

Spectral-360: A Physics-Based Technique for Change Detection

Mohamed Sedky

Faculty of Computing, Engineering
and Sciences
Staffordshire University
Stafford, United Kingdom
m.h.sedky@staffs.ac.uk

Mansour Moniri

Faculty of Computing, Engineering
and Sciences
Staffordshire University
Stafford, United Kingdom
m.moniri@staffs.ac.uk

Claude C. Chibelushi

Faculty of Computing, Engineering
and Sciences
Staffordshire University
Stafford, United Kingdom
c.c.chibelushi@staffs.ac.uk

Abstract— This paper presents and assesses a novel physics-based change detection technique, Spectral-360, which is based on the dichromatic color reflectance model. This approach, uses image formation models to computationally estimate, from the camera output, a consistent physics-based color descriptor of the spectral reflectance of surfaces visible in the image, and then to measure the similarity between the full-spectrum reflectance of the background and foreground pixels to segment the foreground from a static background. This method represents a new approach to change detection, using explicit hypotheses about the physics that create images. The assumptions which have been made are that diffuse-only-reflection is applicable, and the existence of a dominant illuminant. The objective evaluation performed using the ‘changedetection.net 2014’ dataset shows that our Spectral-360 method outperforms most state-of-the-art methods.

Keywords— image formation models; change detection; colour constancy, background modelling

I. INTRODUCTION

Robust change detection techniques with high detection probabilities and low-false alarm rates are a subject of extensive research. Developments of such techniques have become important in the past three decades. This is due to the fact that change detection is a key computer vision component which contributes to a wide spectrum of applications in diverse disciplines [1].

The task of change detection is to extract pixels which are changing due to motion of objects in image sequences of a dynamic scene and labelling such pixels which belong to the moving (foreground) objects or regions from the rest of the scene (background). The output of these techniques is a foreground mask, a set of connected pixels which correspond to moving objects (foreground). The mask provided by such techniques can then be considered as a valuable low-level visual cue to perform high-level object analysis tasks such as object recognition, tracking, and activity classification.

Apart from the motion of object(s) relative to the background or disappearance of background object(s), foreground mask errors may result from a combination of different physical problems, being a consequence of the image formation (e.g. illumination variation, shadows, reflections, sensor noise, in addition to camouflage foreground, similar background-foreground colored objects),

which result in two types of errors, false alarms and missed detections.

The research community has addressed, with various degrees of success, the technical problems associated with change detection and foreground segmentation techniques. However, very few of the proposed techniques are up to the challenges of a real-world operating environment, where robustness, adaptability, and flexibility are essential.

Change detection approaches based on background modelling are widely used. A detailed modelling of all type of changes is required to segment relevant changes for a given application.

Background modelling may be decomposed into four major steps, image representation, statistical manipulation, threshold selection and background model updating. The literature suggests that the best combination between these steps is a key success to develop a real-time, robust and adaptive change detection algorithm.

An important factor in change detection is the choice of the image representation and the transformation that is applied to the raw data in order to obtain the information that is relevant to the specific application domain.

Change detection approaches may be classified depending on the image representation used as physics-based or non-physics-based. Non-physics based change detection methods use one of the known color spaces as a cue to model the scene. While the word physics refers to the extraction of intrinsic features about the materials contained in the scene based on an understanding of the underlying physics which govern the image formation.

Most of the proposed change detection techniques use well-known color spaces to represent the image, trying to model the variation in illumination and camera noise using statistical modelling techniques. The gap appears to be in the image representation used which does not give any consideration to the elements which govern the camera output. Few solutions have been proposed in the literature with regard to the use of physical models of image formation to solve the change detection problem.

This paper proposes a novel physics-based change detection approach. First, we introduce image formation models in Section II. In Section III, our Spectral-360 approach is discussed in detail. In Section IV, we compare our results to other change detection methods using the ‘changedetection.net 2014’ dataset.

