MVA '92 IAPR Workshop on Machine Vision Applications Dec. 7-9,1992, Tokyo

TOWARDS A VERSATILE FRAMEWORK FOR INTERMEDIATE-LEVEL COMPUTER VISION

Luciano da Fontoura Costa

Grupo de Instrumentação e Informática, IFQSC,
Univ. de S. Paulo, Caixa Postal 369, S. Carlos, SP, 13560 Brazil.

ABSTRACT

The paper presents and discusses the development and implementation of an effective and versatile framework for intermediate-level computer vision. Like the primary visual cortex, this framework is able to effectively provide the description of edge-detected images in terms of bars.

INTRODUCTION

The achievement of computer vision systems with ability similar to that of primates has proved to be an elusive goal. The principal obstacles which have been preventing such achievement includes: (a) relatively little is known about the primate vision system whose emulation have so often been tried and (b) the current hardware is not powerful enough to allow real-time execution of the involved techniques and effectively implement the necessary data structures [1]. In order to give an idea of the complexity of natural vision systems, it is reminded that about half of the primate cortex is related to vision, which implies a total of at least 2^{130} binary states, an amount larger than the total number of elementary particles (such as electrons and protons) in the known universe.

Although neurophysiological investigations, usually in cats and monkeys, have so far revealed great part of the low-level and intermediate structures and processes in advanced natural vision; only psychological hypotheses are currently available on the high-level visual processing.

Each human eye has about 130 million receptors while conventional digital images rarely have more than 512x512=262144 image elements and assuming that each image element can be accessed in a machine cycle, about 13ms are needed for a relatively powerful 20MHz processor just to read a 512x512 image. It is thus hardly surprising that so much effort has been spent on developing improved vision algorithms and parallel/dedicated hardware for their respective implementations.

This paper presents the development of an effective framework for intermediate vision which emulates in great part the function of the primate visual cortex, i.e. it produces a description of the edge-detected image in terms of bars. The following sections start by addressing the biological motivation of the approach and follow by characterizing bars in terms of digital straight line segment – DSLSs, presenting and discussing the adopted techniques, commenting on their implementation and presenting five applications of a tranputer-based implementation of the framework.

THE NEED FOR AN EFFECTIVE LINE DETECTOR

After the visual information transmitted by the optic nerves has left the eyes (whose retinas perform edge-element detection), re-arranged in the optic chiasm and passed through the lateral geniculate nuclei, it reaches the visual cortex to be further processed. In a series of neurophysiological investigations which led to a Nobel prize, Hubel and Wiesel figured out that most of the neural structures in the primary visual cortex can be classified according to their function into one of the following basic types [2]: (a) *Simple cell-* sensitive to a bar with a specific orientation placed in a specific position of the visual field; (b) *Complex cell-* sensitive to bars with a specific orientation and/or movement placed anywhere in the visual field and (c) *Hypercomplex cell* - similar to the complex cell but responding only to bars with specific lengths.

It is thus clear that most of the primary visual cortex is dedicated to bar detection, which explains why we can so precisely identify straight lines while being more reluctant about the nature of other curves. It should also be recalled that any curve in a quantized space can be exactly described in terms of DSLSs. So, one of the principal lessons nature has to teach us about intermediate vision is that the fast and effective detection of bars is specially relevant.

DIGITAL BARS AND STRAIGHT LINE SEGMENTS

Although the concept of straight line is trivial for continuous geometric spaces, the characterization of its digital equivalent has not proved to be straightforward and motivated a series of works [3]. Digital straight lines are usually defined as the result of the quantization of some ideal straight line into a discrete geometric space; the present work adopts the grid-intersect quantization scheme [4].

Digital bars can be defined and characterized in terms of DSLSs [5]. This can be done by defining a generic digital bar as the result of the grid-intersect quantization of a band with intercept width W, projected endpoints P_1 and P_2 and having as its upper boundary a straight line segment. Algorithms for the generation of such digital bars can be found in [4]. Figure 1 illustrates these conventions for a digital bar with $W = 2$, $P_1 = 3$, $P_2 = 7$ and having as upper boundary the straight line defined by $y = 0.5x + 2.5$

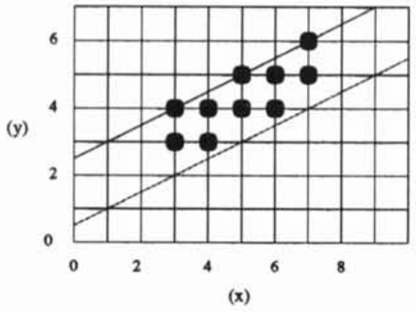

Figure 1: Example of a digital bar (see text for more detail).

