Preliminary subjective focus assessment results

R T Shilston and F W M Stentiford

University College London, Adastral Park Campus

Abstract: For both cameras and humans to see the world, they need to focus, yet they do not use the same focus strategies. This paper introduces these various strategies, in addition to several mathematical focus measures. An experiment to measure the human focus measure is proposed, and initial results are presented. These early results indicate that an attention based focus measure, using an RGB colour space matches the human focus opinions significantly more strongly than other focus measures.

1 Introduction

Whilst it is entirely possible to calculate the correct lens position for focussing at a given distance, real world scenes are not at a uniform distance from the camera. As such, when looking at a scene, it is necessary to determine the focus position such that the portion of the scene that is of interest to the observer is in focus. This paper explores how focus measures can be compared against the human experience in a real-world image.

1.1 Cameras

The first autofocus camera was released by Minolta in 1985. Since then, technology has advanced considerably, processing power has become increasingly mobile, and camera manufacturers have capitalised on these developments [1]. Whilst the early autofocus systems relied upon maximising the contrast between two horizontally adjacent pixels located at the centre of the image, many improvements have been made. These include multi-point focus, where the camera assesses the focus at a number of positions, and then focusses the camera at the position where the most of the image is in focus. Such systems can allow the photographer to override the selected position, and request the camera focus at an arbitrary point in the frame [2].

Recently, both Canon and Nikon released cameras which look for faces within the scene, and focus and meter upon them [3, 4]. Whilst these strategies do work, there are many caveats and circumstances when they fail – such as if faces "appear small, large, dark or bright relative to the overall composition, [or] if the subjects are looking sideways, lying down, or their faces are partially obscured" [3, p45].

1.2 Humans

The human vision system has been shown to use a number of different sources of information to control the ciliary muscles in the eye. Firstly, the angle of convergence of the two eyes can be used to determine the distance to the subject, and hence know the appropriate signal to send to the ciliary muscles. Other common sources of information include the 3D arrangement of the scene and object size. It has also been shown that some subjects can use chromatic aberration as a cue to defocus [5]. In addition to these methods, the eye also works to minimise blur.

1.3 Focus measures

Many mathematical focus measures have been proposed [6, 7], including average FFT, summing a thresholded gradient measure (Tenengrad), double differentiating the image (high pass filtering) or summing the two-dimensional contrast, many of which have been used successfully in microscopy applications. The focus measures used by camera manufacturers are not available in the public domain.

2 This work

With the exception of Crane's work in the 1960s [7], no-one has explored the mathematical definition of the human concept of blur. The authors have previously shown [8] that visual attention models such as those of Stentiford [9] and Itti [10] can be used as a mathematical focus measure. Whilst the various focus measures can be compared against each other, this does not provide any information as which most accurately matches the human perception. The following experiment and results show one way in which human focus can be compared to mathematical measures.

In our initial experiments to investigate the human opinion of focus, one real-world scene was used. It was captured using an Olympus SP500UZ camera, controlled by Pine Tree Computing's Camera Controller, which took 76 images of the scene, each focused at slightly different distances.

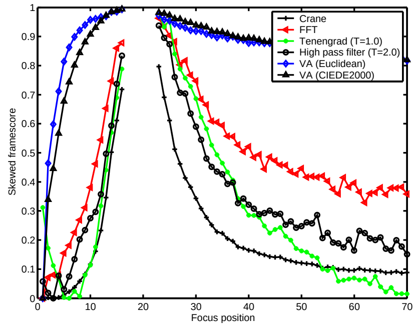
(a) Position 10 (b) Position 20 (c) Position 30

Figure 1: Source images were of a still-life image captured using an Olympus SP500UZ digital camera at 76 different focus distances.

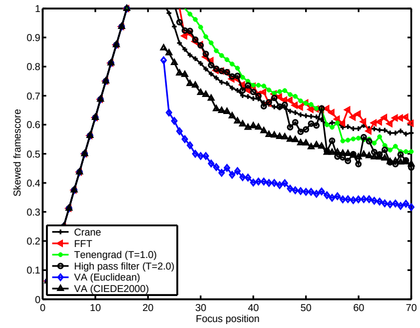
A computer was configured to use a Griffin Powermate USB dial to scroll through the images of each scene. The images were sorted by focus distance, and were pre-loaded into RAM so that there was no lag between turning the dial and the subject being presented with a new images on screen. The images were displayed full screen, and in such a way that changing between images caused no perceptible flickering of the screen.

The subjects were shown each scene five times in a random order, and asked to select the image they believed to be the most in focus. Once the optimum focus position was established, the computer split the screen to show two images horizontally adjacent. The left hand image was fixed to display a target photo, and the subject was requested to find the image that was equivalently defocused by using the dial. The target was selected randomly from the set of photos taken with a focus distance *shorter* than that of the image the subject determined to be most in-focus during stage 1 of the experiment. The candidate images were the set of photos with a focus distance *greater* than the in-focus image. The computer was configured to supply random target images for five minutes.

