Target Recognitions in Multiple Camera CCTV Using Colour Constancy

U Soori[1], *PWT Yuen[1,2], I Ibrahim[1], JW Han[3], K Hong[1], WT Chen[1], C Merfort[4], D James[1] & M Richardson[1]

Abstract — *People tracking using colour feature in crowded scene through CCTV network have been a popular and at the same time a very difficult topic in computer vision. It is mainly because of the difficulty for the acquisition of intrinsic signatures of targets from a single view of the scene. Many factors, such as variable illumination conditions and viewing angles, will induce illusive modification of intrinsic signatures of targets. The objective of this paper is to verify if colour constancy (CC) approach really helps people tracking in CCTV network system. We have testified a number of CC algorithms together with various colour descriptors, to assess the efficiencies of people recognitions from real multi-camera i-LIDS data set via Receiver Operating Characteristics (ROC). It is found that when CC is applied together with some form of colour restoration mechanisms such as colour transfer, the recognition performance can be improved by at least a factor of two. An elementary luminance based CC coupled with a pixel based colour transfer algorithm, together with experimental results are reported in the present paper.*

Keywords— CCTV surveillance; Illumination invariance detection; Colour constancy; Colour transfer; Colour feature

*corresponding author: email p.yuen@cranfield.ac.uk

1 Dept of Informatics & system Engineering, Cranfield University, Defence Academy of UK, Shrivenham SN6 8LA

2 Nanchang Hang Kong University, Nanchang, PRC

3 InnovTech Solutions Ltd, The Croft, Tewkesbury, Gloucester GL20 5TE

4 University of Siegen, 57068 Siegen, Germany

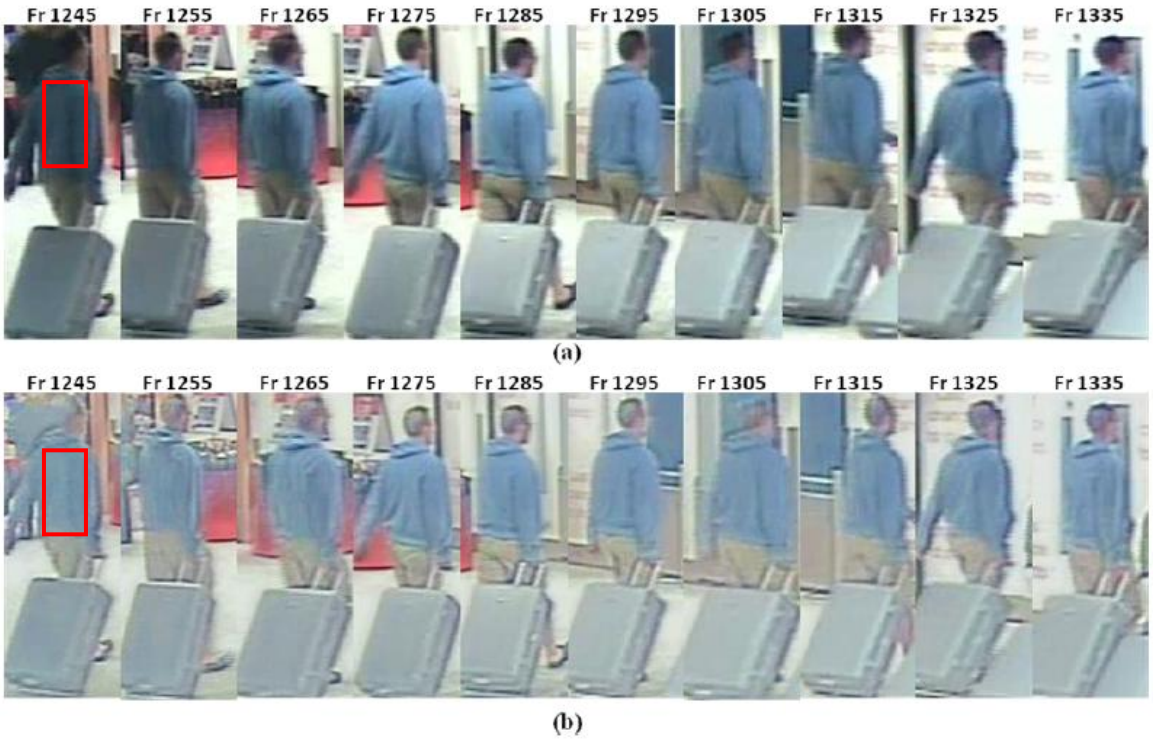
I. INTRODUCTION

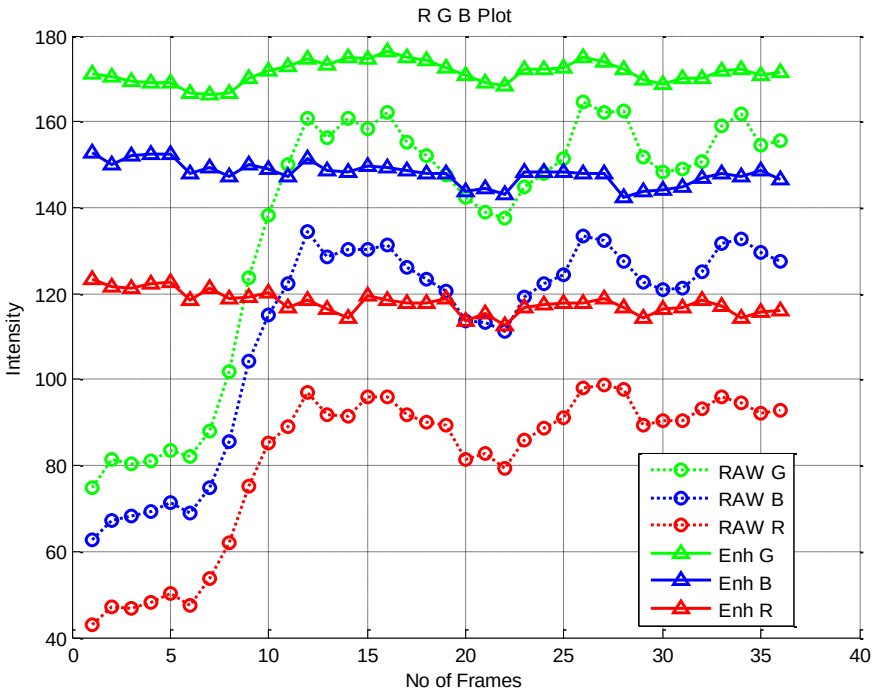
People tracking in crowded place using closed circuit television (CCTV) have been widely deployed for surveillance and security purposes particularly in strategic public places [22]. Regardless of the advances in machine vision technology, it is in fact still a daunting task to track a person from the CCTV footage. Part of the problem is due to the occlusions of targets in the crowded scene, and partly it is the variable lighting conditions that make target recognition very difficult in real scenarios [24, 25].

Conventional methods like face recognition cannot be realistically deployed in the street scene because of the low resolution of the CCTV vision for sensing large views from a distance which can hardly capture any useful facial features for recognitions. Alternatively, the use of the colour of people's dress may offer a more effective means for target recognitions. Unlike in the human visual system [22, 23], most of today's machine vision technology is incapable to perceive the actual colour of a target irrespective of illumination conditions. For instance, an object appears to be in different colours when it is under shade. Moreover, the same object may be seen very differently from the CCTV when it is viewed from different angles or through another TV in the network. Figure 1 highlights the effects of this illumination/viewing angle issues which exhibit undesired apparent colour variations when the target was walking towards the exit where the illumination intensity was a lot stronger than other parts of the room.

