

IMAGE & GRAPHIC READER

Ingo Redeke

University of Bremen, TZI

ABSTRACT

The Image & Graphic Reader (IGR) is an assisting tool to retrieve information from business charts and present it to blind computer users. Different business charts occurring in newspapers and web sites are analyzed by standard image processing methods. Common segmentation algorithms produce homogeneous areas and contour lines that characterize specific chart types. Since the component's geometric features encode the tabular content of the chart, their dimensions are measured against a certain scale. This scale is either explicitly defined in the image, like scales of a previously identified axis or is intrinsic to the type of the chart (e.g. angles in pie charts).

After the primal classification the extracted tabular data is analyzed to produce a representation that subdivides a chart into predefined levels which are then verbally described. A blind user can thus navigate through the levels by keyboard and the spoken feedback of a speech synthesizer.

1. INTRODUCTION

Blind people have been widely excluded from a majority of computer based information devices by the introduction of graphical user interfaces. Graphical environments are dominated by mouse controlled applications and selectable icons are meant to provide an overview for quick execution of applications. Furthermore, textual references to images and diagrams might occur for "visualizing" of what is meant in the text. Although there are text-based solutions available like screen readers [1, 2] or text only web browsers [3], there is no tool to systematically present image contents to a blind computer user. Different approaches, like the Kevin project [4], apply tactile print-outs, sometimes integrating verbal descriptions. These approaches are generally expensive both in costs for hardware and paper and they are limited to a resolution that can be understood haptically. The development of methods to automatically identify and annotate images like the IRIS project of the TZI [5] does, has started only a few years ago.

To facilitate the blind users situation, the TZI develops the IGR prototype that describes business charts and

the contained information thoroughly and comprehensive. While understanding charts is mainly achieved by means of navigation through the semantics of the graph, the image processing stage is to retrieve the type of a chart and its data. One major aspect of the project is to define the link between the analysis and the presentation. I.e. how need the data be processed to create a concept of a chart that a blind user can understand.

Business charts like in fig.1 are all common in that they represent an overview of a data table of dimension $n \geq 2$. The kind and range of the data predominates the chart type used, e.g. a pie chart would be a bad choice (less "telling on the first view") to display changes of stock values.

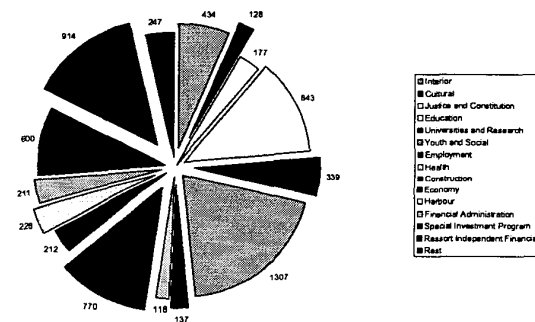


Fig. 1. Pie chart

The main idea behind the IGR prototype is to have a technology at hand that can be easily added to an existing application. A common screen reader might include the IGR as a plug-in. Once an image is detected (either as a link to an image file or as an area a scanner reads), the analysis is started by a key stroke or is automatically triggered.

This paper is built on a description of the low level image processing unit in section 1 and a hierarchical concept of what needs being concentrated on when dealing with a chart's data in section 2. Section 3 notes down how the charts intrinsic imprecision should be dealt with. Finally some results of the image analysis are displayed in section 4 and the paper closes with an outlook to future work (section 5).

Thanks to the BMA for funding. Ingo Redeke is with the TZI, redeke@tzi.de

2. IMAGE ANALYSIS

2.1. Low Level Analysis and Basic Geometrical Objects

The image analysis starts with the segmentation of areas of homogeneous color and/or texture. Since the images in this case are artificial, it can be assumed that the segmentation produces objects with noiseless geometrical features. The segmentation and the subsequent capture of contour lines (fig.2) are low level image processing tasks and can be performed by standard techniques [6, 7, 8]. The 1 pixel wide

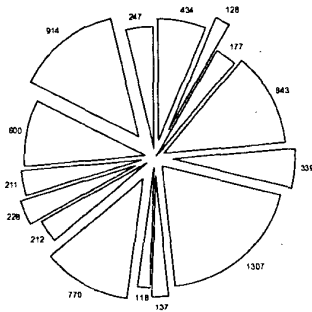


Fig. 2. Contour lines extracted from segments in fig.1

contours of the segmented regions are produced into polygons by detection of high curvature points. The polygons are further divided into straight lines and arcs. Combinations of these lines produce T-, Y-, X-, or L junctions [9, 10]. The assembly of lines, arcs, and crossings generate geometric objects, such as rectangles and segments. We have defined a set of simple objects that can be constructed this way. E.g. a pie segment may consist of two straight non parallel side lines of similar lengths s connected at their starting points. Their end points are linked by a line or a convex arc having a radius that is approximately s . Exceptions must be considered, e.g. half-circles which consist of only one straight line of length s' whose end points are linked by an arc with radius $s'/2$.

