Colour Space Effect on Tracking in Video Surveillance

Patrick Sebastian[1], Yap Vooi Voon[2], and Richard Comley[3]

[1]Electrical and Electronic Engineering Dept. Universiti Teknologi PETRONAS
Bandar Seri Iskandar 31750 Tronoh, Perak, Malaysia
[2]Electronic Engineering Dept. Universiti Tunku Abdul Rahman
Jalan Universiti Bandar Barat, 31900 Kampar, Perak, Malaysia
[3]Middlesex University The Burroughs,
Hendon, London NW4 4B, United Kingdom

Abstract: The utilization of video surveillance systems is becoming common and is expected to more widespread as societies become more complex and the population continues to grow. The implementation of these video surveillance systems has provided valuable information and assistance in monitoring large areas. These systems typically use human operators to determine human behavior and to manually track people or objects of interest over an array of cameras. With the application of computers with video surveillance, real time surveillance of large public areas, people and their activities has been made possible for monitoring and security. The main contribution of this paper is to show the effect of color space on tracking methods in video surveillance. Results from evaluations on different tracking methods have indicated that YCbCr and HSV color spaces have better tracking ability compared to grayscale and RGB color spaces. In addition, the results from evaluations have also indicated that data from selected layers in some color spaces can be used for the purpose of tracking namely the Cb and Cr layers from the YCbCr color space and the H layer from the HSV color space.

1. Introduction

The utilization of video surveillance systems is becoming common and is expected to more widespread as societies become more complex and the population continues to grow. In utilizing video surveillance in a large scale environment, it leads to problems in recognizing human behavior and tracking people over an array of cameras. The task of recognizing human behavior is difficult by itself and in adding the task of tracking a particular person or persons over an array of cameras makes the task even more difficult for a human operator.

The motivation of this paper is to show how different color spaces effects the tracking performance of a video surveillance system. The paper is divided into 6 sections. Section 2 covers the fundamental background information related to colour spaces of images. Section 3 covers recent and related research work in tracking methodologies. Section 4 covers the experimental setup for evaluating the effect on tracking performance on different colour spaces. Section 5 covers the results of the experiments done in determining tracking performance in different colour spaces.

2. Background

Current implementations of tracking and surveillance systems have provided valuable information and assistance in monitoring of large areas. These systems typically utilize human operators to determine human behavior and to manually track people of interest over an array of cameras in the surveillance system. The limitation or difficulty for a human operator to constantly monitor cameras or to track people over an array of cameras leads to the merging of technologies of cameras and computers. The merging of cameras and computers in monitoring

areas of interest or tracking people of interest leads to growth of intelligence in the surveillance systems developed. The introduction of intelligence in video surveillance systems gives the ability to the system to determine if certain areas are being trespassed and that specific people can be tracked over an array of cameras. The current application of intelligent surveillance systems is where real time surveillance is done to determine the number of people that have entered a specific area. With the application of computers into video surveillance, it has given a rise to real time surveillance of areas, people and their activities [1]. This allows video surveillance to be used as a tool for monitoring and security.

In the field of video surveillance, the object of interest, that includes people and cars, would be used to be monitored or tracked. People are typically the object of interest in video surveillance tracking especially monitoring a person walking through a secluded or security sensitive area and even when monitoring people in a public area. With differing objects of interest, tracking methods vary from a single camera configuration to multiple camera configuration [2-6].

Tracking methods in video surveillance utilize different parameters such as short term biometrics like color of clothing [7] , path and velocity of person tracked [8], head motion of the person [4], trajectory and position of person tracked [9] and motion tracking and modeling set of colors for person tracked [3]. Apart from single parameter tracking, these tracking methods using single parameters can be combined for more robust tracking that overcomes typical problems in tracking people in a crowd such as occlusion [1, 3, 10-12] and noise in the surveillance videos [11].

3. Colour Space Fundamentals

This section presents the fundamentals of colour space where all video image formats are based on different colour spaces. This section shows the different images available ranging from monochrome images to colour images. The data from the different image types can be stored in different colour spaces ranging from grayscale, RGB, YCbCr and HSV colour spaces.

Different types of image formats are available when displaying images that range from grayscale to colour images. Grayscale images can be classified as intensity type images where the images in this format are images without colour. These images are also known as achromatic images or one-colour images. The data used in the representation of grayscale images is the measurement of the intensity or amount of light [13]. The number of bits used in each pixel determines the number of levels of brightness [14]. For example if 8 bits are used to represent each pixel, that would provide 256 levels of brightness or gray levels.

Apart from grayscale or monochrome images, other image formats to be looked at were colour images. Colour images are images that contain colour information for each pixel. Colour images can be modelled as a set of multiple layers of grayscale images where each layer of the colour image corresponds to a certain band in the visible light spectrum [14]. The information that is stored in each layer in a colour image is the brightness in each spectral band. Typical colour images are represented as red (R), green (G) and blue (B) layers or RGB images [14]. Where the red, green and blue spectral colour bands represent the primary colours in the visible range of the electromagnetic spectrum [15]. The RGB colour bands correspond to the absorption characteristics of the human eye where the cones in the eyes have different colour sensitivities. Approximately 65% of all cones are sensitive to red light, 33% are sensitive to green light and 2% are sensitive to blue [15]. With the corresponding absorption characteristics of the eye, colours are seen as variable combinations of the primary colours of red, green and blue. In RGB colour space, each colour appears in its primary spectral components of red, blue and green and this can be illustrated in Figure 1.
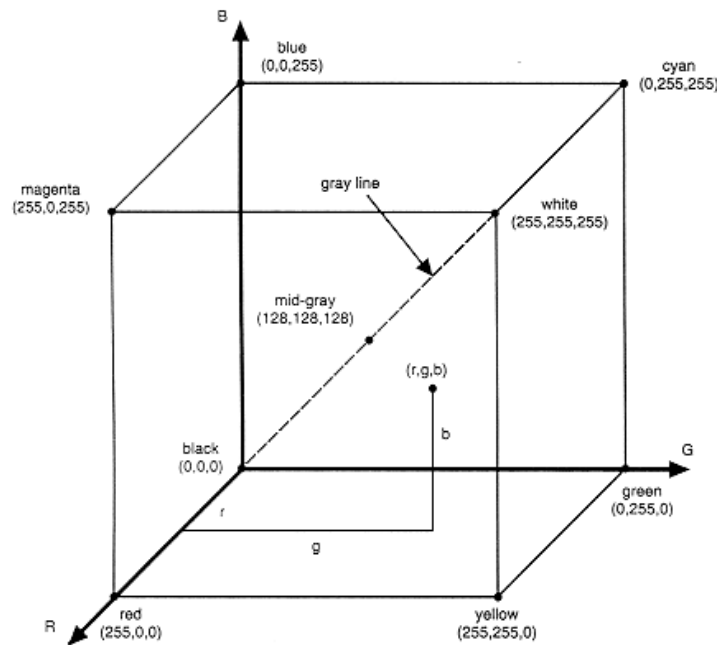
Figure 2. [16]