II. IMAGE FORMATION MODELS

Appearances of scenes depend on four fundamental elements: an illuminant, a medium, a material and an imaging device. The illuminant represents the source of visible electromagnetic energy and is characterized by its Spectral Power Distribution (SPD).

The electromagnetic waves travel from the light source through a medium, and eventually reach a physical object in the scene. The surface material of this object reflects the light; it effectively modulates the incident electromagnetic energy. This property is represented by the Surface Spectral Reflectance (SSR), the fraction of incident radiation reflected by the surface. The reflected electromagnetic energy will eventually reach the imaging device. The imaging device is characterized by the spectral sensitivities of its photosensitive sensors, which represent the response of such elements to the received electromagnetic waves.

Apart from the aforementioned spectral features which characterize the elements of the scene appearance, another important parameter is the set of geometrical features, which represent the scene structure, the illuminant orientation, the surface roughness and the viewing geometry. These features combine non-linearly to yield a digital image.

Recovering these features from images is an important problem in image processing; however, this recovery is generally hard with the limited amount of information provided by standard commercial imaging devices.

The dichromatic model represents the light reflected by an inhomogeneous dielectric material as a linear combination of diffuse and specular reflections. Each of these parts is further divided into two elements, one accounting for the geometry and another purely spectral.

There is a great number of reflection models, most of them developed in the field of computer graphics. Among these methods, the dichromatic reflection model in [2] is a usual choice for those algorithms employing a physical model to represent color images, as shown in Figure 1, and the equation below:

$$I_c = w_d \int_{\Omega} E(\lambda) S(\lambda) Q_c(\lambda) d\lambda + w_s \int_{\Omega} E(\lambda) Q_c(\lambda) d\lambda \quad (1)$$

Where Ω is the visible range from 400nm to 700nm, I_c is the measured color intensity of the reflected light, w_d and w_s are geometrical parameters for diffuse and specular reflection respectively, $E(\lambda)$ is the spectral power distribution function of the illumination, $S(\lambda)$ is the SSR of the object, $Q_c(\lambda)$ is the camera sensor spectral sensitivities characteristic, and c represents the color channels (Red, Green and Blue).

Assuming Lambertian surfaces, theoretically, an image taken by a digital color camera (for diffuse only reflection) can be described as:

$$I_c = w_d \int_{\Omega} E(\lambda) S(\lambda) Q_c(\lambda) d\lambda \quad (2)$$

For what concerns the surface properties, this model assumes that: the surface is opaque; not optically active (no fluorescence) and the colorant is uniformly distributed.

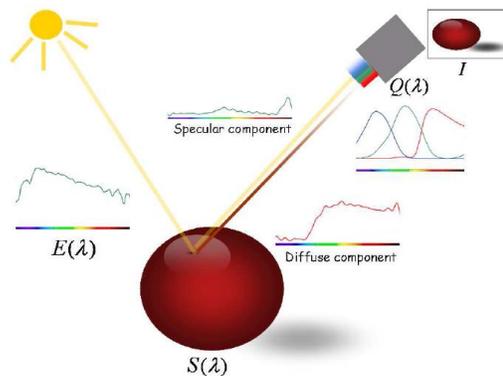


Figure 1 Schematic diagram of image formation

III. SYSTEM DESCRIPTION

Spectral-360 adopts physics-based image formation models to denote the image representation. In particular Spectral-360 uses the dichromatic reflection model [2]. The assumptions made by the dichromatic reflection model are: there is a single light source that can be a point source or an area source; the illumination has a constant SPD across the scene and the amount of illumination can vary across the scene.

Our approach is different from all the previous work, in that it relies on models, which can represent wide classes of surface materials. It makes use of the pre-trained linear SSR models, i.e. Parkkinen basis functions [3], shown in Figure 3, to represent the SSR of the objects in the scene.

