

AN ATTENTION BASED FOCUS CONTROL SYSTEM

Robert Shilston and Fred Stentiford

University College London, Adastral Park Campus; r.shilston, f.stentiford@ee.ucl.ac.uk

ABSTRACT

Whilst the physical mechanism used by the eye to accommodate has been known since the mid 19th century, there have been few suggestions as to the method by which the human visual system controls the ciliary muscles. This paper proposes and explores an even function which, by building upon existing models of visual attention, could be used to model the eye's focusing mechanism. Specifically, by performing a two dimensional summation of the saliency score for each pixel in the image, it is possible to give the image an overall score. Focus can then be changed to maximise this score. It is anticipated that such an approach could be used to perform other parameter optimisation in image acquisition applications.

Index Terms— Focusing, machine vision

1. INTRODUCTION

Even in the mid 19th century, it was still unclear as to how the eye actually performed accommodation. It was only the suggestion that the lens was elastic and held under tension by various muscles by Helmholtz in 1856 that really allowed this to become the accepted theory of accommodation [1]. Figure 1 lists key individuals and the theories of accommodation they proposed over the years:

Proponent	Focusing mechanism
Home (1795)	Cornea curvature changes
Magendie (1816)	Eye has universal focus
Burow (1841)	Lens moves back and forth
Donders (1846)	Pupil changes in size
Listing (1853)	Eye ball elongates
Helmholtz (1856)	Lens changes shape

Fig. 1. Proposed accommodation methods (adapted from [1])

2. ACCOMMODATION CONTROL

In 1951 the first real evidence of the mechanisms used by the eye to control accommodation was found [2]. The eye appears to use several cues for driving the ciliary muscles, the simplest of which is to use the information available from the convergence mechanisms in binocular vision.

An alternative approach, proposed by Crane in 1966, involves maximising the “two-dimensional summation of the magnitude of the two-dimensional spatial derivative” [3]. In 1970, Toates proposed a model of accommodation in the form of a classic negative feedback proportional control system [4], which relies upon an odd function to assess blur but did not propose such a function.

A last main cue used by the ocular system is chromatic aberration [5], where the lens refracts different wavelengths by different amounts, causing coloured fringes to be present at object edges. Recent work has shown that this can be used to provide an odd function signal for ciliary control as Toates' model requires [6].

However, these are by no means the only methods that the visual system can use. For example, a change in apparent distance (that is, by varying angular size of an object) can also cause trigger a change in accommodation [7]. Other work has shown that subjects can even use audio biofeedback to exert voluntary control of accommodation, and that this can reduce myopia [8].

Yet, whilst it is clear that a multitude of mechanisms can be used by the eye, they do not explain all possible circumstances. Convergence-related information can be removed by simply closing one eye and chromatic aberration can be eliminated by using narrow line-width light sources etc.

3. AUTOFOCUS STRATEGY

The problem of deciding how to focus is obviously not limited to biological systems. Cameras have to solve exactly the same problem, and they too use a number of the methods listed in Figure 1, by relying on a fixed focus, having a small aperture, or by moving lenses. However, the signal to control the lens movement still needs to be generated.

There are two classes of solutions to this problem. One is to use some form of active metering, for example with an infra-red range sensor. The second is passive - performing some operation to determine the clarity of the image given the current lens position, and trying to optimise it. Typically, this relies on assessing the contrast at one or more points in the image, maximising the intensity difference between adjacent pixels (eg [9]).

Unfortunately, such contrast seeking can easily be misled. For example, if the subject of the photo is off-centre, or does

not occupy a large portion of the image, then the camera operator typically must indicate to the camera what they wish to be in focus. Recent developments from Nikon have helped to make life easier for camera users. Their face-priority focus system looks for face-like objects in the frame, and then maximises the contrast in those particular areas. However, even this is not fool-proof. Nikon explain “the camera may not recognize a face ... when a person is wearing sunglasses” [10].