Figure 1(a) depicts the raw RGB images of the i-LIDS [21] video clip for every ~0.5sec intervals recorded by camera 1, and it is seen that the colours of the dress (overcoat) and the suit case appear in different shade of tone when the target is approaching towards the exit. Figure 1.1(b) shows the same video images but after colour constancy (CC) transformation

using a luminance based algorithm (see below for more information). Figure 1.1(c) shows the means of the red (R), green (G) and blue (B) bands extracted from the region of interest (ROI) for every 10 frames of the video data. In the raw data (circled plot) the colour attributes of the ROI exhibits an abrupt step change of values after ~2 seconds of the travels (FR1275) when the illumination intensity is much stronger from that spot. Colour constancy is the method which attempts to reduce effects due to non-uniform illumination artefacts. In the same plot also shows the result of the same data after transformed by the proposed colour invariance algorithm (triangle plot) which will be described in section III.





(c)

Fig. 1. Outline one of the most demanding issues in machine vision which still experiences great difficulty for tracking targets from CCTV footage. (a) Highlights the apparent colours of the target's overcoat which is seen changing with the illumination conditions. (b) Same as (a) but the video images are transformed by one of our colour constancy algorithm. (c) Depicts the mean RGB attributes of the target extracted from the ROI (red box in (a)) of the raw data and after colour constancy transformation in circled and triangle plots respectively.

This paper provides more evidence of how colour constancy (CC) can improve people tracking in real CCTV surveillance applications. During this course of the work, various forms of Retinex based CC algorithms have been assessed using real CCTV data. By using a luminance based CC together with a pixel based colour transfer algorithm, it is found that the target recognitions performance has been improved by at least a factor of two. Note that the colour characteristic is the only feature used for the target detection throughout this work.

II. OVERVIEW (PREVIOUS WORK)

There are a number of algorithms and methods proposed for colour constancy within the past two decades. Elementary methods like gamma adjustment, logarithmic compression and histogram equalization were found unable to produce colour constancy particularly when the image is taken under complex illumination condition. More advanced techniques based on Retinex [1,2,7,8] and its derivatives such as single scale Retinex (SSR) [3] and Multi-Scale Retinex (MSR) [3,4,5,6] have been found more useful for complex scenes having large dynamic range of pixel attributes. These algorithms tend to improve the colour perception of the scene through the 'gray-world' principle. While these algorithms manage to maintain the colour constancy of the scene, the resulting images tend to turn 'gray' and lose the original colour integrity, which is seen as one of the biggest drawbacks in machine vision particularly for target detections[25]. Some measures along this line have been using colour restoration (CR) such as in the MSRCR algorithm [3] which involves the estimation of colour factors from the raw data and it is then used to enhance local contrast after colour constancy transformation [9-11]. One problem of this approach is that the result is sensitive to the proportions of the 'colour distorted' pixels in the scene. Alternative approach has been the estimation of the luminous of the scene [12, 13] and the method has been applied in conjunction with a non-symmetric adaptive Gaussian function for the correction of 'halo-effects' in MSR [14].

Parameterisation in all Retinex based algorithms has been non-trivial and all free parameters such as iteration cycles for each spatial scale, Gaussian surround function parameters and the gains and offsets for the colour restoration, are needed to tune

manually. There are attempts in parameterisation using edge sharpness [15] but it lacks robustness and more research in this direction is needed. Alternative approach to the Retinex theory is the luminance perception based [16, 25] which estimates the reflectance of objects without the need of calibration standards to be present in the scene. This is by no means a more robust method and its performance is evaluated here together with the well-tried Retinex base colour constancy method in this paper.

III. ENHANCED LUMINANCE REFLECTANCE COLOUR CONSTANCY ALGORITHM [ELRCC] FOR CCTV NETWORK

One common way to synchronise spectral characteristics between multiple cameras is through camera calibration. Camera calibration has been an important factor for maintaining the white balance of the recorded images in each camera view within the multi-camera surveillance system. When this method is not available or when calibrations cannot be performed to achieve high degree of accuracy, then alternative method has to be employed for maintaining consistent colour characteristics over multi- camera network. In this paper we use a modified form of colour transfer principle [17, 25] for achieving multi-camera spectral synchronisation.

A. Colour Transfer

Colour transfer is the method for the correction of colour differences in two sets of images using statistical means. This method can be very useful for applications such as image analysis in multi-camera system. In real situations there are factors such as various illumination conditions, view angles and camera settings which all can induce colour distortions not intrinsic to the image. Initial efforts in this area have been using gain or exposure compensation for image intensity normalisation. Recent work by Erik Reinhard [17] has developed a colour transfer method for reducing illumination induced artefacts.

The method utilised simple statistics of two images and introduces a relationship between the colour of the target image and that of the source image through a transform as shown in equation (1).

$$I'_{j(x,y)} = \mu_s + \left(\frac{\sigma_s}{\sigma_t}\right)\left(I_{j(x,y)} - \mu_t\right) \tag{1}$$

Where *I'* is the intensity map of the scene which receives the characteristics from another camera *I* (designated as source) within the network, *j* represents the R-G-B components, *s* and *t* depict the information to be extracted from the source and target scenes respectively, σ and μ are the standard deviation and mean respectively. This method is termed as **CT** hereafter in this paper.

I is the one-band intensity value at pixel location (x,y) which can be evaluated using for example the RGB value components:

$$I(x, y) = \max[r(x, y), g(x, y), b(x, y)] \tag{2}$$

where *r*, *g* and *b* are the RGB components of colour images in RGB colour space. The intensity map I is the function of the luminance (L) and its reflectance (R) in equ. (3):

$$I(x, y) = L(x, y)\, R(x, y) \tag{3}$$

There are several ways to estimate the luminance L of the scene and one approach uses a low-pass filtering [13, 14] of the intensities I at (m,n) through a 2D discrete Gaussian G as shown in equation (4):

$$L(x, y) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m, n)G(m + x, n + y) \tag{4}$$

G is the 2D Gaussian at pixel location (x,y):

$$G(x,y) = q.e^{\left(-\left(\frac{x^2+y^2}{c^2}\right)\right)} \tag{5}$$

where *c is* the surround neighbourhood constant (values of the neighbourhood constant c (4~8)) and *q* is the normalization constant computed via $\iint q.e^{\left(-\left(\frac{x^2+y^2}{c^2}\right)\right)} dxdy = 1$. Thus the reflectance R can be obtained through equation (3).

B. Adaptive dynamic range compression of luminance

To achieve a balance of contrast enhancement and colour constancy, researchers in the field commonly use a non-linear transfer function, such as the Gaussian or the Windowed Inverse Sigmoid (WIS) function [16], to condition the luminance such that the dynamic range of the image is compressed into a user defined moderate range. Typical WIS transfer function is in the form of:

$$f(v) = \frac{1}{1+e^{-v}} \tag{6}$$

In practise the luminance is moderated into a range of user defined *vmin* and *vmax* through the transfer function:

$$L_n' = L_n\left[f(v_{\max}) - f(v_{\min})\right] + f(v_{\min}) \tag{7}$$

$$L_n'' = \ln\left(\frac{1}{L_n'} - 1\right) \tag{8}$$

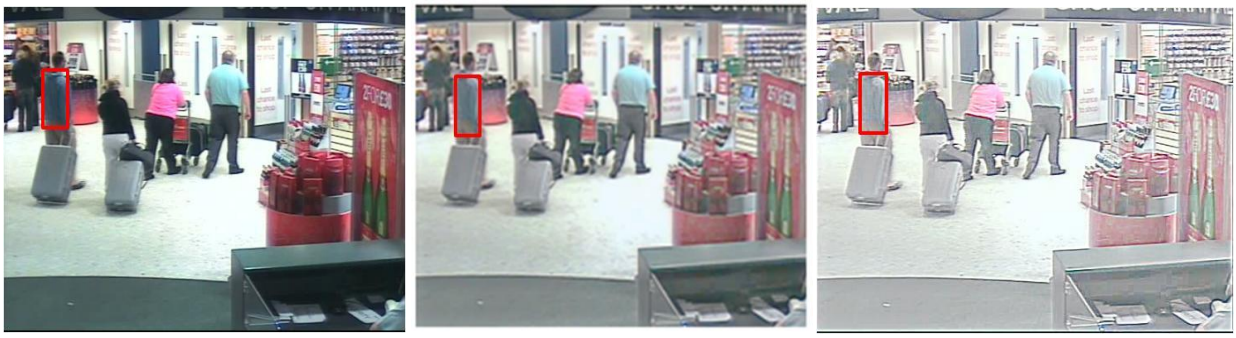
$$L_{n,enh} = \frac{L_n'' - v_{\min}}{v_{\max} - v_{\min}} \tag{9}$$

