Braille Vision Using Braille Display and Bio-inspired Camera

Roman Graf, Ross King and Ahmed Nabil Belbachir
AIT Austrian Institute of Technology, Vienna, Austria

Keywords: Disabled Users, Braille Display, Bio-inspired Camera, Computer Vision, Image Processing.

Abstract: This paper presents a system for Braille learning support using real-time panoramic views generated from the novel smart panorama camera 360SCAN. The system makes use of the modern image processing libraries and state-of-the-art features extraction and clustering methods. We compare the real-time frames recorded by the bio-inspired camera to the reference images in order to determine particular figures. One contribution of the proposed method is that image edges can be transformed to the presentation on Braille display directly without any image processing. It is possible due to the bio-inspired construction of camera sensor. Another contribution is that our approach provides Braille users with images recorded from natural scenes. We conducted several experiments that verify the methods that demonstrate learning figures captured by the smart camera. Our goal is to process such images and present them on the Braille Display in a form appropriate for visually impaired people. All evaluations were performed in the natural environment with ambient illumination of 200 lux, which demonstrates high camera reliability in difficult light conditions. The system can be optimized by applying additional filters and features algorithms and by decreasing the rotational speed of the camera. The presented Braille learning support system is a building block for a rich and qualitative educational system for the efficient information transfer focused on visually impaired people.

1 INTRODUCTION

Though visually impaired people have quite good developed computer interfaces for Braille code (Jiménez et al., 2009) they have very limited access to image information. Smart cameras provide application-specific information out of a raw image or video stream. Real-time smart camera operation areas can be extended for Braille education task. Until now, reasoning has not been applied to smart camera output. Currently the verification and analysis of camera output are usually carried out manually and are not used for further analysis or processing. The rich information from images is not evaluated and there are no automated methods to estimate the content of current natural scene, to detect objects observed during camera use, or to validate the camera output. There is a need for digital preservation methods to handle such information and make it useful for visually impaired people.

In this work, we described a system for the automated Braille educational support with a specific smart camera (the AIT 360SCAN presented in Figure 1) that could be used to enhance the accessibility to graphical information for visually impaired people and efficiency of its understanding and process-

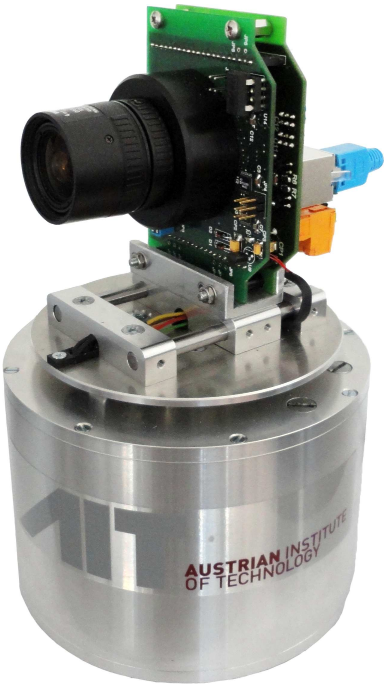

Figure 1: The bio-inspired panoramic smart camera 360SCAN.

Graf R., King R. and Belbachir A..
Braille Vision Using Braille Display and Bio-inspired Camera.
DOI: 10.5220/0004949302140219
In *Proceedings of the 6th International Conference on Computer Supported Education* (CSEDU-2014), pages 214-219
ISBN: 978-989-758-022-2

ing. The paper (Belbachir et al., 2012) provides a detailed description of the newly designed smart camera 360SCAN for real-time panoramic views. The main contribution described in this paper is a real-time image analysis approach focused on needs of visually impaired people with algorithms that provide automatic information extraction and support methods for graphical information processing and presentation. In order to implement image processing for natural scenes we apply duplicate detection and object recognition to recorded images of smart camera output. For example, our system carries out an initial image processing by object detection and searching for duplicates against a set of reference images like triangle or circle before making presentation of the detected form in a Braille appropriate format using dot matrix.

