Object Colour Extraction for CCTV Video Annotation

Muhammad Fraz, Iffat Zafar and Eran Edirisinghe
Computer Science Department, Loughborough University, Loughborough, U.K.

Keywords: CCTV Videos, Colour Extraction, Video Indexing, Video Annotation.

Abstract: In this paper, we have addressed the problem of object colour extraction in CCTV videos and proposed a frame work for efficient extraction of object colours by minimizing the effect of variable illumination. CCTV videos are generally very low quality videos due to significant presence of factors like noise, variable illumination, colour of light source, poor contrast, camera calibration etc. The proposed frame work makes use of conventional Grey World (GW) Colour Constancy (CC) method to reduce the effect of variable illumination. We have proposed a novel technique for the enhancement of colour information in video frames. The framework improves the results of colour constancy system while maintaining the actual colour balance within the image. Colour extraction has been done by quantizing HSV space into bins along 'Hue', 'Value' and 'Saturation'. A novel set of procedures has also been proposed to fine tune the extraction of white colour. Finally, temporal accumulation of results is performed to increase the accuracy of extraction. The proposed system achieves accuracy up to 93% when tested on a comprehensive CCTV test dataset.

1 INTRODUCTION

The use of CCTV cameras has enormously increased for security surveillance Networks of CCTV cameras are in operation to assist security agencies to keep an eye at suspicious individuals and record unusual events. With this increase in the use of CCTV surveillance, need for appropriate organization of growing collection of video data has also increased. It is an extremely hectic task to manually index the video data by assigning meaningful annotations to rapidly search through these videos in case of a certain event. In this case, events are typically derived either a vehicle passing or a person walking through the scene. Colour extraction in this type of applications is operated as a task of finding persons with clothing of a particular colour or vehicles of particular colours.

This work has investigated the extraction of colour information in poor quality video frames for its effective use as an indexing attribute for annotation of video databases. The focus is to extract colour information of the moving objects such as humans and vehicles. The proposed system works as a part of a complex video annotation framework that is capable of extracting foreground moving objects and classifies them as a human, vehicle, baggage etc. It executes queries like: bring all video frames containing: a human, a human with baggage or a vehicle. The search can be more meaningful if the capability of system is enhanced to process more specific queries like: bring all video frames containing: a human with red top/shirt or a vehicle with green and white colour.

We have devised a method for enhancing the true colours of objects. The method works in conjunction with colour constancy algorithm. Secondly, a histogram peaks based method is proposed to calculate the colour of moving object. In addition a procedure is proposed to perform a confirmation check for white colour. Lastly, the temporal information is used to increase the confidence of calculated colour for every object.

The paper is organized as follows: After a brief review of related work in section 2, Section 3 explains the proposed methodology. Experimental results and discussion is presented in section 4 followed by the conclusion in section 5.

2 LITERATURE REVIEW

Colour is an important attribute for efficient visual processing. According to Schettini et al., the problem of colour extraction is based on either predefined conversant colours or on a query

Fraz M., Zafar I. and Edirisinghe E..
Object Colour Extraction for CCTV Video Annotation.
DOI: 10.5220/0004279404550459
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2013), pages 455-459
ISBN: 978-989-8565-47-1

illustration. Our work addresses only the latter case and therefore requires a query input. Brown proposed a very similar system for retrieval of predefined familiar object colour. Her framework accumulates a histogram of coloured pixels for a small number of human perceived colours. Parameterization of this discretization is performed to determine the dominant colour of the object. Swain provided the initial idea of colour recognition based on colour histograms, which are matched by histogram intersection. Modifications in this idea contain improvements upon histogram measurements, incorporating information about the spatio-temporal relationships of the colour pixels. Our method is also based on histogram binning technique along with temporal accumulation of results on various frames to enhance the accuracy of true colour extraction.

Wui et al. addressed the task of colour classification into pre-specified colours for tracked objects. Weijer et al. and Zhang et al. proposed probabilistic latent semantic analysis (PLSA) and Co-PLSA based approaches for object colour categorization in videos. These methods rely on complex features like SIFT and MSER to articulate the objects into various parts, i.e. tyres and windows for vehicles, and separate them to reduce the effect of their colour in categorization of vehicle's main colour. These methods require extensive processing which make them less suitable for real time applications.