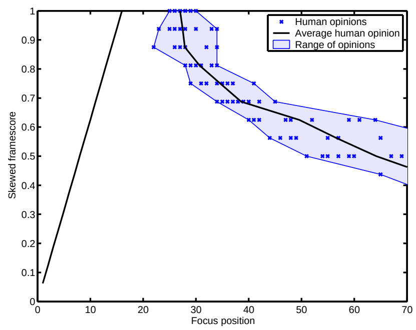
All subjects had normal, or corrected-to-normal sight, and did not suffer from colour blindness. Seven experimental runs were conducted, involving five different subjects. In total, over 28,500 images were examined by the subjects in the process of finding equivalent images for 361 target images.



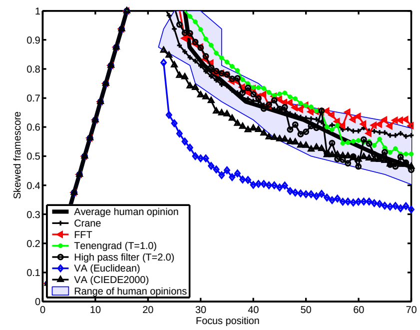
(a) Original focus measures



(b) Normalised focus measures



(c) Human focus opinions



(d) Human focus opinions with mathematical measures

Figure 2: Focus score curves for a variety of measures

3 Results

Each data point in the experiment represents a focus equivalence – that a given image was perceived to have the same amount of defocus as the target image. However, it does not provide any indication as to the relative focus quality between target images. As such, to analyse the data, all focus measures were normalised such that the focus score varied linearly from 0 to 1 for images focussed closer to the observer than the subject, thereby allowing the measures' response for far distances to be compared. Figures 2(a) and 2(b) shows the original and normalised curves respectively.

Each experimental result, which is of the form of a pair of images (the target, and the subject-selected equivalent image) was plotted by determining the y-value for the target, and using that y-value to plot the score for image the subject selected. This means that the variability in results is represented by the range of focus positions that equate to a given score – effectively the error bars are horizontal. Figure 2(c) shows the experimental results, whilst Figure 2(d) overlays the mathematical focus measures for comparison.

The data series were then compared to determine which mathematical focus measure most closely matched the human response. Figure 3 tabulates these results, and shows that the Visual Attention using a conventional RGB colourspace is significantly closer to the experimental results than other measures.

4 Conclusions

The experimental technique allowed human focus opinions to be compared in a quantitative fashion, without requiring an absolute focus measure to be determined. The results broadly match the existing mathematical focus measures, and have been shown to most closely correlate with a measure based upon Stentiford's Visual Attention algorithm when operating in the RGB

Measure	r^2	*Correlation*
Crane	94.4	97.2
FFT	92.2	96.0
Tenengrad	94.9	97.4
High pass filter	90.3	95.0
VA (Euclidean)	96.1	98.0
VA (CIEDE2000)	95.8	97.9

Figure 3: Percentage correlation between focus measures and the average human focus opinion

colourspace. Further experiments are necessary to explore these results with other source images, and a wider range of subjects, and will involve eye-tracking data so that the precise regions contributing to the focus determination can be investigated.

Acknowledgements

This research has been conducted within the framework of the European Commission funded Network of Excellence "Multimedia Understanding through Semantics, Computation and Learning" (MUSCLE) [11].

References

[1] P. Greenspun, "History of photography timeline," `http://photo.net/history/timeline`, Jan 2007, [Online; accessed 02-May-2007].

[2] Canon, "Canon launches first powershot a-series camera with optical image stabilization," August 2006.

[3] Canon, "Digital ixus i7 advanced camera user guide," 2006.

[4] Nikon Corporation, "Nikon face-priority AF," `http://www.nikonimaging.com/global/news/2005/pdf/FacepriorityAFnr.pdf`, Feb 2005.

[5] P. B. Kruger and J. Pola, "Stimuli for accommodation: blur, chromatic aberration and size," *Vision Research*, vol. 26, pp. 957–971, 1986.

[6] E. Krotkov, "Focusing," *Intl. Journal of Computer Vision*, vol. 1, no. 3, pp. 223–237, October 1987, Reprinted in Physics-Based Vision, Jones and Bartlett, 1992.

[7] H. D. Crane, "A theoretical analysis of the visual accommodation system in humans," Tech. Rep. NASA CR-606, NASA - Ames Research Centre, Sep 1966.

[8] R. Shilston and F. W. M. Stentiford, "An attention based focus control system," in *Proc. IEEE International Conference on Image Processing, Atlanta 2006*, October 2006, pp. 425–428.

[9] F. W. M. Stentiford, "An estimator for visual attention through competitive novelty with application to image compression," in *Proc. Picture Coding Symposium, Seoul*, April 2001, pp. 101–104.

[10] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[11] EC 6th Framework Programme, "Multimedia understanding through semantics, computation and learning, network of excellence," `http://www.muscle-noe.org/`.