The paper is structured as follows: The next section gives an overview of related work and concepts. In the third section we describe the Braille education system concepts and our algorithm exposed in a workflow. In the fourth section experimental setup and evaluation results are presented. In the last section concluding remarks and the outlook on planned future work is given.

2 RELATED WORK

Educational devices for Braille users currently are limited by several Braille displays due to very high price of Braille display. One of the graphical Braille displays that is in production is described in (Matschulat, 1999). The display area consists of a matrix 16 to 16 dots with 1mm distance between taxels. The advantage of this piezoelectric VideoTIM3 device is its high speed (24 frames per second), strong stroke, robustness and ability to demonstrate different shapes even if its resolution is not very high. The device consists of a hand-held conventional video camera and the main unit with the tactile display. But this technology is focused on document reading with your fingers and is not appropriate for another kind of images like natural scenes. This device is pretty large, heavy and expensive and such technique requires training.

In HyperBraille project was developed a graphics-enabled display for blind computer users (Prescher et al., 2010), (Erp et al., 2010). This display is designed to increase the amount of information percievable to blind computer user through both hands and enables graphical information to them. Besides high cost of this device it requires development of special software and focuses of standard Office and Internet applications.

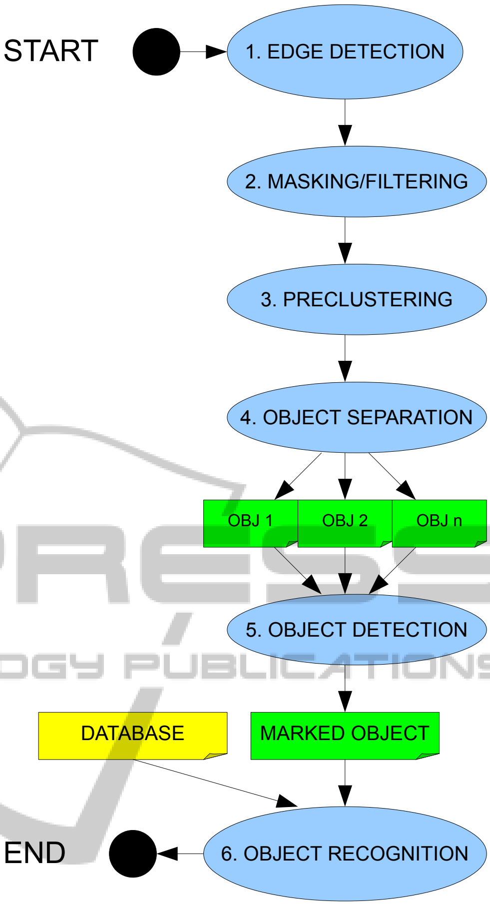

Figure 2: A workflow for object detection from recorded image.

The developers (Zeng and Weber, 2011), (Zeng and Weber, 2010) of project Tangram make use of HyperBraille device for graphical presentations. They developed education concepts and standardised document formats and use device extended by audio information. Therefore the main focus of their research is in creation of standards, educational concepts and software for existing device.

The advantage of our approach is that images from natural scenes are produced by a similar to human eye low cost smart camera with low data rate and high dynamic range. The output of this camera can be directly presented by the small Braille module in form of dot matrix but also can be processed, analysed (see Figure 2) an presented as required by Braille user.

