STUDY ON FOREGROUND SEGMENTATION METHODS FOR A 4D STUDIO

CORINA BLAJOVICI[1], ZSOLT JANKÓ[2], AND DMITRY CHETVERIKOV[2]

ABSTRACT. A 4D reconstruction studio is an intelligent environment that enables 3D modelling of moving actors and deformable objects. The visual quality of the final 3D model, in terms of both geometry and texture, is greatly influenced by the precision and accuracy of the segmented foreground object. This paper extends our previous work on the image segmentation methods developed for the 4D studio at MTA SZTAKI, Budapest, Hungary. The studio uses a three-step approach for extracting the foreground silhouette: (i) background subtraction using spherical coordinates, (ii) foreground post-processing using a colour filtering approach based on the background colour, (iii) detection and removal of casted shadows. We give an overview of these techniques and perform a comparative evaluation in terms of both quantitative measures and qualitative analysis. We discuss our results in various imaging conditions, such as illumination variations.

1. INTRODUCTION

Obtaining models of dynamic 3D objects plays an important role in content generation for computer games, film production, interactive media, motion analysis and other applications. Several multi-camera systems for dynamic 3D modeling, also called 4D reconstruction studios, have been proposed. A brief survey of advanced 4D studios is presented in [14]. A 4D reconstruction studio is a room with characteristic background, usually green or blue, equipped with lighting sources, multiple calibrated, synchronised, high-resolution video cameras and appropriate computing facilities. The main objectives of a 4D reconstruction studio are: capturing a scene from multiple viewpoints and

automatically building time-varying 3D models of real objects or actors moving in the scene.

Dynamic 3D models of actors can be used to populate virtual worlds or models of large-scale real-world scenes obtained by other sensors such as LI-DAR. Such a system for reconstruction and visualisation of complex scenes that integrates both outdoor 4D data measured by a rotating multi-beam LIDAR sensor, and 4D models obtained in a 4D studio is presented in [2]. Furthermore, real-time 4D reconstruction can enhance classical means of remote communication by adding realistic sensations to visual perception. With a suitable facility, such as a 3D monitor or a CAVE, 3D models can create an immersive and realistic experience.

The 4D Studio at MTA SZTAKI [14, 12] is, to the best of our knowledge, the first 4D Reconstruction Studio in Eastern Europe. The basic hardware and configuration of the studio are presented in [14]. The studio can operate in off-line [14] and real-time modes [12]. The main steps of the reconstruction process are: extracting colour images from captured data, segmentation into foreground and background, creating volumetric model using the Visual Hull algorithm [15], creating the triangulated mesh from the volumetric model using the Marching Cubes algorithm [16] and finally adding texture to the model.

For both off-line and real-time modes, image segmentation is a key step, since the visual quality of the 3D model highly depends on the accuracy of the extracted silhouette. In real-time mode, a simplified segmentation procedure is implemented. This is less robust and precise, but much faster [12]. For the off-line version, a three-step approach aimed at removing the background, is developed. The foreground image is obtained by background subtraction [4]. The result is then filtered using the constraint that the background colour of the studio is green [6]. In the end, shadows are detected and removed to obtain the final foreground silhouette [4].

This paper extends our previous work on foreground segmentation for the 4D studio at MTA SZTAKI, presented in [6]. First, we perform a comparative evaluation of the segmentation methods proposed for removing misclassified background regions using the background colour of the studio. Second, we analyse how using different colour representations for the shadow model can improve shadow detection in 4D studio images. We compare different colour representations, in terms of both quantitative measures and qualitative analysis, and discuss our results.

The rest of this paper is structured as follows. In Section 2, we present relevant background-foreground segmentation approaches proposed for other 4D studios around the world, and review shadow detection methods based on different colour spaces. We give an overview of the background removal methods developed for the 4D studio at MTA SZTAKI. The solutions for detecting

green regions belonging to the background that may still be present after background subtraction are presented in Section 3. Section 4 describes the shadow physical model and how it is used in different colour spaces. Experimental results for evaluation of the shadow detection approaches and exploiting the background colour are discussed in Section 5. Our conclusions are presented in Section 6.