The RGB colour model illustrated in Figure 2 is based on Cartesian coordinate system. In Figure 2, the colours are indicated on the corners of the cube. Where the RGB values are on three corners of the cube, cyan, magenta and yellow are at another three corners, black is at the origin of the cube and white is at the farthest corner from the origin [15]. Grayscale values, where each point has equal RGB values, is located on the line that connects the black point to the white point in the RGB colour model. Figure 2 illustrates that each red, green and blue colour components have an 8-bit depth that indicates that the maximum value is 255 for each colour. Each colour pixel in the RGB colour space, where each pixel is said to have three values of (R,G,B), would have a pixel depth of 24 bits which is the number of bits used to represent a colour in a particular colour space.

Different colour spaces are also available due to some issues related to limitations in the RGB colour space. The RGB colour space has luminance or brightness information embedded into each layer of the RGB colour space [17]. The varying levels of brightness in an image causes the RGB values in an image to shift [18] and introduces instability in the image [19]. The presence of varying brightness or illumination in an image causes the values in the RGB colour space to change accordingly to the level of brightness and thus cause instability in the RGB values of the image [19]. The susceptibility of the RGB colour space to brightness makes this colour space also be affected on each layer of the RGB colour space. This indicates that each layer in the RGB colour space is correlated to each other [19, 20]. The correlation of each layer in the RGB colour space would show that any effect or change on one of the layers would also have the same effect on the other layers.

In order to remove the brightness effect, the RGB colour space is transformed into a different colour space that would decouple the brightness effect from the colour information [14]. The decoupling of the brightness information from the colour information in the RGB colour space results in different colour spaces that have one layer of brightness and 2 layers of colour information [14]. Other typical colour spaces used in the area of video tracking and video surveillance are YCbCr [18] and HSV [21]. The application of different colour spaces other than the RGB colour space was due to properties in the RGB colour space that has illumination data embedded into the colour layers of the colour space and that RGB colour

layers are correlated between each other [19, 20]. With the effect of the RGB colour layers highly correlated to each other, it has greater tendency that an effect on any one layer would affect the other layers in the RGB layer and thus affect the overall RGB image.

The application of the YCbCr and HSV colour spaces in image and video processing for the purpose of tracking was due to the properties of these colour spaces that have separated the brightness information from the colour information [17]. In the YCbCr colour space, the information is separated into luminance at the Y layer and chrominance in the Cb and Cr layers. In the HSV colour space, the luminance information is place in the V layer and the chromaticity information is placed in the H (hue) and S (saturation) [14, 15]. The separation of the brightness information from the chrominance or chromaticity reduces the effect of uneven illumination in an image. The utilization of chrominance or chromaticity in the YCbCr and HSV colour space provides the opportunity for robust tracking. For robust tracking in the RGB colour space, the RGB colour space would be normalized to obtain chromaticity [22-25]. The chromaticity data obtained through the normalization of RGB values is still dependent on the data obtained from the RGB colour space that is easily affected from varying illumination. This differs for the YCbCr and HSV colour spaces that have different layers of information that has the luminance information separated from the colour information.

The YCbCr colour space was defined by the International Radio Consultative Committee and is widely used in digital video and image compression schemes. The information stored in this image colour space is stored in the luminance (Y) and chrominance layers (CbCr) where Y stores the light intensity and Cb and Cr layers store the colour difference information. The Cb layer is the difference between the blue layer information and a reference value. The Cr layer is the difference between the red layer information and a reference value. The relationship between the RGB and YCbCr colour spaces can be seen in the following equations.

$$Y \quad = \quad 0.299 \, R + 0.586 \, G + 0.114 \, B \tag{1}$$

$$Cb \quad = \quad -0.169 \, R - 0.331 \, G + 0.5 \, B \tag{2}$$

$$Cr \quad = \quad 0.5 \, R - 0.419 \, G - 0.081 \, B \tag{3}$$

Another colour space investigated was the HSV that has information stored in the hue (H), saturation (S) and value (V) layers respectively. The HSV colour space is a colour space that describes colour the way the human eye observes colour. Typically the human eye describes a colour by its hue, saturation and brightness. Hue (H) is an attribute that describes the pureness of a colour, whereas saturation (S) describes the measure of the colour that is diluted by white light. Value (V) is brightness that is a measure of the intensity of the colour. The HSV colour space can be further explained by HSV hexagonal cone in Figure 3.

