Improved motion segmentation based on shadow detection

M. Kampel* and H. Wildenauer+ and P. Blauensteiner* and A. Hanbury*

* *Pattern Recognition and Image Processing Group, Vienna University of Technology, Favoritenstr.9, A-1040 Vienna, Austria*
+ *Automation and Control Institute, Vienna University of Technology, Gusshausstr.27, A-1040 Vienna, Austria*

Abstract

In this paper, we discuss common colour models for background subtraction and problems related to their utilisation are discussed. A novel approach to represent chrominance information more suitable for robust background modelling and shadow suppression is proposed. Our method relies on the ability to represent colours in terms of a 3D-polar coordinate system having saturation independent of the brightness function; specifically, we build upon an Improved Hue, Luminance, and Saturation space (IHLS). The additional peculiarity of the approach is that we deal with the problem of unstable hue values at low saturation by modelling the hue-saturation relationship using saturation-weighted hue statistics. The effectiveness of the proposed method is shown in an experimental comparison with approaches based on RGB, Normalised RGB and HSV.

Key Words: Motion detection, shadow detection, background subtraction, colour spaces.

---

1 Introduction

The underlying step of visual surveillance applications like target tracking and scene understanding is the detection of moving objects. Background subtraction algorithms are commonly applied to detect these objects of interest by the use of statistical colour background models. Many present systems exploit the properties of the Normalised RGB to achieve a certain degree of insensitivity with respect to changes in scene illumination.

Hong and Woo [1] apply the Normalised RGB space in their background segmentation system. McKenna et al. [2] use this colour space in addition to gradient information for their adaptive background subtraction. The AVITRACK project [3] utilises Normalised RGB for change detection and adopts the shadow detection proposed by Horprasert et al. [4].

Beside Normalised RGB, representations of the RGB colour space in terms of 3D-polar coordinates (hue, saturation, and brightness) are used for change detection and shadow suppresion in surveillance applications. François and Medioni [5] suggest the application of HSV for background modelling for real-time video segmentation. In their work, a complex set of rules is introduced to reflect the relevance of observed and background colour information during change detection and model update. Cucchiara et al. [6] propose a RGB-based background model which they transform to the HSV representation in order to utilise the properties of HSV chrominance information for shadow suppression.

Our approach differs from the aforementioned in the way that we build upon the IHLS colour space, which is more suitable for background subtraction. Additionally, we propose the application of saturation-weighted hue statistics [7] to deal with unstable hue values at weakly saturated colours. Also, a technique to efficiently classify changes in scene illumination (e.g. shadows), modelling the relationship between saturation and hue has been devised.

The remainder of this paper is organised as follows: Section 2 reviews the Normalised RGB and the Improved Hue, Luminance and Saturation (IHLS) colour space. Furthermore it gives a short overview over circular colour statistics, which have to be applied on the hue as angular value. Section 3 presents how these statistics can be applied in order to model the background in image sequences. In Section 4 we describe metrics for the performance evaluation of our motion segmentation. The conducted experiments and their results are presented in Section 5. Section 6 concludes this paper and gives an outlook.

2   Colour Spaces

In this section, the Normalised RGB and IHLS colour spaces used in this paper are described. It also gives a short overview over circular colour statistics and a review of saturation weighted hue statistics.

2.1   Normalised RGB

The Normalised RGB space aims to separate the chromatic components from the brightness component. The red, green and blue channel can be transformed to their normalised counterpart by using the formulae

$$l = R + G + B, \;\; r = R/l, \;\; g = G/l, \;\; b = B/l \tag{1}$$

if $l \neq 0$ and $r = g = b = 0$ otherwise [8]. One of these normalised channels is redundant, since by definition r, g, and b sum up to 1.

Therefore, the Normalised RGB space is sufficiently represented by two chromatic components (e.g. r and g) and a brightness component l. From Kender [9] it is known that the practical application of Normalised RGB suffers from a problem inherent to the normalisation; namely, that noise (such as, e.g. sensor or compression noise) at low intensities results in unstable chromatic components. For an example see Figure 1. Note the artefacts in dark regions such as the bushes (top left) and the shadowed areas of the cars (bottom right).

Figure 1: Examples of chromatic components. Lexicographically ordered - Image from the *PETS2001* dataset, it's normalised blue component b, normalised saturation (cylindrical HSV), IHLS saturation.

2.2 IHLS Space

The Improved Hue, Luminance and Saturation (IHLS) colour space was introduced in [10]. It is obtained by placing an *achromatic axis* through all the grey ($R = G = B$) points in the RGB colour cube, and then specifying the coordinates of each point in terms of position on the achromatic axis (brightness), distance from the axis (saturation s) and angle with respect to pure red (hue θ^H). The IHLS model is improved with respect to the similar colour spaces (HLS, HSI, HSV, etc.) by removing the normalisation of the saturation by the brightness. This has the following advantages: (a) the saturation of achromatic pixels is always low and (b) the saturation is independent of the brightness function used. One may therefore choose any function of R, G and B to calculate the brightness.