2. Background

2.1. **Related Work.** Precise and robust foreground segmentation is a key element in many computer vision applications such as 3D reconstruction from silhouettes, object tracking and vision-based motion capture systems. Background subtraction is one of the most common approaches used for separating moving foreground objects from the scene [21]. Chroma-keying approaches [26] have also been used for extracting the foreground object from the background in controlled environments with uniform, possibly constant-coloured background, usually blue or green. The real-time 3D modeling system proposed by Petit et al. [20] uses background subtraction in YUV colour space, based on a combination of a Gaussian model for the chromatic information and an interval model for the intensity information. Vlasic et al. [27] use in a multi-view reconstruction studio a combination of background subtraction and chroma-keying to obtain the foreground silhouettes, which are further used to create visual hulls for an initial phase of geometry reconstruction. Wu et al. [28] setup a high-speed multi-view performance capture environment with a green screen background, where background subtraction is performed by simple thresholding in the HSV colour space using the background image captured by each camera prior to recording.

Foreground objects obtained by different segmentation methods might still be noisy due to shadows that might be misclassified as foreground. The reader is further referred to [22], [24], and [1] for comprehensive surveys on methods for detecting moving shadows developed over the past decade. Most shadow detection approaches are based on the assumption that pixels in the shadow region are darker but have the same chromaticity. Cucchiara et al. [9, 8] used HSV colour space for shadow detection. They show that the value component is darker in the shadow regions, while hue and saturation vary within certain limits. Horprasert et al. [13] proposed a method that uses the RGB representation to separate the colour into chromaticity and brightness components. Other exploited colour spaces are the normalised RGB [18, 7], c1c2c3 [23], YUV [17, 25]. Petit et al. [20] use for shadow detection in their real-time 3D reconstruction the method based on the approach of Horprasert et al. [13]. Methods based on chromaticity have been preferred in computer vision due

to their simplicity and speed. However, such methods are more sensitive to noise, since they are pixel-based.

Some studies evaluate the efficiency of colour representation for shadow detection. Benedek et al. [3] proposed a framework for evaluating the HSV, RGB, c1c2c3, CIE L*u*v colour spaces in simple indoor environments, highways with dark shadows and outdoor scenes with difficult illumination. For all these cases, they reported that CIE L*u*v was the most effective.

2.2. **Overview of foreground segmentation for the 4D studio at MTA SZTAKI.** The foreground-background segmentation developed for the 4D Studio at MTA SZTAKI has been recently presented in [4, 5]. The method is based on background subtraction using spherical colour representation that is known to improve robustness to illumination changes and shadows in optical flow calculation [19]. Scalar difference between the reference background image and the processed image is calculated pixel-by-pixel and stored in a greyscale difference image. Foreground pixels are select as outliers by a robust outlier detection algorithm, assuming that background pixels form a majority. In [5] a comparison of three colour spaces is performed for background subtraction, namely the RGB, the normalised RGB, and the spherical colour representation. The authors concluded that the spherical colour representation makes the background subtraction algorithm more robust and thus less sensitive to illumination changes.

In the second step, foreground extraction exploits the background colour of the studio to detect green regions that might have been assigned to the foreground during background subtraction. Although using background colour for segmentation removes also shadows casted on the ground, this is only available when illumination conditions are good and thus shadows are not too dark. In addition, when the foreground object contains such green regions, the method for background removal based on green colour cannot be applied.

Shadows that are casted by the object itself tend to be classified as part of the foreground and can cause object shape distortion. Therefore, shadows are detected and removed from the silhouettes obtained in the previous steps. Shadow removal starts by detecting the ground region of the studio, using a simple single seeded region growing algorithm in the HSV colour space and then applying shadow detection inside this area [4, 5]. This improves robustness of the overall process since shadows are detected only in the expected areas.

At the end, the morphological post-processing cleans noise and removes small holes in the locations where the object colour is similar to that of the background. Holes in the silhouettes lead to holes in the 3D shape since the

Visual Hull algorithm calculates the spatial intersection of the back-projected silhouettes. Therefore small size holes are filled.

3. Segmentation based on background colour

The studio is equipped with green walls and green ground, simulating a green-screen environment. The background colour is essential information for segmentation in order to obtain high-quality silhouettes. The background colour can be used with the goal of identifying and reducing green parts belonging to the background that are not removed by background subtraction.

There are a number of factors that make the task of detecting such mis-classified green regions more difficult. Due to limitations of lighting, there exist darker regions in the background, where the green colour cannot be easily identified. Reflections of the green background on the foreground objects change the colour of the borders in areas where the colour is brighter or tends to be more bluish. Another factor that can negatively impact the quality of the segmentation is motion blur resulted from fast activities in the scene.