As Figure 2 shows, our technique starts with the training phase where N training images are acquired ($N=60$) which represent relatively static background frames. The illumination estimation stage calculates the Correlated Color Temperature (CCT) of the illuminant using McCamy's formula [4]. In order to build a computational physical model, we have represented the SSR using 3 basis functions of Parkkinen. Representative sensitivities of camera sensors are obtained from a camera manufacturer data sheet [5]. The weights of Parkkinen basis functions, which represent the SSR, are then recovered from the RGB values. Then, the mean SSR of the background training images is calculated and taken as a reference SSR. In order to build the background model, the technique passes through all background frames and calculates the correlation between each SSR of the N background frames and the reference SSR. Then, the maximum and minimum correlations for each pixel are added to the background model.

After building the background model, the testing phase starts by capturing a new frame, and then the SSR recovery module converts the new frame to its SSR weights using the estimated CCT. Then the statistical manipulation module calculates the correlation between the new SSR with the reference SSR. A decision stage uses an adaptive threshold to generate a binary foreground mask, using the maximum and minimum values of the correlation measured from the pool of background frames.

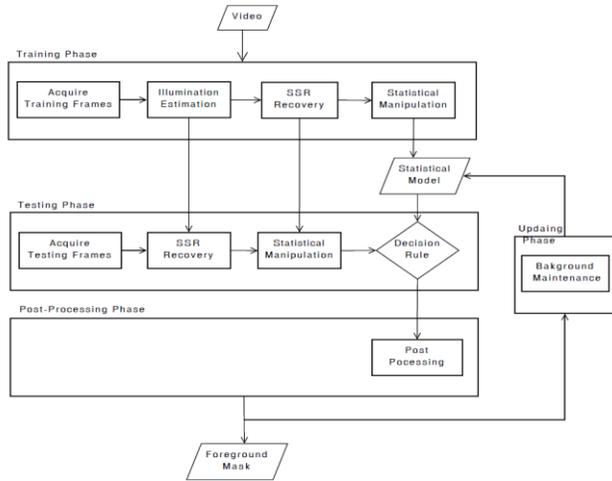


Figure 2 Proposed system block diagram

After the segmentation of the foreground mask, a post-processing phase applies a post-processing step which performs a 9x9 median filter to smooth the changed regions within the segmented mask.

The background model, mean SSR as well as maximum correlation and minimum correlation, is evaluated by observing the scene during a training period; this model should be updated continuously.

Continuing updating of the model makes the foreground extraction more robust to illumination change or objects that become part of the background. This occurs when a moving object originally enters the field of view and then stops moving. The system then updates the background model through a model maintenance phase, using an update rate of 0.005. Static objects for more than 300 frames are gradually updated to the background model with an update rate of 0.01. A condition has been added to accommodate the Pan-Tilt-Zoom (PTZ) scenario, by resetting the background model if the number of pixels in the foreground mask is greater than half the frame size and the difference in CCT in current frame and previous frame is greater than 100°K.

A. Illumination estimation

This stage aims to calculate the Correlated Color Temperature (CCT) as an illumination parameter of the illuminant. McCamy's method [4] is used in our approach, which is a simple equation to compute the CCT from CIE 1931 chromaticity coordinates x and y :

$$n = \frac{x - x_e}{y_e - y} \quad (3)$$

Where $x_e = 0.3320$ and $y_e = 0.1858$ then:

$$CCT = 449n^2 + 3525n^3 + 6823.3n + 5520.33 \quad (4)$$

This equation proves useful for implementation in real-time applications. This method has a maximum absolute error of less than 2°K for colour temperatures ranging from 2856°K to 6500°K.

B. Surface Spectral Reflectance Recovery

Assuming diffuse-only materials, and using finite-dimensional linear models to represent the SSR, the physical model can be re-written as:

$$I_c = w_d \int_{\Omega} E(\lambda) \left(\sum_{i=1}^3 w_i \phi_i(\lambda) \right) Q_c(\lambda) d\lambda \quad (5)$$

where $\phi_i(\lambda)$ is the i^{th} reflectance linear basis function, and w_i is its corresponding weight. Using this equation, an *RGB* pixel could be linearly transformed to its equivalent SSR represented by three basis function weights.