A number of approaches deal with object detection from images and video. Many of these approaches are limited to text detection. For example, the efficient algorithm proposed in (Epshtein et al., 2010) uses the idea of detecting the width of character strokes. This method is very efficient for text

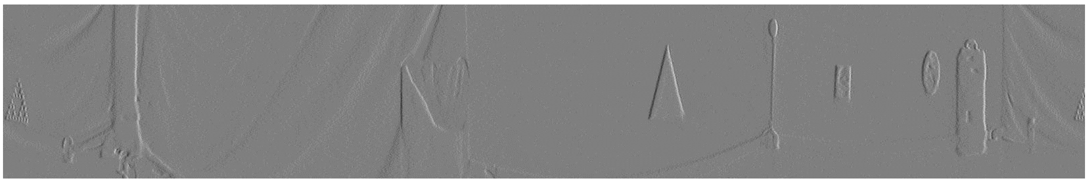

Figure 3: Event map generated by 360SCAN.

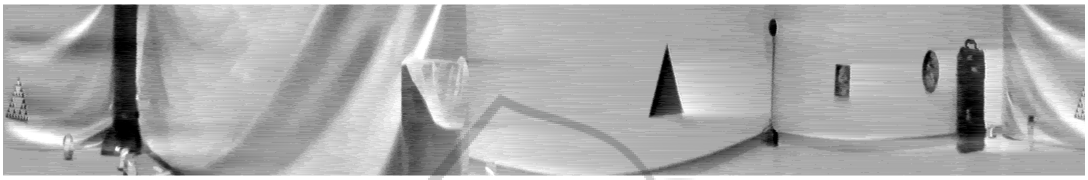

Figure 4: Reconstructed grayscale image from the event map.

Figure 5: Detected objects.

recognition but do not solve the problem of duplicate search. A multiresolution approach (Cinque et al., 1998) for page segmentation is able to recognize text and graphics by applying a set of feature maps available in different resolution levels. This method breaks down an image into several blocks that represent text, line-drawings and pictures. However, too much additional analysis effort is required to compare the shapes provided by our smart camera. The AIT *matchbox* tool (Huber-Mörk and Schindler, 2012), developed for the SCAPE project[1], is based on the OpenCV library and implements image comparison for digitized text documents. The similarity computation task is based on the SIFT (Lowe, 2004) feature extraction method. This tool demonstrates high accuracy, good performance and provides duplicate search and comparison of digitized collection but is limited for text in images and is too slow to be applied to the smart camera domain.

In order to meet the requirements of quality control for our smart camera, we use image processing application based on the ImageMagick (Salehi, 2006) tool and PSNR metric. The application compares the camera output (see Figure 3-4) to a reference image collection in order to detect objects (see Figure 5) and to analyse the surrounding environment. This tool extends the functionality of *matchbox* for the domain of smart cameras with the ability to analyse images and video frames including segmentation. This application could be reused for the domain of digital preservation for Braille domain e.g. detection of shapes, forms, text blocks and for similar tasks.

3 BRAILLE EDUCATION SYSTEM CONCEPTS

With increasing volumes of graphical data produced by smart cameras, data analysis plays an increasingly important role. The figures of interest should be detected. In order to carry out this detection, we conduct an image analysis according the object detection workflow shown in Figure 2. Our hypothesis is that we can detect figures of interest applying image processing algorithms. We assume that the reference images stored in database will have significant similarity with objects extracted from recorded image. In order to meet these expectations, we analyze images for duplicates using the PSNR metrics and applying a workflow for duplicate search to ensure the correct object detection. In database we store figures of interest for particular educational task.

We started workflow with edge detection using event map 3 created by camera. In the second step we employ filtering for required pixels or masking if required for the reconstructed greyscale image 4. Then we applied pre-clustering employing parameter like pixel distance and cluster distance. The object separation occurred using clustering with parameters like