It is interesting that this normalisation of the saturation by the brightness, which results in the colour space having the shape of a cylinder instead of a cone or double-cone, is usually implicitly part of the transformation equations from RGB to a 3D-polar coordinate space. This is mentioned in one of the first papers on this type of transformation [11], but often in the literature the equations for a cylindrically-shaped space (i.e. with normalised saturation) are shown along with a diagram of a cone or double-cone (for example in [12, 13]). Figure 1 shows a comparison of the different formulations of saturation. The undesirable effects created by saturation normalisation are easily perceivable, as some dark, colourless regions (eg., the bushes and the side window of the driving car) reach higher saturation values than their more colourfull surroundings. Also, note the artefacts resulting from the singularity of the saturation at the black vertex of the RGB-cube (again, the bushes and the two bottom right cars).

The following formulae are used for the conversion from RGB to hue θ^H, luminance y and saturation s of the IHLS space:

$$
\begin{aligned}
s &= \max(R, G, B) - \min(R, G, B) \\
y &= 0.2125R + 0.7154G + 0.0721B \\
cr_x &= R - \frac{G+B}{2}, \; cr_y = \frac{\sqrt{3}}{2}(B - G) \\
cr &= \sqrt{cr_x^2 + cr_y^2} \\
\theta^H &= \begin{cases} \text{undefined} & \text{if } cr = 0 \\ \arccos\left(\frac{cr_x}{cr}\right) & \text{elseif } cr_y \leq 0 \\ 360° - \arccos\left(\frac{cr_x}{cr}\right) & \text{else} \end{cases}
\end{aligned}
\tag{2}
$$

where cr_x and cr_y denote the chrominance coordinates and $cr \in [0, 1]$ the chroma. The saturation assumes values in the range $[0, 1]$ independent of the hue angle (the maximum saturation values are shown by the circle on the chromatic plane in Figure 2). The chroma has the maximum values shown by the dotted hexagon in Figure 2. When using this representation, it is important to remember that the hue is undefined if $s = 0$, and that it does not contain much useable information when s is low (i.e. near to the achromatic axis).

2.3 Hue statistics

In a 3D-polar coordinate space, standard (linear) statistical formulae can be utilised to calculate statistical descriptors for brightness and saturation coordinates. The hue, however, is an angular value, and consequently the appropriate methods from circular statistics are to be used.

Now, let θ_i^H, $i = 1, \ldots, n$ be n observations sampled from a population of angular hue values. Then, the vector $\mathbf{h}_i$ pointing from $\mathbf{O} = (0,0)^T$ to the point on the circumference of the unit circle, corresponding to θ_i^H, is given by the Cartesian coordinates $(\cos\theta_i^H, \sin\theta_i^H)^T$. *

---

*Note that, when using the IHLS space (Eq. 3), no costly trigonometric functions are involved in the calculation of $\mathbf{h}_i$, since $cos(\theta_i^H) = cr_x/cr$ and $sin(\theta_i^H) = -cr_y/cr$.
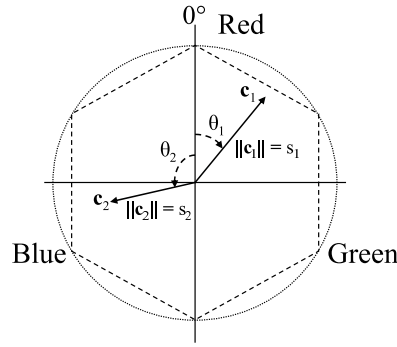
Figure 2: The chromatic plane of the IHLS color space.

The mean direction $\overline{\theta}^H$ is defined to be the direction of the resultant of the unit vectors $\mathbf{h_1}, \ldots, \mathbf{h_n}$ having directions θ_i^H. That is, we have

$$\overline{\theta}^H = \arctan 2 \left(\mathcal{S}, \mathcal{C} \right), \tag{3}$$

where

$$\mathcal{C} = \sum_{i=1}^{n} \cos \theta_i^H, \quad \mathcal{S} = \sum_{i=1}^{n} \sin \theta_i^H \tag{4}$$

and $\arctan 2(y, x)$ is the four-quadrant inverse tangent function.
The mean length of the resultant vector

$$\overline{\mathcal{R}} = \frac{\sqrt{\mathcal{C}^2 + \mathcal{S}^2}}{n}. \tag{5}$$

is an indicator of the dispersion of the observed data. If the n observed directions θ_i^H cluster tightly about the mean direction $\overline{\theta}^H$ then $\overline{\mathcal{R}}$ will approach 1. Conversely, if the angular values are widely dispersed $\overline{\mathcal{R}}$ will be close to 0. The *circular variance* is defined as

$$\mathcal{V} = 1 - \overline{\mathcal{R}} \tag{6}$$

While the circular variance differs from the linear statistical variance in being limited to the range $[0, 1]$, it is similar in the way that lower values represent less dispersed data. Further measures of circular data distribution are given in [14].

