Customizing Graphics for Tiny Displays of Mobile Devices

Thomas Rist, Patrick Brandmeier
DFKI Saarbrücken, Germany. Email: {rist, brandmei}@dfki.de

Abstract

Advances in mobile devices and wireless telecommunication infrastructure already provide mobile users with access to online information sources and services. Compared to the PC-world, however, mobile access is still quite restricted especially with regard to the display of graphical representations, such as images, drawings, diagrams, maps, and logos. Since graphical representations are increasingly used in the world-wide-web for the purpose of information presentation the adaptation of graphics for tiny displays is a challenge that should not be neglected. The current contribution discusses several transformation approaches which might be employed to accomplish this adaptation task.

1. Introduction

During the last few years, we have seen a variety of mobile appliances such as mobile phones, micro PDA's, and also first working prototypes of the next generation's wrist watches that - in addition to their original functionality - provide wireless access to the Internet and the World-Wide-Web. While it is debatable whether web-browsing is amongst the most useful applications for the users of such mobile devices, there is no doubt that these appliances provide a high potential for a broad range of new information services that can accommodate for the specific needs of mobile users. For example, think of the commuter who whishes to get the newest travel info, whereas the stock jobber may wish to inspect the development of shares and perform transactions while being on the move. It is our believe that many of these information services should be brought to the user by means of clever combinations of written text, voice and sound, and graphics.

With a focus on information presentation, DFKI has explored automated adaptation mechanisms to serve mobile users in a number of different areas of application. This contribution briefly reports on the objectives and achievements of this work, and finally sketches some ideas towards future research.

2. WAP-Access to Information Sources and Services on the WWW

For the time being, the so-called Wireless Application Protocol (WAP) together with its page description language WML (Wireless Markup Language) are part of the enabling technology for mobile access to the Internet – at least in Europe. Thus, accessing a WWW service through a WAP-enabled device requires some kind of "www2wap" transformation.

Currently, several attempts are being made to develop transformation mechanisms that take as input arbitrary information sources (e.g. html pages) and deliver information presentations that can be displayed on mobile devices with limited display capabilities, such as lack of screen real estate and lack of colors. In case of a textual source, a straightforward approach is to fragment the text into yet displayable chunks. However, this strategy can easily result in huge stacks of WML pages which are too unpractical to read on a mobile device. Other approaches [1, 2] provide filtering and parsing technology that enables a user to specify which textual contents of a web page shall be extracted and converted into a format displayable on a mobile device.

In the case of visual media, such as graphics, animation and video, however, partitioning is often not possible at all, and even the transformation of static graphics is yet an almost untouched challenge. Neglecting graphical representations completely may be acceptable if these representations serve as decorations only. However, there are many examples of services which can hardly do without using graphical representations to encode information. For the purpose of illustration, we refer to DFKI's AiA Travel Agent [3] and to DFKI's Personal Picture Finder [4].