where Equation (7) linearly maps the normalized luminance to the magnitude range *f(vmin) f(vmax)* and equation (8) is the non-linear inverse sigmoid function. Equation (9) is the normalisation after the intensity transfer to bound the range of the luminance. Parameters *vmax,* and *vmin* controls the exact line shape of the transfer function which effectively smooth out the extremes of the intensity map. The range *vmin* and *vmax* affects the contrast enhancement, and the *vmax* can be set arbitrary while *vmin* can be extracted from the scene. Other authors [16] have been using different methods and in our case it is proposed to use a pixel-wise evaluation for the *vmin*. This will give more control over the details of colour correction: the dark pixels should be given a smaller *vmin* value whereas bright pixels should have larger value of *vmin,* and one form to achieve the objective is illustrate in Equation (10):

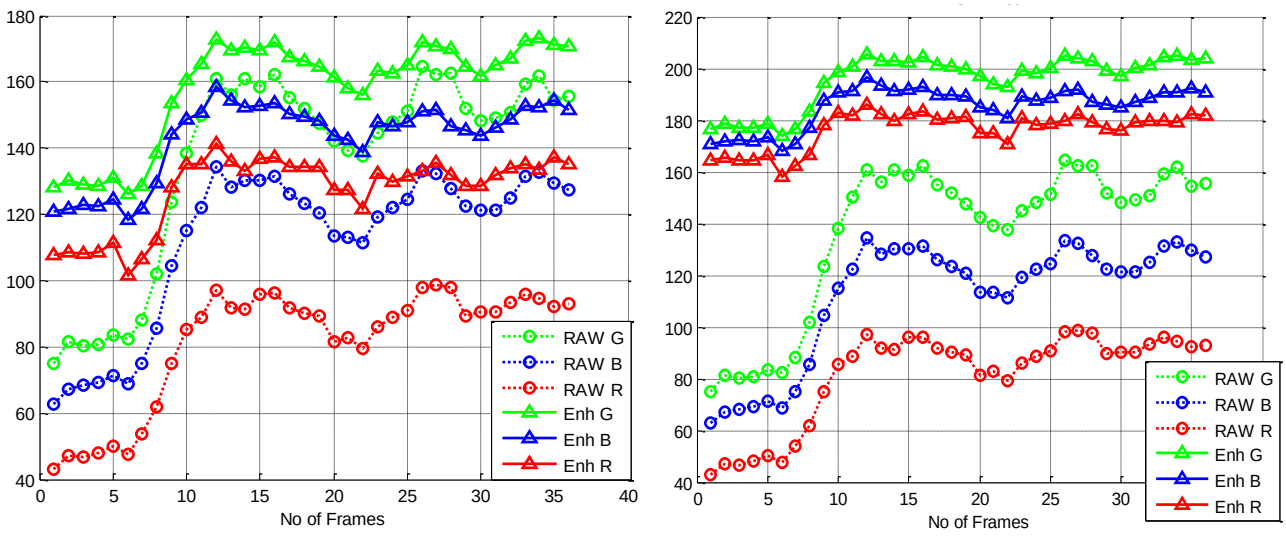
$$v_{\min} = \left(\left(\frac{I(x,y) - 1}{254} \right) 5 \right) - 6 \qquad (10)$$

C. Evaluation of Vmin in ELRCC: pixel-wise vmin

In the above section it is seen that the value of *Vmin* is a sensitive parameter for the colour correction through the sigmoid function. Conventionally this is a user's set fixed parameter, and most often the global mean of the scene has been commonly used to set this constant *Vmin* value for whole image. Figure 2 illustrates the effects of *Vmin* setting which directly affects the colour constancy performance. Figures 2 (b) & (c) are the processed results of the raw image in (a) by ELRCC using *Vmin*= -5 and -10 respectively. The images shown are the first clip of the video sequence. The spectral characteristics of the target extracted from the ROI of the red rectangle as depicted in the images (a-c) are shown in figure 2 (d) & (e) for the complete video clip using constant *Vmin* =-5 and -10 respectively. The sampling frequency is 5 frames/sec, and the colour attributes are in R,G,B. It is quite clear that the colour constancy is rather poor when *Vmin* of -5 is employed (figure 2 d), and the constancy is seen to improve a little when *Vmin* =-10 is used. However, the larger value of the *Vmin* also increases the colour bleaching. .



(a) (b) (c)

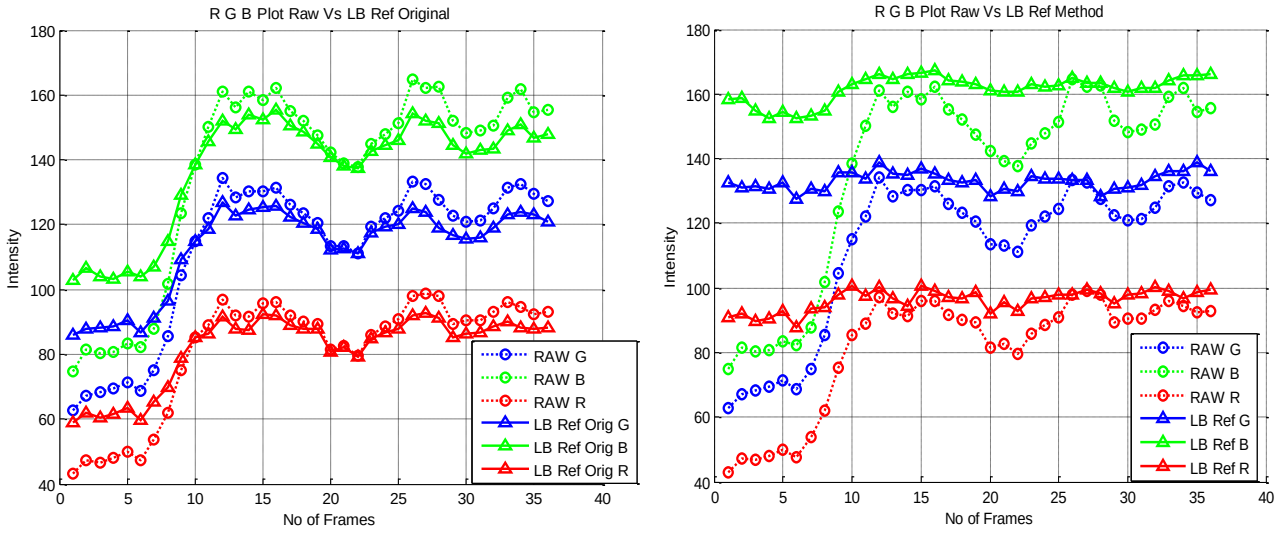


(d) (e)

Fig 2: Illustrates the effects of vmin settings on the colour constancy of the ELRCC algorithm. (a) raw data, (b) & (c) are the sample of processed results by ELRCC using *Vmin*= -5 and -10 respectively. The images shown are the first clip of the video sequence. (d) & (e) are the spectral characteristics (in RGB) of the target extracted from the ROI of the red rectangle as depicted in the images (b&c) for the complete video clip using *vmin*=-5 and -10 respectively. The ELRCC processed of the ROI is compared with that of the raw in both (b&c) and data sampling frequency is 5 frames/sec.

