Special Issue: SAICSIT '99
## Contents

**Preface**  
P. Machanick ............................................................... 1

### Research Articles

- **Active Learning: Issues and Challenges for Information Systems and Technology**  
  RD Quilling, GJ Erwin and O Petkova ............................................................... 5

- **A Generic Modelling Framework for Interactive Authoring Support Environments**  
  Paula Kotze ............................................................... 15

- **An Application of a Framework for Evaluation of the Factors Affecting Software Development Productivity in the Context of a Particular Organisational Environment**  
  O Petkova and JD Roode ............................................................... 26

- **An Information-Theoretic Semantics for Belief Change**  
  T Meyer ............................................................... 33

- **A Complexity Metrics Model for Software Correction**  
  A Törn, T Andersson and K Enholm ............................................................... 40

- **A Conceptual Design for High-Volume Data Processing of Warehouse Database into Multidimensional Database**  
  Paisarn Trakulsuk and Vichit Avatchanakorn ............................................................... 49

- **A Pragmatic Approach to Bitemporal Databases: Conceptualization, Representation and Visualisation**  
  Chiyaba Njovu and WA Gray ............................................................... 58

- **A Building Recognition System**  
  SP Levitt and B Dwolatzky ............................................................... 68

- **Computer Programming and Learning to Write**  
  John Barrow ............................................................... 77

- **Co-operating to Learn using JAD Technologies**  
  TA Thomas ............................................................... 87

- **Critical Success Factors for the Implementation of DSS at a Selection of Organisations in Kwazulu/Natal**  
  URF Averweg and GJ Erwin ............................................................... 95

- **Enhancing the Predictability of Two Popular Software Reliability Growth Models**  
  Peter A Keiller and Thomas A Muzzuchi ............................................................... 105
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalised Unification of Finite Temporal Logic Formulas</td>
<td>Scott Hazelhurst</td>
<td>110</td>
</tr>
<tr>
<td>Harmonizing Global Internet Tax: A Collaborative Extranet Model</td>
<td>E Lawrence and B Garner</td>
<td>119</td>
</tr>
<tr>
<td>Improving Object Oriented Analysis by Explicit Change Analysis</td>
<td>Lui Yu, Siew Chee Kong, Yi Xun and Miao Yuan</td>
<td>128</td>
</tr>
<tr>
<td>Reconciling the Needs of New Information Systems Graduates and Their Employers in Small, Developed Countries</td>
<td>Rodney Turner and Glenn Lowry</td>
<td>136</td>
</tr>
<tr>
<td>Shortest Delay Scheduling Algorithm for Lossless Quality Transmission of Stored VBR Video under Limited Bandwidth</td>
<td>Fei Li, Yan Liu, Jack Yiu-Bun Lee and Ishfaq Ahmad</td>
<td>146</td>
</tr>
<tr>
<td>Software Croma Keying in an Immersive Virtual Environment</td>
<td>Frans van der Berg and Vail Lalioti</td>
<td>155</td>
</tr>
<tr>
<td>Some Automata-Theoretic Properties of $\cap$-NFA</td>
<td>Lynette van Zijl and Andries PJ van der Walt</td>
<td>163</td>
</tr>
<tr>
<td>The CILT Multi-Agent Learning System</td>
<td>Hema L Viktor</td>
<td>176</td>
</tr>
<tr>
<td>The Development of a Generic Framework for the Implementation of Cheap, Component-Based Virtual Video-Conferencing System</td>
<td>Soteri Panagou and Shaun Bangay</td>
<td>185</td>
</tr>
<tr>
<td>The Role of Experience in User Perceptions of Information Technology: An Empirical Examination</td>
<td>Meliha Handzic and Graham Low</td>
<td>194</td>
</tr>
<tr>
<td>What are Web Sites Used for: Cost Savings, Revenue Generating or Value Creating?</td>
<td>Man-Ying Lee</td>
<td>201</td>
</tr>
<tr>
<td>New Ideas Papers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approaches to Video Transmission over GSM Networks</td>
<td>Bing Du and Anthony Maeder</td>
<td>210</td>
</tr>
<tr>
<td>From Information Security Baselines to Information Security Profiles</td>
<td>Rossouw von Solms and Helen van der Haar</td>
<td>215</td>
</tr>
<tr>
<td>Grounded Theory Methodology in IS Research: Glaser versus Strauss</td>
<td>J Smit</td>
<td>219</td>
</tr>
<tr>
<td>Introducing a Continuum of Abstraction-Led Hierarchical Search Techniques</td>
<td>Robert Zimmer and Robert Holte</td>
<td>223</td>
</tr>
</tbody>
</table>
Multimedia as a Positive Force to Leverage Web Marketing, with Particular Reference to the Commercial Sector
Stan Shear ................................................................. 229

Understanding HCI Methodologies
Peter Warren .............................................................. 234

Electronically Published Papers

Experience Papers
A Java Client/Server System for Accessing Arbitrary CANopen Fieldbus Devices via the Internet
Dieter Bühler, Gerd Nusser, Gerhard Gruler and Wolfgang Küchlin ........................................... 239

An Object-Oriented Framework for Rapid Client-side Integration of Information Management Systems
Ralf-Dieter Schimkat, Wolfgang Küchlin and Rainer Krautter ...................................................... 244

Distributed Operating Systems: A Study in Applicability
Jürgen Prange and Judith Bishop ........................................... 249

Formal Verification with Natural Language Specifications: Guidelines, Experiments and Lessons so far
Alexander Holt .......................................................... 253

Introducing Research Methods to Computer Science Honours Students
Vashti Galpin, Scott Hazelhurst, Conrad Mueller and Ian Sanders .................................................. 258

Visualising Eventuality Structure
ST Rock .................................................................. 264

Electronically Published Papers
SAICSIT’99

South African Institute of Computer Scientists and Information Technologists

Annual Research Conference 17-19 November 1999

Prepare for the New Millennium

Is there life after y2k?