This paper proposes a new approach to focusing in image forming devices which is not dependent upon specific features that are assumed to be present in the image such as sharp edges, contrast, or particular objects. Instead optimisation is carried out on the basis of “informativeness” or the effective saliency of the entire image.

4. SALIENCY BASED APPROACH

It is proposed that, rather than using such non-semantic measures such as contrast or chromatic aberration, instead we ensure that the image has maximum saliency. In this method, the saliency maps of a number of photos of the same scene, each taken with a different focal length, are scored. The photo with the maximum score should correspond to the frame with the greatest saliency. This approach extracts global information from saliency maps which are more typically used to provide local information. The source of saliency information is not critical to this approach; two prominent models of visual attention [11][12] have been used with equal success. Details of the saliency measures can be found in [11] and [12].

The proposed scoring method makes a two-dimensional summation of the saliency of each pixel in the image, ie: $score = \sum_x \sum_y Saliency(x,y)$, then determine the focal distance that corresponds to the maximum score.

Indeed, such a hunting approach has been observed in human vision. [3] suggests that “lens vibrations appear to be ... normal accommodation correction cycles”, and terms them “accommodation saccades”. It would also be anticipated that, in the absence of other cues, the eye would make a considerable number of incorrect initial responses, and a 50% error rate has been observed [13].

5. EXPERIMENTS

To assess the acuity of the proposed scoring system, a series of photos were taken of the same object, but with varying focal length. Carefully controlled still life scenes with subject material present at different depths were used in these experiments, and influences that might confound the results, such as movement during capture sequences, were excluded. Each photo was then scored using three measures:

a) 2D summation of saliency (Stentiford’s algorithm).

b) 2D summation of saliency (Itti’s algorithm). This was done with [14] using a command line of `ezvision -X input.ppm`.

c) 2D summation of vertical contrast plus 2D summation horizontal contrast. Contrast was computed using Matlab’s `diff` function (similar to [9]).

For each measure, the values were normalised across all candidate images, such that the score was in the range $0 \leq x \leq 1$.

5.1. Tigger

A series of photos were taken of a scene in which a furry toy was placed 2m from the camera, against a plain background. These were focused at distances of 50cm, 70cm, 1m, 2m, 3m and 5m. The results show that all three measures produced a clear peak (Fig. 2), corresponding to the image taken with the appropriate focal distance (Fig. 3).

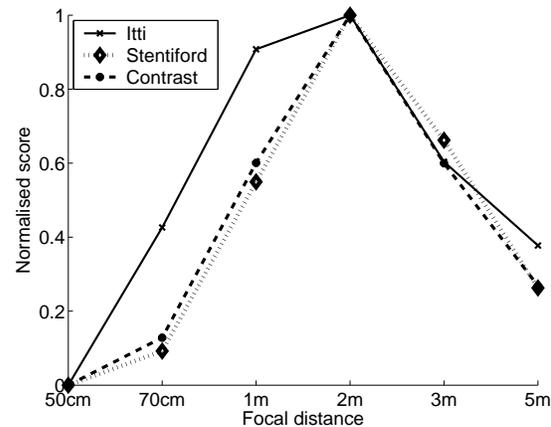


Fig. 2. Score vs focal distance for Tigger



Fig. 3. Tigger at 2m

5.2. Cactus

A more complex scene, with two clear focal distances was then chosen. A small cactus was placed 21cm in front of a rush mat, and a series of photos taken with varying focus. Each image was then scored using both Stentiford's Visual Attention algorithm, and contrast. Figure 5 shows these results, with 0cm focal distance corresponding to an image correctly focused on the cactus.

The image corresponding to the peak in score obtained from the Stentiford algorithm is shown in Figure 6, and the image corresponding to the other peak (the image with maximal contrast) is shown in Figure 7. The contrast algorithm concentrates on the background, whereas the saliency algorithm focuses on the subject.

Future experiments are planned that will vary the principal subject size relative to the background and may reveal performance differences between the saliency measures.

