Effects of camera aperture correction on keying and compositing of broadcast video

Oliver Grau and Jim Easterbrook

BBC Research
Development
Centre House, 56 Wood Lane, London, W12 7SB, UK
phone, email: Oliver.Grau@bbc.co.uk

Abstract

This contribution discusses the effects of camera aperture correction in broadcast video on colour-based keying. The aperture correction is used to 'sharpen' an image and is one element that distinguishes the 'TV-look' from 'film-look'. If a very high level of sharpening is applied, as is the case in many TV productions then this significantly shifts the colours around object boundaries with hight contrast. This paper discusses these effects and their impact on keying and describes a simple low-pass filter to compensate for them. Tests with colour-based segmentation algorithms show that the proposed compensation is an effective way of decreasing the keying artefacts on object boundaries.

Keywords: Image processing, colour-based segmentation, keying, matting.

1 Introduction

The process of classifying the pixels of a digital image into foreground and background is called segmentation or matting (in film industry) and keying (in broadcast). The result of this process is stored as the alpha-channel of an image. The main application for keys or mattes is in compositing, for example to exchange the background of a scene with a different image. In some applications (including image understanding) binary values of the alpha channel 0,1 are acceptable. To achieve high quality results in compositing a continuous value of alpha in the interval [0..1] is obtained. These intermediate values for alpha occur for example in transparent objects or motion blur and in particular at the border of foreground objects when only a part of a pixel belongs to the object (mixed pixels).

A number of approaches have been developed to create a key or matte automatically. Chroma-keying is a long established technique for special effects in film- and TV-productions (see [SB96] for an overview) and relies on the subject being filmed in front of a screen with known colour (usually blue or green). This technique is very robust, but limited to controlled environments. Difference keying is another technique that works in two steps: First an image of the background (the background plate) is acquired. The alpha value is then estimated from the difference of an image with subject to the background plate. More recent work is addressing keying of natural images, i.e. images with varying foreground and background colour. In order to solve the segmentation problem, the user is required to provide a tri-map (for example see [CCSS01, HHR05, RKB04, FJR05]) to indicate regions of foreground, background and uncertain areas.

Although different approaches have been described to generate a key or matte under various assumptions about either the background or foreground object, little or no work has been done that takes into account specific characteristics of broadcast cameras.

First presented at the 5th European Conference on Visual Media Production 2008, extended and revised for JVRB.

1.1 Background on Segmentation Methods

A common application for keyed images is to recombine them with different background. This process is called 'compositing' and can be described in the *compositing equation*:

$$C = \alpha F + (1 - \alpha)B \tag{1}$$

The combined colour C is a weighted sum between the foreground colour F and the background colour B. This equation is only correct in linear colour space. However, our implementation is working with gamma-corrected colour and neglects the effects of gamma correction.

The generation of a key α also includes the task to estimate F and B from the combined colour C and therefore the problem is underdetermined.

The problem to compute alpha can be formulated as a function $f(C)$ that segments the colour space (RGB in our case) into foreground and background areas.

Most methods used for chroma-keying define a *global* segmentation function, for example see Vlahos (summarised in [SB96]) or [Mis92, GTAHS07].

Methods for natural image keying are extending this framework to *local* segmentation functions (for example [CCSS01, HHR05]).

Common to both approaches is the assumption that foreground and background have a different colour distribution that can be used to separate them.

In this paper we are investigating the effects of aperture correction and are testing these with two global segmentation methods 'fast green' and a 'k-nearest-neighbour classifier', as described in [GTAHS07] and a local operating difference keying method. After the segmentation we compute the foreground / background colours for 'mixed pixels' with the method described by Hillman in [HHR05].

Fast Green: This method is implemented in a similar form in commercial chroma-keying devices and is based on the difference between the green channel intensity value for a given pixel and the maximum of the red and blue channel values:

$$D_{fg} = g - \max r, b \tag{2}$$

The segmentation S_{fg} is computed using threshold σ_{fg} :

$$S(x,y) = \begin{cases} 0 & , d_{fg} < \sigma_{fg} \\ 1 & , otherwise \end{cases} \tag{3}$$

The threshold $\sigma_{fg} isadjustedmanuallybyanoperator$.

