PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Introduction to New geometric techniques in computer vision, a Discussion Meeting held at the Royal Society of London

The Royal Society

Phil. Trans. R. Soc. Lond. A 1998 **356**, 1067-1069
doi: 10.1098/rsta.1998.0210

The Royal Society

Introduction

By H. Christopher Longuet-Higgins

Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, UK

According to any reasonable metric, computer vision is undergoing an explosive expansion. We should have found it quite impossible to cover the whole subject in a mere two days. But the need to be selective has enabled us to concentrate on a few growth areas without in any way implying that others are less important. It is not surprising that our final choice of themes was the result of a process of evolution rather than strictly logical planning. First, I shall present an overview of the meeting as a whole.

The papers in Session I deal with the visual appearance of curved surfaces. Professor Koenderink is primarily concerned with the internal representation of grey-level images, and his arguments lead one to suspect that even robots might be subject to optical illusions. Dr Giblin's paper develops a systematic account of apparent contours, calculated to appeal to a geometrically minded hill walker making his way across the South Downs. Dr Cipolla, building on this work, describes techniques for reconstructing a smooth surface from the way in which its outline is affected by a known change of viewpoint. His paper is recommended reading for the engineers in charge of the Mars rover, and forms a natural bridge to Session II.

Session II, on structure and motion, is concerned with the three-dimensional interpretation of sets or sequences of two-dimensional images. This has been a recurring theme in computer vision for many years, but the subject is still a rich mine of research problems.

Perhaps a personal reminiscence is in order, dating from the early 1980s. An obvious condition to be satisfied by two separate two-dimensional images of the same three-dimensional scene is that the lines of sight from the two camera centres to any feature point must intersect in space, at the point in question, needless to say. This leads to the so-called *epipolar constraint*, which is discussed further in this volume. It can be expressed as a bilinear relation of the form,

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}' = 0,$$

where $\boldsymbol{x}$ and $\boldsymbol{x}'$ are the homogeneous coordinates of the two images of the given feature, and $\boldsymbol{E}$ is a 3×3 matrix determined by the relative orientation of the camera positions. So, given two images, with as many as eight feature points visible in both, one should be able to obtain the entire geometry by solving eight homogeneous linear equations for the elements of $\boldsymbol{E}$, and using these to derive the relative orientation. This expectation was fulfilled, so I discovered, if the feature points were chosen at random; but when I tried the *8-point algorithm* on the most rugged configuration one could imagine, the eight vertices of a cube, it gave manifestly absurd results. God the Geometer obviously had a wry sense of humour, or was the ghost of my old maths teacher trying to teach me a lesson? Readers will not, of course, have forgotten that the eight equations for the elements of $\boldsymbol{E}$ will have a unique solution only if they are linearly independent; but it took me years to discover the precise

Phil. Trans. R. Soc. Lond. A (1998) **356**, 1067–1069

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

geometrical equivalent of this condition. The 8-point equations fail, it transpires, if and only if the eight feature points *and the two camera viewpoints* all lie on the same quadric surface. This condition is automatically satisfied if (but not only if) the eight feature points constitute a *quadric octet* as is the case with the vertices of a cube, and with almost any other neat arrangement of eight points you might care to think of (Longuet-Higgins 1981, 1984, superseded by 1987). Moral: in order to verify a mathematical result one may need the resources not only of arithmetic and algebra but also of geometry.

Such are the barriers to communication between disciplines that it was not until the late 1980s that the concepts of projective geometry, or should one say algebraic geometry, began to be applied to real-life problems in computer vision. Steve Maybank, one of the pioneers in its application (Maybank 1989), recommended Semple & Kneebone's *Algebraic projective geometry* for enlightenment on such recondite matters as the circular points at infinity, the absolute conic and the twisted cubic (Semple & Kneebone 1979). As an amateur who has gone through fire and water in pursuit of these concepts I can heartily recommend the book to those who still have to complete the journey. Semple & Kneebone make it plain why the use of homogeneous coordinates, essential to the subject, positively demands the presence of such remarkable constructs. There is often a problem of communication between those familiar with complex entities and those who find them perplexing. The contributors to this volume have all borne this in mind as they introduce us to their more sophisticated mathematical ideas. The summary below may serve to inoculate non-geometers against some of the more baffling terminology.