C. Spectral Similarity Measure

Our statistical manipulation sub-module measures the spectral similarity between the spectral signatures of two SSRs. Cha [6] has carried out a comprehensive survey on similarity/distance measures. He has categorized various similarity/distance measures as Minkowski family, L_1 family, intersection family, inner product family, fidelity family, Squared L_2 family or χ^2 family, Shannon's entropy family as well as combination of these families. In our proposed algorithm, the similarity between spectral pixels are measured using the inner product (correlation) similarity measure, other similarity measures could be used.

D. Decision Rule

The threshold operation creates an initial binary foreground mask defined as FG_{Mask} , using the maximum C_{max} and the minimum C_{min} values of the correlation measured from the pool of background frames.

$$FG_{Mask} = \begin{cases} 0 & \tau_{min} C_{min} < C < \tau_{max} C_{max}, \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

where the thresholds are chosen experimentally as follows:

$$\tau_{max} = 1.0315 \text{ and } \tau_{min} = 0.9563$$

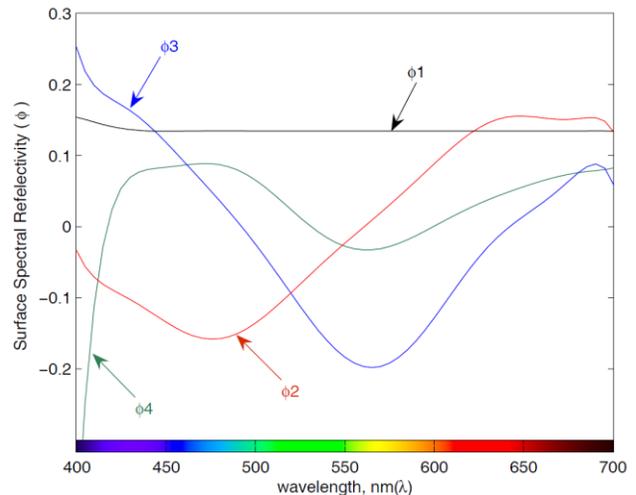


Figure 3 Parkkinen's first 4 basis functions [3]

Table 1: Results of Spectral-360 on all eleven scenarios using all seven measures.

	Recall	Specificity	FPR	FNR	PBC	F-measure	Precision
Baseline	0.9616	0.9968	0.0032	0.0384	0.4265	0.9330	0.9065
Camera Jitter	0.6696	0.9906	0.0094	0.3304	2.0855	0.7142	0.8387
Dynamic Background	0.7819	0.9992	0.0008	0.2181	0.3513	0.7766	0.8456
Intermittent Object Motion	0.5878	0.9835	0.0165	0.4122	5.3734	0.5609	0.7374
Shadow	0.8898	0.9893	0.0107	0.1102	1.5682	0.8519	0.8187
Thermal	0.7238	0.9939	0.0061	0.2762	1.6337	0.7764	0.9114
Bad Weather	0.7032	0.9977	0.0023	0.2968	0.6804	0.7569	0.8211
Low Frame Rate	0.7515	0.9941	0.0059	0.2485	0.8964	0.6437	0.5946
Night Videos	0.6237	0.9739	0.0261	0.3763	4.4642	0.4832	0.4610
PTZ	0.5047	0.9416	0.0584	0.4953	6.0771	0.3653	0.3265
Turbulence	0.8815	0.9859	0.0141	0.1185	1.4375	0.5429	0.4982

Table 2: Comparison of Spectral-360 to several state-of-the-art methods using all seven proposed performance measures.

	Recall	Specificity	FPR	FNR	PBC	F-measure	Precision
Spectral-360 (our method)	0.7345	0.9861	0.0139	0.2655	2.2722	0.6732	0.7054
KNN [8]	0.6650	0.9802	0.0198	0.3350	3.3200	0.5937	0.6788
GMM Stauffer & Grimson [9]	0.6846	0.9750	0.0250	0.3154	3.7667	0.5707	0.6025
KDE - ElGammal [10]	0.7375	0.9519	0.0481	0.2625	5.6262	0.5688	0.5811
Mahalanobis distance [11]	0.1644	0.9931	0.0069	0.8356	3.4750	0.2267	0.7403
GMM Zivkovic [12]	0.6604	0.9725	0.0275	0.3396	3.9953	0.5566	0.5973
Euclidean distance [11]	0.6803	0.9449	0.0551	0.3197	6.5423	0.5161	0.5480

IV. PERFORMANCE EVALUATION

The Changedetection.net (CD.net) challenge [7] aims to initiate a rigorous and comprehensive academic benchmarking effort for testing and ranking existing and new algorithms for change and motion detection.