Mount Amanzi Lodge, Hartebeespoort, South Africa

Sponsors

http://www.cs.wits.ac.za/

PHDS research/PHDS.html

Think different.
Preface

Philip Machanick, Overall Chair: SAICSIT’99

Running SAICSIT’99, the annual research conference of the South African Institute for Computer Scientists and Information Technologists, has been quite an experience.

SAICSIT represents Computer Science and Information Systems academics and professionals, mainly those with an interest in research. When I took over as SAICSIT president at the end of 1998, the conference had not previously been run as an international event. I decided that South African academics had enough international contacts to put together an international programme committee, and a South African conference would be of interest to the rest of the world.

I felt that we could make this transition at relatively low cost, given that we could advertise via mailing lists, and encourage electronic submission of papers (to reduce costs of redistributing papers for review).

The first prediction turned out to be correct, and we were able to put together a strong programme committee.

As a result, we had an unprecedented flood of papers: 100 submitted from 21 countries. As papers started to come in, it became apparent that we needed more reviewers. It was then that the value of the combination of old-fashioned networking (people who know people) and new-fashioned networking (the Internet) became apparent. While the Internet made it possible to convert SAICSIT into an international event at relatively low cost, the unexpected number of papers made it essential to find many additional reviewers on short notice. Without the speed of e-mail to track people down and to distribute papers for review, the review process would have taken weeks longer, and it would have been much more difficult to track down as many new reviewers in so little time.

Even so, the number of referees who were willing to help on short notice was a pleasant surprise.

The accepted papers cover an interesting range of subjects, from management-interest Information Systems, to theoretical Computer Science, with subjects including database, Java, temporal logic and implications of e-commerce for tax.

In addition, we were very fortunate in being able invite the president of the ACM, Barbara Simons as a keynote speaker. Consequently, the programme for SAICSIT’99 should be very interesting to a wide range of participants.

We were only able to find place in the proceedings for 36 papers out of the 100 submitted, of which only 24 are full research papers. While this number of papers is in line with our expectation of how many papers would be accepted in each category, we did not have a hard cut-off on the number of papers, but accepted all papers which were good enough, based on the reviews. Final selection was made by myself as Programme Chair, and Derrick Kourie, as editor of the South African Computer Journal. Additional papers are published via the conference web site.

We believe that we have put together a quality programme, and hope you will agree.

Acknowledgments

I would like to thank the South African Computer Journal production team, Andries Engelbrecht and Herna Viktor, respectively from the Department of Computer Science and Informatics, University of Pretoria, for their work on producing the proceedings.

The reviewers listed overleaf did an excellent job: many wrote very detailed reports, sometimes after being called in on very short notice. Inevitably, there were some glitches resulting from the unexpected workload, but the buck stops with the programme chair: I promise to do better next time.

I would also like to thank my own department for putting up with the extra work and expense that running a conference entails. I tried not to burden them with too much extra work, but our secretaries, Zain Gowar and Leanne Reddy, inevitably had to take on some extra work. John Ostrowick provided valuable assistance with design of our web pages and call for papers poster. Carol Kernick, who handles our finances and membership records, did a fine job of keeping up with the demands of the conference.

Finally, I would like to thank our sponsors, whose contribution made this conference been possible:

- PricewaterhouseCoopers – sponsored generous prizes and the conference banquet
- National Research Foundation (NRF) – provided financial support
- University of the Witwatersrand – provided financial support
- Programme for Highly Dependable Systems, University of the Witwatersrand – provided financial support
- Standard Bank – provided financial support
Editorial

- Apple Computer – provided equipment for the conference
- Qualica – provided technical support including helping with the conference web site

Web Site

For more information about SAICSIT, including a pointer to the conference site, see <http://www.saicsit.org.za>.