2.4   Saturation-weighted hue statistics

The use of statistics solely based on the hue has the disadvantage of ignoring the tight relationship between the chrominance components hue and saturation. For weakly saturated colours the hue channel is unimportant and behaves unpredictably in the presence of colour changes induced by image noise. In fact, for colours with zero saturation the hue is undefined.

As one can see in Figure 2, the chromatic components may be represented by means of Cartesian coordinate vectors $\mathbf{c}_i$ with direction and length given by hue and saturation respectively. Using this natural approach, we introduce the aforementioned relationship into the hue statistics by weighting the unit hue vectors $\mathbf{h}_i$ by their corresponding saturations s_i.

Now, let $(\theta_i^H, s_i), i = 1, \ldots, n$ be n pairs of observations sampled from a population of hue values and associated saturation values. We proceed as described in Section 2.3, with the difference that instead of calculating the resultant of unit vectors, the vectors $\mathbf{c}_i$, which we will dub *chrominance vectors* throughout this paper, have length s_i.

That is, we weight the vector components in Eq. 4 by their saturations s_i

$$\mathcal{C}_s = \sum_{i=1}^{n} s_i \cos \theta_i^H, \quad \mathcal{S}_s = \sum_{i=1}^{n} s_i \sin \theta_i^H, \tag{7}$$

and choose the mean resultant length of the chrominance vectors (for other possible formulations see, e.g. [7]) to be

$$\overline{\mathcal{R}}_n = \frac{\sqrt{\mathcal{C}_s^2 + \mathcal{S}_s^2}}{n}. \tag{8}$$

Consequently, for the mean resultant chrominance vector we get

$$\overline{\mathbf{c}}_{\mathbf{n}} = (\mathcal{C}_s/n, \mathcal{S}_s/n)^T. \tag{9}$$

Here, the length of the resultant is compared to the length obtained if all vectors had the same direction and maximum saturation. Hence, $\overline{\mathcal{R}}_n$ gives an indication of the saturations of the vectors which gave rise to the mean of the chrominance vector, as well as an indication of the angular dispersion of the vectors. To test if a mean chrominance vector $\overline{\mathbf{c}}_{\mathbf{n}}$ is similar to a newly observed chrominance vector, we use the Euclidean distance in the chromatic plane:

$$D = \sqrt{(\overline{\mathbf{c}}_n - \mathbf{c}_o)^T (\overline{\mathbf{c}}_n - \mathbf{c}_o)}, \tag{10}$$

with $\mathbf{c}_o = s_o \mathbf{h}_o$. Here, $\mathbf{h}_o$ and s_o denote the observed hue vector and saturation respectively.

3   The IHLS Background Model

With the foundations laid out in Section 2.4 we proceed with devising a simple background subtraction algorithm based on the IHLS colour model and saturation-weighted hue statistics. Specifically, each background pixel is modelled by its mean luminance μ_y and associated standard deviation σ_y, together with the mean chrominance vector $\overline{\mathbf{c}}_n$ and the mean Euclidean distance σ_D between $\overline{\mathbf{c}}_n$ and the observed chrominance vectors (see Eq. 10).

On observing the luminance y_o, saturation s_o, and a Cartesian hue vector $\mathbf{h}_o$ for each pixel in a newly acquired image, the pixel is classified as foreground if:

$$|(y_o - \mu_y)| > \alpha \sigma_y \quad \vee \quad \|\overline{\mathbf{c}}_n - s_o \mathbf{h}_o\| > \alpha \sigma_D \tag{11}$$

where α is the foreground threshold, usually set between 2 and 3.5.

In order to decide whether a foreground detection was caused by a moving object or by its shadow cast on the static background, we exploit the chrominance information of the IHLS space. A foreground pixel is considered as shaded background if the following three conditions hold:

$$y_o < \mu_y \quad \wedge \quad |y_o - \mu_y| < \beta \mu_y, \tag{12}$$

$$s_o - \overline{\mathcal{R}}_n < \tau_{ds} \tag{13}$$

$$\|\mathbf{h}_o \overline{\mathcal{R}}_n - \overline{\mathbf{c}}_n\| < \tau_h, \tag{14}$$

where $\overline{\mathcal{R}}_n = \|\overline{\mathbf{c}}_n\|$ (see Eq. 8,).

These equations are designed to reflect the empirical observations that cast shadows cause a darkening of the background and usually lower the saturation of a pixel, while having only limited influence on its hue. The first condition (Eq. 12) works on the luminance component, using a threshold β to take into account the strength of the predominant light source. Eq. 13 performs a test for a lowering in saturation, as proposed by Cucchiara et al. [6]. Finally, the lowering in saturation is compensated by scaling the observed hue vector $\mathbf{h}_o$ to the same length as the mean chrominance vector $\overline{\mathbf{c}}_n$ and the hue deviation is tested using the Euclidean distance (Eq. 14).