For completeness we give an overview of the segmentation methods aimed at removing such green parts: the Adaptive Green Filtering [6] and the Basic Green Filtering [4] methods.

The basic background colour filtering described in [4] is based on the assumption that all background pixels have the green component larger than the red one and the blue one. Thus the green pixels belonging to the background are detected based on the following condition:

$$(1) \qquad F^G - F^R > \tau_1 \land F^G - F^B > \tau_2$$

Here F^k are the colour components in the foreground image, $k = R, G, B$ and τ_1 and τ_2 are manually set thresholds.

In order to exploit the green colour information in a more robust manner, less sensitive to lighting, the adaptive green filtering method [6] introduces the green factors, which are defined as follows:

$$(2) \qquad \varphi_1 = \frac{I^G}{I^R}, \varphi_2 = \frac{I^G}{I^B}$$

Here I^k, $k = R, G, B$ are the red, green and blue channels for each pixel in a given image I. For each pixel, the green factors are computed by taking the average of the green factors over a 5×5 window.

Using these factors, green regions belonging to the background are detected when following conditions are satisfied:

$$(3) \qquad \varphi_i^B > \alpha_i \land \varphi_i^F > \alpha_i, \quad i = 1, 2.$$

$$(4) \qquad (1 - t \cdot \varphi_i^B) < \frac{\varphi_i^B}{\varphi_i^F} < (1 + t \cdot \varphi_i^B), \quad i = 1, 2.$$

Here φ_i^B and $\varphi_i^F, i = 1, 2$, are the green factors for a given pixel in the background image and the foreground image, respectively; α_i are thresholds set manually to 1.1 throughout all our tests. The threshold t represents the tolerance to green and is usually set to $0.2 - 0.3$, depending on the greenness of the foreground, i.e. if parts of the foreground are green, selecting a smaller threshold value makes the algorithm more restrictive, assuring that parts of the foreground are not misclassified as background. Condition 4 makes the method more reliable, as it can distinguish between green colour variations, while it is adapting to the background colour. Since the method does not make use of a global threshold, it can better handle situations when green reflection and other green parts add ambiguities in deciding whether a pixel belongs to the background or not.

4. Shadow Detection

Shadow detection methods can be formulated based on the physics-based model of illumination [11]. According to this illumination model, the colour components of the reflected intensity reaching the sensors at a given point are defined as follows:

$$(5) \qquad \rho^k(e) = \int_\Omega (i^T n) \int_{\lambda=400}^{\lambda=700} E(i, \lambda) R(i, e, \lambda) S^k(\lambda) \mathrm{d}\lambda \mathrm{d}i$$

Here Ω is the set of all possible directions of incoming lights, i and e the unit vectors towards the directions of the incident, respectively, outgoing rays, and n the surface normal in the given point. $E(i, \lambda)$ is the illumination function, $R(i, e, \lambda)$ the reflectance factor, $S^k(\lambda)$ the spectral sensitivities of the sensors of a colour camera for each channel $k = R, G, B$ and λ the wavelength parameter.

Assuming the camera sensors $S^k(\lambda)$ to be Dirac delta functions, i.e. $S^k(\lambda) = s^k \delta(\lambda - \lambda^k)$ [10], shadow effects in the sensor response can be described by the shadow factor:

$$(6) \qquad c^k = \frac{F^k(x)}{B^k(x)}$$

F^k and B^k are the components of a shadow pixel in the frame F and the background image B, respectively. The shadow factor $c^k \in [0; 1]$ since shaded pixels are darker than the background.

This physics model is exploited in a number of colour spaces. Colour representations such as HSV, YUV, CIE L*u*v have been used for shadow

detection, since they provide a natural separation into chromatic and luminance components. Also, photometric colour invariants that are invariant to changes in imaging conditions are exploited. Such photometric invariants are the normalised RGB and c1c2c3, which are defined only on chromaticity components. Benedek et. al. [3] show that the constant ratio rule is a good approximation for the luminance components, while the chromaticity components vary slightly for the shadowed surface.

The ground region of the studio is first detected [4, 5] and shadow detection is applied in this area. Additionally, too dark pixels are excluded from shadow analysis, since they could mislead the detection process. Next sections explore various colour spaces for shadow detection.