The colour information in the HSV colour space are stored in the hue (H) and saturation (S) layers. The hue (H) parameter is expressed as angle around the colour hexagon. The hue of a colour refers to the spectral wavelength that closely matches to the colours of the rainbow. Typically, a hue of 0° is red, which also is used as the reference value, 120° is green and 240° is blue. The saturation (S) parameter is the measure of the purity of the colour which is the distance from the V axis [13]. The edges of the cone has the colours at its purest where the saturation values are at unity. The concept of saturation can be further explained by looking at the colour of bright red that would be located at H=0o and S=1. By mixing white colour to the red colour, the red colour would be less intense and thus reduce its intensity. If more of the white colour is mixed with red colour, the red colour would become paler and the saturation decreases [26]. As such, the saturation is measured by the distance of the coordinate from the V axis. The value (V) parameter indicates the brightness levels or the grayscale values that range from black at the bottom of the axis to white at the top of the V axis.
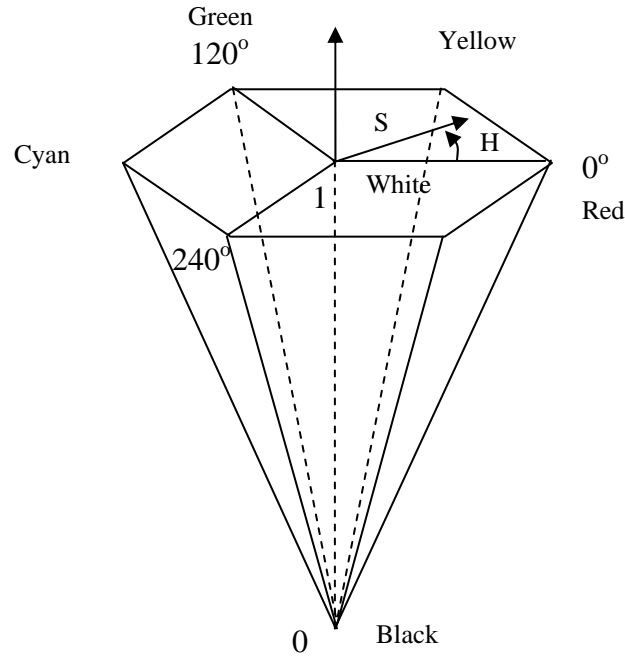
Figure 3. [13]

The relationship between the RGB and HSV colour space is given in the following equations:

$$H = \cos^{-1} \frac{\frac{1}{2}[(R-G)(R-B)]}{\sqrt{[(R-G)^2 + (R-B)(G-B)]}} \qquad (4)$$

$$S = \frac{\max - \min}{\max} \qquad (5)$$

$$V = \max \qquad (6)$$

$$\max = \max imum(R,G,B); \min = \min imum(R,G,B)$$

4. Related Work on Tracking Methodologies

This section will briefly review different tracking methodologies which range from simple target tracking methodologies to more complex target tracking methods [27]. Tracking techniques developed range from window tracking, feature tracking, planar rigid shapes to visual learning.

Among the simplest methods of tracking is window tracking. This technique utilizes a small reference image or rectangular region for tracking the reference image within an image frame. The windows tracking method can be tracked frame by frame by a correlation type matching method [27, 28]. The windows tracking method is the simplest method of tracking where it quantifies the similarity of two group of pixel values that are located in the reference image and the input image [27]. This method determines the highest similarity index between the two groups of pixels to determine the correct tracking of a target. The windows tracking method has certain issues such as image distortion between frames due to change in viewpoint,

302

effects of changing patterns in an image and number of pixels used in tracking. The windows tracking method utilizes a tracking method that tracks all the pixels in the reference image which gives a large number of points to track. A smaller number of points to track is preferable compared to the large number of points done in window tracking [27].

The issues faced by windows tracking have led to another technique of tracking which utilizes detected features in an image. These features can be further classified as local features and extended features [27]. Features can be defined as detectable parts of an image which can be used to support a vision task that includes corners, lines, contours or specially defined regions. Feature tracking initially locates features in subsequent image frames and then matches the detected features accordingly between the frames. The manner of feature tracking is similar in method where the detected features are matched to each other. The difference of feature tracking compared to window tracking is that feature tracking only tracked certain identified features as compared to window tracking that tracked all pixel points available in the window.

Feature tracking can be further classified either as local features and extended features. Local features are image locations with specific properties such as edges, lines and corners [27]. Local features offer invariant features for better tracking where uneven illumination can cause inaccurate tracking. Local features are not only detected on the target image to be tracked but also on the overall image. This gives a possibility of wrongly matching the local features [27].

With the possibility of mismatching local features in an image, extended features was used to have better or more robust tracking. Extended features use a larger part of the image compared to local features. Extended features cover basic shapes such as ellipses, rectangles, free form contours and image regions [5, 27, 29]. The advantage of extended features was its robustness to clutter compared to local features. Regions are defined as connected image parts with intensity or color properties. Image contours are another feature used for tracking purposes where the contour information is used as the boundary to the whole object that would be tracked [27, 30]. For robust tracking of an object it utilizes features available on the tracked object together with additional parameters obtained from tracking the extended features detected in an image. The additional parameters used in tracking range from velocity of features [2, 5, 31], path of features [8, 9, 31], deformable shape of region [1, 32, 33], region color [29, 34] and region sizes [35-37].

In object tracking, the color space utilized determines the features and parameters used in tracking an object. Grayscale tracking typically is done on a blob or a specific region that utilizes additional or accompanying parameters such as distance [35], velocity [2, 5, 38], trajectory [8, 9, 31] and blob dimensions such as height [35, 39]. The application of the accompanying parameters such as distance, velocity and trajectory are used to maintain short term tracking of a target object. In addition to the combination of utilizing features and parameters, other methods of tracking are utilized for tracking purposes. The different measures used for tracking include properties of the tracked object which cover Manhattan Distance [40], Mahalanobis Distance [31, 34], shape tracking [1, 32], histogram similarity index [41] and consistent labeling [42]. Other methods of tracking utilize complex statistical and intelligent methods which were Bayesian Network [11, 12] and neural network [43].