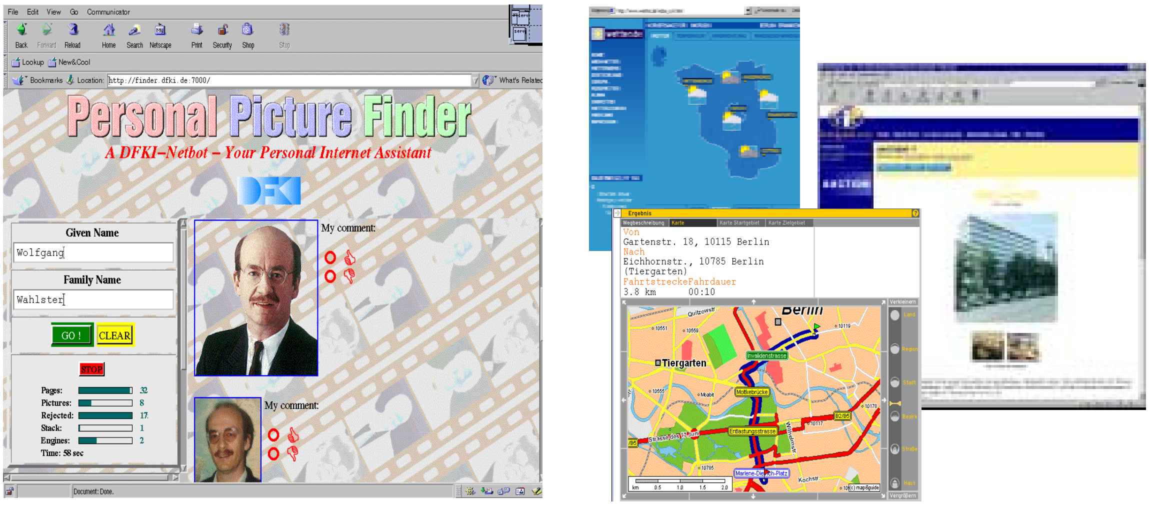


Figure 1. Left: WWW interface of DFKI's Personal Picture Finder. For a given name the system searches the web for a picture of the person. Right: A typical collection of web pages (weather, hotel information and route descriptions) consulted by the AiA travel agent(s) in response to a user query.

The travel agent assists a user in preparing for a journey. For a given travel destination a swarm of agents gather information that might be of interest for the travel. The resulting pool of gathered information typically contains local weather news/forecasts, recommendations where to stay and where to eat, and an overview of local cultural events and entertainments. In a number of cases, the discovery of certain bits of information will trigger further agents to acquire complementary bits. For instance, if a hotel or a restaurant is found relevant, a dedicated "map agent" is triggered to search for a map showing how to get to the place. The right hand side of Figure 1 shows a typical collection of different pieces of textual and graphical material retrieved by the information agents in response to a user query. The Personal Picture Finder is a service that takes as input the name of a person and searches the web for a portrait picture of the person (cf. left part of Fig.1). Such a service can be especially useful when preparing for a meeting with persons for the first time, or when one has to collect someone not met face-to-face before at the airport or the train station.

While in their original versions both services have been designed for PC users who are assumed to interface with the services via an ordinary WWW browser, it is quite obvious that users on the move could greatly benefit from these services as well if they could access the services by mobile devices, such as WAP-enabled mobile phones. At DFKI we started to tackle the problem of how to transform graphical representations so that they can be displayed on very small displays, such as a 100*60 pixel display of a mobile phone. In particular, we are currently investigating different approaches to solve the graphical transformation problem: *uniformed transformations*, *informed transformations*, and *re-generation of graphics*.

3. Image Transformations

In the context of this paper an image transformation is an operation that can be applied to a source graphics[1] in order to obtain a target graphics that meets the particular requirements of the target display screen. Transformation may be complex in the sense that they can be described as sequences or configurations of several less complex operators that all contribute to the result of the total

[1] The term graphics is here used in a broad sense covering all kind of graphical representations that can be displayed on a 2D screen.

transformation process. Targeted towards an adaptation of graphics for tiny displays an image transformation usually comprises a down-scaling operation and a reduction of colors. Also, there may be transformations that take into account a number of additional parameters (e.g., height and width of the target graphics) to adjust the transformation process.

3.1 Uniformed Transformations

A transformation can be uninformed (or blind) in so far, that only little information about the source graphics is taken into account when selecting and adjusting the transformation. Unfortunately, it is very difficult to find a general-purpose transformation that reliably produces suitable results for the large variety of graphics found on the WWW. Figure 2 illustrates the problem. While the applied transformation produces an acceptable result for the source graphics in the first case, the result is less acceptable when applying the same transformation to the source graphics in the second case.

Figure 2. Applying the same transformation T to two different source pictures in order to obtain a small black and white target picture that can be displayed on a mobile device.

3.2 Informed Transformations

A more promising approach starts with an analysis of the source graphics in order to inform the selection and adjustment of transformation parameters.