Colour constancy has supreme importance for accurate extraction of object colours in videos and images. The most recent work on colour constancy uses a Bayesian approach to solve for the illumination conditions (Manduchi, 2006); (Shaefer et al., 2005); (Tsin et al., 2001). Renno et al. evaluated the advantages of two classic colour constancy algorithms (grey world and gamut mapping) for surveillance applications and found both algorithms to improve colour with gamut mapping resulted in small error than grey world. However, we are relying on computational colour constancy and hence using GW because of its simplicity and least processing time among all CC techniques.

3 METHODOLOGY

The process of object colour recognition is carried out in four major steps; colour correction, colour space conversion from RGB to HSV, pixel clustering and fine tuning.

3.1 Colour Correction

The colour correction of input video frames is carried out in two major steps. Conventional colour constancy technique is followed by a set of processing procedures to achieve true colours of the object present in the video frames.

Colour constancy (CC): Colour constancy is extremely important to reduce the effect of illumination and surroundings. It is impossible for a colour recognition system to perform well without colour constancy. We took advantage of existing techniques and tested computational algorithms on a large number of CCTV videos. Grey World, Max RGB and Grey Edge algorithms found to perform approximately similar. We decided to use Grey World because of its less computational complexity which makes it a suitable to a real time application.

Post CC Enhancements: The output of colour constancy system is processed in HSV space to boost contrast and brightness. CCTV videos are generally poor in quality that even after colour constancy procedure the actual colours of objects appear dull and indistinguishable. The proposed method applies modification in 'Saturation' and 'Value' components of HSV pixels as shown in equations 1 and 2.

The original Value represented as V_o of all pixels is scaled in a way that the lower V_o values have higher scaling factor while the scaling factor for higher values decreases gradually. Modified pixel values have been represented using a Quadratic Bezier Curve as shown in figure 1a. The saturation 'S' of all those pixels that have the values less than a threshold t_s is scaled up by a factor f_s.

$$V_n = \alpha^2 P_0 + 2\,\alpha\, P_1 + (1-\alpha)^2 \quad \textit{if } 0 < (1-\alpha) < 1 \qquad (1)$$

$$S_n = S_o * f_s \qquad \textit{if} \quad S_o < t_s \qquad (2)$$

where, $\alpha = (1 - V_o)$, S_o is the Original Saturation 'S' of pixels, V_n is the newly calculated Value of pixels, S_n is the modified Saturation of pixels and t_s is the Upper threshold for saturation 'S'.

P_o, P_1, P_2 are the constants in Quadratic Bezier Curve. The values, $P_o = 0$, $P_1 = 1$ and $P_2 = 1$ worked best for videos present in our dataset. The quality of video gets remarkably enhanced by this method. Figure 2 shows the effect of these enhancements on a video frame.

3.2 HSV Quantization

The next step after colour space conversion is the quantization. The HSV space is quantized into 450

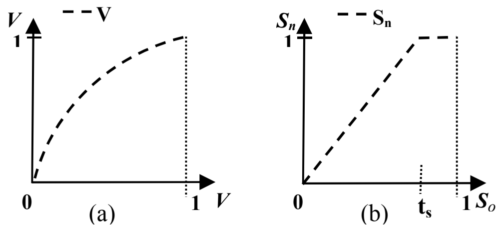

Figure 1: (a) Old vs New 'V' (b) Old vs New 'S' with factor f_s = 1.25 and upper threshold $t_s = \gamma$ where, $0 \leq \gamma \leq 1$.

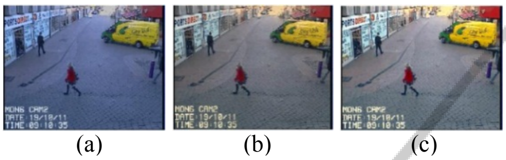

Figure 2: (a) Input video frame. (b) Colour Constancy (c) Enhanced Colour Constancy in HSV domain.

(18x5x5) bins and the object colour is decided upon the highest colour peak in the HSV histogram.