Thus instead of using a global *vmin* for the whole image, it is intuitively to evaluate the effect of using pixel wise *vmin* for the colour constancy behaviour. Figure 3 compares the effects of using a constant *vmin* for the complete image with that using a pixel wise evaluation according to equation (10). The colour characteristics of the video clips processed by the pixel wise *vmin* in the triangle trace of figure 3(d), which exhibits a much better colour constancy performance than that using the global *vmin* for the whole image (see figure 3 (c)).

(a) (b)



(c) (d)

Fig 3: (a) & (c) showing the effect of using a global vmin and to compare it with that using pixel wise vmin as shown in (b) & (d). Note that the image processed by the global vmin in (a) is found unable to rectify local colour non-uniformity due to variable illumination conditions.

D. Adaptive mid tone frequency components enhancement

Like all the Retinex algorithms, the luminance as depicted above exhibits a mid-tone and low frequency components which can be degraded by the dynamic range compression. A centre-surround type of contrast enhancement method can be utilized to help compensate this degradation:

$$L'_{n,enh}(x,y) = L_{n,enh}(x,y)^{E(x,y)} \tag{11}$$

where the exponent $E_{(x,y)}$ is defined by:

$$E(x,y) = \left(\frac{I_{conv}(x,y)}{I(x,y)} \right)^{p} \tag{12}$$

Where *I(x, y)* and the *Iconv(x, y)* are the luminances evaluated using larger values of the neighbourhood constant c (10~20) in equations (4-5). The $L'_{n,enh(x,y)}$ is the luminance after mid-tone frequencies enhancement. The exponent *P* is chosen to be a function of the global standard deviation in the I(x,y) which measures the extremeness of the intensity map. The exact value P is determined by experiments and it can be scene dependent. Typical values of P is (0.1~0.5) for the data set utilised in this study.

Given the knowledge of the luminance *L*, then the reflectance *R* can be achieved using the relationship from equation (3). Once we obtain final luminance $L'_{n,enh(x,y)}$, now the method is ready to combine reflectance *R* (from equ 3) using following equation to produce the enhanced one-band image I_{enh} as shown in equation (13).

$$\mathrm{I}_{\mathrm{enh}}(\mathrm{x},\mathrm{y}) = \mathrm{L'}_{\mathrm{n,enh}}(\mathrm{x},\mathrm{y})\,\mathrm{R}(\mathrm{x},\mathrm{y}) \tag{13}$$

E. Colour restoration algorithm

It is noted from the above sections that almost all colour constancy algorithm exhibit side effects of colour bleaching (see figure 2(c) as example). This is not desirable in target detection and to remedy this shortcoming, colour restoration is needed. The principle of colour transfer is to adjust the mean and standard deviation of each colour band from the target image to match them with that of the source image which can be any other scene taken by any other cameras.

It is found from this work that it is necessary to perform a colour transfer after the application of the colour constancy algorithm. Given a raw image *I* and the processed image I_{enh} obtained after the colour constancy luminance based algorithm the colour bleach in the I_{enh} can be somewhat compensated through the colour transfer coefficient.

$$\mathrm{C}_{\mathrm{CT}(\mathrm{x},\mathrm{y}),\mathrm{i}} = \mu_{\mathrm{s}} + \left(\frac{\sigma_s}{\sigma_t} \right) \left(\mathrm{C}_{\mathrm{Raw}(\mathrm{x},\mathrm{y}),\mathrm{i}} - \mu_t \right) \tag{14}$$

Where *i* depicts the R/G/B channels, C_{CT} is the R-G-B band after colour transfer, C_{Raw} is the raw R-G-B band at pixel (x,y) before the ELRCC, μ is the mean and σ is the standard deviation in each colour band. Colour information is stored in a single channel image I_{enh} via equation (15).

$$\begin{aligned} I_{enh(R)}(x,y) &= \left(\mathrm{I}_{\mathrm{enh}}(\mathrm{x},\mathrm{y})/\mathrm{I}(\mathrm{x},\mathrm{y})\right)\mathrm{C}_{\mathrm{CT\,r}}(x,y) \\ I_{enh(G)}(x,y) &= \left(\mathrm{I}_{\mathrm{enh}}(\mathrm{x},\mathrm{y})/\mathrm{I}(\mathrm{x},\mathrm{y})\right)\mathrm{C}_{\mathrm{CT\,g}}(x,y) \\ I_{enh(B)}(x,y) &= \left(\mathrm{I}_{\mathrm{enh}}(\mathrm{x},\mathrm{y})/\mathrm{I}(\mathrm{x},\mathrm{y})\right)\mathrm{C}_{\mathrm{CT\,b}}(x,y) \end{aligned} \tag{15}$$

where $I_{enh\ R}$, $I_{enh\ G}$, $I_{enh\ B}$ are the enhanced R-G-B components and $C_{CT\ r}$, $C_{CT\ g}$, $C_{CT\ b}$ are the output after colour transfer of the RGB components according to equation (14).

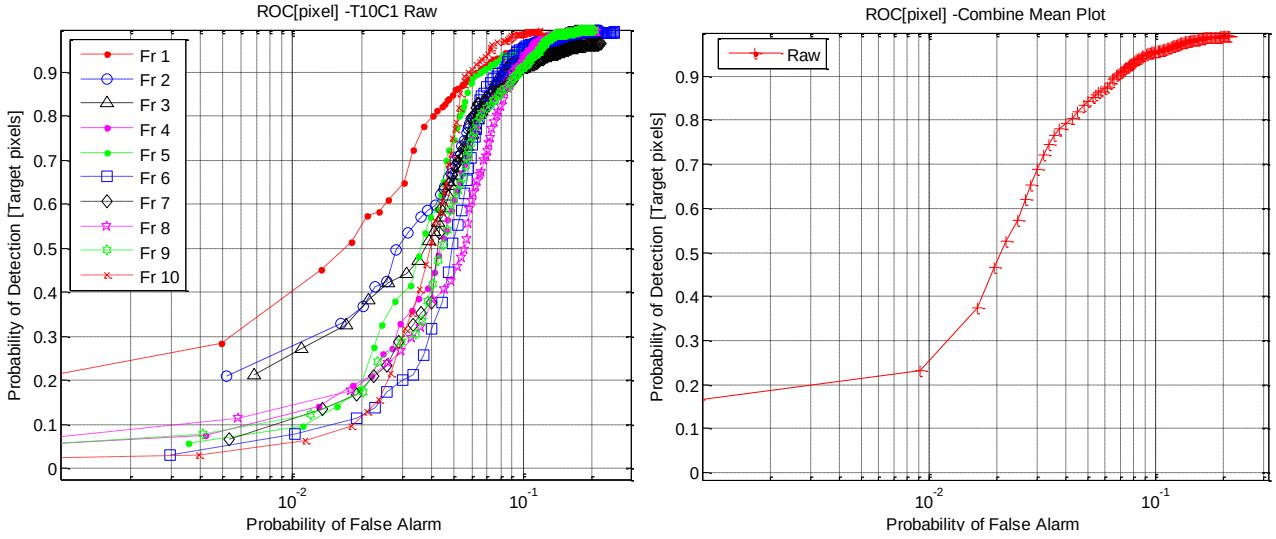
IV. EXPERIMENT

A. Data set and signature acquisition

The objective of this work is to evaluate whether colour constancy (CC) algorithm really helps people tracking from CCTV footage. To this end the i-LIDS multi-camera data set MCT-TR(1001-1005) g and h serials [25, 21] have been exclusively used in this study. Throughout all the detections, all the spectral features of target are acquired from the first THREE frames of the clips from CAMERA 1 (i.e. from MCT-TR1001) and they are then stored in the memory as the signatures of each individual targets. Note that the acquired spectral signatures from the first Frame 1 may NOT necessary fully represent the true characteristics of the targets due to the substantial illumination variations across the view in the scene of camera 1(TR1001 serial).