Referees

- Department of Computer Science, University of Pretoria
  - Derrick Kourie
  - Bruce Watson
  - Vali Lalioti
  - Andries Engelbrecht
  - Ivan Mphahlele
- German National Research Center for Information Technology - GMD
  - Gernot Goebbels
- School of Math Stat and IT, Natal University Pietermaritzburg
  - Peter Warren
- Graduate School of Business, University of Cape Town
  - Kurt April
- CSIR
  - James Jardine
- Department of Computer Science, Rhodes University
  - Shaun Bangay
  - Peter Clayton
  - John Ebden
  - Richard Foss
  - George Wells
  - Peter Wentworth
- Information Technology Division, Rhodes University
  - Caro Watkins
- Department of Information Systems, Rhodes University
  - Brenda Mallinson
  - Dave Sewry
- Department of Informatics, University of Pretoria
  - Carina de Villiers
  - Herna Viktor
  - Niek du Plooy
  - Elsje van Rooyen
  - Machdel Matthee
  - Alan Abrahams
  - Jackie Phahlamohlaka
- School of Information Technology & Engineering, University of Ottawa
  - Stan Szpakowicz
  - Dwight Makaroff
- Department of Information Systems, Victoria University
  - Glenn Lowry
  - Peter Shackleton
  - Tas Adam
  - Alastair Wallace
  - Stephen Burgess
  - Dave Burgess
  - Julie Fisher
  - Jerzy Lepa
  - Geoff Sandy
  - Rod Turner
  - Alastair Wallace
  - Andrew Wenn
- Peninsula School of Computing and Information Technology, Monash University
  - Ainslie Ellis
- School of Information Management and Systems, Monash University
  - Angela Carbone
- School of Computer Science & Software Engineering, Monash University
  - Dianne Hagan
  - Judy Sheard
- School of Management Information Systems, Deakin University
  - John Lamp
  - David Mackay
  - Philip Joyce
- Marie van der Klooster
- Bill Hewett
- Jennie Carroll
- Rodney Carr

• Computer and Automation Research Institute, Hungarian Academy of Sciences
- Ferenc Vajda

• Department of Computer Science, University of the Witwatersrand
- Ian Sanders
- David Lubinsky
- Conrad Mueller
- Yinong Chen
- Bob Baber
- Vashti Galpin
- András Salamon
- Scott Hazelhurst
- Philip Machanick
- Zoltán Fazekas

• Department of Electrical Engineering, University of the Witwatersrand
- Barry Dwolatzky
- Farzin Aghdasi

• Department of Mathematics, University of the Witwatersrand
- Jonathan Burgess

• School of Law, University of the Witwatersrand
- Victoria Bronstein

• Edward Nathan & Friedland Attorneys, Johannesburg
- Justine White

• Department of Information Systems & Technology, University of Durban-Westville
- Geoff Erwin

• Department of Computer Science, Clemson University, Clemson, SC
- Karen Hay

• Logical SA
- Philip Green

• Department of Computer Science, University of Bristol
- Alan Chalmers

• Department of Systems and Computer Science, Howard University, Washington, D.C
- Todd Shurn
- John Trimble

• Department of Computer Science, University of Cape Town
- Sonia Berman

• King's College, London
- Peter Wood

• Department of Information Systems, University of Cape Town
- Paul Licker
- Steve Erlank
- Alemayehu Molla

• Department of Department of Mathematics and Applied Mathematics, University of Cape Town
- Ingrid Rewitzky
- Peter Jipsen
- Renkuan Guo

• Department of Information Systems, University of the Western Cape
- Andy Bytheway

• Department of Electrical Engineering and Computer Science, The George Washington University, Washington, DC
- Rachelle Heller

• School of Biophysical Sciences and Electrical Engineering, Swinburne University of Technology
- Ian Macdonald

• Electrical and Electronic Systems Engineering, Queensland University of Technology
- Anthony Maeder

• Computer Science Department, Åbo Akademi
- Aimo Törn
- Kaisa Sere
- Mats Aspnas

• Division of Informatics, University of Edinburgh
- Graeme Ritchie
- Arturo Espinosa-Romero
- Jon Oberlander
Editorial

- Chris Brew
- Alexander Holt

• Department of Computer Science and Engineering, Helsinki University of Technology
  - Tommi Junttila
  - Nisse Husberg

• Technical University of Budapest, Department of Measurement and Information Systems
  - Andras Pataricza

• NASA-Goddard Space Flight Center
  - Nigel Ziyad

• Department of Computer Science & Information Systems, University of Port Elizabeth
  - Janet Wesson
  - Leon Nicholls

• Computer Technology Department, Indiana University Purdue University Indianapolis
  - Tim Price

• Department of Electrical Engineering and Computer Science, The University of Michigan
  - Trevor Mudge

• Department of Electrical Engineering and Computer Sciences, University of California, Berkeley
  - David Forsyth

• School of Computer Science, University of Birmingham
  - Mark Ryan

• Faculty of Mathematics, Computer Science, Physics & Astronomy, University of Amsterdam
  - Carlos Areces

• Department of Computer Science, Universidade do Vale do Rio dos Sinos - UNISINOS Rio Grande do Sul
  - Marcelo Walter

• School of Information Technology & Mathematical Sciences, University of Ballarat
  - Binh Pham
BuRS: A Building Recognition System

S.P. Levitt and B. Dwolatzky

Department of Electrical Engineering, University of the Witwatersrand, Johannesburg, WITS 2050, South Africa

Abstract

In this paper the design of a novel building extraction system, entitled BuRS, is proposed. It is capable of identifying man-made structures that are visible in monocular greyscale aerial photography of rural settlements and sparsely built-up areas. The system design is texture measure driven; integrates both edges and regions; and does not utilise an explicit model of the structures being sought for. One specific component of BuRS is investigated in detail, namely the region formation component. The region formation component incorporates a region growing operator devised by the authors and entitled "the homogeneous operator". We demonstrate that this operator can successfully locate structures exhibiting a homogeneous appearance and has utility in being the first primary processing stage of a building recognition system.