---

[1] http://www.scape-project.eu/

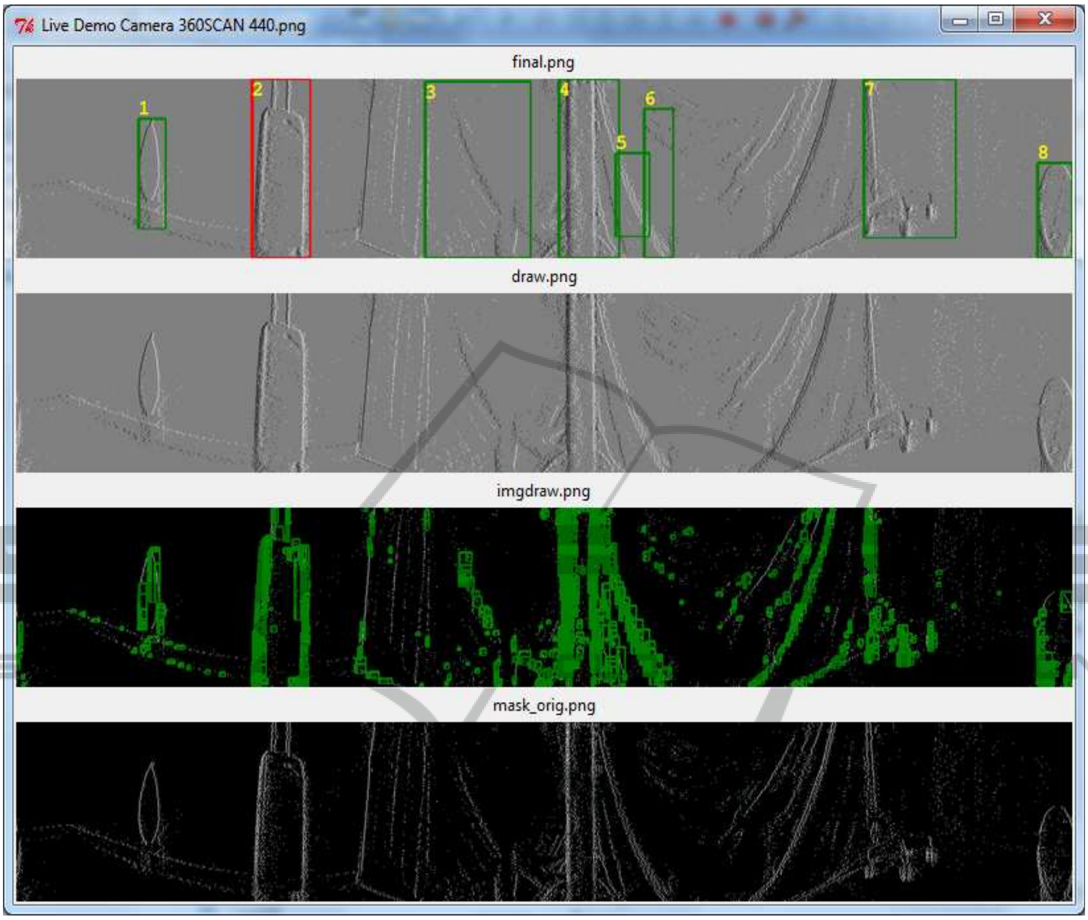

Figure 6: The bag recognition sample.

pixel distance, cluster distance, edge density, evaluation steps on X and Y axis, min cluster size, cluster dimension filter. The output of object detection step are objects marked by green rectangle 5. We alse depicted object number by yellow colour. Among the detected objects we searched for objects of interest and performed object recognition applying image comparison. E.g. PSNR metric and ImageMagick.

4 EVALUATION OF THE 360SCAN IMAGE PROCESSING

A sample application of our approach for bag detection is depicted in 6. The algorithm for that was written in Python 2.7 and given figure demonstrates output of the program.

In this case according to the above described workflow we masked the original raw image (draw.png). In mask orig.png white dots demonstrate dark areas of the raw image and black colour shows hell areas in the raw image. Then in imgdraw.png we filter and cluster detected pixels in order to separate possible objects. Finally in final.png we mark detected objects with green rectangles, depict object number by yellow number and recognized objects by red rectangles. Therefore found bag is marked by red rectangle and has number 2 in this case.