This, in comparison to a check of angular deviation (see Eq. 31 or [6]), also takes into account the model's confidence in the learned chrominance vector. That is, using a fixed threshold τ_h on the Euclidean distance relaxes the angular error-bound in favour of stronger hue deviations at lower model saturation value $\overline{\mathcal{R}}_n$, while penalising hue deviations for high saturations (where the hue is usually more stable).

4 Metrics for Motion Segmentation

The quality of motion segmentation can in principle be described by two characteristics. Namely, the spatial deviation from the reference segmentation, and the fluctuation of spatial deviation over time. In this work, however, we concentrate on the evaluation of spatial segmentation characteristics. That is, we will investigate the capability of the error metrics listed below, to describe the spatial accuracy of motion segmentations.

- Detection rate (DR) and false alarm rate (FR)

$$DR = \frac{TP}{FN + TP} \tag{15}$$

$$FR = \frac{FP}{N - (FN + TP)} \tag{16}$$

  where TP denotes the number of true positives, FN the number of false negatives, FP the number of false positives, and N the total number of pixels in the image.

- Misclassification penalty (*MP*)
  The obtained segmentation is compared to the reference mask on an object-by-object basis; misclassified pixels are penalized by their distances from the reference objects border [15].

$$MP = MP_{fn} + MP_{fp} \tag{17}$$

  with

$$MP_{fn} = \frac{\sum_{j=1}^{N_{fn}} d_{fn}^j}{D} \tag{18}$$

$$MP_{fp} = \frac{\sum_{k=1}^{N_{fp}} d_{fp}^k}{D} \tag{19}$$

  Here, d_{fn}^j and d_{fp}^k stand for the distances of the jth false negative and k^{th} false positive pixel from the contour of the reference segmentation. The normalised factor D is the sum of all pixel-to-contour distances in a frame.

- Rate of misclassifications (*RM*)
  The average normalised distance of detection errors from the contour of a reference object is calculated using [16]:

$$RM = RM_{fn} + RM_{fp} \tag{20}$$

  with

$$RM_{fn} = \frac{1}{N_{fn}} \sum_{j=1}^{N_{fn}} \frac{d_{fn}^j}{D_{diag}} \tag{21}$$

$$RM_{fp} = \frac{1}{N_{fp}} \sum_{k=1}^{N_{fp}} \frac{d_{fp}^k}{D_{diag}} \tag{22}$$

N_{fn} and N_{fp} denote the number of false negative and false positive pixels respectively. q D_{diag} is the diagonal distance within the frame.

- Weighted quality measure (*QMS*)
  This measure quantifies the spatial discrepancy between estimated and reference segmentation as the sum of weighted effects of false positive and false negative pixels [17].

$$QMS = QMS_{fn} + QMS_{fp} \tag{23}$$

with

$$QMS_{fn} = \frac{1}{N} \sum_{j=1}^{N_{fn}} w_{fn}(d_{fn}^j) d_{fn}^j \tag{24}$$

$$QMS_{fp} = \frac{1}{N} \sum_{k=1}^{N_{fp}} w_{fp}(d_{fp}^k) d_{fp}^k \tag{25}$$

N is the area of the reference object in pixels. Following the argument that the visual importance of false positives and false negatives is not the same, and thus they should be treated differently, the weighting functions w_{fp} and w_{fn} were introduced:

$$w_{fp}(d_{fp}) = B_1 + \frac{B_2}{d_{fp} + B_3} \tag{26}$$

$$w_{fn}(d_{fn}) = C \cdot d_{fn} \tag{27}$$

In our work for a fair comparison of the change detection algorithms with regard to their various decision parameters, receiver operating characteristics (ROC) based on detection rate (DR) and false alarm rate (FR) were utilised.

5 Experiments and Results

We compared the proposed IHLS method with three different approaches from literature. Namely, a RGB background model using either NRGB- (*RGB+NRGB*), or HSV-based (*RGB+HSV*) shadow detection, and a method relying on NRGB for both background modelling and shadow detection (*NRGB+NRGB*).

All methods were implemented using the *Colour Mean and Variance* approach to model the background [18]. A pixel is considered foreground if $|c_o - \mu_c| > \alpha\sigma_c$ for any channel c, where $c \in \{r, g, l\}$ for the Normalised RGB and $c \in \{R, G, B\}$ for the RGB space respectively. o_c denotes the observed value, μ_c its mean, σ_c the standard deviation, and α the foreground threshold.

The tested background models are maintained by means of exponentially weighted averaging [18] using different learning rates for background and foreground pixels. During the experiments the same learning and update parameters were used for all background models, as well as the same number of training frames.

For Normalised RGB (*RGB+NRGB*, *NRGB+NRGB*), shadow suppression was implemented based on Horprasert's approach [3, 4]. Each foreground pixel is classified as shadow if:

$$\begin{aligned} l_o &< \mu_l \quad \wedge \quad l_o > \beta\mu_l \\ |r_o - \mu_r| &< \tau_c \quad \wedge \quad |g_o - \mu_g| < \tau_c \end{aligned} \tag{28}$$

where β and τ_c denote thresholds for the maximum allowable change in the intensity and colour channels, so that a pixel is considered as shaded background.

