^i(t) \vee y_s^j(t)$$

end if

end for

until $y_r^i(t+1) = y_r^i(t), \forall r \in S_i$

Above algorithm gives the total method for proposed system (DDA).

We run this algorithm at each camera and at every moment time. We begin all cameras with the equivalent earlier qualities.

Since it distributes average algorithm convergence to exact average in finite time, the assembly of the proposed calculation to concentrated calculation is guaranteed.

Table 2: Dataset: Z is number of time instants

DataSet	N_c	Z	Scene Type
EPFL Terrace[10]	4	275	Outdoor
EPFL Lab6 [10]	4	121	Indoor
UCSB Hallay [12]	5	201	Indoor

Finding pairwise local matches

Every camera I foresee its own particular estimations on to ground plane either using homo-graphy or camera projection organize. The ground plane estimations are passed on among neighbouring cameras. All things

considered that each camera has ground plane estimations of its own and of its neighbours. Prior to we discover association between cameras i and the majority of its neighbours.

We relate an estimation of camera i with an estimation of neighbours camera j if the estimations are close on ground. We use Euclidean separation (at any rate some other separation measures can be used) to edge and pick whether estimations are close or not. In the event that camera i has n_i estimations, and cameras j and k are neighbours, at that point camera i figures two pairwise affiliation networks of size and, where and are the quantity of estimations in camera j and k separately.

In like way, every camera in the system processes pairwise affiliation networks

Finding global associations

Given pair bond relationship at every camera register overall relationship by spreading the area matches. Algorithm 1 in [16], concludes that each camera knows the complete number of estimations m_{sum} in system in a distributed manner using the normal consensus [17]. Camera i introduces n_i vectors everything about m_{sum} . These vectors are introduced by the area matches. The segments take esteem 1 or 0.

On the off chance that there is a neighbourhood coordinate the esteem is 1 or 0 generally. For instance (see Figure-5), three cameras related in line correspondence topology, and Cameras 1, 2 and 3 have estimations 3, 2 and 2 individually.

Expect both 1st and 2nd estimations at Camera 1 are connected with first estimation of Camera 2 and third estimation at Camera 1 is connected with second estimation of Camera 2, at point Camera 1 has $n_i = 3$ introductory vectors, everything about $m_{sum} = 7$. The underlying vectors of Camera 1 are: $\{1, 0, 0, 1, 0, 0, 0\}$, $\{0, 1, 0, 1, 0, 0, 0\}$ and $\{0, 0, 1, 0, 1, 0, 0\}$. Note that each estimation is connected with itself.

At each emphasis, the cameras trade the vectors with their neighbours.

On the off chance that two estimations are connected, at that point by the transitivity of the affiliations, they are additionally related with the different estimations related with them. In the above precedent, 1st vector of Camera 2 is $\{1, 1, 0, 1, 0, 1, 0\}$, in this way the 1st vector of Camera 1 gets refreshed as $\{1, 1, 0, 1, 0, 1, 0\}$. This updating continues till there are no modifications in the vectors. After culmination of the updates, every related estimation is assembled. In the above precedent, at Camera 1, the estimations are converged as $\{1, 2\}$ and $\{3\}$.

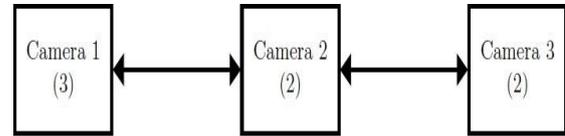
Table 3: MSE Values

Dataset	EMTIC	JPDAEKC	CSPIF	CEIF	DDA
EPFL Terrace [10]	0.9216	1	0.9107	0.9216	0.9130
EPFL Lab6 [10]	0.8005	1	0.5217	0.8005	0.546
UCSB Hallay [12]	0.8261	1	0.6540	0.8261	0.7400

Table 4: VOC results

Dataset	EMTI	JPDAEKC	CSPI	CEIF	DDA
t	C	F	F		
EPFL Terrac	0.8079	0.7755	0.8.93	0.807	0.805
				9	0

e [10]					
EPFL Lab6 [10]	0.6832	0.5661	0.756	0.683	0.694
			5	2	0
UCSB Hallay [12]	0.8097	0.8001	0.824	0.809	0.814
			7	7	2



Camera network: Three cameras connected in line topology. Number in parenthesis represents number of measurements in camera.

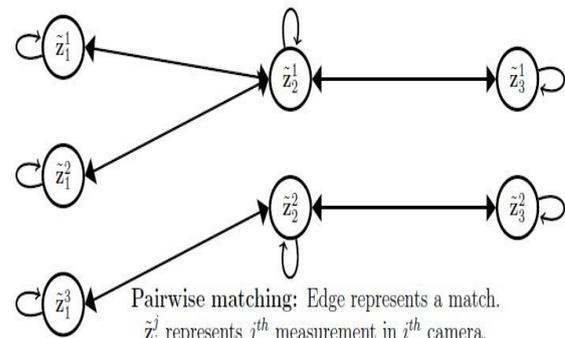


Figure 5: Camera Network: Three cameras connected in line topology

V. EXPERIMENTAL RESULT ANALYSIS

We performed experiments with different freely accessible multi-camera person on foot datasets. The dataset parts are given in Table 1. In datasets, people enter the locale one by one, walk around, stroll around, abandon one by one, causing regular and relentless impediments. The video is administered by number of targets. This urges us to survey the calculations as number of targets changed. We expected ring topology as correspondence graph. For each dataset we have made ground truth for one of the camera using Video Annotation Tool VATIC [18]. We got target positions by running discriminatively organized part based model calculation [19].

Step: 1 Initialize all detections of each frame in each camera using HOG model. 2. Obtain the matrix rows with distance calculated of camera of each target location projected on ground planes using homography. 3. Repeat the above process.