4.1. **Shadow model in the RGB colour space.**
In case of real shadows, the shadow factor must be the same at each sensor, i.e. $c^R = c^G = c^B$. This assumption is valid if the spectral distribution is the same at all light sources and when illuminants of the same type are used. In [4, 5] shadow pixels are selected based on following condition:

$$(7) \qquad |c^R - c^G| < \tau_C \wedge |c^R - c^B| < \tau_C \wedge |c^G - c^B| < \tau_C$$

Here τ_C is a threshold whose value is empirically chosen once per sequence of images. Selecting a smaller threshold value makes the algorithm more sensitive to small changes in shading. Shadow factors are computed for each pixel over a 5×5 window in order to make the method less sensitive to noise.

4.2. **Shadow model in the HSV colour space.**
The shadow model in the HSV colour representation is formulated based on the assumption that when illumination changes, the value component changes, but hue and saturation changes are small. Thus, the presence of a shadow is verified when following condition is satisfied [9]:

$$(8) \qquad \beta_1 \leq \frac{F_p^V}{B_p^V} \leq \beta_2 \wedge F_p^S - B_p^S \leq \tau_S \wedge |F_p^H - B_p^H| \leq \tau_H$$

Here F^k and B^k are the components of the pixel p in the foreground image F and background B, $k = H, S, V$. β_1, β_2, τ_S and τ_V are empirically selected thresholds.

4.3. **Shadow model in the c1c2c3 colour space.**
Using just a chromatic model such as c1c2c3 and not taking into consideration the luminance can result in limitations in accuracy, especially for bright and dark objects, or objects similar in colour to the background. Therefore, candidate shadow pixels are selected based on:

$$(9) \qquad c^k < \tau_c, \quad k = R, G, B.$$

The threshold τ_c is set to 0.9 to make the method more sensitive to changes in shading. Similarly to [23], shadow is then detected by making the assumption that the chromaticity components differ slightly in the shadow pixels:

$$(10) \qquad |c_1^F - c_1^B| < \tau \wedge |c_2^F - c_2^B| < \tau \wedge |c_3^F - c_3^B| < \tau,$$

Here c_i^F, c_i^B are the c1c2c3 components for the foreground image and background image, respectively; τ is a threshold empirically chosen for defining the fluctuations in chromaticity. The differences are computed for a 5×5 window to make the method more robust.

4.4. **Shadow model in the normalised RGB colour space.** The normalised RGB representation is often selected because of its fast calculation. Similarly to the c1c2c3 model, the method selects a set of candidate shadow pixels using the condition 9. Finally, the shadow is detected when the following assumption is satisfied:

$$(11) \qquad |r^F - r^B| < \tau_r \wedge |b^F - b^B| < \tau_b$$

Here r^F, r^B, b^F, b^B are the normalised RGB components of the image F and the background B.

5. Results and Discussions

5.1. **Qualitative Evaluation.** In order to test the post-processing algorithms, we have selected 8 sets of 13 images each, acquired by the 13 cameras of the 4D studio. The results obtained using both the adaptive green filtering (AGF) and the basic green filtering (BGF) are shown in figure 1. Since BGF is more sensitive to lighting conditions and to similarities between background and foreground, some tuning of parameters is needed for some of the tested sequences, in order to optimise the visual appearance of the silhouette. On the other hand, the results obtained with AGF have been generated using the thresholds $t = 2.5, \alpha_1 = 1.1, \alpha_2 = 1.1$, which are the same throughout all the tests. AGF performs better in darker green regions, since it depends on the background colour at each pixel, rather than on global thresholds such as BGF. In addition, AGF is more robust to green reflections of the background that appear on brighter regions of the actor's cloths, while BGF removes such regions, visually distorting the contour of the foreground object shape.

Since shadows casted on the ground are also green, the proposed approach successfully removes such green shadow regions in areas where the shadows are not so strong. However, dark shadows cannot be completely removed and shadow detection is still necessary. Figure 2 gives examples of results obtained by the RGB, HSV, c1c2c3 and the normalized RGB (NRGB) representations for the shadow models. The shadow detection based on RGB colour space