Keywords: Computer Vision, Building Extraction, Homogenous Operator

Computing Review Categories: I.4, I.4.8, I.4.6

1 Introduction and Background

A significant amount of work that has been done in the field of aerial image understanding has concentrated on the detection of buildings and other man-made structures (such as roads). Building recognition strategies for monocular images fall into one of two categories - edge-driven or region-driven approaches.

An edge-driven approach uses an edge map as its starting point. An edge in an image is defined by Jain, Kasturi and Schunck [6] as a “significant local change in image intensity usually associated with a discontinuity in either the image intensity or the first derivative of the image intensity”. Edges are found through the use of edge operators or edge detectors. Such operators are local in that they operate only on a small area or local neighbourhood of an image. Edge detectors normally produce an edge magnitude (representing the strength of the edge) as well as an edge direction or orientation (representing the direction in which the maximal intensity change is occurring). Applying an edge operator results in a gradient image or edge map which may consist of one or both of the above mentioned components.

In a region-driven building recognition strategy, the source image is initially segmented entirely into different regions. A region is defined as a group of connected pixels that have similar properties. Two of the most important criteria for grouping pixels are value similarity and spatial proximity. This comes from the intuitive assumption that neighbouring points on an object’s (building’s) surface will be similar in greyscale value and will project to neighbouring points in an image of the object. The first part of the assumption is often violated - in cases of uneven illumination, when the object’s surfaces vary greatly in greyscale value and so on. After segmentation, an attempt is made to determine which regions correspond to building components.

Most edge detectors and region growers are low-level techniques in that they do not make use of scene-domain knowledge [1]. As a result of this they tend to perform very poorly. The remainder of the interpretation process is devoted to correcting the errors and refining the results of these low level techniques. In edge-based approaches, straight line extractors [2, 15] are used in order to try and reduce the numerous spurious and insignificant edges that are found. The remaining edges are then grouped to form building shapes. Typically, buildings are modelled as rectangles or combinations of rectangles [5, 11, 8]. In region-based approaches, domain dependent knowledge and heuristics are used to identify regions forming building components [13]. Alternatively, shadow is used in order to identify regions corresponding to building roof [10].

A number of strategies have been presented in the literature which combine edge finding and region growing. These combined strategies offer a superior solution because the integration of complementary methods helps to overcome deficiencies of each individual method [10, 14]. Note that strategies which combine edges and regions can still be classified as edge-driven or region-driven according to the method that forms the primary process. The primary process is that method which determines the locations of the building hypotheses (areas in the image that are assumed to be “building”). The secondary process is used for optimisation of hypotheses’ boundaries and for verification of hypotheses.

Relatively little research has been conducted into developing low-level techniques that have the ability to discriminate directly between man-made and natural features in monocular photographs, and to the authors’ knowledge, no such techniques have been incorporated into an aerial image interpretation system. Techniques which have been developed thus far are fractal-based.
2 System Design

2.1 Assumptions and Limitations

BuRS is limited with regard to the images that it is capable of interpreting. Huertas and Nevatia [5] identify three "dimensions of complexity" with regard to aerial images containing man-made structures:

1. **Structure Density.** Rural scenes have low density housing, whilst urban scenes have a high density of man-made structures. Generally for lower housing densities, a successful interpretation is more attainable. This is due to the fact that most houses in rural scenes stand alone and therefore they can be more easily distinguished from background scenery.

2. **Structure Shape.** Buildings may appear as many different two dimensional shapes from the air. The most common shape is naturally a rectangle or a combination of rectangles (such as in "E", "T" and "L" shaped buildings). However many other shapes also occur, such as circles which would correspond to huts or rondavels. The actual three dimensional shape of a structure is also important as it affects its two dimensional appearance in an image. For instance a flat-topped rectangular structure and a simple A-frame house will both have a rectangular border in an aerial view. However the different planes of the A-frame's roof will be found separately if one plane is facing the sun while the other is not. This is a result of one plane having a significantly lighter colour than the other. Consequently, an edge appears along the apex of the roof, dividing the roof into two rectangles.

3. **Image Quality.** Images vary in terms of contrast, resolution and visibility of man-made structures. Having defined the dimensions of complexity above, the assumptions pertaining to these dimensions can be stated:

   1. All images to be interpreted by BuRS depict sparsely built-up areas.

   2. Man-made structures may be of any shape.

   3. Images must be of a quality such that human photo-interpreters can easily identify all the man-made structures. BuRS can be applied to images at scales of around 1:2,500. These images must be scanned at an appropriate resolution (200 to 300 dpi) so that house roofs are constituted by at least 200 pixels.

An additional and important assumption is that man-made structures appear in the images with surfaces of fairly constant intensity. This assumption is reasonable for the scale of photography for which the system has been designed. At this scale, roofs which have been constructed from a single type of material appear as having a uniform intensity because the fine detail and texture of the rooftop is not visible.

2.2 Design Overview

The vision system architecture is shown in Figure 1. The system consists of three major processes/components:
dictate where object edge searching should take place. Snakes (section 4.2). Region formation using a texture measure is the primary process. This idea is central to the system being presented.

Thus, the regions found dictate where object edge searching should take place.

Figure 2: The Homogeneous Operator

Edge Detection, Region Forma tion and the Snake Process. At this point in time only the region formation process has been researched and implemented.