6. CONCLUSIONS

It has been shown that optimal focus information can be obtained from saliency-based measures, with no reliance upon particular features being present in the image. In some circumstances, a saliency based measure can outperform traditional passive focus methods that typically rely on contrast. It is likely that a saliency based approach could be used to optimise image parameters other than focus, such as illumination brightness and spectral content. Further experiments are necessary on a range of different scenes to test the robustness to artifacts arising from movement and to assess the accuracy of the approach.



(a) 1m



(b) 2m



(c) 3m

Fig. 4. Detail of Tigger

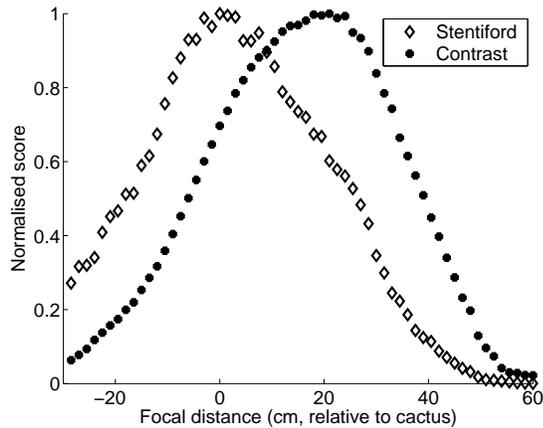


Fig. 5. Score vs focal distance for cactus



Fig. 6. Focus determined by saliency algorithm



Fig. 7. Focus determined by contrast algorithm

7. REFERENCES

- [1] E. G. Boring, *Sensation and perception in the history of experimental psychology*, New York : Appleton-Century-Crofts, 1942.
- [2] E. F. Fincham, "The accommodation reflex and its stimulus," *British Journal of Ophthalmology*, vol. 35, pp. 381–393, 1951.
- [3] H. D. Crane, "A theoretical analysis of the visual accommodation system in humans," Tech. Rep. NASA CR-606, NASA - Ames Research Centre, Sep 1966.
- [4] F. M. Toates, "A model of accommodation," *Vision Research*, vol. 10, pp. 1069–1076, 1970.
- [5] P. B. Kruger and J. Pola, "Stimuli for accommodation: blur, chromatic aberration and size," *Vision Research*, vol. 26, pp. 957–971, 1986.
- [6] P. B. Kruger, S. Mathews, M. Katz, K. R. Addarwala, and S. Nowbotsing, "Accommodation without feedback suggests directional signals specify ocular focus," *Vision Research*, vol. 37, no. 18, pp. 2511–2526, 1997.
- [7] W. H. Ittelson and A. Ames Jr, "Accommodation, convergence, and their relation to apparent distance," *The Journal of Psychology*, vol. 30, pp. 43–62, 1950.
- [8] J. N. Trachtman, V. Giambalvo, and J. Feldman, "Biofeedback of accommodation to reduce functional myopia.," *Biofeedback and self-regulation*, vol. 6, no. 4, pp. 547–562, 1981.
- [9] A. Misawa, "Focusing apparatus for adjusting focus of an optical instrument," *US Patent 6,864,474*, Feb 2002.
- [10] Nikon Corporation, "Nikon face-priority AF," http://nikonimaging.com/global/news/2005/0216_06.htm, Feb 2005.
- [11] F. W. M. Stentiford, "An estimator for visual attention through competitive novelty with application to image compression," *Picture Coding Symposium, Seoul*, Apr 2001.
- [12] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, no. 10-12, pp. 1489–1506, May 2000.
- [13] A. Troelstra, B. L. Zuber, D. Miller, and L. W. Stark, "Accommodative tracking: A trial-and-error function," *Vision Research*, vol. 4, pp. 585–594, 1964.
- [14] University of Southern California iLab, "iLab Neuromorphic Vision C++ Toolkit," <http://ilab.usc.edu/toolkit/>, Jan 2006.