Basically the analysis phase performs a classification of the source amongst syntactic or even semantic features. For instance, in our current work, the set of implemented semantic classifiers comprises classifiers that distinguish between portrait and non-portrait images, outdoor versus indoor images, outdoor images that show a scene with blue sky, clouds, sunset, water, forest or meadows, and snow-covered landscapes. In the ideal case, each image class can be associated with a certain transformation that produces acceptable results for the vast majority of instances of that class.

For the purpose of illustration let's assume a classifier *strong-inner-outer-contrast(g)* $\rightarrow b \in$ *{true, false}* performing a region-based analysis of pixel brightness in the source graphics according to the following simple heuristics: The pixels of a graphics are partitioned into pixels belong to a "inner-region" versus pixels belonging to the "outer-region" which is just the complement of the former. Then an average brightness value is computed for each of the two regions as the sum of the brightness values of pixels divided by the number of pixels per region. In case the difference between the average pixel brightness of the inner-region and the average brightness of the outer region exceeds a certain threshold, the classifier yields the value true whereas the value false will be returned in all other cases. The underlying rational of this heuristic is to have a means to identify images with a strong figure background contrast.

Let' further assume the availability of the two different transformations T1 and T2 configured as follows:

T1: GIF x WBMP with

T1 transforms a source graphics stored in the GIF-format[2] into a target graphics stored in the WBMP-format[3] through the following sequence of operators: [(i) scaling, (ii) colour-reduction, (iii) WBMP-conversion];

T2: GIF x WBMP with

T1 transforms a source graphics stored in the GIF-format into a target graphics stored in the WBMP-format through the following sequence of operators: [(i) determine the set of background pixels (ii) set the colour value of all background pixels to either white or black in order to maximise the contrast between inner- and outer-region, (iii) scaling, (iv) colour-reduction, (v) WBMP-conversion];

Given a source image we can now use the classifier *strong-inner-outer-contrast* introduced above to decide whether to apply transformation T1 or T2. As illustrated in Figure 3, applying T2 yields a better result than applying T1 to the same source image. However, in cases where no significant figure ground contrast is found, the application of the computationally more expensive transformation T2 is not motivated and therefore T1 may be used as a fall-back solution.

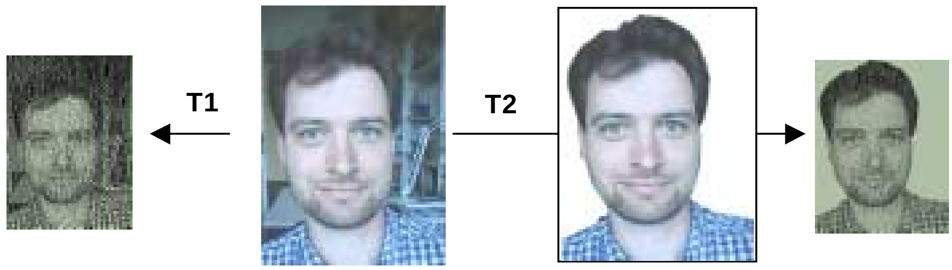


Figure 3: Applying T1 and T2 to a portrait image of a person (2nd from left) with a strong figure-background contrast. The result of T1 (1st from left) appears worse than the result of T2 (4th from right). Part of T2 is a background masking operation to enhance further the figure background contrast (3rd from left) before scaling, colour reduction and wbmp-conversion is performed.

3.3 Using Machine Learning Techniques to Select Appropriate Transformations

In the previous section we have sketched how an image classifier can be used in order to choose among different available transformations. However, it is still difficult to make an assignment between recognized features of an image on the one hand, and available transformations and their parameter adjustments on the other hand. We are currently investigating in how far this problem can be solved by deploying machine learning techniques. That is, in a training phase, a graphics design expert manually assigns images to transformations and thereby allow the system to recognize and generalize correspondences between image features and transformation parameters. In our current test setting we use some 430 features to characterize images. In contrast, our repertoire of transformations is yet quite small. By means of a software package for machine learning [5] we trained the system with a set of 130 images. So far, the achieved scores for adequate selections of transformations are in the range of 60-70 %. While this result is encouraging, further refinements of the approach are required.