The reason for this is that region formation process is the system's primary process while the edge detection and snake processes are secondary. The secondary processes are used only to refine the results of the region formation process. As such, these processes are enhancements to the system rather than core functionality. The output resulting from the primary process is, therefore, critical in determining the success of the entire system while the outputs of the secondary processes are not.

The secondary components of BuRS are briefly documented for completeness, as well as to demonstrate how the various system elements interrelate.

BuRS has been designed to incorporate techniques for both edge detection and for region formation. Edges and regions can be regarded as duals of each other [6]. In an ideal image they provide identical information. It should be possible to obtain edges from regions by detecting the region boundaries, and regions from edges by using a filling algorithm. However for real world images this is not the case because the image edges and region boundaries hardly ever coincide. This is due to noise present in the image and inadequacies in the region and edge finding algorithms. This means that although edges and regions will not give the same information, they will give similar information, thus they can be used to support each other in hypothesis formation. This idea is central to the system being presented.

The edge detection and region formation processes take place in parallel (Figure 1). The original image undergoes smoothing and then edge detection. In parallel, a region growing algorithm is applied, which if combined with connected component extraction, will result in the image being segmented into regions. The edges and regions found are then integrated using active contour models or snakes (section 4.2). Region formation using a texture measure is the primary process. Thus, the regions found dictate where object edge searching should take place.

3 Region Formation

The region formation process involves the sub-processes of region growing followed by connected component extraction and reduction. These sub-processes are described below.

3.1 Region Growing using the Homogeneous Operator

3.1.1 Theory

The "homogeneous" operator has been designed by the authors to highlight uniform regions within an image. This is useful as the roofs of man-made structures appear very uniform in intensity at the scale and resolution of the imagery being dealt with. The operator simply averages the absolute difference in pixel intensity between every pixel in the window and the centre pixel. The homogeneous operator (HO) represents a non-linear, spatially invariant filter. The operator is illustrated for a 3 x 3 window in Figure 2.

The homogeneous operator is a local neighbourhood operator applied globally to an image. The window or mask depicted in Figure 2 is moved over the image and the HO value is calculated for each pixel position.

A suitable threshold needs to be chosen in order to adequately segment the image. In order for the homogeneous operator to be applied successfully the window size must be smaller than the smallest house roof. A good heuristic is that ten non-overlapping windows should be able to fit onto the smallest roof panel. Thresholding the operator output results in an image foreground and background. All the pixels forming homogeneous regions in the image (having an HO value less than the threshold) are marked as foreground.

3.1.2 Results

All images presented here have been intentionally brightened for improved visual appearance. The foreground represents the output of the texture measure and is formed by regions of pure black.

Image 1 is the original source image. Note that it depicts a sparsely built-up area. There are a total of 14 houses. Two of these houses have been marked with letters. The longitudinal axis of the house marked "A" does not lie on either the horizontal or vertical axes as do the longitudinal axes of the other houses. The house marked "B" is interesting because the right-hand side (sun-facing) roof panel exhibits very little contrast with the ground. The rest of the image consists of a segment of road, some natural vegetation, and large barren sand patches.

Image 2 represents the ground truth for the original image. All houses have been identified and form the image foreground (coloured black). The ground truth was created by manually interpreting the image.

Image 3 shows the results of applying the homogeneous operator to the source image using a window size of 7 x 7 and a threshold of 12.
From Image 3 it can be seen that the homogeneous operator successfully finds all homogeneous regions in the image. The roads, houses and barren sand patches are all identified because of their uniform appearance.

The majority of houses are formed by three homogeneous regions - two roof panels and the house’s shadow. Most roofs are split up into two panels because one panel is sun-facing whilst the other is not. This leads to one panel being lighter in intensity than the other and high values in the gradient image occurring along the apex of the roof.

The right-hand roof panel of house B has “leaked” into the surrounding background. This is because of the minimal contrast that exists between the edge of that roof panel and the ground.

Only the interior area of uniform regions are found. As an object’s boundary is approached, the neighbourhood of the pixels close to the boundary begins to become less and less uniform, eventually surpassing the homogeneous threshold.

Areas in the image corresponding to natural vegetation do contain homogeneous regions but these regions tend to be relatively small and fragmented.

Image 4 depicts the HO values (refer to Figure 2) generated by the homogeneous operator. This image was generated by converting the HO values to pixel intensity values and then equalising the histogram of the resulting image. Histogram equalisation is performed in order to display the image reasonably as the homogeneous operator output is confined to a narrow range.

Black corresponds to the lowest HO value calculated while white corresponds to the highest. Therefore, black in the image represents areas where little change in intensity is occurring, while white represents areas having a high degree of change. From this image, it is possible to see strong edges (white) and homogeneous regions (black) as well as intermediate areas.

3.1.3 Evaluation

The advantages and disadvantages of the homogeneous operator are discussed below.

Aerial images have a highly variable structure that
Image 4: HO Values

does not allow for stringent system goal definitions [1]. The buildings being sought for may be distorted in various ways. These include RST (rotation-scaling-translation) distortion, the fact that buildings may vary in shape, and that they may be obscured. The homogeneous operator is impervious to all of these distortions when it is correctly applied to an image. For the operator to be correctly applied, a window should be used which has an area that is approximately a factor of ten smaller than the smallest roof panel in the image. If this is the case, then the window will fit many times within the borders of each building roof. This simply and effectively results in the operator being invariant to changes in rotation, scaling and translation. This is shown by the different buildings detected by the operator (Image 3). The operator is also unaffected by different building shapes as long as an appropriate window size is used. Finally, even if a house is obscured, the operator is still capable of recognising the unobscured parts.