The grayscale color space allow for object tracking using the brightness information embedded in grayscale images. With color images having more information embedded in the images, it would be expected that tracking using color would be more robust, accurate and overcome problems associated to grayscale color space. One the primary problems of grayscale color space was the brightness influence on grayscale images which affects tracking [30]. Brightness also plays a part in tracking effectiveness in color image where illumination was required to be uniform for effective tracking [10]. One method of minimizing the effect of brightness on color image tracking would be to convert the typical RGB color images to different color spaces such as YCbCr [17, 18, 44]and HSV [21] for robust tracking. The usage of different color spaces other than RGB would be to overcome the correlation between layers

of the RGB color space [20]. The color spaces of YCbCr and HSV can be used for more accurate tracking due to these color spaces having less noise [19] and the ability of using different layers without losing data [21, 44] and tracking ability [19].

Table 1: Listing of Tracking Methods used against Color Space

Authors	Color Space	Tracking Method
Lewis [52]	Grayscale	Normalized cross correlation
Sebastian et al. [53]	Grayscale	Normalized cross correlation
Nillius et al. [54]	Wavelet	Normalized cross correlation
McErlean [55]	Wavelet	Phase correlation
Atsushin et al. [5]	Grayscale	Multiple cameras tracking on binary image of human blob. Tracking of target is based on ellipsoid shape and velocity of blob.
Lu et al. [2]	Grayscale	Tracking is done from a top view and uses heads as a feature to be tracked. Tracking is based on the apparent speed of a tracked object.
Bodor et al. [8]	Grayscale	Tracking done based on path and velocity of tracked object.
Niu et al. [32]	Grayscale	Tracking based on relative trajectories and position of each tracked object.
Cai et al. [31, 38, 39]	Grayscale	Tracking on multiple camera using velocity and position of tracked object for tracking. Using additional information of statistical measure for tracking (Mahalanobis Distance)
Sato et al. [35]	Grayscale	Tracking based on parameter of blob being tracked such as relative distance, velocity and size.
Han et al. [43]	Grayscale	Tracking uses neural network to maintain tracking of target object.
Wahl et al. [41]	Grayscale	Tracking is based on similarity index calculated from image histogram.
Chang et al. [6]	Color	Multiple cameras track based on apparent height and color of tracked target.
Iocchi et al. [3]	Color	Tracking from a plan view using a set of colors to represent each target tracked.
Li et al. [10]	Color	Uses color invariance for tracking in a light varying condition
Siebel et al. [4]	Color	Multiple track algorithms used for robust tracking based on head, motion and region track.
Bird et al. [7]	Color	Tracking based on short term biometrics of clothing color.
Park et al. [29]	Color	Tracking done based on blob size, color, location, perimeter, shape and orientation.
Matsuzawa et al. [34]	Color	Tracking uses neural network for shape representation. Utilizes Mahalanobis Distance with color information for tracking
Teknomo et al. [36]	Color	Tracking based on feature such as area, perimeter box width and color.
Moritomo et al. [37]	Color	Tracking based on area, width and height of blob, position and color of blob.
Zivkovic et al. [49]	Color	Uses Bhatacharrya Distance as similarity index for tracking purpose.
Ali et al. [28]	Color	Uses multiple methods of tracking. Uses template matching for tracking. Uses blob features such as size, motion vector and centroid location Uses Euclidian and histogram comparison for matching.
Greenspan et al. [51]	Color	Uses similarity index between distributions for tracking such as Gaussian Mixture distributions. Similarity index calculated based on intersection, Chi-Squared and Kullbeck-Leiber measures for tracking.
Wang et al. [50]	Color	Uses Bhattacharya Distance to measure similarity between color regions.
Wang et al. [33]	Color	Shape information used as parameter for tracking
Yang et al. [30]	Grayscale ; Wavelet	Tracking is based on wavelets with parameters such as trajectory, size, grayscale distribution and texture
Amiri et al. [56]	Wavelet	Tracking is based on template matching for block matching.
Ellis [57]	Color	Tracking based on centroid, height, width, chromaticity mean and trajectory
Chen et al. [18]	Color	Tracking based on color detected against reference color

As in grayscale color space tracking, color tracking utilizes similar tracking methods such as shape tracking [4, 34, 45] and blob [35, 40] or region tracking [1, 46]. In addition to the existing tracking methods, color information is also used in tracking [7, 29, 36, 37]. For robust tracking in color, the typical tracking method uses color [3, 7, 10] and region properties such as color of head region, height of head region, area, width, length, perimeter, shape and position [6, 29, 33, 36, 37, 47]. Although color information adds robustness to object tracking it is also susceptible to brightness variations [10]. Color information is used together with motion parameters such as velocity , position [29] and motion vector [28, 33, 48]. Apart from the tracking methods described, other methods of tracking utilize color information as a means for determining the similarity between objects that is being tracked. The measure of similarity between a reference object and a tracked object can be measured by measures of Mahalanobis distance [34], Bhatacharrya distance [49, 50], color distribution [50] and histogram similarity index such as histogram intersection, Squared Euclidian Distance, Chi-Squared and Kullback-Leiber [41, 49-51]. Based on the different color spaces, different tracking methods were developed based on the different information available on the color spaces. Some of the tracking methods utilized in surveillance can be seen in the following Table 1 that lists the research completed by different authors on the tracking method used for tracking targets in video surveillance and the color space used.

Table 1 has listed a number of papers listing the image type used in video surveillance against the tracking method used. Each tracking method developed is based on the information or image used in video surveillance. Color information would seem to be a factor in tracking for video surveillance. This can be seen in some tracking results done in different color spaces.