(A) Input   (B) Adaptive Filtering   (C) Basic Filtering
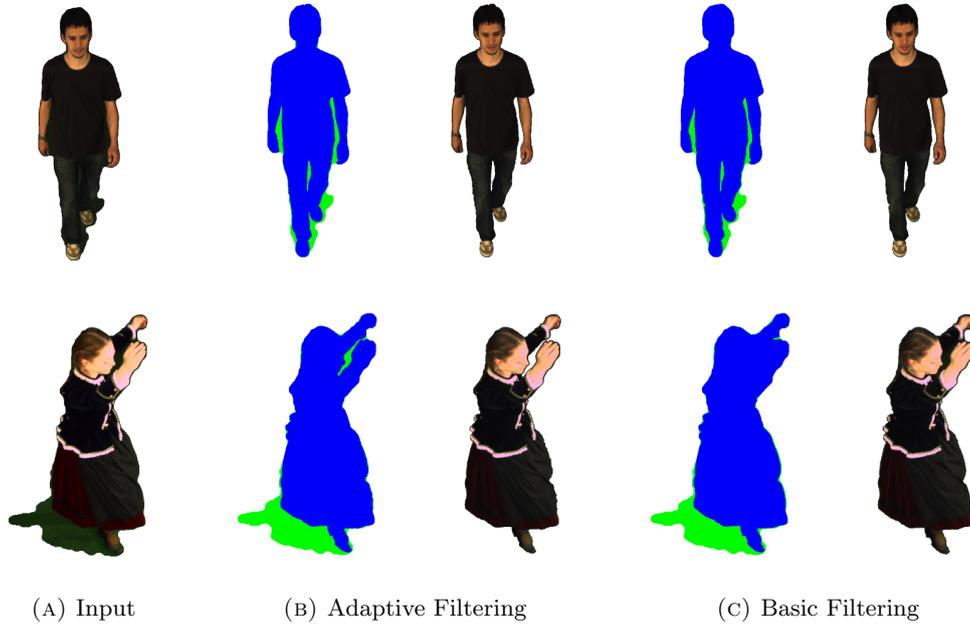
FIGURE 1. Sample results of filtering based on background colour using the two described approaches. The foreground object is shown in blue, the background regions in green.

performs slightly better as it removes the darker shadow regions. Although all the methods show promising results, they have a few limitations and we analyse them in the specific context of the 4D studio.

Due to the assumption that the chromaticity is preserved in the shadow area, they are sensitive to colour similarities between the foreground object and background, such as green reflections and motion blur. However, the shadow detection is applied only in the ground region, which improves the robustness of the detection with respect to such similarities. In addition, all the methods are sensitive to dark regions, such as black clothing or dark hair, which tend to mislead the detection process. Different viewing angle provide different illumination conditions, and some sequences have darker frames with strong shadows created on the ground. For such sequences, results were obtained with different settings for the thresholds extracting the dark pixels, in order to improve the accuracy of the segmentation. However the thresholds are set once for the entire sequence, since it is not possible to change them for individual frames in the 4D studio pipeline.

(A) Input   (B) RGB   (C) HSV   (D) c1c2c3   (E) NRGB
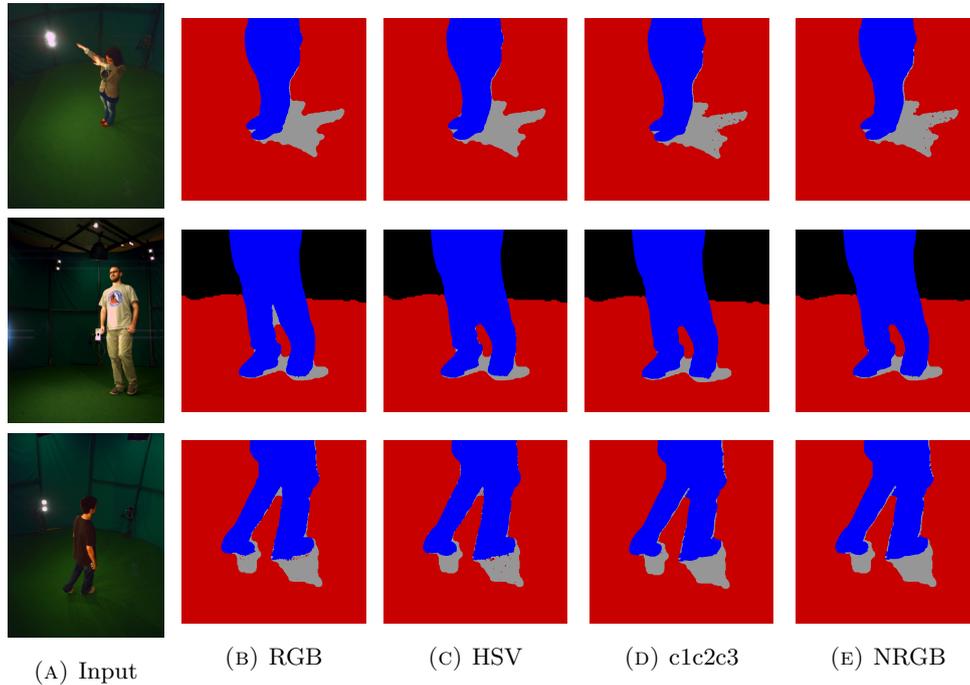
FIGURE 2. Sample results of the four shadow detection methods. The foreground object is shown in blue, the shadow in grey, the ground region in red.