The limitations of the homogeneous operator are as follows. Firstly, it is only able to identify uniform regions in an image. Therefore, if the scale of imagery is such that houses do not appear as uniform regions then the operator will be unable to identify them. If a house roof is made up of disparate materials (an example of this is given in section 3.3) or has many chimneys or slopes which lessen its homogeneous appearance then once again the operator will fail.

Secondly, the scope of the homogeneous operator is very narrow as it is a local operator. Consequently, it is unable to distinguish between a uniform region representing a sand patch, for example, and a uniform region representing a house. The operator finds all uniform regions in an image to which it is applied. Further processing of these regions is required in order to differentiate between them.

However, even with these limitations it is felt that the homogeneous operator is a valuable contribution to the set of methods that can be used in aiding a vision system in the identification of buildings and houses. It is also worth noting that comparative studies [9], conducted by the primary author and Aghdasi, show that the homogeneous operator substantially outperforms Cooper et al’s fractal error operator [4] on images of the kind presented in this paper.

Image 5: Filtering Regions by Size

Image 6: Finding Buildings Using Shadow

3.2 Connected Component Extraction and Reduction

After the application of a texture measure/region grower, an image foreground is produced. Connected components are extracted from the foreground according to the definition of 4-connectivity. Usually, the image foreground is divisible into a large number of connected components. Each of these components represents a region of fairly uniform intensity (if the homogeneous operator was used to derive the foreground). However, not every uniform region in the image will correspond to a structure or part of a structure. It is necessary to try and eliminate the connected components representing false-positives from the set of all connected components that have been found. Eventually, the only connected components that should remain are those representing building shadows, and building components. The following methods attempt to do this.

3.2.1 Size Filter

The connected component list is filtered on the basis of size, using a double threshold. The upper and lower limits define a range of acceptable region sizes. The lower limit is
set to exclude small regions (which are regarded as noise),
while the upper limit excludes very large regions that can
be due to roads or areas of sparse vegetation.

Image 5 shows the results of filtering the homogeneous
operator output (Image 3) based on region size. All small
regions and very large regions have been removed. Nearly
all of the roof panels and most of the house shadows re­
main. A number of false positives remain as well. Note
that the sun-facing roof panel belonging to the house B has
been removed. This is because the panel and some of the
surrounding background have merged to form a single re­
region, having a size exceeding the allowable range.

Overall, the results of size filtering are encouraging.
The number of false hypotheses are few and no false nega­
tives have occurred.

3.2.2 Use of Shadow

The homogeneous operator (and other region growers) in­
advertently find the interior regions of shadows. This is
due to the fact that shadows appear in aerial images as
regions of uniform intensity. However, shadows are nor­
mally darker than most objects in a scene and therefore
they can be relatively easily detected by considering the
lower portion of a histogram of the image. The average
intensity of each connected component extracted is calcu­
lated. Components with an average intensity that is lower
than a specified threshold are regarded as shadows. Most
shadow regions found by the homogeneous operator can
be identified in this manner. Each shadow found forms the
basis for a building hypothesis.

Shadows are an important cue for detecting buildings
and other raised structures. For this system, shadows are
used in a very elementary manner. A user has to manually
identify the 90 degree quadrant (North-East, South-East,
South-West or North-West) in which buildings can be ex­
tected to be found relative to their shadow location.

A search is then conducted for the building compo­
nents responsible for creating the shadows found. The
search takes place in the vicinity of the centroid of each
shadow region that has been found, in an area called the
shadow box. The shadow box is a square area, with the
length of the sides being determined by the user. For each
shadow region, the box is placed so that it has one corner
on the centroid of the shadow region and lies in the quad­
rant in which the structure causing the shadow is expected
to be located.

Any connected components (which have not been
classed as shadows) having centroids falling within a par­
ticular shadow’s shadow box are deemed to be the build­
ing components that are causing that shadow. These build­
ing components are lumped together with their respective
shadow to complete the building hypothesis. Shadows for
which no corresponding building components can be found
are regarded as false-positives and eliminated. Connected
components which have been classed as neither building
nor shadow are also regarded as false positives and elimi­
nated. Note that once a building component has been al­
located to a particular building hypothesis, it may not be
allocated to any other building hypotheses.

The use of shadows in the manner described helps to
reduce the number of misclassifications, as all homoge­
nous regions which do not correspond to raised structures
can be removed. It is important to realise that shadows can
be successfully used in identifying man-made structures
without requiring explicit models of such structures.

Image 6 shows the results of employing shadow to
reduce the number of false hypotheses. Only the actual
building components of each hypothesis are marked as
foreground. The shadow component which also forms part
of each hypothesis has been relegated to the image back­
ground. Parameters used here are as follows:

- Window size: 3 x 3
- Homogeneous threshold: 10
- Range of acceptable region sizes (in pixels): 70 - 350
- Shadow threshold: 45
- Length of shadow box: 20
- Quadrant in which buildings are expected: 1

Comparing Image 6 with the ground truth (Image 1),
one notices that ten out of fourteen houses have been suc­
cessfully recognised - a recognition success rate of 71.42
percent. In addition, one roof panel of an eleventh house
(house "B") has been found. Houses with axes of elonga­
tion aligned with the principle axes, as well as the house
that was marked "A" in Image 1, have been detected.