In Table 1, our Spectral-360 results using the seven performance measures, used by CD.net, are shown for all eleven scenarios. The results of our Spectral-360 is compared to several state-of-the-art methods (made available by CD.net prior to the change detection workshop) in Table 2. Our Spectral-360 method performs better than all other methods in terms of ‘Percentage of Bad Classification’ (PBC) and F-measure. El-Gammal scores the best Recall and False Negative Rate (FNR) at the cost of low Precision. On the other side, Mahalanobis distance scores the best Specificity, False Positive Rate (FPR) and Precision at the cost of very low Recall.

V. CONCLUSION

This paper argues that image formation models offer interesting new alternative physics-based cues for change detection compared with other representations used in conventional methods. Features such as surface spectral reflectance have not been applied yet in the field of change detection. The reason is the computational complexity of such models, and hence possible unfeasibility of real-time implementation. This challenge was tackled, firstly, by choosing an appropriate reflection model, the dichromatic reflection model. Secondly, by setting a feasible set of assumptions for such model which best match the reduction of model complexity and does not contradict with real-world operational conditions. In this paper we have proposed a physics-based change detection approach which estimates the full spectrum of surface spectral reflectance from the video and then segments the moving objects.

REFERENCES

- [1] M. Sedky, M. Moniri and C. C. Chibelushi, “Classification of smart video surveillance systems for commercial applications,” Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance, pp. 638 – 643, 2005.
- [2] G. J. Klinker, S. A. Shafer, and T. Kanade, “A physical approach to colour image understanding,” in: International Journal on Computer Vision, vol. 4, no. 1, pp. 7-38, 1990.
- [3] J. Parkkinen, J. Hallikainen and T. Jaaskelainen, “Characteristic spectra of Munsell colours,” in: Journal of the Optical Society of America A, vol. 6, no. 2, pp. 318-322, 1989.
- [4] McCamy and S. Calvin, “Correlated colour temperature as an explicit function of chromaticity coordinates,” in: Journal of Colour Research & Application, vol. 17, no. 2, pp. 142-144, 1992.
- [5] <http://www.eureca.de/datasheets/01.xx.xxxx/01.02.xxxx/01.02.0185/CX098BQ.pdf> (31th of March 2014, date last accessed)
- [6] S. H. Cha, “Comprehensive survey on distance/similarity measures between probability density functions,” International Journal of Mathematical Models and Methods in Applied Sciences, City, 1(2), 1, pp. 300–307, 2007.
- [7] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, “changedetection.net: A new change detection benchmark dataset,” in Proc. IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, 2012.
- [8] Z. Zivkovic , F. van der Heijden, “Efficient adaptive density estimation per image pixel for the task of background subtraction” Pattern Recognition Letters, vol. 27, no. 7, pp. 773-780, 2006.
- [9] C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking,” in Proc. Int. Conf. on Computer Vision and Pattern Recognition, Vol. 2, 1999.
- [10] A. Elgammal, D. Harwood, and L. Davis, “Non-parametric model for background subtraction,” in Proc. Eur. Conf. on Computer Vision, Lect. Notes Comput. Sci. 1843, 751-767 2000.
- [11] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, and C. Rosenberger. “Comparative study of background subtraction algorithms” in J. of Elec. Imaging, vol., no. 3, pp. 1–12, 2010.
- [12] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction,” in Proc. Int. Conf. Pattern Recognition, pp. 28-31, IEEE, Piscataway, NJ 2004.