K-nearest neighbour classifier: This classifier is controlled by a simple GUI: The user clicks on positions in an image that represent background. The RGB colour values of that pixel are stored as a prototype $P_i = I$ into a list. All pixels in the image that are within a radius in RGB space r_1 of the colour prototype are then marked as background as well. The user continues to choose background pixels until the resulting segmentation is satisfying.

The segmentation $S_{k-nearest}$ is computed by finding the nearest colour prototype p_{best} from the list. With the distance d of the pixel RGB values I:

$$d = Distance_{inRGB}(P_{best}, I) \tag{4}$$

the crisp segmentation the result is:

$$S_{k-nearest} = \begin{cases} 0 & , d < r \\ 1 & , otherwise \end{cases} \tag{5}$$

In order to get continuous values a soft key can be obtained using a second radius r_2. See [GTAHS07] for details.

Difference keying: This is another often used segmentation technique. It is based on the difference in colour space between a pixel I of the image and the corresponding pixel I_{bg} in the background plate. The background plate can be created by either taking a picture of the scene without any foreground objects or if this is not possible then alternatively a background plate can be generated by applying a temporal median filter to remove moving foreground objects.

The difference between I and I_{bg} can be computed in any colour space. We used the difference in RGB space here:

$$\delta = Distance_{inRGB}(I, I_{bg}) \tag{6}$$

$$S_{diff} = \begin{cases} 0 & , \delta < \sigma \\ 1 & , otherwise \end{cases} \tag{7}$$

It is obvious that segmentation S_{diff} is sensitive to noise or variation in colour. The threshold σ can be increased only very limited, otherwise foreground objects are not be separated robustly from the background colour.

The rest of this paper is structured as follows: The next section describes the aperture correction used in broadcast cameras and analyses the resulting problems for colour segmentation. Section 3 describes a filter to compensate these effects. In section 4 aspects of how the techniques described in the paper can be used in compositing are discussed.

The paper finishes with some results and conclusions.

2 Aperture correction in broadcast cameras

Broadcast cameras have a control known as aperture correction or sometimes 'detail'. The aperture correction is used to 'sharpen' an image and is one element that distinguishes the 'TV-look' from 'film-look'. Effectively the correction emphasises high-frequency image components and is therefore a high-boost filter. Figure 1 shows an example of a broadcast image. The image was taken during a rugby match with a Sony HDC-1500 high definition camera.

Figure 1: Image of a sport scene from a broadcast camera (top) and detail (bottom).

In particular in sport productions it is quite common to add a lot of 'detail'. The effect can be seen in the detail picture (figure 1 bottom) as over- and undershoots of the image signal (visible as black 'haloes' of white shirts against the background).

In broadcast this feature is intended to get a better 'contrast' of objects. For segmentation the effect is a challenge since significant colour changes take place at the edge of objects.

We tested different settings of a Sony HDC-1500 camera equipped with a Fujinon HA 16x6.3 lens with a colour chart, as depicted in figure 2. The images were capture in $Y'C_bCr$ format (commonly knwon as YUV) and then converted to RGB. The aperture correction (called 'detail' on the Sony cameras) was varied between -100 to +100 in steps of 10.

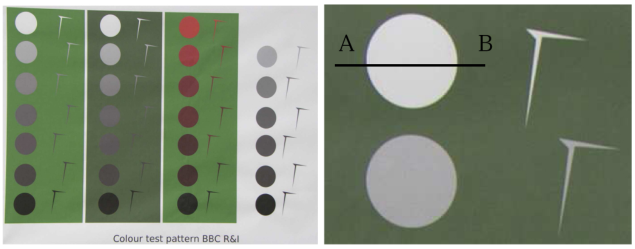

Figure 2: Test chart captured with a Sony 1500 camera.

In figure 3 four selected settings of the aperture correction are shown. In broadcast this setting is left to the camera operator and takes into account the type of lens used, but also creative and stylistic decisions. On sport events such as depicted in figure 1 it is usually between 0 and +40.