The aim of our application is to find three dominant colours of an object. Therefore, three highest peaks are considered as the most dominant colour of the object. The reason for extracting more than one dominant colour is to tackle the inaccuracies in object segmentation algorithm.

So far, no object segmentation algorithm has been able to achieve 100% accuracy. It is a quite common problem with object segmentation algorithm to pick the shadow or some area of background as part of object. The shadow appears longer than object during afternoon or early in the morning, therefore, the colour of shadow region forms highest peak in HSV histogram and the actual object colour form second or third highest peaks.

3.3 Fine Tuning of White Colour

The chances of error in searching a white colour object are quite high. The reason being, white colour always reflects its surroundings and it may appear in the shades of blue or off white.

For example, consider a scenario where there is a white car on the highway on a sunny day. The white car reflects sunlight and most of its pixels have values closer to light yellow or light blue colour of the sky. In this case, when the pixel clustering is performed, most pixels will fall in the category of light blue or yellow colour and the vehicle's dominant colour will be extracted as light blue or off white. In order to reduce the chances of missing out a white object, some mechanism is required to fine tune the results of colour extraction algorithm.

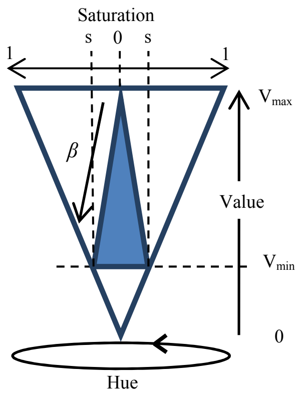

Figure 3: The dark triangle is highlighting the region (on HSV plane) the pixels of which are more likely to be the pixels of a white colour object.

A new method has been proposed here to reduce the chances of missing a white colour. The average colour of a candidate object is calculated and used to identify whether the object is likely to be a white object. The average Saturation 'S' and average Value 'V' of candidate object is checked according to the criteria shown in figure 3.

If the average Value 'Vavg' and Saturation 'Savg' of the candidate object lies within the dark triangular region as shown in figure 3, chances are high that the object is a white coloured object. This criterion is tested by equating 'Veq' of pixels against average 'S'. If, 'Vavg ≤ Veq' then the object is a white coloured object. In this case, the third dominant colour of the object is amassed. The value of 'Veq' is calculated by using equation 3.

$$V_{eq} = \beta\,(S_{avg}) + \phi \qquad (3)$$

where,

β = slope of the cone that selects 'V' against 'S'

ϕ = constant value of y-intercept.

S_{avg} and V_{avg} of the pixels in the biggest cluster are calculated. If S_{avg} is less than 0.45 then V_{eq} is calculated and compared with the average Value of the pixels in the clusters. If V_{avg} of the pixels in the cluster is smaller than V_{eq} then chance are really high that the candidate object is of white colour. Therefore, the third biggest cluster colour is replaced by white colour. The probability of missing a white object due to poor illumination has greatly reduced through this approach.

3.4 Temporal Accumulation

The colour of every single object is extracted in each frame until the object goes out of the frame. We are taking the advantage of presence of candidate objects in more than one frame. The colour of each object is extracted in every consecutive frame. As soon as an object leaves the frame the extracted colour of that object in analysed in each frame by majority voting and decision is finalized accordingly. The chances of wrong extraction that may occur because of occlusion or inaccuracy of foreground extraction algorithm are greatly reduced.

4 EXPERIMENTS AND RESULTS

The proposed framework has been applied to a collection of foreground objects from 28 CCTV videos each of 2 minute duration captured from 7 different cameras. These videos have been recorded at local CCTV control room for experiments and analysis during various phases of video annotation project. The objects have been segmented by foreground extraction module of the system. Segmented foreground objects are shown in figure 4.

Table 1 presents a confusion matrix of the system after integrating all the proposed functionalities. It is evident from the results that the accuracy of colour extraction is better when the frames are processed through proposed framework. The accuracy of white colour extraction has remarkably increased. As shown in figure 5 that previously all dominant colours were detected wrong for the white vehicle in a frame because it appears bluish due to poor illumination. A great deal of improvement occurs when the frame is processed through the proposed colour enhancement framework. First two dominant colours are not white but the tuning filter stage extracts the correct colour of the van as third colour.