AN EFFICIENT TECHNIQUE FOR DETECTION OF DIGITAL STRAIGHT LINE SEGMENTS

Koplowitz et al. [6] showed that the total number of digital straight lines in an NxN grid is of $O(N^4)$, which implies a substantial demand for techniques for digital straight line detection. Amongst the several techniques which have been attempted in order to solve this issue, the Hough transform [7] has proved to be one of the simplest and yet relatively most effective alternatives.

The Hough transform for DSLS detection consists of mapping every image feature element (usually its edge-elements) into a curve in a discrete parameter space, i.e. the accumulator(s) array(s), in such a way that peaks in the latter indicate *possible* instances of DSLSs in the image. Its advantages include: suitability for parallelization, can detect multiple digital straight lines at the same time, present some tolerance to noise, distortion or occlusion in the image and, most remarkably can simultaneously perform image segmentation and curve fitting. On the other hand, the Hough transform presents some shortcomings for digital straight line detection which have been mostly circumvented by the adoption of a connectedness analysis and a merging stage, which will be presented later in this section.

Before proceeding to discuss **edge-element detection**, it is important to distinguish between *edge-elements* and *edges* or *borders*: while the first arises from abrupt variations of some physical characteristic of the image, usually the brightness, the second is herein understood as related to actual boundaries of the objects in the image. It should be noticed that not every group of edge-elements corresponds to an actual border (e.g. shades can lead to false borders).

Despite all the efforts which have been spent on the problem of edge-element detection, there is no such a thing as an effective general edge-element detector; moreover, the versatility of edge-element detectors usually follows the complexity of the respective techniques. The frequent demand for real-time execution implies that the edge-element detection technique should be able to be executed fastly, simple 3x3 operators such as the Sobel and Laplacian have thus been adopted in the current approach.

The **binary Hough transform** – BHT – [5,8] is a variation of the standard Hough transform for digital straight line detection, proposed as a patent by P. V. C. Hough in 1962, which is based on a quadruple slope-intercept parameterization which consists of a single expression and a set of four coordinate transformations [5]. The BHT generates a DSLS in each of four accumulator arrays, for each feature point in the image. Provided that the image and the accumulator arrays have dimensions which are integer powers of two, the BHT presents interesting features such as optimal compaction of the accumulator arrays [5] and suitability for parallel execution in integer arithmetic without any rounding error, products or table look-ups – in fact it demands less additions than typical standard Hough transforms [5]. Such features favour its effective implementation in software and hardware. Moreover, experimental evaluations [5] have indicated that the BHT presents accuracy comparable to that of other Hough transforms. The effective technique for simplification of the accumutor arrays proposed by Gerig and Klein [9] may be necessary for dense images.

Since the Hough transform does not take into account the connectedness of the image feature elements, some post-Hough-transform **connectedness analysis** is necessary to confirm DSLSs and determine their endpoints. In the present approach, for each detected peak (assumed to be any cell in the accumulator arrays with count larger or equal to a given threshold), a unidimensional histogram is built by projecting the image feature elements which contributed to the peak onto the x- or y-axis accordingly to the line slope (for slopes between 1 and -1 project onto x and otherwise onto y); the connectedness of such histograms can be easily verified by an algorithm which looks for consecutive runs of feature elements [5]. Every group of connected feature points with projection larger than a given threshold P is confirmed as an actual DSLS; a maximum gap, G, can also be allowed to compensate for discontinuities and eventual distortions or noise in the image.

An interesting strategy to reduce the number of replicated lines and at the same time join the broken segments produced by the Hough transform and connectedness analysis processing consists of **merging** the detected DSLSs according to their similarity, which is expressed in terms of their proximity and parameters (i.e. slope and intercept) in the present approach.

The developed merging technique consists of finding for each detected DSLS D all the other segments which have their endpoints close to the endpoints of D and then find the set of three segments whose parameters have the greatest similarity; if these three segments are similar enough, they are merged into a single one. A special data structure has been employed in this technique in order to improve the handling of the DSLSs [5].