TWO different forms of the same data for the target detection performance assessment have been utilised here: A) raw data recorded by cameras (1,3,5) and B) after colour transfer from camera 1 into others (cameras 3,5). All targets presented in this work do not make appearances in cameras 2 and 4 and therefore target detection have been performed for the data recorded by camera 1, 3 and 5 only.

Fig. 4. Shows the representative pictures of three targets exploited in this study: (left to right) T10, T10B and T1



(a) (b)

Fig. 5. (a) Typical ROC results for the detection of a target in every 10 frames of a video clip when the target transverse across the scene. There are ~10 detections for a particular target and each produces one ROC. Note the spread of the ROC results within this short video clip as the result of strong illuminations across the scene. (b) To better visualise the detection performance for this particular target within this video clip, all ROC results as seen in (a) is averaged into one.

B. Targets and presentations of detection results

The purpose of the present work is to validate whether colour constancy helps target detections in multi-camera CCTV scenario. Two targets have been chosen, namely, the man with pale blue shirt (T10) and the lady with pink coat (T10B and representative pictures of these targets are respectively depicted in figure 4. Short duration of clips are selected whenever these targets appear within the camera field of views (views 1, 3 & 5), and the detection is made for every 10 frames of video clips and the detection result is presented in receiver operating characteristics (ROC) before and after the data set is processed by the ELRCC algorithm as described in the above sections. Within each video clip there is ~10 detections with ~10 ROC results (see figure 5a), which is not easy to visualise the effectiveness of the detection. Therefore, all these ~10 ROC results are averaged into one as shown in figure 5b. In all cases, the ratio colour descriptor (see next section) has been employed as the colour feature:

Ratio Feature (F2)

The ratio feature (F2) has been commonly exploited in machine vision as it is insensitive to the change of illumination intensity:

$$\begin{aligned} FR_1 &= R/G \\ FR_2 &= B/R \\ FR_3 &= G/B \end{aligned} \tag{16}$$

All processing involves colour transfer (CT), couple with the enhanced luminance reflectance colour constancy algorithm (ELRCC) and again colour transfer (CT) has been employed here prior to target detection (thereafter denoted as CT LB CT). Covariance based matched filter algorithm [19] has been used for the target detection throughout this work.

C. *Optimum colour descriptor*

It is well known that the selection of appropriate colour feature descriptor is detrimental to object recognitions. In this study we have employed 6 different colour descriptors in 3 colour spaces for the target recognitions over multiple camera views. Figure 6 present the averaged results of 10 frames of images for the detection of target T10 in camera 1 view using six different colour descriptors of RGB Feature (F1) [18], Ratio Feature (F2), Sum Feature (F3), Double Opponency (F4), L1L2L3 descriptor (F5) [18] and C1C2C3 descriptors F(6) [18]. It is quite clear that only the ratio (F2) and the C1C2C3 (F6) descriptors have exhibited rather good colour invariance properties with better detection performances (figure 6). Due to the popularity of the ratio descriptor (F2) and also it is the fact that it exhibits almost the best detection performance and spectral invariance, the ratio descriptor (F2) and C1C2C3 (F6) have been selected for assessing the effectiveness of colour constancy approach for target detections in the CCTV footage.

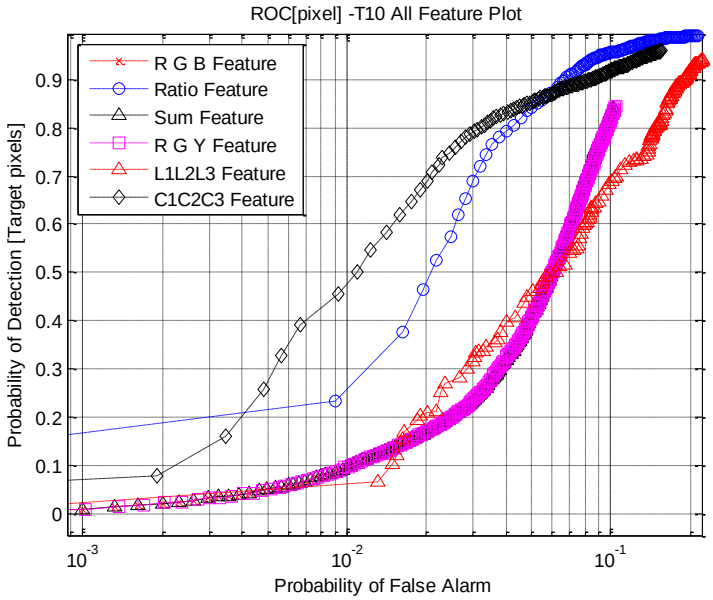


Fig. 6. Shows the averaged detection ROC of T10 raw data in camera view 1. Note that some descriptors, such as double opponency (F4) or sum feature (F3), exhibit very poor detection performances.

D. *Detection performances of various CC algorithms*

The effectiveness of CC algorithms can be seen from their object recognition performances

after the same video clip is processed by various CC algorithms. The experiment involves the averaging of 10 detection results for target (T10) while he was walking approaching to the exit of scene 1(camera 1). The mean of the ~10 detection ROC results is presented in figure 7. The ratio colour descriptor (F2) has been employed throughout this experiment. In the legend of fig 7 it depicts various CC algorithms employed in this experiment: (1) raw data (no CC), (2) after proposed ELRCC and colour transfer CT (denoted by LB CT), (3) after ELRCC but no CT (denoted by LB) (4) after sub-band CC [20] (denoted by subband MSR) (5) after LB MSR. Two main results have been observed from figure 7: (1) all CC algorithms produce WORSE detection result than that using the raw data, (2) ONLY the CC algorithms that enhanced by using the proposed LB CT improves target detection particularly in the low probability of false alarm (PFA) region.

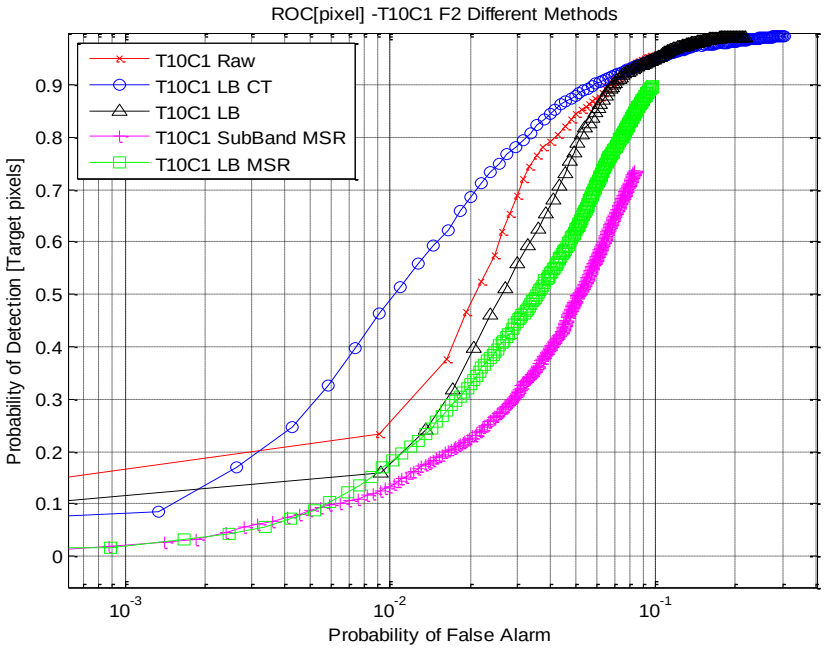


Fig. 7: shows the average of 10 detection results for target T10 in camera view 1 after processed by all the CC algorithms. Note that all CC algorithms, except for the one enhanced by the proposed method (in blue circle), exhibit WORSE detection result than that using the raw data (in red dot).