5. Experiment Setup

This section describes the setup of the experiments in determining the effect of colour space in tracking a target in video surveillance. Two video types were used in determining the effect of colour space in the performance of a tracking algorithm. The set of videos used in Experiment 1 is a set of 4 videos where the target was a moving face to be tracked. The first tracking experiment was run using normalized cross correlation as the method to track the face within the video. Experiment 2 was run with a set of 5 videos where the target was a moving whole person which included the head, torso and limbs of the target. The tracking in Experiment 2 was done using histogram comparison between a reference histogram against the tracked target histogram. The success of a tracking methodology is based on the similarity of the reference histogram compared to the input image histogram.

6. Results and Discussion

This section presents the results of a face tracking and histogram tracking algorithm to evaluate the tracking performance in different colour spaces. The performance of a tracking algorithm is measured in terms of Tracker Detection Rate (TDR). The TDR measure is the ratio of number of correct tracks to the total number of actual appearance of tracked target [58]. In comparing the performances of tracking algorithms in different color spaces, the tracking results from Experiment 1 can be seen in Figure 1.

The TDR tracking result displayed in Figure 1 shows the result of tracking a face in different color spaces. The results indicate that tracking in the grayscale color space has lower TDR compared to the other color spaces of YCbCr and HSV. Based on the TDR results indicated in Figure 1, the TDR results for the grayscale and RGB color space are similar due to the similarity in the information embedded in each color space which is the brightness information. The TDR results for the YCbCr and HSV color spaces have better results due to the brightness information being stored on a different layer compared to the color information. The results also indicate that the brightness information is not necessary in tracking a target as can be seen in Figure 1 that shows information from the CbCr layers from the YCbCr color spaces and HS layers from HSV color spaces were sufficient for tracking.
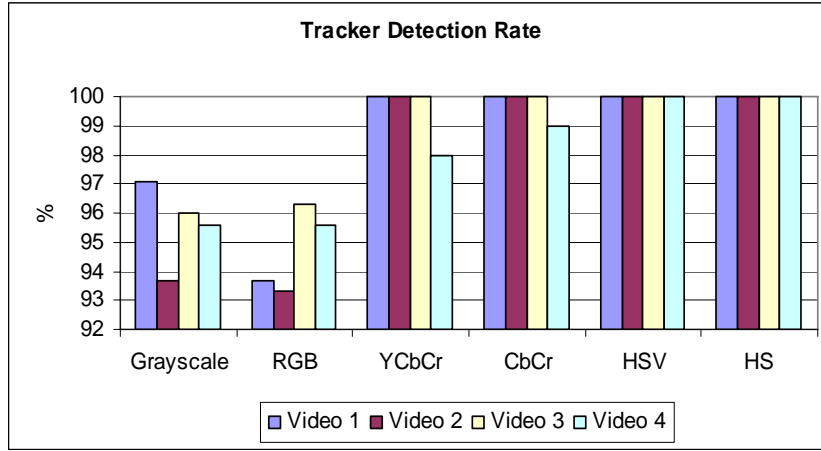
Figure 1

The effect of color space on tracking performance can also be seen in the tracking results from Experiment 2 that uses histogram similarity for tracking a target. Histogram similarity is a method of measuring the similarity between the histogram of a reference image and the histogram of the tracked image. The histogram in this case would be the histogram or distribution of the pixels in the single blob of the tracked object. In this instance the tracked object was a whole person with multiple regions. Each run or sample has a different target or a variation of the tracked target. Each graph shows the tracking ability in terms of TDR for each histogram comparison method such as histogram intersection, Euclidian Distance, Chi-Square one, Chi-Square Two and Kulbeck-Leibler methods. The results from Experiment 2 are shown in Figure 2 where histogram tracking is based on histogram similarity [41, 49, 59] indices where the histogram intersection, Euclidian distance, Chi-Squared One, Chi-Squared Two and Kullbeck-Leibler histogram equations are listed as follows:
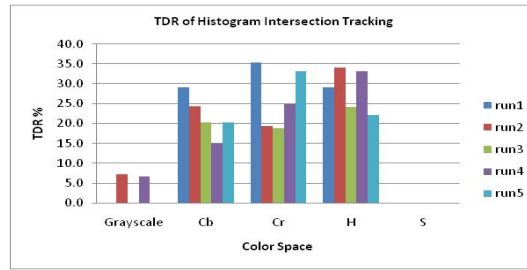
$$\bigcap (Ho, Ho') = \sum_{i=1}^{d} \min (Ho(i), Ho'(i)) \tag{7}$$

$$\varepsilon (Ho, Ho') = \sum_{i=1}^{d} (Ho(i) - Ho'(i))^2 \tag{8}$$
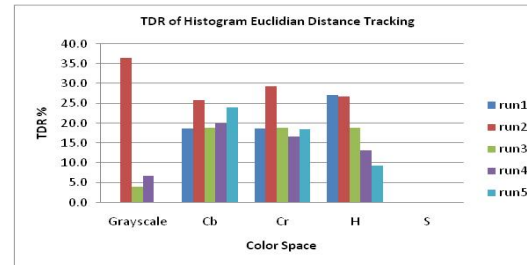
$$\chi^2 (Ho, Ho') = \sum_{i=1}^{d} \frac{(Ho(i) - Ho'(i))^2}{Ho(i)} \tag{9}$$