5.2. **Quantitative Evaluation.** The quantitative evaluation is performed with respect to a ground-truth segmentation of the foreground object. We have manually created ground-truth segmentations of the foreground object for 6 frames, selected for evaluation in various conditions, such as different viewing angle and different clothing. In addition, some illumination changes occur between selected frame from different sets despite the fixed cameras. The selected images are shown in figure 3.

There are two types of errors that can appear in the segmentation: false positives and false negatives. False positives are background pixels classified as belonging to the foreground, while false negatives are foreground pixels that have been misclassified as background. To quantify the results of post-processing, we use the weighted false positives and false negatives measurements defined in [5]. Each pixel incorrectly classified is assigned a weight,

FIGURE 3. Images selected for quantitative comparison of the discussed methods

defined by the following formula:

$$(12) \qquad w(u) = \frac{1}{1 + \sum_{v \in N(u)} S_v},$$

where $N(u)$ is the 8-neighbourhood of the pixel u. For false positives, $S(v)$ is 1 if v belongs to the foreground in the ground truth image, 0 otherwise. Smaller weights are assigned to false positive pixels close to silhouette edges, since it is often impossible to define the sharp edge of a silhouette. On the other hand, more distant pixels are assigned with larger weights. For false negatives, its vice versa.

Table 1 summarises the results of the comparison for the two segmentation methods based on the background colour. We observe that the overall performance of the adaptive green filtering is better than of the basic filtering. The number of false positives is smaller, since AGF successfully detects larger green regions, which otherwise are misclassified as belonging to the foreground. The false positive pixels are usually located on darker and shadow areas, which cannot be detected as green regions. Both methods tend to remove some border pixels of the shape, resulting in false negatives. Figure 4 gives examples of false positives and false negatives for the two methods.

For an initial quantitative comparison of the discussed shadow detection methods, we perform the same evaluation based on the weighted false positives and false negatives. The same selected ground-truth images are used. Figure 5 shows examples of false positives and false negatives obtained using the four shadow models. False positives are located in dark shadow regions and on

TABLE 1. Weighted False Positives and False Negatives for AGF and BGF

	False Positives		False Negatives	
	AGF	BGF	AGF	BGF
Image 1	1074.1	1067.9	137.2	184.1
Image 2	2375.0	2936.6	14.3	78.6
Image 3	6408.6	8488.4	48.5	43.4
Image 4	3102.6	6584.2	3.7	2.0
Image 5	1457.4	2624.1	871.6	1501.9
Image 6	983.1	2284.5	42.0	15.9

FIGURE 4. Sample false positives and false negatives for the AGF for (first column) and BGF (second column) for *Image* 1 and *Image* 5, respectively. False positive pixels are shown in red, while false negative pixels are shown in cyan. For *Image* 5 the pixels around the leg are classified as false negatives for both methods, due to motion blur.

borders around the shape. Such borders result from background subtraction, since it tends to extend the shape. The shadow detection methods in the RGB and HSV colour spaces are better at removing the strong shadow in the darker area. While both remove false positives, RGB performs slightly better in some small areas that are nevertheless important for the final 3D model. In the c1c2c3 colour space, the shadow detection performs slightly better than in the normalised RGB. As illustrated in figure 5, both fail to remove some small dark shadow regions. We also show an example where the shadow detection methods are sensitive to colour similarity with the background. This is illustrated for the second row in figure 5, where the RGB method has the

(A) RGB   (B) HSV   (C) c1c2c3   (D) NRGB

FIGURE 5. Sample false positives and false negatives for the four shadow detection methods for *Image* 2 (first row) and *Image* 5 (second row), respectively. False positive pixels are marked in red, while false negative pixels are in cyan.

TABLE 2. Weighted False Positives for the four Shadow Detection Methods

	RGB	HSV	c1c2c3	NRGB
Image 1	1068.9	963.1	1000.7	1044.4
Image 2	3057.4	2417.2	2861.9	2843.3
Image 3	6249.3	5964.0	6267.2	6309.6
Image 4	2962.1	2872.3	2851.4	2776.3
Image 5	3148.8	3884.2	3410.5	3200.7
Image 6	2661.2	2718.6	3498.4	3509.7

biggest number of false pixels. Here HSV performs better, but it removes a smaller part of the shadow region. Tables 2 and 3 present the number of false positives and false negatives, respectively, computed for each method.