A vertical resolution of 360SCAN camera is 256 pixel. Therefore for 16 to 16 pin-matrix we should compress raw information to required size applying Equation 1.

$$\begin{aligned} G_{n,m} &= \{a_n, b_m\}, n \in \{0,1,\ldots,15\}, m \in \{0,1,\ldots,15\}, \\ p &= 16, \qquad (1) \\ a_n &= \frac{1}{p}\sum_{k=0}^{p-1} x_{pn+k}, \\ b_m &= \frac{1}{p}\sum_{k=0}^{p-1} y_{pm+k}. \end{aligned}$$

Where $G_{n,m}$ represents the graphical Braille dots grid computed over the dimension 16 x 16 dots. This functions are dependent from camera settings. E.g. for evaluation set vertical resolution of the camera was set to 256 pixels and vertical resolution was 1600 pixels. The n and m represent the dot number for X and Y axis. A dot value for pixels around a grid point is computed as an average value over all evaluated pixels which are located in acceptable distance to the grid point. For one grid point in our evaluation case we regard $p = 16$ neighbouring values $p/2$ to the left and to the right $p/2$ from the current value for X and Y axes. The set $G_{n,m}$ is dependent on a_n and b_m functions.

Equation Equation 2 calculates dot values.

$$
\begin{aligned}
S_{n,m} &= \sum_{n,m} \delta_{n,m} \cdot d(P_{n,m}, G_{n,m}) \qquad (2)\\
\delta_{n,m} &= \begin{cases} 1 & \text{if } d(P_{n,m}, G_{n,m}) < d_{max}, \\ 0 & \text{else.} \end{cases}\\
d(P_{n,m}, G_{n,m}) &= \sqrt{(G_x - P_x)^2 + (G_y - P_y)^2}.
\end{aligned}
$$

Where $S_{n,m}$ represents the value of matching detected pixels $P_{n,m}$ with coordinates P_x and P_y computed over the dimensions n and m around correspondent grid point $G_{n,m}$ with coordinates G_x and G_y. d represents the distance between a grid point and an evaluated current pixel point coordinates. d_{max} is a threshold value for decision between black (0) and white (255) colour e.g. 127. $\delta_{n,m}$ is a coefficient with value 1 for pixel range $< d_{max}$ or 0 otherwise.

Therefore presentation appropriated for Braille user can be calculated using this method and is shown in Figure 7 for triangle and in Figure 8 for circle.

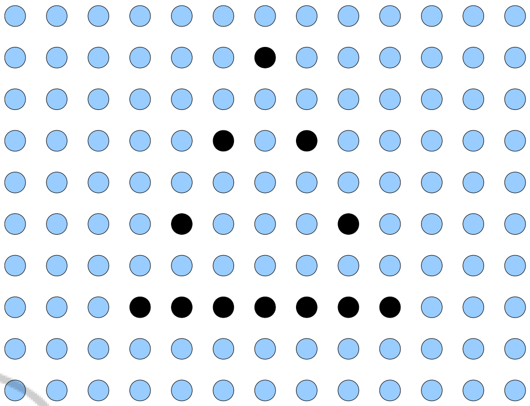

Figure 7: A triangle demonstration for Braille user.

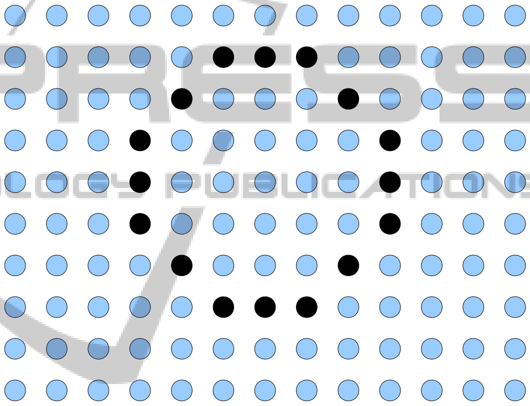

Figure 8: A circle demonstration for Braille user.