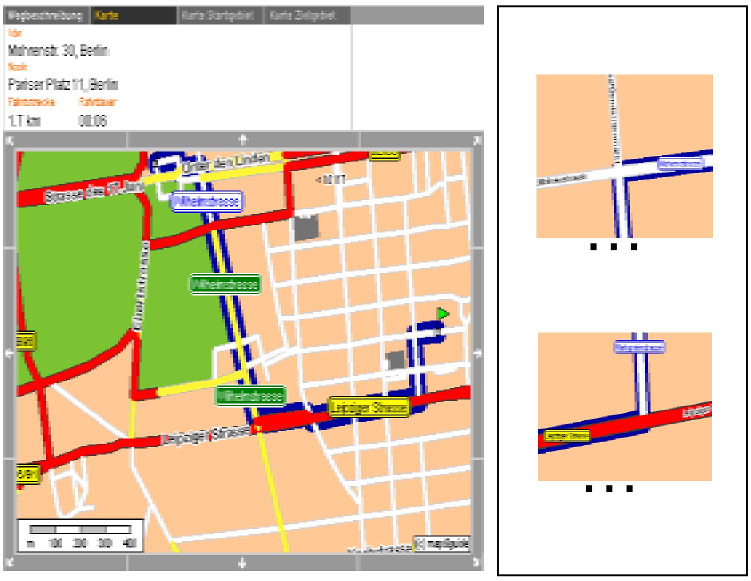
---

[2] GIF (CompuServe Graphics Interchange Format file), a highly compressed format for colour-mapped images with less than 8 bit colour. GIF is commonly used to upload images on webpages.

[3] WBMP (Wireless BitMaP), an uncompressed format for B/W bitmaps that can be included in WML pages. The term “WBMP conversion” is used to refer to the transformation of a given image format into the WBMP format.

4. Re-generation of Graphics

The third approach, *re-generation of graphics*, does not aim at a modification of a source graphics at the picture level. Rather, the idea is to generate a new picture from a content description obtained from a deeper semantic analysis of the source graphics. We illustrate this approach by a transformation of a route description as it may be obtained from a routing service available on the WWW (cf. left frame of Figure 4).




```
<node
  <destination>
  <left-landmark>
    <landmark type="Attraction“>
      <label Leipzigerstr</label>
      <location>
        <x-coord 704.0</x-coord>
        <y-coord324.0</y-coord>
      </location>
    </landmark>
  </left-landmark>
&lt;/segment&gt;
&lt;segment direction=“West"&gt;
  <label Wilhelm Str></label>
    <destination>
        <node>
          <location>
              <x-coordinate>690.0
            </x-coordinate>
         ...............
```


Figure 4. Analyzing a map in order to construct a symbolic description (here in XML format) of the core contents (here a route description).

Once a semantic representation of the picture content has been extracted, it is possible to generate a variety of new graphical representations which adopt different graphical styles in order to meet resource limitations of the output device and/or a user’s personal style preferences. Figure 5 presents several variants of graphical displays showing a routing maneuver.

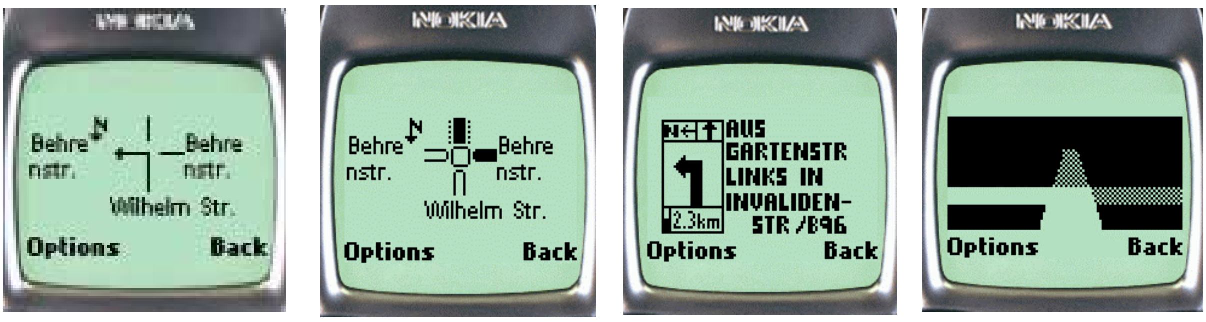


Figure 5. Design variants for the graphical display of a routing maneuver (“turn left on next crossing”) shown on a WAP-enabled mobile phone (Nokia 7110).