TABLE 3. Weighted False Negatives for the four Shadow Detection Methods

	RGB	HSV	c1c2c3	NRGB
Image 1	152.60	104.02	134.70	202.67
Image 2	6.72	24.00	29.82	13.37
Image 3	59.85	74.28	117.03	67.65
Image 4	2.42	2.73	10.97	6.88
Image 5	1680.30	873.08	993.42	1131.50
Image 6	2.77	23.43	2.32	2.32

6. CONCLUSION

We have presented our efficient segmentation techniques developed for the 4D studio at MTA SZTAKI. Obtaining precise and clean silhouettes is a critical step in the 4D studio reconstruction process, as it improves the quality of the final 3D model. After obtaining an initial foreground segmentation using background subtraction, the 4D studio uses a segmentation method that exploits the background colour of the studio. This is aimed at removing smaller green regions that might have been misclassified as belonging to the background. Two approaches have been evaluated, namely the Adaptive Green Filtering and the Basic Green Filtering methods. We have shown the robustness of the Adaptive Green Filtering approach by performing qualitative and quantitative comparisons with respect to the manually created ground-truth images. For the quantitative evaluation, a measure of performance was used based on the weighted sums of false positive and false negative pixels.

Additionally, we have examined the problem of colour space representation for shadow detection in 4D studios. We have analysed shadow removal using four colour representations, namely RGB, HSV, c1c2c3 and normalised RGB. The shadow detection method based on RGB presented in [4, 5] eliminates strong shadows better, but it is more sensitive to similarities between background and foreground than the other methods. In the HSV colour space, the shadow detection is more sensitive to dark areas, but because luminance is separated from chrominance, the detection is less sensitive to colour similarities. c1c2c3 gives better results than the normalised RGB, but both leave some small regions for strong shadows in dark areas and are sensitive to similarities in chromaticity. Our conclusions for the shadow detection methods are drawn based on a relatively small ground-truth dataset. In future, we plan to create a larger dataset and evaluate other colour spaces such as YUV and

CIE L*u*v, in order to experiment with more representations that separate the luminance from the chromaticity.

References

[1] N. Al-Najdawi, H. E. Bez, J. Singhai, and E. Edirisinghe, *A survey of cast shadow detection algorithms*, Pattern Recognition Letters, vol. 33, no. 6, pp. 752–764, 2012.

[2] C. Benedek, Z. Jankó, C. Horváth, D. L. Molnár, D. Csetverikov, and T. Szirányi, *An integrated 4d vision and visualisation system*, Lecture Notes in Computer Science, C. Mei, L. Bastian, and N. Bernd, Eds., no. 7963.   Wien: Springer, 2013, pp. 21–30.

[3] C. Benedek and T. Szirányi, *Study on color space selection for detecting cast shadows in video surveillance*, International Journal of Imaging Systems and Technology, vol. 17, no. 3, pp. 190–201, Oct. 2007.

[4] C. Blajovici, D. Chetverikov, and Z. Jankó, *4D studio for future internet: Improving foreground-background segmentation*, Proc. IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom), 2012, pp. 559–564.

[5] C. Blajovici, D. Chetverikov, and Z. Jankó, *Enhanced object segmentation in a 4D studio*, Proc. Conference of the Hungarian Association for Image Processing and Pattern Recognition (KEPAF), 2013, pp. 42–56.

[6] C. Blajovici, J. Zsolt, and D. Chetverikov, *Robust background removal in 4D studio images*, Proc. International Conference on Intelligent Computer Communication and Processing (ICCP), 2013.

[7] A. Cavallaro, E. Salvador, and T. Ebrahimi, *Detecting shadows in image sequences*, Proc. First European Conference on Visual Media Production, 2004, pp. 15–16.

[8] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, *Detecting moving objects, ghosts and shadows in video streams*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 10, pp. 1337–1342, Oct. 2003.

[9] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, *Improving shadow suppression in moving object detection with hsv color information*, Proc. IEEE Intelligent Transportation Systems, Aug. 2001, pp. 334–339.

[10] G. Finlayson, S. Hordley, C. Lu, and M. Drew, *On the removal of shadows from images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 1, pp. 59–68, 2006.