$$\chi_2^2 (Ho, Ho') = \sum_{i=1}^{d} \frac{(Ho(i) - Ho'(i))^2}{Ho(i) + Ho'(i)} \tag{10}$$
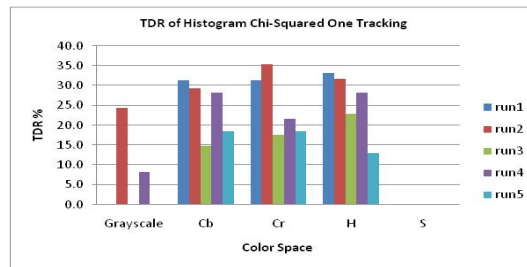
$$\kappa (Ho, Ho') = \sum (Ho'(i) - Ho(i)) \ln \frac{Ho'(i)}{Ho(i)} \tag{11}$$
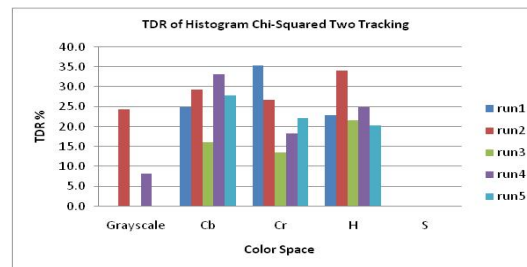
306

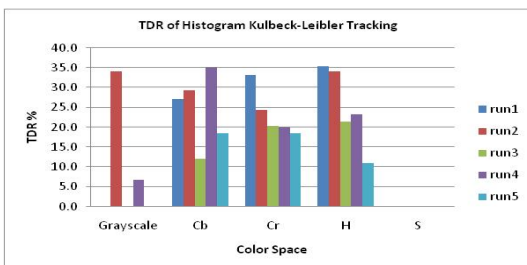(a) Histogram intersection

(b) Euclidian Distance

(c) Chi-Squared One

(d) Chi-Squared Two

(e) Kullbeck-Leibler

Figure 2

The results in Figure 2 indicate that the grayscale color space for the different videos has consistently lower tracking ability compared to other color spaces that has the brightness information and the color information on different layers. By looking at the overall results shown in Figure 2, grayscale color space would not be able to provide reliable tracking ability using histogram similarity as a method for tracking. TDR tracking results shown in Figure 2 indicate that histogram tracking in YCbCr and HSV color spaces can be done using either one of the layers that has the color information such as the Cb, Cr and H layers. The S layer in the HSV color space as indicated in Figure 2 has been able to provide consistent tracking ability using histogram similarity. Figure 2 also indicates a similarity between different histogram similarity tracking methods where tracking in the Cb and Cr layers in the YCbCr color space and tracking in the H layer in the HSV color space has better tracking ability compared to grayscale color space and S layer in the HSV color space. Comparing between the different tracking method of normalized cross correlation and histogram similarity, grayscale color space was shown to be a color space that had low or inconsistent tracking performance. YCbCr and HSV color spaces had consistently higher tracking performance compared to grayscale color space. Tracking results also indicate that utilization of a single layer where the color information was stored could be used for tracking. The color information in the YCbCr color space are stored in the Cb and Cr layers whereas in the HSV color space the color information is stored in the H layer.

Conclusion

This paper looks at the different parameters that affect the area of video tracking that cover different color spaces, tracking methodologies and tracking performance. By examining different color spaces, it can be seen that different color spaces have different information embedded in the images. The grayscale and RGB color spaces have been seen to be easily affected by uneven brightness levels within a specified area. The brightness information is embedded into the grayscale and RGB images that make tracking in these color spaces inaccurate due to varying levels of brightness. Varying brightness levels would result in inaccurate tracking. To overcome varying levels of brightness, other color spaces used were color spaces that have the brightness information separated from color information which are the YCbCr and HSV. The YCbCr color space has the brightness information in the Y layer of the color space and the brightness information is located in the HSV color space.

In tracking a target object, tracking methods utilize data available from the color spaces used in the tracking methods. These tracking methods utilize information available from the color spaces to develop metrics for tracking purposes such as Mahalanobis Distance, Bhatacharrya Distance, color distributions and similarity indices. These metrics are used as a means of maintaining a constant track on a target object. This paper lists the different techniques used in the tracking of a specific object.

This paper also shows the tracking performance of two different tracking methods in different color spaces. The results from two different tracking methods showed that there was similarity in terms of the tracking performance where the grayscale color space had consistently lower tracking performance compared to YCbCr and HSV color spaces. The results obtained from the two experiments as illustrated in Figure 1 and Figure 1 have given an indication on the effect of colour space on tracking performance. The results indicated that colour spaces such as grayscale and RGB that utilize brightness information in the colour layers have consistently lower tracking performance compared to the YCbCr and HSV colour spaces. The YCbCr and HSV color space have the brightness and color information on separate layers. Based on the results shown in Figure 1, colour spaces YCbCr and HSV have better tracking performance compared to the grayscale colour space. Tracking done in the YCbCr and HSV color spaces could also be done by using either one of the layers where the color information is stored for tracking purposes. In the two different tracking methods used in the experiments, it indicated that colour space has a consistent effect in tracking performance.

References

[1] Haritaoglu, I., D. Harwood, and L. S. Davis, W4: "Real-time surveillance of people and their activities", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000. 22(8): p. 809-830.

[2] Lu, H., Y. Wang, and Z. Wang, "A crowd detecting and tracking algorithm based on projection of grayscale image", *Third International Conference on Image and Graphics, 2004*. 2004.

[3] Iocchi, L. and R. C. Bollees, "Integrated Plan-View Tracking and Color-based person Models for Multiple People Tracking", *International Conference on Image Processing 2005*. 2005.

[4] Siebel, N. T. and S. Maybank, "Fusion of Multiple Tracking Algorithms for Robust People Tracking", *In Proceedings of the 7th European Conference on Computer Vision (ECCV 2002)*. 2002.

[5] Atsushi, N., et al, "Tracking multiple people using distributed vision systems", *IEEE International Conference on Robotics and Automation, 2002*. 2002. Washington, DC, USA.

[6] Chang, T. H. and S. Gong, "Tracking multiple people with a multi-camera system", Proceedings 2001 *IEEE Workshop on Multi-Object Tracking*. 2001. Vancouver, BC, Canada.