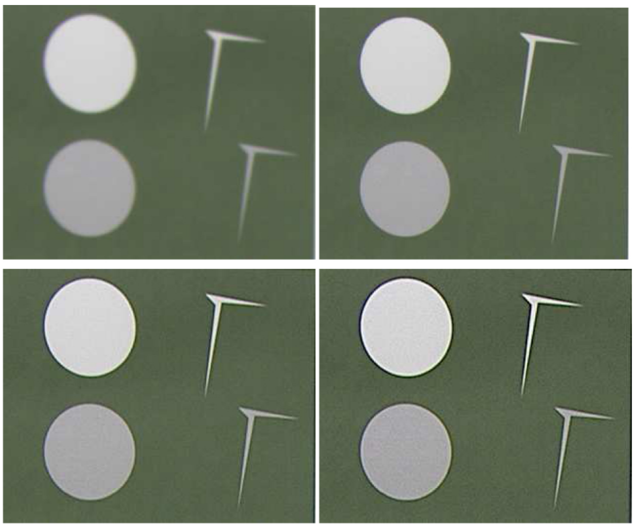

Figure 3: Test chart with different aperture settings (top left -80, top right -40, bottom left 0, bottom right +40).

In figure 4 the luminance profile along the line A-B (depicted in figure 2 right) is shown for the different 'detail' settings. Everything below approximately -40 gets effectively low-pass filtered, -40 was roughly considered to be 'neutral' in our setup and settings above showed high-boost characteristics.

The effects in RGB-colour space are depicted in figures 5 7. The colour distribution for detail = -40, depicted in figure 5 is as expected: The green background and the light grey circle manifest themselves in two clear clusters. Between these clusters a number of 'mixed colours' can be found. Interestingly these are not on a single line, but on two curves. This is a result of lens chromatic aberrations (see conclusions).

The colour distributions of figures 6 and 7 show a significant degradation of the ideal colour distribution and this would potentially lead to overlapping and interfering distributions for more complex scenes.

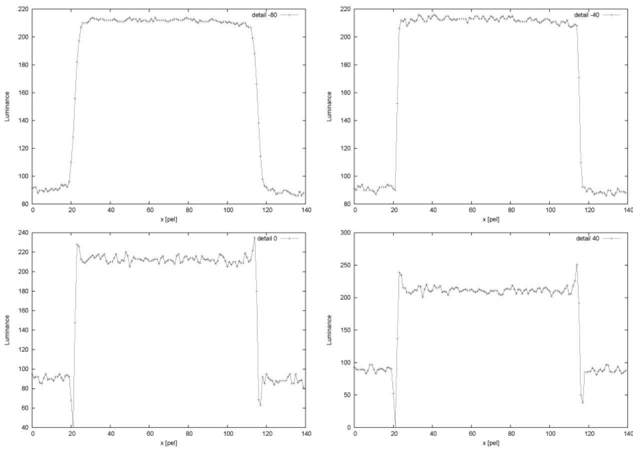

Figure 4: Luminance profiles detail: -80 (top left), -40 (top right), 0 (bottom left) and +40 (bottom right).

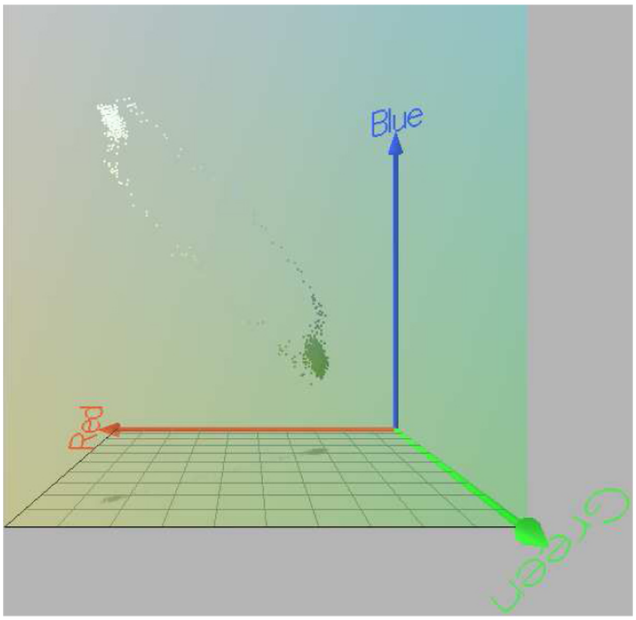

Figure 5: Colour space detail -40.