5. Summary and Future Work

Mobile access to information services needs to be customized in order to benefit the mobile users. In particular this task includes the customization of graphical representations. In this contribution we have sketched three different approaches for transforming graphical representations suitable for display on a small screen of a mobile device. Practical experience has shown that oftentimes reasonable results can only be achieved when different target images are processed by different transformations, too. In order to automate the selection of transformations and settings for the parameters of transformations we started to investigate in how far a correlation between image features and transformations can be learned using machine learning methods. In addition to image transformations, we also addressed the issue of graphics re-generation from a deep semantic representation that has been gained from an analysis of the source image. While such an approach can be implemented for a restricted application, a general solution is yet far beyond the state of the art.

In our ongoing research we also address multimodality as a further important aspect of future mobile interfaces. Especially the combination of voice input together with graphical selections as well as the combined display of graphical and audio output is very likely to become a predominant interaction paradigm for users of tiny mobile devices – eventually enabled by upcoming technologies like GPRS[4] and UMTS[5]. Envisaged scenarios include: (a) the user formulates information requests likewise: verbally, by a graphical selection, or by a combination of both; (b) the user can flexibly switch between several information display modes including the display of text and graphics, audio display, and effective combinations of multiple modes (see also [6]). While the concept of multimodal interfaces is not a new one [7], its instantiation for the case of users of mobile devices, however, reveals a number of new challenges to be dealt with including technical issues (e.g., how to perform robust speech analysis on a mobile device, how to flexibly translate contents given in one medium into another format), and usability issues (e.g., how to decide on which media combinations are most appropriate considering a mobile user's current task and situation). For instance, mobile users are often situational disabled in contrast to users sitting in front of a stationary PC. While walking or driving, users might prefer to receive requested information verbally instead of graphical displays. Consequently, the design of services and user interfaces must take into account the specific conditions that result from non-stationary system usage. Intelligent mechanisms for on-the-fly adaptation of input/output modalities offer a promising perspective for the support of mobile users who want to access information services or participate in multi-user applications.

References

[1] GoSMS.com Ltd., TIP Reference Manual. GoSMS.com Ltd, Tel-Aviv, Israel, Nov. 2000.

[2] Bergström, A., Jaksetic, P. and Nordin, P. Enhancing Information Retrieval by Automatic Acquisition of Textual Relations Using Genetic Programming. In: Proceedings of Intelligent User Interfaces (IUI) 2000, ACM Press, 2000.

[3] André, E., Rist, T. and Müller, J. (1999). Employing AI Methods to Control the Behavior of Animated Interface Agents. *Applied Artificial Intelligence* 13:415-448.

[4] Endres, C., Meyer, M., and Wahlster, W.: Personal Picture Finder: Ein Internet-Agent zur wissensbasierten Suche nach Personenphotos, Online'99 (in German). Service temporarily available: *http://finder.dfki.de:7000/*

[5] Witten, I.H. and Frank, E.: Data mining: Practical machine learning tools and techniques with Java implementations. Morgan Kaufmann, San Francisco. 2000. Software temporär verfügbar: http://www.cs.waikato.ac.nz/~ml/weka/

[6] Rist, T.: Towards Services that Enable Ubiquitous Access to Virtual Communication Spaces. In Proc. of UAHCI 2001, to appear.

[7] Maybury, M. and Wahlster, W. editors. 1998. Readings in Intelligent User Interfaces. Morgan Kaufmann Press.

[4] GPRS (General Packet Radio Service)
[5] UMTS (Universal Mobile Telecommunications System) cf. http://www.umts-forum.org/