Low-level vision – a tutorial

Professor Roy Davies

Overview and further reading

Abstract

This tutorial aims to help those with some experience of vision to obtain a more in-depth understanding of the problems of low-level vision. As it is not possible to cover everything in the space of 90 minutes, a carefully chosen series of topics is presented. This section of the notes is a commentary to accompany the slides in the presentation.

1. The nature of vision

Vision is a complex process and has long been modelled, rather loosely, as a cascade of low, intermediate and high-level stages following image acquisition. However, this model is somewhat impoverished in that it ignores the possibility of downward as well as upward flow of information which can help with interpretation. In addition, there is need for considerable complexity and sophistication – so much so that the problems of low-level vision are often felt to be unimportant and are forgotten. Yet it remains the case that information that is lost at low level is never regained, while distortions that are introduced at low level can cause undue trouble at higher levels [1]. Furthermore, image acquisition is equally important. Thus, simple measures to arrange suitable lighting can help to make the input images easier and less ambiguous to interpret, and can result in much greater reliability and accuracy in applications such as automated inspection [1]. Nevertheless, in applications such as surveillance, algorithms should be made as robust as possible so that the vagaries of ambient illumination are rendered relatively unimportant.

2. Low-level vision

This tutorial is concerned with low-level vision, and aims to indicate how some of its problems and limitations can be solved. This is a huge subject covering image acquisition, noise suppression, segmentation, colour and texture analysis, shape analysis, object recognition, and many other topics. Hence it is not possible to do justice to it in the space of one and a half hours. It is therefore assumed that participants have a reasonable concept of the subject area, and the tutorial therefore focusses on a number of topics that have been chosen because they involve important issues. The topics fall into five broad categories:

- Feature detection and sensitivity
- Image filters and morphology
- Robustness of object location
- Validity and accuracy in shape analysis
- Scale and affine invariance.

In what follows, references are given for the topics falling under each of these categories, and for other topics that could not be covered in depth in the tutorial.

3. Feature detection and sensitivity

General [1].
Edge detection [2, 3, 4].
Line segment detection [5, 6, 7].
Corner and interest point detection [1, 8, 9].
General feature mask design [10, 11].
The value of thresholding [1, 12, 13].

4. Image filters and morphology

General [1].
Effect of applying rank-order filters [1, 14].
Mathematical morphology [1, 15].

5. Robustness of object location

General [1, 16].
Centroidal profiles v. Hough transforms [1, 17–19].
Fast object location by selective scanning [1, 20].
Robust statistics [1, 21, 22].
The RANSAC approach [23].

© E.R. Davies, 2011
6. Validity and accuracy in shape analysis

Shape analysis [1, 24].
Object labelling [1].
Distance transforms and skeletons [1, 25, 26].
Boundary distance measures [27, 28].
Symmetry detection [29, 30].

7. Scale and affine invariance

At this point it will be useful to consider a topic that has been developing for just over a decade, starting with papers [31, 32]. It arose largely because of difficulties in wide baseline stereo work, and with tracking object features over many video frames – because features change their character over time and correspondences are easily lost. To overcome this problem it was necessary first to eliminate the relatively simple problem of features changing in size, thereby necessitating ‘scale invariance’ (it being implicit that translation and rotation invariance have already been dealt with). Later, improvements became necessary to cope with ‘affine invariance’ (which also covers shear and skew invariance). Lindeberg’s pioneering theory [31] was soon followed by Lowe’s work [32, 33] on the ‘Scale Invariant Feature Transform’ (SIFT). This was followed by affine invariant methods [34–37]. In parallel with these developments, work was proceeding on maximally stable extremal regions [38] and other extremal methods [39, 40] (the latter concept has also been followed in a different context [13]).

Much of this work employs the interest point methods of Harris and Stephens [9], and is underpinned by careful in-depth experimental investigations and comparisons [37, 41, 42]. Finally, there are signs that the tide may be turning in other directions, specifically the design of feature detectors that concentrate on especially robust forms of scale invariance – as in the case of the ‘Speeded Up Robust Features’ (SURF) approach [43, 44]. See also the fuller review article [45].

8. Concluding remarks

The topics presented in this tutorial necessarily give an incomplete picture because of the short span of time available. Nevertheless, they provide interesting lessons, and demonstrate important factors – specifically, the need for sensitivity, accuracy, robustness, reliability and validity. Further factors enter into the equation when practical systems are being designed – speed and cost of real-time hardware being amongst the most relevant.

Overall, it is clear that low-level vision is an essential ingredient of the vision hierarchy and that its problems are fundamental to the whole of vision. Although the focus of attention in the subject may have moved on to more modern topics, it is definitely not the case that this side of the subject is played out and that everything is now known about it. While data sets change and specifications for low-level algorithms remain incomplete, workers will have to go on developing algorithms to cope with the specific idiosyncrasies of their data to make the most of it in all the ways outlined above.

9. Further reading

Participants may wish to further explore the material covered in this tutorial. The author’s book [1] covers many of the topics in fair depth; particular attention is drawn to the following chapters:

- Chapter 3, for median and mode filters, including edge shifts;
- Chapter 5, for edge detection;
- Chapters 6 and 7, for shape analysis – especially thinning and centroidal profiles;
- Chapter 10, for Hough-based methods of circular object location;
- Appendix on robust statistics.

The book also contains a considerable amount of detailed information on 3D interpretation, invariants, texture analysis, automated inspection and practical issues.

Acknowledgements

The author would like to credit the following sources for permission to reproduce material from his earlier publications:

- Elsevier for permission to reprint text and figures from [1, 2, 20, 26];
- IET for permission to reprint text and figures
from [5, 7, 8, 11, 13, 14, 16, 28];
- IFS Publications Ltd. for permission to reprint figures from [18].

References

ACM 24, no. 6, 381–395