Looking at the horizontal luminance profiles, as depicted in figure 4, we can see that the edge, with its overshoots, is five or six pixels wide. In theory, this finite aperture filter requires an infinite aperture, or recursive, filter to compensate for it. In practice, however, another FIR filter of similar aperture is sufficient.

So far, we designed the compensation filters by trial and error. Horizontal and vertical compensation processes each use a one dimensional symmetric low pass filter of the following form:

3 Compensating the aperture correction effects

Removing the dark haloes can be achieved with simple filtering. The amount of filtering required has to be adjusted manually, at present, but could be automated based on edge detection and then adjusting the filter to produce maximum flatness in proximity to the edge, coupled with maximum steepness of the edge.

Comparing the chrominance (C_b C_r) values of images captured at different aperture correction settings shows that the aperture correction only affects the luminance (Y) channel. This is to be expected, as the eye does not perceive much detail in chrominance, so applying correction would have little visible effect.

$$y_{out}^{in} = \frac{1}{1 + 2\sum_{i=1}^{j} k_i} \left(y_n + \sum_{i=1}^{j} k_i (y_{n-i} + Y_{n+i}) \right) \quad (8)$$

In this equation the subscripts n, $n-i$ & $n+i$ refer to pixel numbers (in the horizontal filter) or line numbers (in the vertical filter).

Studying the edges in detail it appears that the darkest overshoot pixel is two pixels from the first bright pixel after the transition, so filter coefficient k_2 is adjusted first (setting all other coefficients to zero). Further adjustments are made to all the coefficients until a satisfactory result is obtained. For an aperture correction setting of 0, the compensation filter coefficients are:

$$k_2 = 0.3; k_4 = 0.05 \quad (9)$$

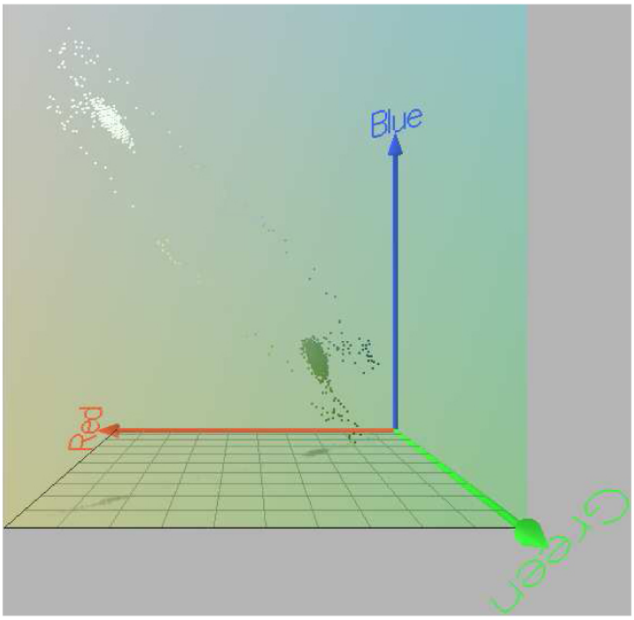

Figure 6: Colour space detail 0.