In the HSV-based approach (*RGB+HSV*) the RGB background model is converted into HSV (specifically, the reference luminance μ_v, saturation μ_s, and hue μ_θ) before the following shadow tests are applied. A foreground pixel is classified as shadow if:

$$\beta_1 \leq \frac{v_o}{\mu_v} \quad \leq \quad \beta_2 \tag{29}$$

$$s_o - \mu_s \quad \leq \quad \tau_s \tag{30}$$

$$|\theta_o^H - \mu_\theta| \quad \leq \quad \tau_\theta \tag{31}$$

The first condition tests the observed luminance v_o for a significant darkening in the range defined by β_1 and β_2. On the saturation s_o a threshold on the difference is performed. Shadow lowers the saturation of points and the difference between images and the reference is usually negative for shadow points. The last condition takes into account the assumption that shading causes only small deviation of the hue θ_o^H [6].

For the evaluation of the algorithms, three video sequences were used. As an example for a typical indoor scene *Test Sequence 1*, recorded by an AXIS-211 network camera, shows a moving person in a stairway. For this sequence, ground truth was generated manually for 35 frames. *Test Sequence 2* was recorded with the same equipment and shows a person waving books in front of a coloured background. For this sequence 20 ground truth frames were provided. Furthermore in *Test Sequence 3* the approaches were tested on 25 ground truth frames from the *PETS2001* dataset 1 (camera 2, testing sequence). Example pictures of the dataset can be found in Figure 3.
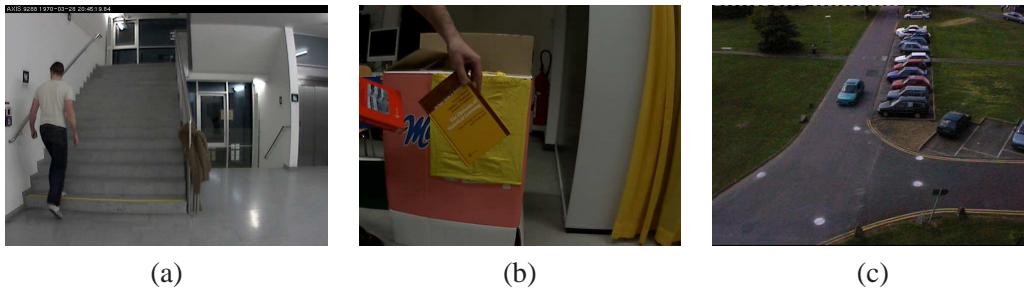

| (a) | (b) | (c) |

Figure 3: Evaluation dataset: *Test Sequence 1* (a), *Test Sequence 2* (b), *Test Sequence 3* (c)

For a dense evaluation, we experimentaly determined suitable ranges for all parameters and sub-sampled them in ten steps. Figure 7 shows the convex hulls of the points in ROC space obtained for all parameter combinations. We also want to point out that *RGB+HSV* was tested with unnormalised and normalised saturation; however, since the normalised saturation consistently performed worse, we omit the results in the ROC for clarity of presentation.

As one can see, our approach outperforms its competitors on *Test Sequence 1*. One reason for this is the insensitivity of the *RGB+NRGB* and *NRGB+NRGB* w.r.t. small colour differences at light, weakly saturated colours. *RGB+HSV*, however, suffered from the angular hue test reacting strongly to unstable hue values close to the achromatic axis. For conservative thresholds (i.e. small values for τ_c or τ_θ) all three approaches either detected shadows on the wall as foreground, or, for larger thresholds failed to classify the beige t-shirt of the person as forground. Figure 4 shows output images from *Test Sequence 1*. We present the source image (a), the ground truth image (b), the resulting image from our approach (c), and the resulting images from the algorithms we compared with. I.a. it is shown that the shirt of the person in image (c) is detected with higher precision as in the images (d), (e), and (f), where it is mostly marked as shadow.