IMPLEMENTATION

This section briefly discusses the transputer implementation of the basic techniques in the DSLS detection framework; higher performance implementations in dedicated/parallel hardware can be found in [5] and [10].

The inherent simplicity of the adopted **edge-element detector** allows its straightforward implementation into transputers arranged in a linear or mesh topologies by assigning a uniform section of the image to each processing element. An one-pixel overlap should be allowed between columns and lines of the image segments for proper operation.

One of the simplest and most effective ways of parallelising the **BHT** execution amongst several transputers consists of assigning a range of the slope parameter [5] to each processing element in such a way that every processing element operates over the same image feature element– it has been verified [5] that, at least for a small number of transputers, an almost linear speed-up can be achieved through this strategy, which indicates that the part of the execution time required to transmit the image feature points coordinates to the processing elements via the serial links does not amount to a substantial overhead.

The **connectedness analysis** can also be effectively accomodated into linear or mesh architectures by keeping a copy of the image feature elements into the memory of each processing element, which can determine and analyse the histograms corresponding to each of the detected peaks resulting in the portion of the accumulator array in its associated memory. Such a strategy is supported by the facts that at least one megabyte of memory is usually provided for each transputer and that the x- and y-coordinates of the image feature elements have anyway to be broadcast to the processing elements during the BHT.

Although the adopted **merging** algorithm implies a relatively complex control and data structure which prevents its effective implementation in dedicated hardware, it usually operates over a relatively compacted input: the parameters and endpoints of the DSLSs. This technique also adapts well to a linear or mesh array of transputers where the image is uniformly distributed, the results of each segment being sent to a single processing element which will perform the final global merging.

APPLICATIONS

A transputer-based implementation of the framework presented in the previous sections has been applied to five actual problems as described in the following:

Image compression: A DSLS can be represented by its slope, intercept and the projected coordinates of its endpoints – it is interesting to observe that the a DSLS can not be precisely described by the full Cartesian coordinates of its endpoints [5]. For most segment lengths, this representation involves less storage than keeping a list of every edge-element belonging to the DSLS, which suggests that this principle can be used as basis of an image compression technique. A complete compression and decompression technique based on this principle and which operates over grey-level thresholded versions of the original 256x256 image [5,11] has allowed a compression rate of about 0.1, a compression time of about one minute and a decompression time of about 3 minutes.

Quality control of ultrasound transducers: Some types of ultrasound transducers produce beam which have a substantial linear region of interest. The described framework has been applied in order to infer quantitative values of parameters of the transducers such as the orientation and position of the propagation axis, the width and extent of the linear region of the beam as well as its symmetry, from which the quality of the transducers can be evaluated [12]. The adopted approach involved the description of the beam in terms of DSLSs from which the above mentioned performance parameters could be easily obtained.

Semi-automated analysis of clay: Several physical properties of a clay mass can be inferred by the analysis of micrographs of its samples. The described framework has been applied to determine the preferred orientation of clays having constituent particles of 'slippery pennies' type [13]. In this application the principal objective was to determine the preferred orientation of the particles in the micrograph, which was achieved by edge-detecting the micrographs and applying the above described framework to produce a description of the straight features in the micrograph in terms of its DSLSs from which the preferred orientation can be easily determined. The results so far obtained have proved to be more accurate than the alternative technique based on local operators which implement an approximation of the gradient.

Automation of the production of cork stoppers: The production of stoppers from the cork bark as typically done in Portugal and Spain involves cutting cork sheets into thin stripes according to the orientation which is perpendicular to the preferred orientation of the grooves in the cork sheet. The bar detection framework has been applied to produce the cutting angle in a similar way as that described in the previous sub-section [14].

Intermediate-level of a model-based polyhedra recognition system: The model-based recognition of objects from the description of the edge-detected image in terms of bars is one of the most natural applications of the framework presented in this paper. This application has been experimentally evaluated for the recognition of polyhedra (3-D objects) from their 2-D images. After being edge-detected, the straight features of the image are described in terms of the projected endpoints and slope/intercept parameters of their corresponding DSLSs. This information is processed by a model-based recognition system [15] in order to detect the type, position and perspective of the objects in the original image. This application has led to particularly encouraging results, which substantiates the suitability of this type of intermediate processing as well as the effectiveness of the employed matching algorithm. Typical execution times for the description of images containing about 200 edge-elements in terms of their segments have been of about 20 s.