E. Does colour constancy really helps target tracking in multi-camera CCTV surveillance

The purpose of the present work is to validate whether colour constancy helps target

detections in multi-camera CCTV scenario. Two targets have been chosen, namely, the man with pale blue shirt (T10) and the lady with pink coat (T10B) (see figure 8). Short duration of clips are selected whenever these targets appear within the field of views of cameras (views 1, 3 & 5), and the detection is made for every 10 frames of video clips and the detection results are presented in ROC which are then averaged to form a representative ROC. The ratio colour descriptor (F2) has been employed and the CC processing involves (a) colour transfer (CT), (b) the ELRCC and (c) colour transfer (CT) which is denoted by CT LB CT prior to target detection.

Fig. 8. Shows representative images selected from the 3 camera views of target T10 (upper panel) raw data and (bottom panel) after transformed by the ELRCC+CT algorithm.

i. Target T10

Typical RGB images of target T10 before (raw image) and after processed by the proposed ELRCC algorithm (CT LB CT) have been shown in figure 8 (a) & (b) respectively. The target detection results for T10 in camera views 1,3 & 5 are presented in figure 9 (a, b & c) respectively. In camera views 1& 3 it is quite clear that the CC processed data (in blue trace)

exhibits consistent and improved target detections over that using the raw data for the detection (in red trace).

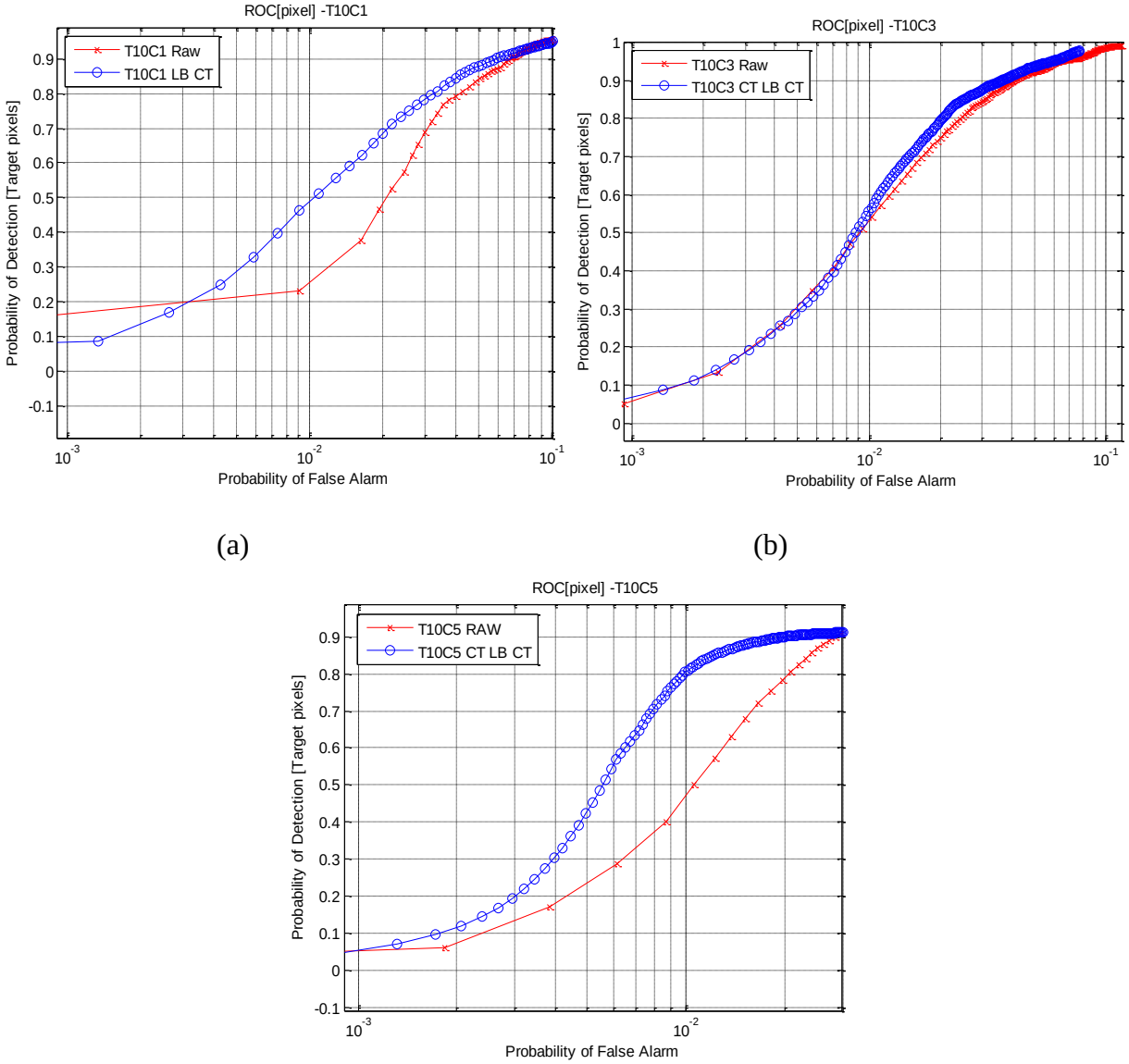
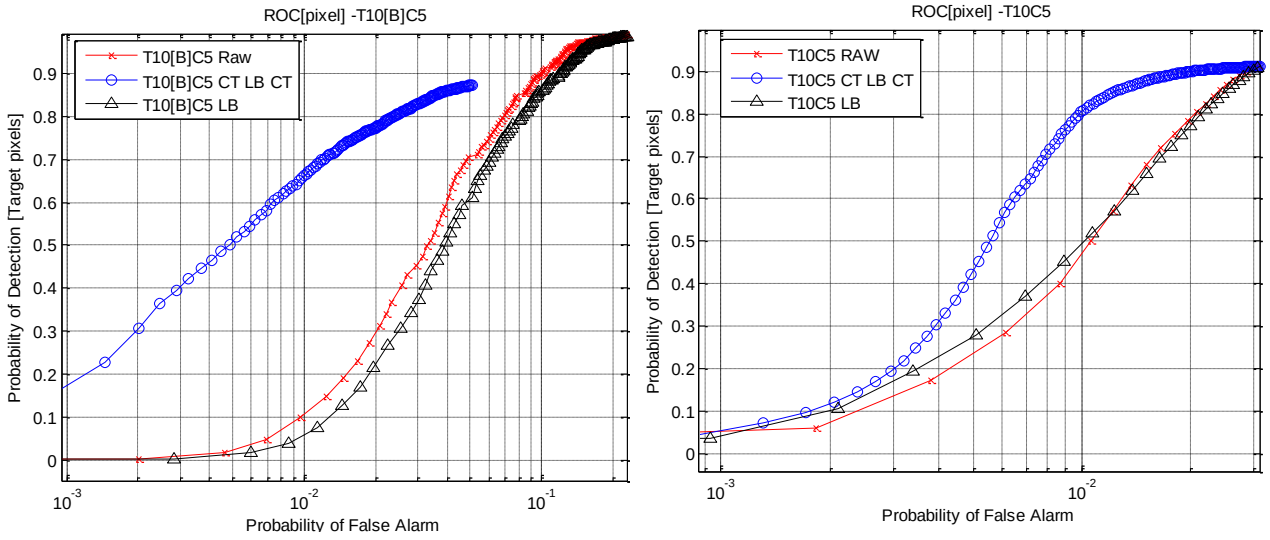


Fig. 9. Shows the mean of the ROC results for the detection of target T10 who has appeared in the three camera views of (a) camera 1, (b) camera 3 and (c) camera 5. The ratio colour descriptor (F2) has been employed in all cases, and colour transfer has been applied to cameras 3 & 5 before the detection.