All simple geometrical objects need to have minimum dimensions. We have applied a θ_a in pixels for the minimum size of areal objects and a θ_l in pixels for the minimum length of a line belonging to a specific geometrical object. Other composite thresholds apply for rectangles, e.g. relations of side lengths, etc.

2.2. High Level Analysis

Once the primal geometrical objects have been identified, their relations can be analyzed. Compositions of objects that are relevant for a specific chart type are modeled by a set of constraints. There exist several ways to classify objects

and their compositions. Knowledge presentation in this area is usually based on graph grammars [11], constraint nets, entity-relationship graphs or Prolog-like descriptions [12]. To categorize relevant parts of a chart, we have developed leveled concepts of chart models that formulate constraints for their components. These models can be visualized by UML-diagrams, such as in fig.4.

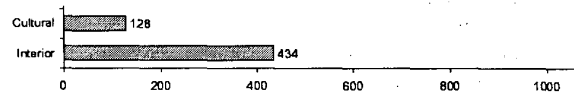


Fig. 3. Axis.candidate

If the analysis detects, for instance a “long” straight line with a set of equally spaced T-, or X-junctions on it, the object is a candidate for an axis (fig.3).

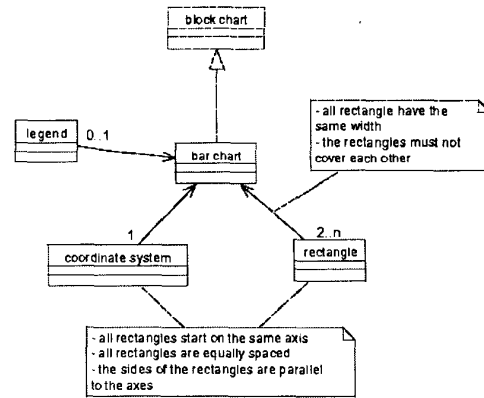


Fig. 4. Model for classification of a bar chart

Although there is no general standard for business charts, all charts that we have used so far can be reduced to a limited set of reproducible constraints for their elements. This yields a limited total amount of different chart types. UML diagrams [13] facilitate modularizing the constraints down to an elementary level.

To identify a specific chart type, all of its minimum conditions must be fulfilled. That means its simple geometrical objects must occur in a predefined amount. Furthermore their relations to one another and to other geometrical objects must fit for the chart type. The classes for the charts are sufficiently dissimilar to guarantee unique identification.

2.3. Associated Text Elements

Textual information must be considered a problem of its own. Apart from OCR mechanisms, text in charts are locally attached to associated graphical information. Text might occur inside or close at extended areal objects. Other text parts lie along axes to denote scales.

Text blocks, rectangular areas that contain text elements (red fields in fig.5), will be extracted by standard OCR. Their center points can be used to quantize distances to formerly classified chart parts. A text that consists of a numerical value will be associated to this part, if its center distance lies closest of all to the text block. Specific objects show special cases of association, such as pie segments, whose enlarged cone defines a primary search area for associated text (green area in fig.5).

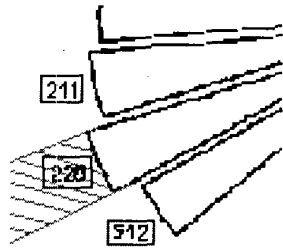


Fig. 5. Associated Text

So far, an optical character recognition analysis has not been added to the IGR. It is planned to implement an algorithm that quickly pre-identifies parts of the image where text is very likely to occur. These parts would be cropped and sent to a standard OCR software that takes over the identification.

3. (IM)PRECISION

An intrinsic risk to our idea is the lack of accurate precision in charts. Unless the data is normalized to pixel units, the resulting chart image is far from being an accurate copy. On the other hand, charts are determined to merely provide a brief overview of the data. So few graphics designers will ever set a high value on visual accuracy. The true handicap is that the image analysis might produce an accuracy that is just not there. For instance, take a pie segment which is described by its angle or area compared to a full circle. The proportion the program computes for the segment might be unwillingly exact (showing a result of 20.8401% when the chart actually denotes 21%). To meet this problem, either the annotated numbers extracted by the OCR are used, ignoring the computational results, or if there are no such annotations, the computed proportions are rounded to integers

and will be presented to the user as "approximate" values.