[7] Bird, N. D., et al, "Detection of loitering individuals in public transportation areas", *IEEE Transactions on Intelligent Transportation Systems*. 2005.

[8] Bodor, R., B. Jackson, and N. P. Papanikolopoulos, "Vision-Based Human Tracking and Activity Recognition", *11th Mediterranean Conference on Control and Automation*. 2003.

[9] Niu, W., et al, "Human Activity detection and recognition for video surveillance", *IEEE International Conference on Multimedia and Expo 2004*. 2004.

[10] Li, J., C. S. Chua, and Y. K. Ho, "Color based multiple people tracking", *7th International Conference on Control, Automation, Robotics and Vision 2002. ICARCV 2002*. 2002.

[11] Diaz de Len, R. and L. E. Sucar, "Continuous activity recognition with missing data", *16th International Conference on Pattern Recognition 2002*. 2002.

[12] Kitani, K. M., Y. Sato, and A. Sugimoto, "Deleted Interpolation Using a Hierarchical Bayesian Grammar Network for Recognizing Human Activity", *2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance 2005*. 2005.

[13] Gonzalez, R. C., R. E. Woods, and S. L. Eddins, "Digital Image Processing using Matlab", 2004: Pearson Prentice Hall.

[14] Umbaugh, S. E., Computer Vision and Image Processing, 1998: Prentice Hall.

[15] Gonzalez, R. C. and R. E. Woods, "Digital Image Processing", 2 ed. 2002, New Jersey: Prentice Hall.

[16] Burdick, H. E., Digital imaging: Theory and Applications, 1997: Mcgraw-Hill. 315.

[17] Kim, W. -S., D.-S. Cho, and H. M. Kim, "Interplane prediction for RGB video coding", *2004 International Conference on Image Processing (ICIP '04)*. 2004.

[18] Chen, Y. -J., et al, "The Implementation of a Stand-alone Video Tracking and Analysis System for Animal Behavior Measurement in Morris Water Maze", *27th Annual International Conference on Engineering in Medicine and Biology Society 2005 (IEEE-EMBS 2005)*. 2005. Shanghai.

[19] Zhao, M., J. Bu, and C. Chen, "Robust background subtraction in HSV color space", *Proceedings of Multimedia Systems and Applications V*. 2002. Boston, USA.

[20] Kobayashi, M., et al, "Lossless compression for RGB color still images", *1999 International Conference on Image Processing (ICIP99)*. 1999. Kobe, Japan.

[21] Stern, H. and B. Efros, "Adaptive Color Space Switching for Face Tracking in Multi-Colored Lighting Environments", *Fifth IEEE Conference on Automatic Face and Gesture Recognition 2002*. 2002. Washington, D.C.

[22] Beetz, M., B. Radig, and M, "Wimmer. A Person and Context Specific Approach for Skin Color Classification", *18th International Conference on Pattern Recognition, 2006 (ICPR 2006)*. 2006. Hong Kong.

[23] Soriano, M., et al, "Skin Detection in video under changing illumination conditions", *15th International Conference on Pattern Recognition 2000*. 2000. Barcelona.

[24] Kawato, S. and J. Ohya, "Automatic skin-color distribution extraction for face detection and tracking", *5th International Conference on Signal Processing Proceedings 2000 (WCCC-ICSP 2000)*. 2000. Beijing.

[25] Park, J., et al, "Detection of Human Faces using skin color and eyes", *2000 IEEE International Conference on Multimedia and Expo 2000 (ICME 2000)*. 2000. New York, NY.

[26] Castleman, K. R., Digital Image Processing. 1996: Prentice Hall.

[27] Trucco, E. and K. Plakas, "Video Tracking: A Concise Survey", *IEEE Journal of Oceanic Engineering*, 2006. **31**(2): p. 520-529.

[28] Ali, M. A., S. Indupalli, and B. Boufama, "Tracking Multiple People for Video Surveillance", *First International Workshop on Video Processing for Security*. 2006. Quebec City, Canada.

[29] Park, S. and J. K. Aggarwal, "Segmentation and Tracking of interacting human body parts under occlusion and shadowing", *Workshop on Motion and Video Computing 2002*. 2002.

[30] Yang, Y., J. F. Doherty, and V. D. R. E, "Moving Object Tracking in Video", *29th Applied Imagery Pattern Recognition Workshop 2000*. 2000. Washington, D.C.

[31] Cai, Q. and J. K. Aggarwal, "Tracking human motion using multiple cameras", *13th International Conference on Pattern Recognition, 1996*. 1996. Vienna, Austria.

[32] Niu, F. and M. Abdel-Mottaleb, "View-invariant human activity recognition based on shape and motion feature", *IEEE Sixth International Conference on Multimedia Software Engineering 2004*. 2004.

[33] Wang, J., G. Bebis, and R. Miller, "Robust Video-Based Surveillance by Integrating Target Detection with Tracking", *2006 Conference on Computer Vision and Pattern Recognition Workshop*. 2006.

[34] Matsuzawa, Y. and I. Kumazawa, "Object tracking with shape representation network using color information", *2000 International Conference on Image Processing*. 2000. Vancouver, BC, Canada.

[35] Sato, K. and J. K. Aggarwal, "Tracking and recognizing two-person interactions in outdoor image sequence", *2001 IEEE Workshop on Multi-object tracking*. 2001. Vancouver, BC, Canada.

[36] Teknomo, K., Y. Takeyama, and H. Inamura, "Frame-based tracing of multiple objects", *2001 IEEE Workshop on Multi-object Tracking*. 2001. Vancouver, BC, Canada.

[37] Moritomo, T., et al, "Object tracking in video pictures based on image segmentation and pattern matching", *IEEE International Symposium on Circuits and Systems 2005*. 2005.