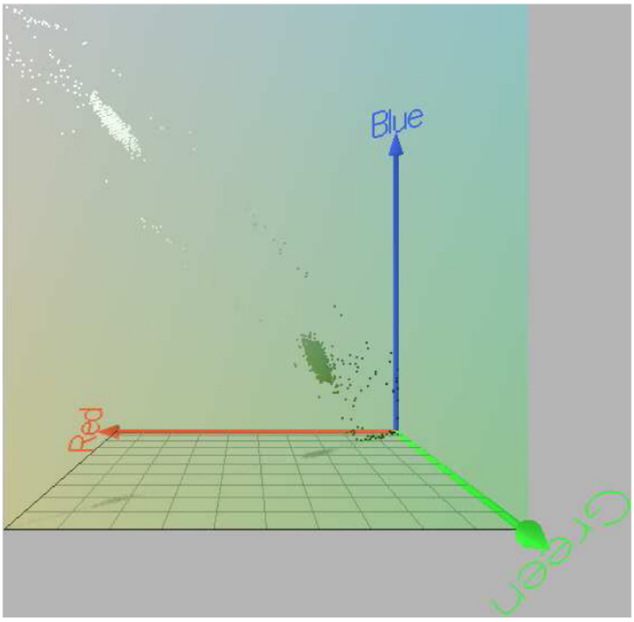

Figure 7: Colour space detail 40.

For an aperture correction setting of 40, the coefficients are:

$$k_2 = 0.6; K_4 = 0.2 \quad (10)$$

Vertical compensation filtering is similar, but only lines from the same field are used. In this case the coefficients for an aperture correction setting of 0 are:

$$k_1 = 0.2; k_2 = 0.05 \quad (11)$$

and for an aperture correction setting of 40 the coefficients are:

$$k_1 = 0.4; k_2 = 0.1 \quad (12)$$

4 Using the technique in a compositing application

Some computer vision applications like image interpretation need only the segmentation as the result of the keying. For compositing applications the image characteristics of the final composited image and the original images have to match.

The ideal procedure to apply the techniques discussed in this paper would be to compensate the input images for the aperture correction, compute the keys and substract the background colour component of mixed pixels. This results in an alpha channel and a pre-multiplied image. In the compositing step this input will be combined with an new background image, that should have the same image characteristics as the compensated image, ie. without high frequency image components. A high-pass boost filter can then be applied to the resulting composed image in order to give it the same characteristics as the aperture corrected image.

Alternatively the segmentation can be computed for the compensated image, but then used as alpha channel of the original (crisp) image. As explained in section 2 bright objects against darker background show an overshoot close to the edge of the object and other hand the darker background shows an undershoot. If the original image + alpha channel would be used to compute a composited image with a darker background, then the undershoot would not be present in the composed image.

An approximation of the undershoot can be computed by mixing a proportion of 'black' to the background image. To achieve this the background pixels in the segmented image are set to zero, ie. $I = (0, 0, 0)$

and an alpha value is computed that adds mixes a proportion of the black to the background making use of the compositing equation eq. 1.

The principle of the computation of the alpha value is illustrated in figure 8: The same filter as for the compensation of the colour image is applied to the segmentation. Values above 1 are clipped and negative values are inverted. The latter component is used to 'darken' the background image.

The method simulates the effect of the aperture correction on the background image quite effectively. In our current implementation it is restricted to bright foreground objects on a darker background. However, it would also be possible to handle dark foreground objects using the same principle.

5 Results

In order to get a quantitative evaluation of the proposed compensation of the aperture compensation, we tested two segmentation algorithms, fast-green and k-nearest-neighbour (NN) against the colour chart pictures. We then compared the segmentation results with a manual segmentation. Although this does not give an absolute reference, since it is still debatable where exactly the object boundaries in the image are, but at least the difference gives a relative measure.

The following tables give the number of misclassified pixels of the test chart for the two methods. Figure 9 shows the results (left) and misclassified pixels (right) for an aperture setting of 40. Figure 10 shows the results after compensating the aperture correction.

detail	N_{diff} fastgreen	N_{diff} corrected
-40	1575	-
0	1542	1265
40	1537	1195

Table 1: Number of misclassified pixel 'fast green'

detail	N_{diff} NN	N_{diff} corrected
-40	586	-
0	1303	743
40	2006	823

Table 2: Number of misclassified pixel 'k-nearest-neighbour'