Table 1: Confusion Matrix.

	Black	Blue	Green	Grey	Silver	Red	Pink	White	Yellow	Brown
Black	**89.17**	3.5	2.37	4.96	0	0	0	0	0	0
Blue	8.3	**79.16**	7.42	3.16	1.94	0	0	0	0	0
Green	6.43	3.21	**83.33**	0	3.91	0	0	3.12	0	0
Grey	7.6	3.8	0	**71.15**	5.7	0	0	11.5	0	0
Silver	5.41	0	2.08	10.41	**70.83**	0	0	10.03	1.22	0
Red	10.41	0	0	3.1	0	**75.0**	9.3	2.08	0	0
Pink	3.21	0	0	0	2.79	11.62	**78.65**	4.21	0	0
White	2.63	0.76	0.44	1.10	1.12	0	0	**93.14**	0.81	0
Yellow	7.77	0	0	3.33	0	0	0	11.11	**71.11**	6.6
Brown	9.96	0	0	1.63	0	0	0	4.34	9.75	**74.32**

5 CONCLUSIONS

We have presented a very effective approach to accurately extract colour information in poor quality CCTV videos. Colour constancy algorithm along with proposed processing worked really well to enhance the actual colour of objects. Fine tuning filter helped a great deal in improving the colour extraction of white objects particularly under poor illumination or poor camera settings. The complete integrated result showed excellent performance especially in the case of white colour where it achieved an average accuracy of 93%.

(a) (b) (c)

Figure 4: Segmented foreground objects in a frame.

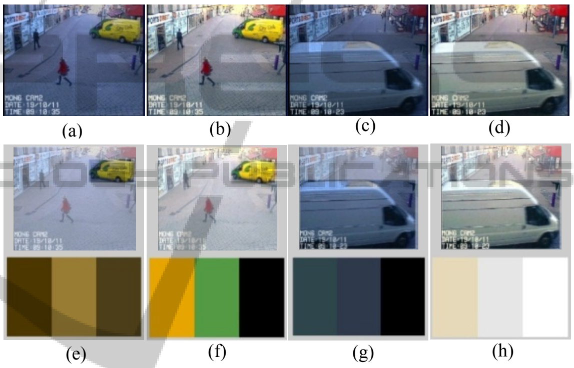

Figure 5: Shows the result of colour extraction by using the proposed framework (a) & (c) are the original video frames. (b) & (d) are the colour corrected version of the frame through proposed frame work. (e) & (g) show three extracted colours of the foreground objects without applying colour correction. (f) & (h) show three dominant colours of the foreground objects extracted after using proposed colour correction mechanism.

REFERENCES

Brown, L. M., 2008. Color retrieval for video surveillance. In *proceedings of 5th IEEE international conference on advanced video and signal surveillance.*

Manduchi, R. 2006. Learning outdoor color classification. In *proceedings of IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol 28, No. 11*, pp1713-23.

Renno, J. P. et al. 2005. Application and Evaluation of Color Constancy in Visual Surveillance. In *proceedings of 2nd joint IEEE international workshop on VS-PETS, Beijing, pp 301-308.*

Schaefer, G. et al. 2005. A combined physical and statistical approach to color constancy. In *proceedings of IEEE Conf. of Computer Vision and Pattern Recognition.*

Schettini, R. et al. 2001. A survey on methods for color image indexing and retrieval in image databases. In Color Imaging Science: Exploiting Digital Media, John Wiley.

Swain, M., and Ballard, D., 1991. Color Indexing. *Internaional Journal of computer Vision.* vol.7, no. 1, pp. 11-32.

Tsin, Y. et al. 2001, Bayesian color constancy for outdoor object recognition. In *proceedings of IEEE Conf. of Computer Vision and Pattern Recognition.*

Wijer, J. V., Schmid, C., Verbeek, J. 2007. Learning color names from real-world images. In *Proceedings of CVPR 2007*.

Wui, G. et al. 2006. Identifying color in motion in video sensors. In *proceedings of IEEE CVPR 2006.*

Zhang, Y., Chou, C., Yu, S., and Chen, T., 2011. Object categorization in surveillance videos. In *proceedings of IEEE International conference on Image processing.*