One false positive remains which indicated by "C"
in Image 6. This false hypothesis occurs because it is
of an acceptable size and it corresponds to the "shadow" formed by the vegetation to the South-west of it. There are
three false negatives or houses that have been completely
missed. A false negative occurs due to one of the following
reasons:

1. The house shadow was not found.
2. The house shadow was removed during the size filter­
ing process. Shadow regions are removed as a result
of being too small.
3. A mistake occurred in allocating building components
to shadow components. This often happens if build­
ings are closely clustered.

It is better to err on the side of false negatives than on
the side of false positives. This because it is easier for a hu­
manship interpreter, who is correcting an interpretation,
to find missed buildings than erroneous hypotheses.

Overall, the results are encouraging and show that the
region formation component designed is capable of per­
forming fairly well on imagery similar to Image 1.

The fact that roads generally have a uniform appear­
ance and are detected by the homogeneous operator means
that it is possible that this operator could be used as a start­
ing point in their identification. Instead of reducing the
number of regions found based on size and shadow, they could be eliminated based on the length of their axis of elongation.

3.3 Region Formation in Complex Images

This section demonstrates the region formation component of BuRS on a non-ideal image or an image not conforming to the system assumptions. In this away it is possible to evaluate the effectiveness of the region formation component over a range of input data.

Image 7 is the original source image. It depicts an informal South African township. The shacks are closely clustered together in an unstructured manner, and some of the shack roofs are composed of disparate materials. A shadow is present to the South-East of each building. Image 8 shows the results of applying both the homogeneous operator (Section 3.1) and size filter (Section 3.2.1) to Image 7. The parameters used are:

- Window size: $3 \times 3$
- Homogeneous threshold: 10
- Range of acceptable region sizes (in pixels): 20 - 550

Shacks which are composed of disparate materials, for example, shack “D” which is indicated in image 8, are completely undetected or are partially detected by the operator.

Image 9 shows the buildings found by using the simple method described for determining building-to-shadow correspondence (section 3.2.2). Additional parameters used are:

- Shadow threshold: 75
- Length of shadow box: 25
- Quadrant in which buildings are expected: 2

4 Secondary Components

4.1 Edge Detection

The original image is first filtered using a Gaussian filter. This essentially involves convolving the image with a two dimensional Gaussian function. Gaussian filtering suppresses small scale structure or fine detail whilst preserving large scale structure. This is desirable as it helps to reduce the amount of spurious edges found by edge detectors. After Gaussian filtering, the Sobel edge operator is used to create a gradient image.

4.2 Integrating Regions and Edges Using Snakes

Two distinct low-level approaches have been applied up to this point. As stated they are co-relative in that an edge detection algorithm will find the edges of (for instance) a house, while the use of a region growing operator will identify the interior of the house. Active contour models are now used to detect the true boundaries of the objects identified through the use of the region grower.

4.2.1 Background to Snakes

Active/deformable contour models or snakes were initially proposed by Kass, Witkin and Terzopoulos [7]. A snake is an energy minimising spline. The energy functional has components of internal energy and external energy. Refer to the equation (1) below.
This means they are initialised on the interior of building that have been grown using the region growing operator. In the contours that it is trying to detect, otherwise the snake has to be initialised in a position which is spatially close to the boundary of the object being sought for. In this approach may latch onto local minima that do not correspond to the snakes are initialised on a gradient image of the original (section 4.1). The gradient image is negated so that edges which previously corresponded to peaks in the gradient now form valleys or troughs. This allows a snake to minimise its energy by moving toward minima in the gradient image. For instance, this allows the contour to be attracted to certain image features such as edges.

Snakes can have a very simple form. They can consist of a set of control points or nodes. Each node is joined to its neighbouring nodes via straight lines. Each node has a position, given by its \((x,y)\) pixel co-ordinates in the image. A snake is entirely specified through the number and co-ordinates of its nodes. A snake deforms as adjustments are made to the positions of each of the snake nodes. The direction of deformation is chosen so as to minimise the snake's energy functional.

For this application closed contours are used. The internal energy of the contour is specified in such a manner as to cause the contour to inflate or balloon outward. Inflating contours have been used by Cohen [3] for the detection of ventricles in medical images. The external energy of the snake is defined so that the snake is attracted to edges in the image.

In order for a snake to be attracted to edges it must be initialised on a gradient image of the original (section 4.1). The gradient image is negated so that edges which previously corresponded to peaks in the gradient now form valleys or troughs. This allows a snake to minimise its energy by moving toward minima in the gradient image.

**4.2.2 Snake Initialisation**

In any approach which uses snakes to find image contours, the initialisation of the snake is very important. A snake has to be initialised in a position which is spatially close to the contours that it is trying to detect, otherwise the snake may latch onto local minima that do not correspond to the boundary of the object being sought for. In this approach the snakes are initialised on the boundaries of the regions that have been grown using the region growing operator. This means they are initialised on the interior of building roofs and they expand towards the building edges (refer to Figure 3).

**4.2.3 Advantages of using Snakes**

Snakes provide an elegant and logical way of combining edge detection and region growing.