5 CONCLUSIONS

We have presented a system for Braille learning support using the smart camera 360SCAN. Learning support for visually impaired people is required for specific information presentation, which in turn relies on standard Braille output device and automatically processing of bio-inspired smart camera images in a natural environment. In order to provide education support, we apply image processing on recorded images and object recognition to the smart camera output. The edge detection methods enhances object recognition by providing accurate object shapes of real-live objects and makes this system unique for object presentation to visually impaired people.

The main contribution of described work is a system including a smart camera, a Braille module and an image processing algorithm implementation that provides methods for shape detection of real-live figures and supports automatic information presentation. Another contribution of the proposed method is that image edges can be transformed to the presentation on Braille display also directly without any image processing. It is possible due to bio-inspired construction of camera sensor. Additional contribution is that our approach provides Braille users with images recorded from natural scenes. This approach makes use of the modern image processing libraries. We employ state-of-the-art features extraction and clustering methods. We conducted several experiments that evaluate the methods we have implemented and demonstrate learning figures captured by the smart camera presented in form of taxels. The experimental evaluation presented in this paper demonstrates the effectiveness of employing the image processing techniques for an education system.

As future work we plan to produce our own Braille display in order to use it as an output device for 360SCAN camera. We plan to extend an automatic education approach of image analysis to new application scenarios also involving information storage and digital preservation. The educational use cases could be extended with specific software for visually impaired people in order to give a method at hand to work with figures and images.

REFERENCES

Belbachir, A., Mayerhofer, M., Matolin, D., and Colineau, J. (2012). Real-time 360 panoramic views using bica360, the fast rotating dynamic vision sensor to up to 10 rotations per sec. In *Proceedings of the Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, pages 727–730.

Cinque, L., Lombardi, L., and Manzini, G. (1998). A multiresolution approach for page segmentation. *Pattern Recognition Letters*, 19(2):217 – 225.

Epshtein, B., Ofek, E., and Wexler, Y. (2010). Detecting text in natural scenes with stroke width transform. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2963 –2970.

Erp, J., Kyung, K.-U., Kassner, S., Carter, J., Brewster, S., Weber, G., and Andrew, I. (2010). Setting the standards for haptic and tactile interactions: Isos work. In Kappers, A., Erp, J., Bergmann Tiest, W., and Helm, F., editors, *Haptics: Generating and Perceiving Tangible Sensations*, volume 6192 of *Lecture Notes in Computer Science*, pages 353–358. Springer Berlin Heidelberg.

Huber-Mörk, R. and Schindler, A. (2012). Quality assurance for document image collections in digital preservation. In *Proc. of the 14th Intl. Conf. on ACIVS (ACIVS 2012)*, LNCS, Brno, Czech Republic. Springer.

Jiménez, J., Olea, J., Torres, J., Alonso, I., Harder, D., and Fischer, K. (2009). Biography of louis braille and invention of the braille alphabet. In *Survey of Ophthalmology*, pages 142–149.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. of Comput. Vision*, 60(2):91–110.

Matschulat, G. D. (1999). Tactile reading device. Number EP0911784.

Prescher, D., Weber, G., and Spindler, M. (2010). A tactile windowing system for blind users. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '10, pages 91–98, New York, NY, USA. ACM.

Salehi, S. (2006). Imagemagick tricks, web image effects from the command line and php. page 226, Olton, Birmingham, B27 6PA, UK. Packt Publishing Ltd.

Zeng, L. and Weber, G. (2010). Collaborative accessibility approach in mobile navigation system for visually impaired. In *Virtuelle Enterprises, Communities and Social Networks, Workshop GeNeMe '10, TU Dresden 07./08.10.2010*, pages 183–192.

Zeng, L. and Weber, G. (2011). Accessible maps for the visually impaired. In *Proc. ADDW 2011, CEUR*, volume 792, pages 54–60.