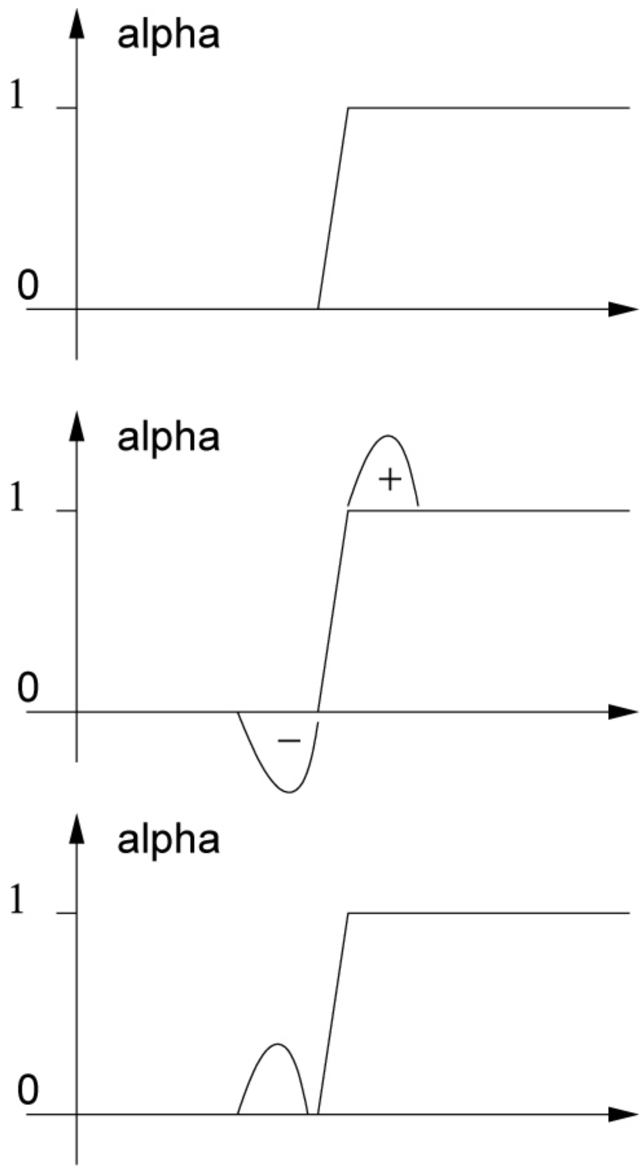

Figure 8: Computation of a correcting high-boost component in the alpha channel. Top: original alpha, middle: highboost filtered, bottom inverted negative component.

We further applied the aperture correction compensation to the sports scenes, as depicted in figure 1. Figure 11 shows some detail of the resulting segmentation using the k-nearestneighbour classifier. On the left the alpha channel is depicted and on the right the original image is keyed and the background replaced by a constant colour. The exact camera setting for that scene is unknown, but is roughly estimated to be between 0 and 40.

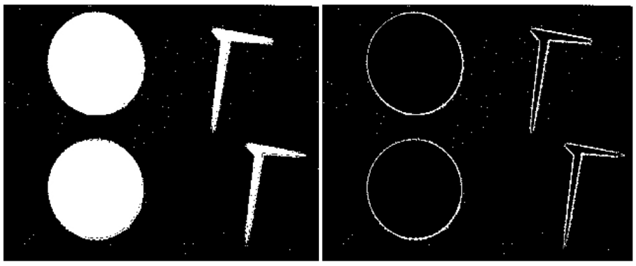

Figure 9: Segmentation detail '40' (left) and difference to reference (right).

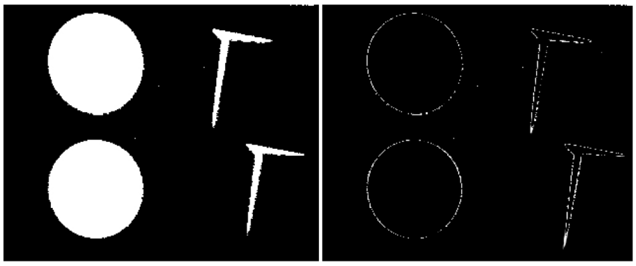

Figure 10: Segmentation detail '40' after correction (left) and difference to reference (right).

Figure 11: Detail of segmented broadcast picture.