[11] D. A. Forsyth, *A novel algorithm for color constancy*, International Journal of Computer Vision, vol. 5, no. 1, pp. 5–36, Sep. 1990.

[12] J. Hapák, Z. Jankó, and D. Chetverikov, *Real-time 4D reconstruction of human motion*, Proc. 7th International Conference on Articulated Motion and Deformable Objects, ser. Lecture Notes in Computer Science, vol. 7378, 2012, pp. 250–259.

[13] T. Horprasert, D. Harwood, and L. S. Davis, *A statistical approach for real-time robust background subtraction and shadow detection*, Proc. International Conference on Computer Vision, vol. 99, 1999, pp. 1–19.

[14] Z. Jankó, D. Csetverikov, and J. Hapák, *4D reconstruction studio: Creating dynamic 3D models of moving actors*, Proc. 6th Hungarian Conference on Computer Graphics and Geometry, 2012, pp. 1–7.

[15] A. Laurentini, *The visual hull concept for silhouette-based image understanding*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 2, pp. 150–162, 1994.

[16] W. E. Lorensen and H. E. Cline, *Marching cubes: A high resolution 3d surface construction algorithm*, SIGGRAPH Computer Graphics, vol. 21, no. 4, pp. 163–169, Aug. 1987.

[17] N. Martel-Brisson and A. Zaccarin, *Moving cast shadow detection from a gaussian mixture shadow model*, Proc. IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, 2005, pp. 643–648.

[18] I. Mikic, P. C. Cosman, G. T. Kogut, and M. M. Trivedi, *Moving shadow and object detection in traffic scenes*, Proc. 15th International Conference on Pattern Recognition, vol. 1, 2000, pp. 1321–1324.

[19] J. Molnár, D. Chetverikov, and S. Fazekas, *Illumination-robust variational optical flow using cross-correlation*, Computer Vision and Image Understanding, vol. 114, no. 10, pp. 1104–1114, Oct. 2010.

[20] B. Petit, J.-D. Lesage, C. Ménier, J. Allard, J.-S. Franco, B. Raffin, E. Boyer, and F. Faure, *Multi-camera real-time 3D modeling for telepresence and remote collaboration*, International Journal of Digital Multimedia Broadcasting, vol. 2010, January 2010.

[21] M. Piccardi, *Background subtraction techniques: a review*, IEEE Transactions on Systems, Man, and Cybernetics, vol. 4, 2004, pp. 3099–3104.

[22] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, *Detecting moving shadows: Formulation, algorithms and evaluation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, pp. 918–923, 2003.

[23] E. Salvador, A. Cavallaro, and T. Ebrahimi, *Cast shadow segmentation using invariant color features*, Computer Vision and Image Understanding, vol. 95, no. 2, pp. 238–259, Aug. 2004.

[24] A. Sanin, C. Sanderson, and B. C. Lovell, *Shadow detection: A survey and comparative evaluation of recent methods*, Pattern Recognition, vol. 45, no. 4, pp. 1684–1695, 2012.

[25] O. Schreer, I. Feldmann, U. Golz, and P. Kauff, *Fast and robust shadow detection in videoconference applications*, Proc. 4th EURASIP IEEE International Symposium on Video Processing and Multimedia Communications, 2002, pp. 371–375.

[26] A. R. Smith and J. F. Blinn, *Blue screen matting*, Proc. 23rd Annual Conference on Computer Graphics and Interactive Techniques, ser. SIGGRAPH '96, 1996, pp. 259–268.

[27] D. Vlasic, P. Peers, I. Baran, P. Debevec, J. Popović, S. Rusinkiewicz, and W. Matusik, *Dynamic shape capture using multi-view photometric stereo*, ACM Transactions on Graphics, vol. 28, no. 5, pp. 174:1–174:11, 2009.

[28] D. Wu, Y. Liu, I. Ihrke, Q. Dai, and C. Theobalt, *Performance capture of high-speed motion using staggered multi-view recording*, Computer Graphics Forum, vol. 31, no. 7, pp. 2019–2028, 2012.

[1] FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, BABEŞ-BOLYAI UNIVERSITY, CLUJ-NAPOCA, ROMANIA
E-mail address: `blajovici@cs.ubbcluj.ro`

[2] MTA SZTAKI, BUDAPEST, HUNGARY
E-mail address: `janko.zsolt@sztaki.mta.hu, csetverikov.dmitrij@sztaki.mta.hu`