Step: 2 For all cameras and by giving threshold we form matrix with distance below threshold as 1's, otherwise 0's, 4. Dimension of matrix depends on the total targets in that frame corresponding to all cameras 5. Send the each camera, their information to neighbours and receive information from neighbour cameras 6; for all the targets of particular camera features/detections are there then 7. If the particular sender information and receiver information of value becomes 1 then 8. Association of "or" operation happen on

the particular column to row vectors. 9. end if 10. If the association of sender of targets in camera associated to the receiver of targets in all other cameras, means in column of sequence targets have 1 value 11. then do again "or" on the association happened and treat that targets as different / at that time increment the target representation number 12 end if 13, Until all the targets in that camera completed. Finally we can find the grouped as similar target number that represents others as different target representation.

In the above demonstrated outcomes, by differing the objectives, we run the algorithm for each dataset and watch the consequences of each objective in the camera tackled by the affiliation issue. Here any of the objective in edge allocated as target1 then this objective in some other camera comparing to a similar edge and demonstrates the equivalent target1 (speak to it in same shade of box). So we can watch the objective in any camera, same focuses in other camera will be coordinated and appeared with a similar shading box generally spoken to with other box. In the above shown results, by varying the targets, we run the algorithm for each dataset and observe the results of each target in the camera solved by the data association problem. Here any of the targets in frame assigned as target1 then these target in any other camera corresponding to the same frame and shows the same target1 (represent it in same color of box). So we can observe the target in any one of camera, same targets in other camera will be matched and shown with the same color box otherwise represented with other box.

VI. CONCLUSIONS

In this paper, we addressed the problem of data association in camera networks. In first part we solved the problem in centralized manner using homography constraint and showed results on real data. As experimental results showed that the error is reduced significantly. Next we developed distributed algorithm for data association. We discover nearby matches between cameras i and its neighbours and spreading these crosswise over system we acquire worldwide associations. Again we demonstrated the performance on various real world datasets and from results we observed that the implemented algorithm is able to assign labels correctly most of the time. The developed algorithm is easy to implement and fast. Although the developed algorithm performs reasonably good there is still scope to improve the performance. We performed experiments on real world and simulation data. The analysis results demonstrate that the DDA technique performs superior to JPDA-EKCF and EMTIC. The execution of our calculation is better when there is more noteworthy nonlinearity in the model. We are taking a shot at stretching out the proposed algorithm to different targets and to enhance its dependability. Our calculation is basic, simple to execute and appropriate. In spite of the fact that our calculation performs superior to the best in class (EMTIC) strategy and it has a decent execution.

REFERENCES

1. W.M. Baca, S. Jendro, M. Miller, J. Ryan, On irregular total labelling. *Discrete Math.*, 307 (2007) 1378-1388.

2. W.M. Baca, M.K. Siddiqui, Total edge irregularity strength of generalized prism. *Applied Mathematics and Computation*, 235 (2014) 168-173.
3. G. Chartrand, M.S. Jacobson, J. Lehel, O.R. Oellermann, S. Ruiz, F. Saba. Irregular networks. *Congr. Numer.*, 64 (1988) 187-192.
4. J.A. Gallian, A dynamic survey graph labelling, *Electron. J. Comb.* 19 (2012) 1-216.
5. A. Ahmad, O.B.S.A. Mushayt, M. Baca, On edge irregularity strength of graphs. *Applied Mathematics and Computation*, 243 (2014) 607-610.
6. H.W. Kuhn, The Hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2) 83-97 1955.
7. M.D. Breitenstein, F. Reichlin, B. Leibe, E.K. Meier, L.V. Gool. Online multiperson tracking-by-detection from a single, uncalibrated camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(9) 1820-1833 2011.
8. R. Mohedano, N. Garca, Robust multi-camera 3d tracking from mono-camera 2d tracking using bayesian association. *IEEE Transactions on Consumer Electronics*, 56(1) 1-8 2010.
9. J. Berclaz, F. Fleuret, E. Turetken, P. Fua. Multiple object tracking using k-shortest paths optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(9) 1806-1819 2011.
10. V. Rudakova, Probabilistic framework for multi-target tracking using multi-camera. applied to fall detection, 2010.
11. M. Andersen, R.S. Andersen. Multicamera person tracking using particle filters based on fore ground estimation and feature points. 2010.
12. X. Dai, S. Payandeh, Tracked object association in multi-camera surveillance network. *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 4248-4253 2013.
13. W. Hu, M. Hu, X. Zhou, J. Lou, S. Maybank. Principal axis-based correspondence between multiple cameras for people tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4) 663 2006.
14. N.S. Thakoor, L. An, B. Bhanu, S. Sunderrajan, B.S. Manjunath, People tracking in camera networks. Three open questions. *Computer*, 48(3) 78-86 2015.
15. P.F. Felzenszwalb, R.B. Girshick, D.M. Allester, D. Ramanan, Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9) 1627-1645 2010.
16. E. Montijano, R. Aragues, C. Sagues, Distributed data association in robotic networks with cameras and limited communications. *IEEE Transaction on Robotics*, 29(6) 1408-1423 2013.
17. A.Y. Kibangou, Finite-time average consensus based protocol for distributed estimation over AWGN channels. *IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, 5595-5600, 2011.
18. C. Vondrick, D. Patterson, D. Ramanan, Efficiently scaling up crowd sourced video annotation. *International Journal of Computer Vision*, 101(1) 184-204 2013.
19. P.F. Felzenszwalb, R.B. Girshick, D.M. Allester, D. Ramanan, Object Detection with Discriminatively Trained Part Based Models. *IEEE Transactions on PAMI*, 32 1627-1645, 2010.