Figure 12 shows the resulting segmentation after compensation of the aperture correction. The alpha channel and the keyed image show fewer artefacts. The image however looks 'softer' than the original (in figure 11 right). Compositing this foreground image with a similarly soft background, and then applying aperture correction appropriate to the application should deliver the best image quality. Applying aperture correction after compositing will allow for situations where the contrast between foreground and background of the composite image is very different from that of the original. Alternatively the approach outline in the previous section can be used to simulate the effect of aperture correction on the background image.

Figure 12: Detail of segmented broadcast picture after correction.

Figure 13: Detail of segmented original broadcast picture using difference keying.

The figures 13 + 14 show results of difference keying applied to the original broadcast image (shown in fig. 13) and after compensation of the aperture correction (fig. 1 14). In contrast to the global methods above the pitch lines are surpressed except in areas with shadows. As expected the segmentation in the compensated image is more precisely aligned to ob-

Figure 14: Detail of segmented broadcast picture after correction using difference keying.

ject edges. That is clearly visible in the right image of figure 14.

6 Conclusions and future work

The compensation of the aperture correction has shown benefits for colour-based keying of the chart images and the sport scenes.

As the results in tables 1 and 2 show the number of misclassified pixels is reduced by compensating the effect of the aperture correction for the two investigated segmentation algorithms. The performance of the 'fast-green' method is overall quite poor in the chosen example. The reason is that the used green tone in the test chart is not an ideal colour for this simple chroma-keying method.

The k-nearest-neighbour classifier shows reasonable overall classification results. However even after careful adjustments the number of misclassified pixels is never reaching zero for the test images. The reason for that is lens chromatic aberrations that are visible as green and blue colour fringes on each side of the objects. The affected areas are roughly 2-3 pixels wide for the two circles investigated here. Generally chromatic aberrations are a function of the lens and are more apparent in wider angle settings and in the corners of pictures. Furture work will be dedicated to compensate these effects.

In work so far the filter coefficients for aperture correction compensation have been determined empirically. This is a time-consuming and inaccurate process that could be automated to make the method easier adaptable to a wider field of application scenarios.

7 Acknowledgments

This work has been funded by the UK TSB as part of the iview project.

References

[CCSS01] Yung-Yu Chuang, Brian Curless, David H. Salesin, and Richard Szeliski, *A bayesian approach to digital matting*, In Proceedings of IEEE CVPR 2001, vol. 2, 2001, pp. 264–271, ISBN 0-7696-1272-0.

[FJR05] G. Friedland, K. Jantz, and R. Rojas, *Siox: Simple interactive object extraction in still images*, Proceedings of the IEEE International Symposium on Multimedia (ISM2005), 2005, pp. 253–259, ISBN 0-7695-2489-3.

[GTAHS07] O. Grau, G.A. Thomas, J. Kilner A. Hilton, and J. Starck, *A robust free-viewpoint video system for sport scenes*, In Proceeding of 3DTV conference 2007, 2007, pp. 1–4, ISBN 978-1-4244-0721-7.

[HHR05] Peter Hillman, John Hannah, and David Renshaw, *Foreground/background segmentation of motion picture images and image sequences*, IEE Transactions on Vision, Image and Signal Processing **152** (2005), no. 4, 387–397, ISSN 1350-245X.

[Mis92] Y. A. Mishima, *Software chromakeyer using polyhedric slice*, Proceedings of Nicograph 92, 1992, pp. 44–52.

[RKB04] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake, *"grabcut": interactive foreground extraction using iterated graph cuts*, ACM Trans. Graph. **23** (2004), no. 3, 309–314, ISSN 0730-0301.

[SB96] Alvy Ray Smith and James F. Blinn, *Blue screen matting*, SIGGRAPH 96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (New York, NY, USA), ACM, 1996, pp. 259–268, ISBN 0-89791-746-4.

Citation
Oliver Grau and Jim Easterbrook, *Effects of camera aperture correction on keying and compositing of broadcast video*, Journal of Virtual Reality and Broadcasting, 7(2010), no. 9, October 2010, urn:nbn:de:0009-6-26693, ISSN 1860-2037.