[38] Cai, Q. and J. K. Aggarwal, "Automatic tracking of human motion in indoor scenes across multiple synchronized video streams", *Sixth International Conference on Computer Vision 1998*. 1998. Bombay, India.

[39] Cai, Q. and J. K. Aggarwal, "Tracking human motion in structured environments using a distributed camera system", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999. **21**(11): p. 1241-1247.

[40] Leo, M., et al, "Complex human activity recognition for monitoring wide outdoor environments", *17th International Conference on Pattern Recognition 2004*. 2004.

[41] Wahl, E., U. Hillenbrand, and G. Hirzinger, Surflet-Pair-Relation Histograms: "A statistical 3D-shape Representation for Rapid classification", *Fourth International Conference on 3-D Imaging and Modelling 2003*. 2003.

[42] Khan, S. and M. Shah, "Consistent labeling of tracked objects in multiple cameras with overlapping fields of view", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003. **25**(10): p. 1355-1360.

[43] Han, M., et al, "A detection-based multiple object tracking method", *2004 International Conference on Image Processing*. 2004.

[44] Chippendale, P., "Towards Automatic Body Language Annotation", *7th International Conference on Automatic Face and Gesture Recognition 2006 (FGR 2006)*. 2006.

[45] Koschan, A., et al, "Video Object Tracking Based on Extended Active Shape Models with Color Information", *1st European Conference on Color in Graphics, Imaging, and Vision CGIV2002*. 2002. Poitiers, France.

[46] Bakowski, A. and G. A. Jones, "Video Surveillance Tracking Using Colour Region Adjacency Graphs", *Image Processing and its Applications*. 1999.

[47] Yan, W. and D. A. Forsyth, "Learning the Behavior of Users in a Public Space through Video Tracking", *Seventh IEEE Workshops on Application of Computer Vision 2005, WACV/MOTIONS '05*. 2005.

[48] Dahmane, M. and J. Meunier, "Real-time video surveillance with self-organizing maps", *The 2nd Canadian Conference on Computer and Robot Vision, 2005*. 2005. Montreal, Canada.

[49] Zivkovic, Z. and B. Krose, "An EM-like algorithm for color-hostogram-based object tracking", *Proceedings of the 2004 IEEE Computer Vision Society Conference on Computer Vision and Pattern Recognition, 2004, CVPR 2004*. 2004.

[50] Wang, Q. and J. Liu, "The Improved Particle Filter for Object Tracking", *6th World Congress on Intelligent Control and Automation 2006*. 2006. Dalian, China.

[51] 51. Greenspan, H., S. Gordon, and J. Goldberger, "Probabilistic Models for Generating, Modelling and Matching Image Categories", *16th International Conference on Pattern Recognition*. 2002: IEEE Computer Society Washington, DC, USA.

[52] Lewis, J. P., "Fast Normalized Cross-Correlation", in *Vision Interface*. 1995, Canadian Image Processing and Pattern Recognition Society. p. 120-123.

[53] Sebastian, P. and Y. Vooi Voon, "Tracking Using Normalized Cross Correlation and Color Space", *Internation Conference on Intelligent and Advanced Systems (ICIAS 2007)*. 2007. Kuala Lumpur, Malaysia.

[54] Nillius, P. and J. -O. Eklundh, "Fast Block Matching with Normalized Cross-Correlation using Walsh Transforms", 2002, Royal Institute of Technology.

[55] Mc Erlean, M., "Hierarchical Motion Estimation for Embedded Object Tracking", *2006 IEEE International Symposium on Signal Processing and Information Technology*. 2006.

[56] Amiri, M., et al, "A New Wavelet Domain Block Matching Algorithm for Real-Time Object Tracking", *International Conference on Image Processing 2003 (ICIP 2003)*. 2003.

[57] Ellis, T., "Multi-camera Video surveillance", *36th Annual 2002 International Carnahan Conference on Security Technology 2002*. 2002.

[58] Sebastian, P., Y. Vooi Voon, and R. Comley, "The Effect of Color Space on Tracking Robustness", *3rd IEEE Conference on Industrial Electronics and Applications*. 2008. Singapore.

[59] Ferman, A. M., A. M. Tekalp, and R. Mehrotra, "Robust Color Histogram Descriptors for Video Segment Retrieval and Identification", *IEEE Transactions on Image Processing*, 2002. **11**(5).

Patrick Sebastian is currently a lecturer in the Electrical and Electronic Engineering Department at Universiti Teknologi PETRONAS (UTP). Prior to his current appointment, Patrick was a Senior Engineer at Penang Seagate Industries Malaysia. He is currently involved in a funded research project in intelligent video surveillance.

Yap Vooi Voon is currently an Assistant Professor and Head of Department in Electronic Engineering Department at Universiti Tunku Abdul Rahman (UTAR). Prior to his current appointment, Dr.Yap was a senior lecturer at Universiti Teknologi PETRONAS (UTP) where he played a major role in the setting up the Information Technology degree. He received his PhD in Wavelet-based Image Compression for Mobile Devices from Middlesex University, London. He has over 20 years teaching experience in colleges and universities both in the United Kingdom and Malaysia, with special interest in embedded systems and software engineering. He is also involved in a funded research project in intelligent video surveillance.

Richard Comley is Professor of Computer Communications in the School of Computing Science at Middlesex University, UK, where he is currently the Associate Dean for Research. Prior to this, he was Head of the Computer Communications Department at Middlesex University and Associate Dean of Engineering at City University, London. He received his BSc in Electrical and Electronic Engineering in 1974 and his PhD in Computer Aided Measurement in 1979, both from City University. He has over 30 years teaching and research experience, with special interest in mobile communications, computer networks and security, and digital signal processing. His research interests are predominately in the signal processing and computer networks areas.