For *Test Sequence 2* the advantageous behaviour of our approach is even more evident. Although the scene is composed of highly saturated, stable colours, *RGB+NRGB* and *NRGB+NRGB* show rather poor results, again stemming from their insufficient sensitivity for bright colours. *RGB+HSV* gave better results, but could not take full advantage of the colour information. Similar hue values for the books and the background resulted
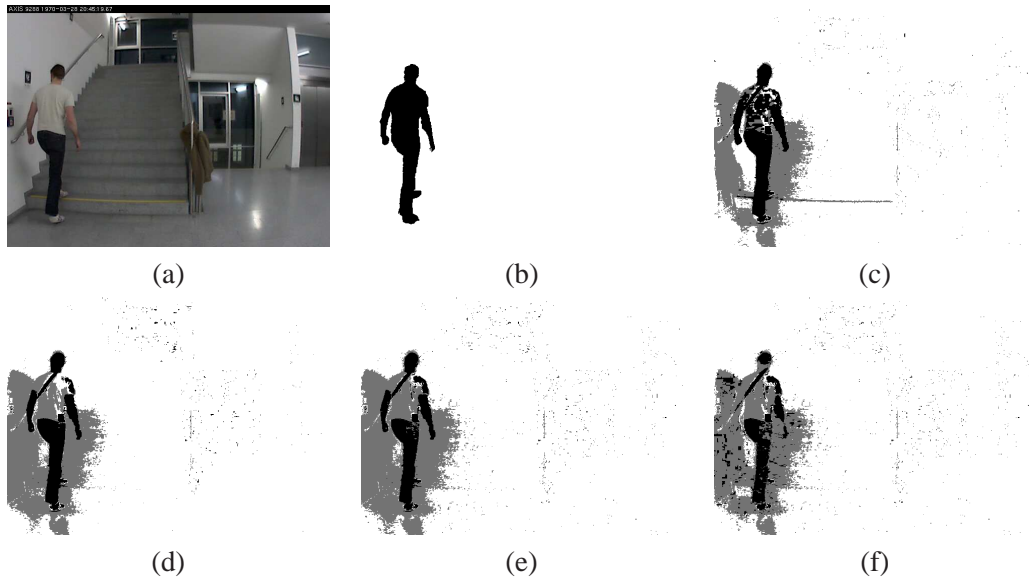
Figure 4: Output images *Test Sequence 1*: *Source Image* (a), *Ground Truth* (b), *Our Approach* (c), *RGB+NRGB* (d), *NRGB+NRGB* (e), *RGB+HSV* (f)

in incorrectly classified shadow regions. Figure 5 shows output images from *Test Sequence 2*. Especially the lower left part of the images (c), (d), (e), and (f) visualizes a better performance of the IHLS approach.
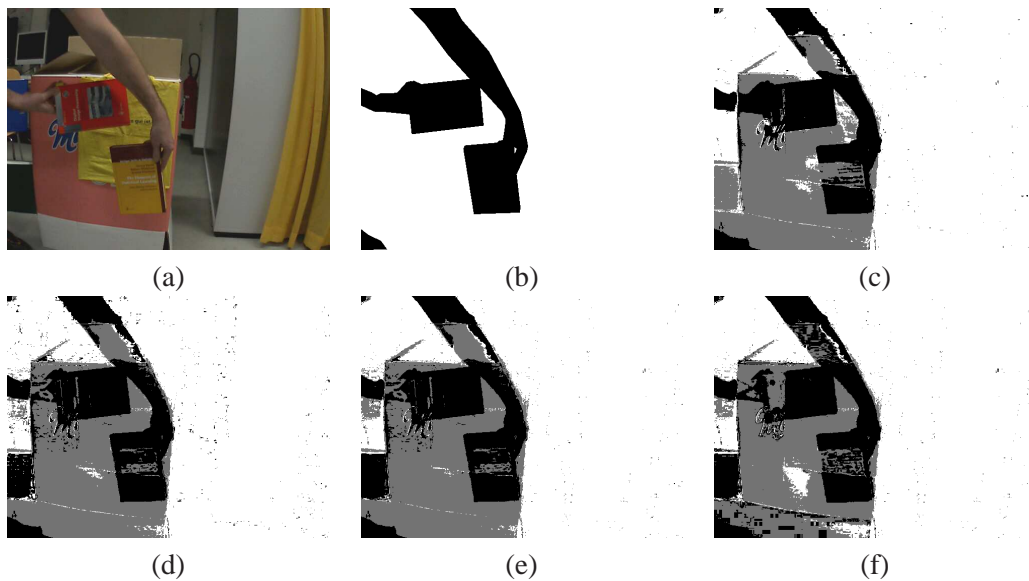

Figure 5: Output images *Test Sequence 2*: *Source Image* (a), *Ground Truth* (b), *Our Approach* (c), *RGB+NRGB* (d), *NRGB+NRGB* (e), *RGB+HSV* (f)