The snake model is defined to be a closed curve and therefore it has the ability to interpolate across missing data. Snakes also contain parameters to enforce smoothness which counteracts the effects of image noise. Snakes are capable of representing a class of objects of varying shape. This is important as it essential that the snake process not compromise our design goal of recognizing man-made structures of any shape. The final output of the system, a vector map containing polygons representing the building hypotheses' boundaries, can be deduced directly from the co-ordinates of the snake nodes. The shape of each boundary detected can be used as independent verification of whether the object found does in fact represent a man-made structure.

**5 Conclusions**

We have demonstrated the utility of our system in interpreting images of sparsely built-up areas and we have shown its limitations when it is applied to complex scenes.

Most building recognition methodologies incorporate an explicit model of the sought objects. Buildings are normally modelled as rectangles or collections of rectangles [5, 8, 11] or as sets of regions with associated geometric, photometric and topological attributes [13]. Liow and Pavlidis's [10] "region growing then edge detection" approach does not incorporate a model of shape. Hypotheses are verified via shadow and shape analysis.

The methodology presented here also does not incorporate any model of shape or place constraints on the shapes of the objects being searched for. This is desirable as there is great variability in building shapes - from circles to pentagons - and thus it is better not to restrict the search to any particular shape/s.

All the edge-driven and region-driven aerial interpretation systems reviewed employ context-insensitive edge detectors and region growers that either produce an edge map of the source image or segment it entirely into regions. Subsequent processing algorithms are applied globally to the image and so the focus of attention is not narrowed.

This system is different in this regard and can be said to be texture measure driven, as our homogeneous operator produces a measure of texture. The homogeneous operator is designed to take advantage of scene-domain knowledge (the fact that the objects being searched for appear with uniform intensity or texture) and find only significant homogeneous regions in the image. This has important consequences in that it reduces the processing time for any subsequent algorithms applied to the image as they can immediately be set to "focus their attention" on the areas discovered by the homogeneous operator.
In this system, we first identify shadows and then attempt to locate the building/s causing each shadow. In other words we use building-to-shadow correspondence. This is in direct contrast to the method documented by Liow and Pavlidis [10]. They segment the entire image into regions and each region is taken to be a building hypothesis. They then perform a shadow-to-building correspondence check on each hypothesis.

The advantage of the method used here is that it involves far less processing. The number of shadow regions that exist in an image will always be far less than the total number of regions that exist. Therefore, it is felt that it is preferable to start with shadows and look for buildings rather than the other way around.

Our utilisation of shadow information is at present rudimentary and our algorithms cannot deal with complex scenes in which buildings are closely clustered.

Finally, we have proposed a novel method for integrating edge detection and region growing with the use of inflating deformable contours.

References


Notes for Contributors

The prime purpose of the journal is to publish original research papers in the fields of Computer Science and Information Systems, as well as shorter technical research notes. However, non-refereed review and exploratory articles of interest to the journal's readers will be considered for publication under sections marked as Communications of Viewpoints. While English is the preferred language of the journal, papers in Afrikaans will also be accepted. Typed manuscripts for review should be submitted in triplicate to the editor.

Form of Manuscript
Manuscripts for review should be prepared according to the following guidelines:

- Use wide margins and 1½ or double spacing.
- The first page should include:
  - the title (as brief as possible)
  - the author's initials and surname
  - the author's affiliation and address
  - an abstract of less than 200 words
  - an appropriate keyword list
  - a list of relevant Computing Review Categories
  - Tables and figures should be numbered and titled.
- References should be listed at the end of the text in alphabetic order of the (first) author's surname, and should be cited in the text according to the Harvard method.

Manuscripts accepted for publication should comply with guidelines as set out on the SACJ web page, http://www.cs.up.ac.za/sacj which gives a number of examples.

SACJ is produced using the $\LaTeX$ document preparation system, in particular $\LaTeX$ 2ε. Previous versions were produced using a style file for a much older version of $\LaTeX$, which is no longer supported.

Please see the web site for further information on how to produce manuscripts which have been accepted for publication.

Authors of accepted publications will be required to sign a copyright transfer form.

Charges
Charges per final page will be levied on papers accepted for publication. They will be scaled to reflect typesetting, reproduction and other costs. Currently, the minimum rate is R30.00 per final page for contributions which require no further attention. The maximum is R120.00, prices inclusive of VAT.

These charges may be waived upon request of the author and the discretion of the editor.

Proofs
Proofs of accepted papers may be sent to the author to ensure that typesetting is correct, and not for addition of new material or major amendments to the text. Corrected proofs should be returned to the production editor within three days.

Letters and Communications
Letters to the editor are welcomed. They should be signed, and should be limited to about 500 words. Announcements and communications of interest to the readership will be considered for publication in a separate section of the journal. Communications may also reflect minor research contributions. However, such communications will not be refereed and will not be deemed as fully-fledged publications for state subsidy purposes.

Book Reviews
Contributions in this regard will be welcomed. Views and opinions expressed in such reviews should, however, be regarded as those of the reviewer alone.

Advertisement
Placement of advertisements at R1000.00 per full page per issue and R500.00 per half page per issue will be considered. These charges exclude specialised production costs, which will be borne by the advertiser. Enquiries should be directed to the editor.