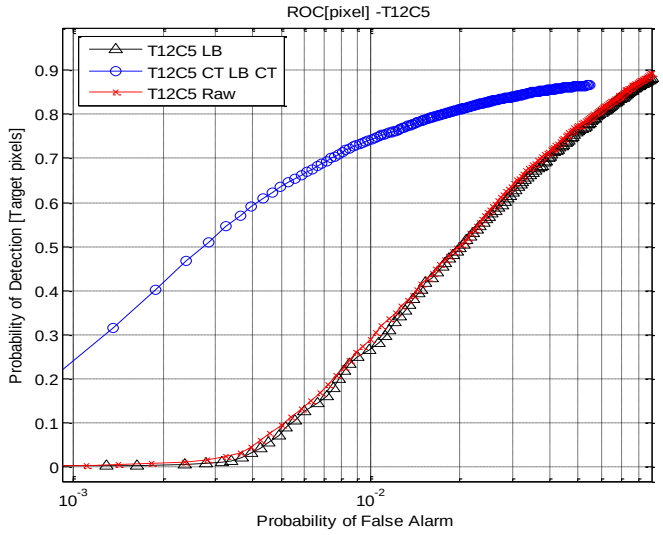
i. Target T10[B]

Figure 10 presents the ROC results for the video clip recorded by camera 5 after CC, with and without the colour transfer from camera 1. It is found that the detection enhancement is in

fact due to the colour transfer (CT) mechanism that we propose here in this work, showing a rigid shift of the ROC for the detection of all three targets with a substantial reduction of the PFA.



(a) (b)

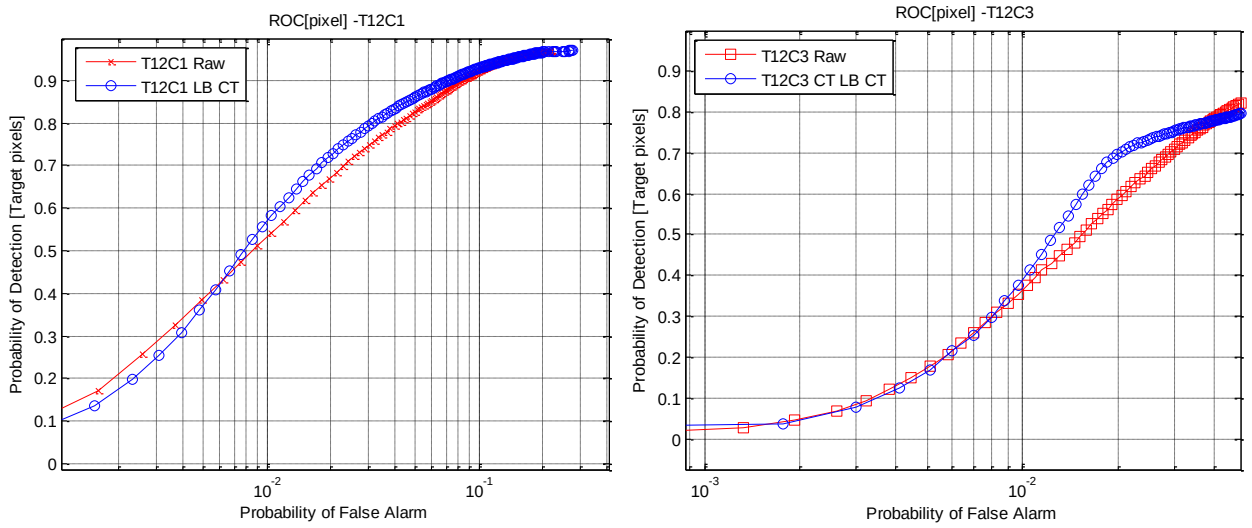


(c)

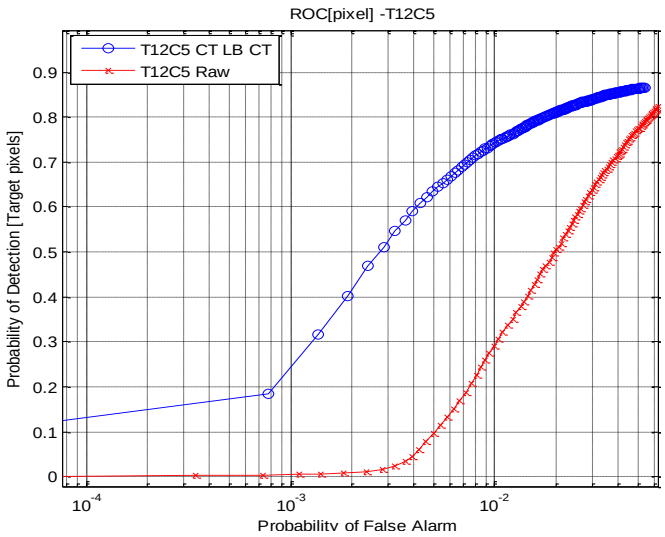
Fig. 10. Show the mean of the ROC results for the detections from camera view 5 for targets (a) T10 [B], (b) T10 & (c) T12 with (in blue) and without (in black) the colour transfer from camera view 1. It is seen that colour transfer scheme serves an important mechanism to help restoring the colour integrity of the image.

Target T12

The RGB images and the detection results for target T12 that presented in figure 11 has exhibited similar conclusion as that of the T10: the images after processed by the CT and ELRCC and CT have exhibited more consistent colours over different camera views, and at the same time it improves the target tracking performances rather substantially (see figure 11c). Note that the detection of targets, particularly in camera view 5 where the objects in the scene (people) appear to be small in pixel sizes, has been improved by at least a factor of two. For the probability of detection (PD) less than 0.8, the detection is enhanced by factors of 2 to 10 (see fig 11(c)).



(a) (b)



(c)

Fig. 11. Shows the mean of the ROC results for the detection of target T12 for three camera views of (a) camera 1, (b) camera 3 and (c) camera 5. The ratio colour descriptor (F2) has been employed in all cases, and again it is shown that great improvement in target detection has been achieved using the colour constancy approach proposed in this work.

V. DISCUSSIONS

The main results in this paper are presented in figures 9, 10 and 11, which have shown a consistent trend of detection enhancements after processed by the proposed colour constancy (CC) algorithm and the colour transfer (CT) treatment. Figures 9 and 11 highlight the improvement of target recognitions after CC and CT processing, while figure 10 demonstrates the importance of CT for restoring the colour integrity of the scene. It is also seen that most of the enhancement is seen from the camera view 5, with negligible detection improvements for the images recorded by cameras 3 and 1. It is speculated without proof that this may be related to the more extreme illumination conditions in camera view 5, where the rear view of the scene was greatly affected by the strong solar illumination through the window (see the right hand panel in figure 8). Furthermore, most targets in camera view 5 appear in the far view of the scene, such as the lady with the pink dress (T10[B]) (see the figure 8), which makes the detection much more difficult particularly when the spectral characteristics of the

very limited number of the image pixels in the targets are heavily distorted by the illumination artefacts.

VI. CONCLUSION

This study involves the development of colour constancy (CC) algorithm with an objective to understand whether target tracking in multi-view camera system is benefited from colour constancy (CC) technique. In this paper, we have reported the target detection efficiency in multi-camera CCTV system using several different forms of CC algorithms, and their detection efficiencies have been critically accessed via real i-LID data set. It is found that *all* CC Retinex algorithms suffer from strong colour bleaching, and compound to the fact that there is a lack of a principle way for the parameterisation of these Retinex based CC algorithms to obtain the optimal result. This paper reports an improved luminance based CC in which it incorporates an enhanced pixel-wise colour transfer mechanism, and together with an adaptive parameterisation for setting the frequency bound automatically using in-scene information. This enhanced CC algorithm has been testified for the detections of a number of targets across multi-camera views, and the result exhibits consistent improvement of target tracking over different camera views. The detection performance is enhanced most significantly when the targets appear to be small in number of pixel sizes, and the people recognition ability is seen to improve by factors of 2 to 10. It is also shown that the selections of appropriate colour space and colour feature descriptors are critically important in the target tracking.