The *Test Sequence 3* sequence shows the problems of background modelling using NRGB already mentioned in Section 2. Due to the low brightness and the presence of noise in this scene, the chromatic components are unstable and therefore the motion detection resulted in an significantly increased number of false positives. *RGB+NRGB* and our approach exhibit similar performance (our approach having the slight edge), mostly relying on brightness checks, since there was not much useable information in shadow regions. *RGB+HSV* performed less well, having problems to cope with the unstable hue information in dark areas. Figure 6 shows output images *Test Sequence 3*.
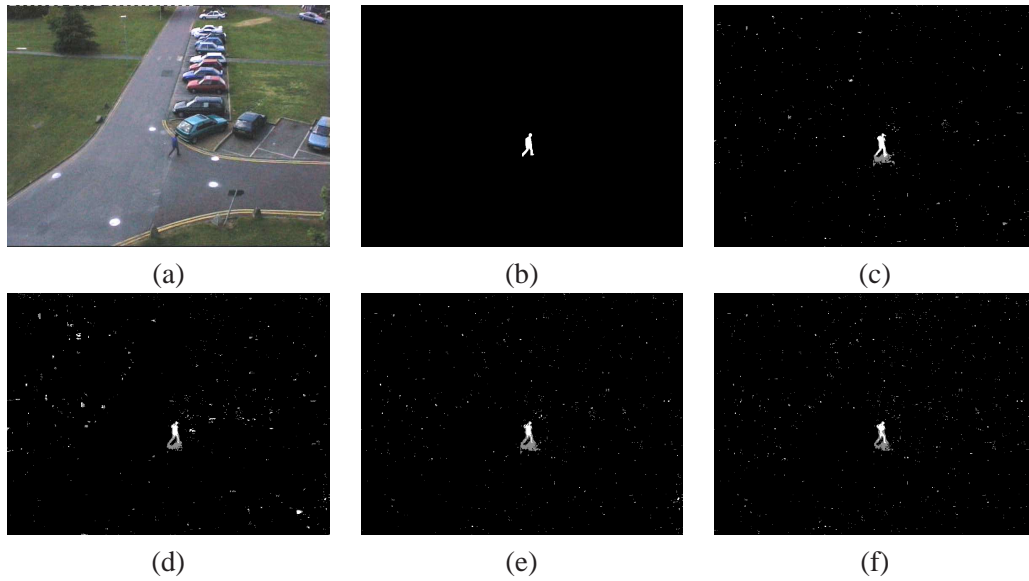
Figure 6: Output images *Test Sequence 3*: *Source Image* (a), *Ground Truth* (b), *Our Approach* (c), *RGB+NRGB* (d), *NRGB+NRGB* (e), *RGB+HSV* (f)

6   Conclusion

We proposed the usage of the IHLS colour space for change detection and shadow suppression in visual surveillance tasks. In the proposed framework, we advocate the application of saturation-weighted hue statistics to deal with the problem of the unstable hue channel at weakly saturated colours.

We have shown that our approach outperforms the approaches using Normalised RGB or HSV in several challenging sequences. Furthermore, our experiments have shown that it is not advisable to use NRGB for background modelling due to its unstable behaviour in dark areas.

One problem of our approach, however, is the fact that due to the use of saturation weighted hue statistics, it is impossible to tell whether a short chrominance vector in the background model is the result of unstable hue information or of a permanent low saturation. Although in the conducted experiments no impairments were evident, it is subject of further research in which cases this shortcoming poses a problem. Other fields of interest are the examination of alternatives to the Euclidean distance for the comparison of the chrominance vectors and an experimental in-depth-investigation of the shadow classification.
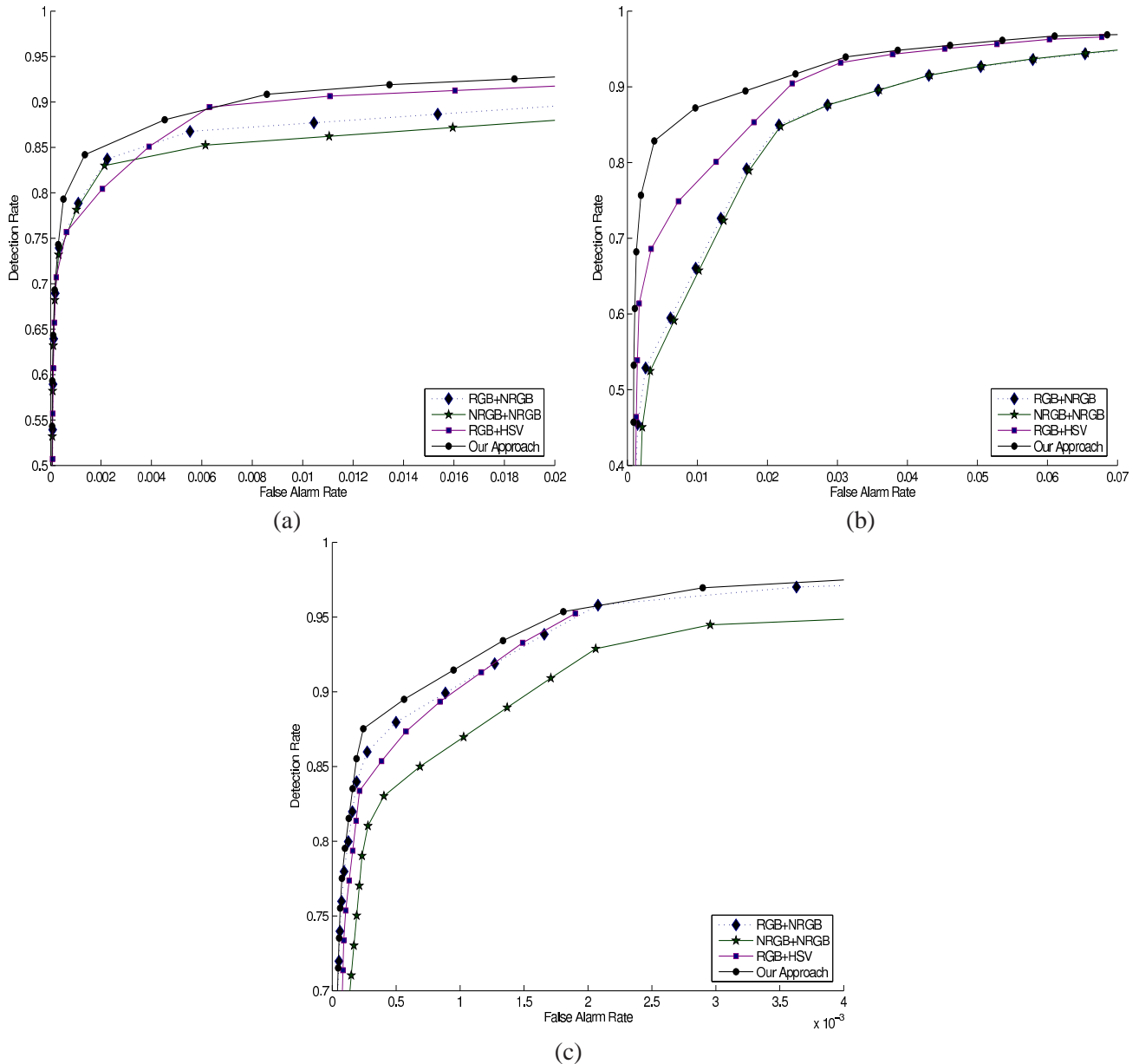
Figure 7: Experimental results: ROCs for *Test Sequence 1* (a), *Test Sequence 2* (b), and *Test Sequence 3* (c).