CONCLUDING REMARKS AND FUTURE DEVELOPMENTS

The present paper has presented the development of a complete framework which is able to produce a description of the edge-detected image in terms of bars, thus emulating in great part the neuronal processing which occurs at the visual cortex. The achieved potential for real-time execution has been allowed by the development and adoption of improved techniques for detection of digital straight line segments and their implementation in concurrent/dedicated hardware, both of which have been discussed in some detail.

The versatility and potential for applications of the described framework was exemplified by a series of five actual cases in image processing (image compression), visual inspection (quality control of ultrasound transducers, clay analysis and control of cork stoppers production) and computer vision (intermediate-level of a polyhedra recognition system).

Intended future developments include the gradual development of a versatile and complete computer vision system having its intermediate level subsystem based on the dedicated concurrent implementation of the framework described in the present paper. Other intended developments include the formal and comparative evaluation of other techniques for digital straight line detection, the derivation of curve detection techniques which uses as input the description of curves in terms of DSLSs, the extension of the proposed techniques to an hexagonal grid as well as new applications such as the automation of wood cutting and automatic interpretation of tomographic images.

ACKNOWLEDGEMENTS

The author is grateful to Dr. Valentin O. Roda for reviewing this paper and for his invaluable suggestions. A substantial part of this work was done during the author's PhD program at the King's College, Univ. of London, under the supervision of Dr. M. B. Sandler. The author is also grateful to all researchers who collaborated in the applications reported in this paper.

REFERENCES

[1] D. Marr *Vision*. W. H. Freeman, 1982.

[2] P. H. Lindsay and D. A. Norman. *An Introduction to Psychology*. Harcourt Brace Jovanovich, 1977.

[3] R. Brons. Theoretical and linguistic methods for describing straight lines. In R. A. Earnshaw, editor, *Fundamental Algorithms for Computer Graphics*, pages 19–57. Springer-Verlag, Berlin/Heidelberg, 1985.

[4] H. Freeman. Computer processing of line-drawing images. *Computing Surveys*, 6(1):57–97, Mar. 1974.

[5] L. da F. Costa. *Effective Detection of Line Segments with Hough transform*. PhD thesis, King's College London, University of London, London, UK, May 1992.

[6] J. Koplowitz, M. Lindenbaum, and A. Bruckstein. The number of digital straight lines on an NxN grid. *IEEE Transactions on Information Theory*, 36(1):192–197, Jan. 1990.

[7] J. Illingworth and J. Kittler. A survey of the Hough transform. *Computer Vision, Graphics, and Image Processing*, 44:87–116, 1988.

[8] L. da F. Costa and M. B. Sandler. The binary Hough transform and its implementation. In *Proc. 1990 SPIE/SPSE Symposium on Electronic Imaging Science and Technology – Curves and Surfaces in Computer Vision and Graphics*, pages 183–193, Sta. Clara, USA, Feb. 1990. Paper No. 1251-21.

[9] G. Gerig and F. Klein. Fast contour identification through efficient Hough transform and simplified interpretation strategy. In *8th International Joint Conference on Pattern Recognition*, pages 498–500, Paris, France, 1986.

[10] L. da F. Costa and M. B. Sandler. A complete and efficient real time system for line segment detection based on the binary Hough transform. In *Proc. Euromicro'90 Workshop on Real Time*, pages 205–213, Horsholm, Denmark, Jun. 1990.

[11] L. da F. Costa and M. B. Sandler. Application of the binary Hough transform to image compression. In *6th International Conference on Digital Processing of Signals in Communications*, pages 56–60, Loughborough, UK, Sep. 1991.

[12] L. da F. Costa, D. R. Andrews, and M. B. Sandler. Quality control of ultra-sound transducers with the binary Hough transform. In *19th International Symposium on Acoustical Imaging*, Bochum, Germany, 1991.

[13] L. da F. Costa, X. Leng, M. B. Sandler and P. Smart. A system for semi-automated analysis of clay-samples. *Review of Scientific Instruments*, 62:2163–2166, Sep. 1991.

[14] L. da F. Costa, M. B. Sandler, and S. Velastin. Applying image analysis techniques to optimize the production of cork stoppers. In *Proc. SIBGRAPI'92*, Lindóia, Brazil, Nov. 1992. Submitted to appreciation.

[15] W. Lin and D. A. Fraser. Algorithms and timing for identification of objects from 2-D images. *Concurrency: Practice and Experience, 3(4):325–331, 1991.*