4. RESULTS

The analysis component of the IGR was successfully tested on a number of simple images each containing one chart of a certain type. The charts were created by a standard software. Since an OCR trigger was not yet added to the IGR, the current test images were reduced to their mere graphical information. The tool in its current state of development can identify geometrical objects like rectangles, arcs, triangles and axes from the contours and junctions. The geometrical objects and their combinations are classified as coordinate systems and information bearing items, like pie segments and bars. Moreover do the combinations of these items yield an identification of the chart's type and the approximated data in it. The tabular overview of the 30 tested bar chart images is given in tab.1. We have tested bar charts of different values, colors, scalings and proportions. Tab.1 shows the percental amount of detected items in the images. A detected bar consists of a rectangle with the proper dimensions and start position on an axis. A detected axis consists of a scaled vertical or horizontal line having the proper scale.

30 bar chart images	
Bars	97 %
Axes	73 %
Type	73 %

Table 1. Detected bar chart elements

The test set for the 25 pie chart images show a slightly different result (tab.2). The charts consisted of different values, explosion characteristics annotations and legends.

25 pie chart images	
Correct and complete	32 %
Correct segments	47 %
Type	93 %

Table 2. Detected pie chart elements

A correct and complete identified pie chart consists of the correct amount of detected segments having the correct percentage. Throughout the 25 images there were 175 individual pie segments of which 47 % were detected properly. With only 2 type misclassifications, the pie charts are better detected than the bar charts, but the quality of the values is less sufficient.

All test runs were performed on a 500MHz PC with the IGR prototype written in Java 1.3. The IGR was also

demonstrated on the RehaCare fair in Düsseldorf/Germany, October 2000.

5. OUTLOOK

A future task for the IGR project is to enlarge its domain. The images and the presented structures will become more complex when technical drawings and schematics are targeted. Editing and reading UML diagrams or flow charts to blind people is another task we are currently working on.

A further idea, detached from the concentration on a single user group, is the adaption of the IGR concept to a more general information retrieval process. Today's tools for summarizing information from multiple sources still cannot use contextual media [17]. An integration of image analysis and identification of the semantics opens a complete new prospect to data mining.

6. REFERENCES

- [1] Catherine Laws, Chieko Asakawa, "IBM Home Page Reader: The Voice of the World Wide Web," in *CSUN's 1999 Conference Proceedings*, 1999.
- [2] "<http://www.dolphinusa.com>,".
- [3] "<http://lynx.isc.org>,".
- [4] P. Blenkhorn, D.G. Evans, "Using speech and touch to enable blind people to access schematic diagrams," *Journal of Network and Computer Applications*, vol. 21, pp. 17–29, 1998.
- [5] P. Alshuth, Th. Hermes, L. Voigt, and O. Herzog, "On Video Retrieval: Content Analysis by ImageMiner," in *IS&T/SPIE Symposium on Electronical Imaging Science & Technology (Storage and Retrieval for Images and Video Databases)*, San Jose, CA, January 1998, vol. 3312, pp. 236–247.
- [6] Gonzalez, Woods, *Digital Image Processing*, Addison-Wesley, 1993.
- [7] Pavlidis, Horowitz, "Segmentation of Plane Curves," *IEEE Trans. on Comp.*, vol. C-23, pp. 860–870, 1974.
- [8] Thorsten Hermes, *Texturen: Analyse, Beschreibung und Synthese*, Ph.D. thesis, University of Bremen, FB3, 1999.
- [9] Ballard, Brown, *Computer Vision*, Prentice-Hall, 1982.
- [10] F. Deschênes, D. Ziou, "Detection of line junctions and line terminations using curvilinear features," *Pattern Recognition*, vol. 21, pp. 637–649, 2000.
- [11] G. Rozenberg, *Handbook of Graph Grammars and Computing by Graph Transformation. Vol I: Foundations*, World Scientific, 1997.
- [12] Wolfgang Bibel, *Wissensrepräsentation und Inferenz*, Vieweg, 1993.
- [13] Booch, Rumbaugh, Jacobson, *The Unified Modeling Language User Guide*, Addison Wesley, 1998.
- [14] Karl P. Dürre, Karl W. Glander, "Design considerations for microcomputer based applications for the blind," in *Human Jobs and Computer Interfaces, proceedings of the IFIP, Tampere, Finland*, 1991.
- [15] Ben Shneiderman, *Designing the User Interface*, Addison-Wesley, 1998.
- [16] "<http://www.ivb.ch/Hilfsmittel.html>,".
- [17] Udo Hahn, Inderjeet Mani, "The challenges of automatic summarization," *Computer, Membership Magazine of the IEEE C.S.*, vol. 33, no. 11, pp. 29–36, 2000.