ACKNOWLEDGMENT

The authors would like to thank the MOD/CPNI, Cranfield University, InnovTech Solutions and the Siegen University for the partial support of the project.

REFERENCES

[1] B. Funt, F. Ciurea, and J. McCann, "Retinex in Matlab," J. Electron. Imaging vol 13, pp. 48-57, 2004.

[2] L. Meylan and S. Susstrunk, "High Dynamic range image rendering with a Retinex based adaptive filter", IEEE Trans. Image Process. vol 15, pp. 2820-2830, 2006.

[3] D. J. Jobson, Zia-ur Rahman, G. A. Woodell, "Properties and performance of a center/surround retinex. IEEE Transactions on Image Processing" 6(3): 451-462 (1997)

[4] D. J. Jobson, Z. Rahman, and G. A. Woodel, "Retinex Image Processing: Improved Fidelity to direct visual Observation", Proc. IS&T 4th Color Imaging Conf. Color Sci. Syst. Appl., pp. 124–126 (1996)

[5] Z. Rahman, D. J. Jobson, and G. A. Woodel, "Multiscale Retinex for Colour Rendition and Dynamic Range Compression" Proceedings of SPIE 2847, 183 (1996)

[6] D. J. Jobson, Z. Rahman, and G. A. Woodel, "A Multi-Scale Retinex for Bridging the Gap between Colour Images and Human Observation of Scenes", IEEE Trans. Image Process. Special Issue Color Process. 6, 965–976 (Jul. 1997)

[7] E. H. Land, and J. J. McCann "Lightness and Retinex Theory", J. Opt. Soc. Am. 61, 1 (1971)

[8] E. H. Land, "An alternative technique of the computation of the designator in the Retinex theory of colour vision", Proc. of the National Academy of Science USA, 83, 2078-3080, May (1986)

[9] Z. Rahman, D. J. Jobson, and G. A. Woodel, "Retinex Processing for Automatic Image Enhancement" Proc. SPIE 4662, 390–401 (2002)

[10] K. Barnard and B. Funt, "Analyzing and improving multi-scale retinex," IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications, Scottsdale, Arizona, November 1997, pp. 221-226

[11] K. Barnard, G. Finlayson, and B. Funt, "Colour constancy for scenes with varying illumination," Computer Vision and Image Understanding, 65, 2, pp. 311-321 (1997)

[12] B. V. Funt, K. Barnard, M. Brockington, and V. Cardei, "Luminance-based multi-scale Retinex," AIC Color 97, Kyoto, Japan, May 25-30 (1997)

[13] B. Sun, W. Tao, W. Chen, H. Li, J. Li, and W. Guo, "Luminance Based MSR for Color Image Enhancement" International Conference on Computational Intelligence and Software Engineering (2010)

[14] B. Sun, W. Chen, H. Li, W. Tao, and J. Li, "Modified Luminance Based Adaptive MSR" Fourth International Conference on Image and Graphics (ICIG 2007)

[15] G. Orsini, G. Ramponi, P. Carrai, and R. Di Federico "A modified retinex for image contrast enhancement and dynamics control" Proceedings 2003 International Conference on Image Processing, Pages: III-393-6

[16] L. Tao, R. Tompkins, and V. K. Asari, "An Illuminance-Reflectance Nonlinear Video Enhancement Model for Homeland Security Applications". AIPR 2005: 28-35

[17] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley. "Colour transfer between images", IEEE Computer Graphics and Applications, 21(5):34-41, 2001.

[18] T. Gevers and A. W. M. Smeulders "Color-based object recognition", Pattern Recognition, 453-464, 1999.

[19] D Manolakis, D Marden, and G. Shaw,' Hyperspectral Image Processing for Automatic Target Detection Applications', Volume 14, Number 1, Lincoln Laboratory Journal, 2003

[20] J. H. Jang, B. Choi, S. D. Kim, and J. B. Ra, "Sub-band decomposed multiscale retinex with space varying gain", Proc. IEEE International Conference on Image Processing 2008,San Diego, U.S.A., Oct. 2008.

[21] http://www.homeoffice.gov.uk/publications/science/cast/ilids-user-guide?view=Binary
[22] Yuen PWT, "People tracking without prior information – a biological cortex-like neural approach" Proc. SPIE, Optics and Photonics for Counterterrorism and Crime Fighting, Vol. 7119, page 711908-711908-11 (2008). DOI: 10.1117/12.800315
[23] Yuen PWT, Tsitiridis A, Hong K, Chen T, Kam F, Jackman J, James D, Richardson M**, "**A cortex like neuromorphic target recognition & tracking in cluttered background", IET Digest, Vol. 2009 (2), 3rd International Conference on Imaging for Crime Detection and Prevention (ICDP-2009), pages 27-34, (2009) ISBN: 978 1 84919 207 1 doi:10.1049/ic.2009.0255
[24] I. Ibrahim; P. W. T. Yuen; A. Tsitiridis; K. Hong; T. Chen; J. Jackman; D. James; M. Richardson, "Illumination independent object recognitions in hyperspectral Imaging" SPIE Defence & Security. Toulouse, France: Proceedings of SPIE, 2010. 783800-1-12.
[25] U Soori; P. W. T. Yuen; I. Ibrahim; J. Han; A. Tsitiridis; K. Hong; T. Chen; J. Jackman; D. James; M. Richardson, "Colour invariant target recognition in multiple camera CCTV surveillance" SPIE Proceedings Vol. 8189 Optics and Photonics for Counterterrorism and Crime Fighting VII, (2011) DOI: 10.1117/12.898079

BIOGRAPHIES

Umair Soori is currently a PhD student at Cranfield University, based in Defence Academy of the UK. He received his BSc in biomedical engineering from Sir Syed University of Engineering and Technology, Pakistan and MSC in Electrical and Electronic Engineering from Universiti Sains Malaysia (USM), Malaysia. His research interests include computer vision and pattern recognition.

Peter Yuen (M'2009) is fellows of Institute of Physics (FInstP) and Institute of Mathematics and its applications (FIMA) since 2001 and 2002 respectively. Peter obtained his Ph.D in semiconductor physics from Imperial College, University of London in 1987 and subsequently he worked in this field till 1997. Then he joined the Asahi Chemical Industry Co. in Japan for two years. In 1999 he worked as a principal scientist for the UK defence agency DERA in the area of submarine detection. In 2003-2007 he was a senior principle scientist with the BAE Systems UK for the development of hyperspectral remote sensing technology. In May 2007 Peter joined Cranfield University as a senior lecturer to lead a group on Electro-optics, remote sensing & computer vision research. Peter has about 75 journal and conference papers in the fields of semiconductor physics, computer vision and defence science.

Izzati Ibrahim is currently a PhD student at Cranfield University in the Sensors Group based at the Defence Academy of the UK. She has a BSc and MSc in physics from the National University of Malaysia. Her current research interest is in hyperspectral image processing.

https://dspace.org/

Target recognitions in multiple camera CCTV using colour constancy

Soori, Umair; Yuen, Peter W. T.; Ibrahim, Izzati; Han, Ji Wen; Hong, Kan ...

2013-06-28T00:00:00Z
SPIE - International Society for Optics and Photonics
http://dx.doi.org/10.1117/1.OE.52.4.047202

Downloaded from a DSpace repository. https://dspace.org/