References

[1] D. Hong & W. Woo, A Background Subtraction for Vision-based User Interface, *Proc. 4th International Conference on Information, Communications and Signal Processing and Fourth IEEE Pacific-Rim Conference on Multimedia*, 2003, 263–267.

[2] S. McKenna, S. Jabri, Z. Duric, H. Wechsler, & A. Rosenfeld, Tracking Groups of People, *Computer Vision and Image Understanding*, *80*(1), 2000, 42–56.

[3] J. Ferryman, M. Borg, D. Thirde, F. Fusier, V. Valentin, F. Brémond, M. Thonnat, J. Aguilera, & M. Kampel, Automated Scene Understanding for Airport Aprons, *Proc. Australian Joint Conference on Artificial Intelligence*, Australia, 2005, 593–503.

[4] T. Horprasert, D. Harwood, & L. Davis, A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection, *Proc. IEEE Conference on Computer Vision, FRAME-RATE Workshop*, 1999.

[5] A. R. J. François & G. G. Medioni, Adaptive Color Background Modeling for Real-Time Segmentation of Video Streams, *Proc. International Conference on Imaging Science, Systems, and Technology*, 1999, 227–232.

[6] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, & S. Sirotti, Improving Shadow Suppression in Moving Object Detection with HSV Color Information, *Proc. Intelligent Transport Systems*, 2001, 334–339.

[7] A. Hanbury & W. G. Kropatsch, Colour Statistics for Matching in Image Databases, *Proc. 27th OEAGM Workshop*, Austria, 2003, 221–228.

[8] G. D. Finlayson, B. Schiele, & J. L. Crowley, Comprehensive Colour Image Normalization, *Proc. 5th European Conference on Computer Vision*, 1998, 475–490.

[9] J. R. Kender, Saturation, Hue and Normalized Colors: Calculation, Digitisation Effects, and Use, Tech. rep., Department of Computer Science, Carnegie Mellon University, 1976.

[10] A. Hanbury, A 3D-polar coordinate colour representation well adapted to image analysis, *Proc. 13th Scandinavian Conference on Image Analysis*, 2003, 804–811.

[11] A. R. Smith, Color gamut transform pairs, *Computer Graphics*, *12*(3), 1978, 12–19.

[12] R. C. Gonzalez & R. E. Woods, *Digital Image Processing* (Englewood Cliffs, NJ: Prentice-Hall, 2002).

[13] D. Hearn & M. P. Baker, *Computer Graphics* (Englewood Cliffs, NJ: Prentice Hall, 1997).

[14] K. V. Mardia, *Statistics of Directional Data* (London, UK: Academic Press, 1972).

[15] C. Erdem & B. Sankur, Performance Evaluation Metrics for Object-Based Video Segmentation, *Proc. 10th European Signal Processing Conference, Tampere, Finnland*, 2000, 917–920.

[16] T. Schlögl, C. Beleznai, M. Winter, & H. Bischof, Performance Evaluation Metrics for Motion Detection and Tracking, *Proc. Proceedings of the International Conference on Pattern Recognition, Cambridge, UK*, 2004, 519–522.

[17] P. Villegas & X. Marichal, Perceptually-Weighted Evaluation Criteria for Segmentation Masks in Video Sequences, *IEEE Transactions on Image Processing*, (8), 2004, 1092–1103.

[18] C. R. Wren, A. Azarbayejami, T. Darrel, & A. Pentland, Pfinder: Real-Time Tracking of the Human Body, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(7), 1997, 780–785.