Two dimensions: $(x/z, y/z)$

Point $\quad x : y : z$

Line $\quad u : v : w$

Incidence $\quad \Longrightarrow\ ux + vy + wz = 0$

Conic
$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

i.e. $\quad ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy = 0$

Circle $\quad x^2 + y^2 + 2fyz + 2gzx + cz^2 = 0$

Line at infinity $\quad z = 0$

A circle and the line at infinity intersect in the point pair,
$$x^2 + y^2 = 0 = z, \qquad \text{i.e.} \quad (1, \pm\mathrm{i}, 0).$$
These points are denoted by (I, J) and known as the circular points at infinity.

Three dimensions $(x_1/x_0, x_2/x_0, x_3/x_0)$

Point $\quad x_0 : x_1 : x_2 : x_3$

Sphere $\quad {x_1}^2 + {x_2}^2 + {x_3}^2 + (\text{terms in } x_0) = 0$

A sphere intersects the plane at infinity $(x_0 = 0)$ in the *absolute conic*:
$${x_1}^2 + {x_2}^2 + {x_3}^2 = 0 = x_0.$$

The reconciliation of different views is the main theme of Session II. Professor Faugeras elegantly expounds the levels of increasing generality at which one may wish to describe the features of a complex scene, using the Euclidean group (the *group of similitudes*), the affine group and the full projective group (Faugeras 1995). Professor Kanade reminds us of the powerful technique of singular-value decomposition and its application to the interpretation of sequences of para-perspective images, a problem for which it might have been 'made to measure' (Tomasi & Kanade 1991). Dr Hartley, inventor of the *trifocal tensor*, which enables one to connect three or more point-and-line images (Hartley 1995), discusses how to minimize the errors associated with the computation of this and related entities. Dr Zisserman's paper brings us up to date on the role of the plane at infinity and the absolute conic in the on-line calibration of a moving camera.

Sessions III and IV are more topic-oriented. In Session III we are reminded that computer vision has its uses in object recognition, for which it would be valuable to have a unified formal approach. There seems to be general agreement that the theory of invariants is an essential ingredient, and that we need more efficient and reliable methods for grouping features together in the manner of Gestalt theory. In Session IV Professor Blake and Professor Kanatani treat computer vision against a noisy background as a problem in statistical estimation, and Dr Torr discusses the interlocking problems of motion segmentation, model selection and parameter estimation. One could hardly disagree that most visual input serves the purpose of updating one's model of the environment, rather than triggering the laborious process of constructing a world model from scratch.

I trust that I shall be forgiven, especially by our distinguished contributors, if this racy overview does much less than justice to the importance of the ideas they describe. And if the programme is less than perfectly balanced, perhaps that is at least partly a result of the untidy and spasmodic growth of the subject itself.

In conclusion I thank, on behalf of the organizing committee, the Royal Society and its staff for their invaluable help with the arrangements for the meeting. It has been an honour and a pleasure to work with them.

References

Faugeras, O. 1995 Stratification of three-dimensional vision: projective, affine and metric representations. *J. Opt. Soc. Am.* **12**, 465–484.

Hartley, R. 1995 A linear method for reconstruction from lines and points. In *Proc. 5th Int. Conf. on Computer Vision*, pp. 882–887.

Longuet-Higgins, H. C. 1984 Configurations that defeat the eight point algorithm. In *Image understanding* (ed. S. Ullman & W. Richards). New Jersey: Ablex.

Longuet-Higgins, H. C. 1987 *Mental processes*, pp. 323–328. Cambridge, MA: MIT Press. (Originally appeared in *Proc. 1st Conf. on Artificial Intelligence Applications* (1984) Silver Spring, MD: IEEE Computer Society Press.)

Maybank, S. 1989 The projective geometry of ambiguous surfaces. Report T(S)T/2012, Project Esprit P 940.

Semple, J. G. & Kneebone, G. T. 1979 *Algebraic projective geometry*. Oxford: Clarendon Press. (Originally appeared 1952.)

Tomasi, C. & Kanade, T. 1991 Shape and motion from image streams: a factorization method. 2. Point features in 3D motion. School of Computer Science, Carnegie-Mellon University, Pittsburg, USA.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES