INSTANT MEDICAL PILL RECOGNITION ON MOBILE PHONES

Andreas Hartl, Clemens Arth, Dieter Schmalstieg
Institute for Computer Graphics and Vision
Graz University of Technology, Austria
{hartl,arth,schmalstieg}@icg.tugraz.at

ABSTRACT
Smartphones have attracted a lot of interest recently as they are becoming more and more computationally powerful platforms for all kinds of tasks in mobile environments. Due to the availability of high quality mobile phone cameras, the use of these platforms even for critical applications leveraging Computer Vision techniques is justified.
In this paper we present an application of Computer Vision techniques for instant and interactive medical pill detection and recognition on mobile phones. Our approach is designed to help trained or untrained personnel outside of medical facilities to perform the task of identifying tablets interactively with or without the help of online resources. By describing our application in detail and giving results on the recognition task we are proofing the plausibility of our approach.

KEY WORDS
Object Recognition, Applications, Mobile Computer Vision

1 Introduction

Medical pill recognition is a tough task for trained, as well as for untrained personnel due to the large number of different medical compounds available and the huge variety in color, shape and size. At the moment, several thousand different medical pills are available worldwide (see the left image in Figure 1 for a snapshot of pills used in our experimental setup). Unfortunately, there is no global standard concerning the properties of medical pills. Although largely assumed, most medical tablets do not have their IDs printed or embossed on their surface. Even if the ID information was generally available, it would still be hard to recognize the compound directly since the identifier would most likely be a logo or an abbreviation. Furthermore, extracting textual information in embossed or printed form from arbitrary surfaces is far from being a classical OCR task. Thus, deploying a generally applicable method for automatic extraction of this information is hardly possible.

From a practical point of view, searching for information about medical compounds in catalogs and medical books is tedious and can only be done if the adequate literature is available. However, in non-clinical environments this is hardly the case, thus we have to rely on alternative methods to gather the information required. Internet databases, together with powerful mobile phone hardware can be used in solving this task in an efficient and user-friendly manner.

One may ask why it is important to solve the task of medical pill identification on a mobile device at all. There are certain scenarios where having a mobile and interactive pill recognition application available is advantageous and, under special circumstances, can save lives. On the one hand, one might consider the scenario of differentiating medical tablets from custom drugs in police inspections, which might influence the decision of releasing or arresting somebody. On the other hand, for paramedics discovering an unconscious person it is pivotal to acquire as much and as precise information about the patient and his potential drug abuse as fast as possible.

Figure 1. A subset of medical pills used in our experiments is depicted on the left. Two smartphones used for our evaluation are depicted on the right, the *Asus m530w* (left) and the *Samsung Omnia Pro* (right).

The main contribution of this work is the presentation of a medical pill recognition application running on modern smartphone hardware. The application enables the user to perform tablet recognition on site instantly and interactively. Additionally, note that all the algorithms used in our approach are not limited to the special object class treated. Rather our methods are also applicable to a wider range of tasks in other areas like entertainment or Augmented Reality (AR), leveraging the same custom modifications for mobile phones that allow them to perform near real-time. Furthermore, compared to earlier evaluations of our system (see [7] for details), we are employing a novel probabilistic query strategy for candidate retrieval, which improves results significantly.

In the following, we first give an overview of our application in Section 2. Then we describe the individual methods involved and give details about the individual

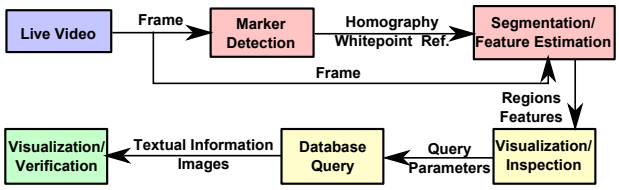

Figure 2. Our high-level application structure. The entire application runs on a mobile phone, using it as image acquisition device and for all Computer Vision processing tasks. After querying an online or onboard database, the results are visualized to the user and can be further analyzed.

adaptations for our special application in Sections 3 and 4. Experimental results are given in Section 5 and concluding remarks are given in Section 6.

2 Application Overview

For the design of any application it is important to consider as many factors involved as possible and condense a clear list of features, requirements and preconditions. In our case this methodology influenced the entire application design as following.

As for any other Computer Vision (CV) based application, the image acquisition conditions are critical and might influence the robustness and overall usability of the entire application. Thus, it is necessary to enforce as much control over this process as possible and reasonable. Given control over this factor, it is further important to keep the entire computational process economical in terms of energy and overall time consumption. For an application this means that the user should not get annoyed by waiting for an actual result for a long time. This implicates that any algorithms have to be tailored towards the actual platform used. Note that the quality of the method still has to be reasonable to fulfill the entire task as diligently as possible. In our case we have to additionally consider the availability of certain external resources required, such as network connectivity, memory resources available onboard or maintainability. Finally, the presentation of the results has to be in a straightforward and intuitive way, giving the user the possibility to interact with the preliminary outcome of the methods and, if necessary, refine it. For reasons of safety the final decision on a medical matter is left to the user. This means that the system can give useful assistance in the process but is not meant to entirely replace human judgement of the situation.

The overall system design was developed according to these considerations and is depicted in Figure 2. All CV tasks are solved right on the device to keep the application as autonomous and responsive as possible. After image acquisition, the medical pills are segmented and characteristic features are extracted (see Sections 3 and 4 for details). The user is provided with an initial result of the feature extraction step, so that the temporary results can be verified. If necessary, the preliminary results can be changed or additional information, such as textual information, can be added.

In our case we use the *Identa* database for identification, which is a freely available online database for medical compounds [9]. Although intended for queries using a web form, we were able to use it for prototyping through some custom middleware. By using a publicly available database, the tedious task of updating and maintaining all the necessary information for pill recognition is offloaded to the provider. However, it is still possible to acquire statistical data about the database content using a set of automated scripts. To make the application autonomous it is also possible to include the database content on a memory expansion card. After querying the database, the final recognition results are provided to the user in an interactive manner. The user is able to browse through the results, which mainly consist of a sample image contained in the database and a short summary of the manufacturer and the product name. Additional information about the actual chemical components and the medication can be provided on demand as well.

3 Robust Tablet Segmentation

Considerations For medical pills, the main problem is their high variability in any kind of feature used for identification. Pills might expose a wide range of different shapes and sizes. However, especially the fact that the coloring can be arbitrary is a big issue, making a wide range of segmentation approaches to fail for certain instances. This is a clear indication that the acquisition conditions have to incorporate some mechanism to allow segmentation irrespective of the object color.

Prior to implementing our own method, we evaluated a set of different popular algorithms for segmenting pills in front of a uniform white background using publicly available implementations: *Graph Cuts* [6], *MSER* [14], *Color Canny Edge Detector* [11] and *Mean Shift* [4]. For all algorithms, the main problem is to deliver accurate boundaries for the pill to be segmented. Especially for pills colored similar to the background, segmentation results are quite poor. A second criterion is the usability of the approaches on mobile phone hardware. All algorithms took one or several seconds to deliver a result on a standard desktop computer when processing realistic examples. However, considering the computational resources available on modern smartphones, the use of those algorithms is prohibitive.

In Figure 3 our actual solution for robust segmentation is depicted. By using a marker target with known dimensions and a checkerboard background, almost arbitrarily colored and shaped tablets can be segmented. The choice of a marker target seems to be prohibitive at first, however, this technique has a considerable number of advantages. Firstly, markers can be detected efficiently and robustly using well-known techniques from the field of AR

(see *e.g.* [10]). Secondly, the marker can encode its own geometry. This means that we are able to acquire reference measurements from the real world, from which we can infer the real dimensions of the tablets. Lastly, the target itself can be created in any kind of format, either as a plastic credit-card sized item or as a simple printout. Thus, the choice of such a target pays off immediately and only slightly harms the application handiness.

Implementation For the target detection step, we used an efficient software library for mobile phones called *Studierstube ES*[1]. As outcome we know the exact location of the target and the checkerboard area in the image. By using a homography H we can theoretically undistort the view of the marker (for more information refer to [8]). However, an explicit undistortion step is not required. We use H to define a region of interest and to limit the focus of the segmentation algorithm to a special area in the image. The segmentation algorithm itself uses local adaptive thresholding based on integral images by Shafait *et al.* [15] and a concatenation of morphological operations (see Equation 1).

$$M_{seg} = (\neg(M_{Th} \bullet SE_1) \circ SE_2) \bullet SE_1 \qquad (1)$$

In Equation 1, M_{Th} denotes the result of local adaptive thresholding with a neighborhood size of $(2 \cdot w_h) + 1$. SE_1 and SE_2 denote structuring elements of length $(2 \cdot w_h) + 1$ and $(2 \cdot w_h) + 3$, respectively. The symbol $\neg$ is used for mathematical inversion and the symbols $\circ$ and $\bullet$ denote morphological opening and closing.

By using a method originating from Chang *et al.* [3], we are able to extract the contours from the segmented object in linear time, making up the final output from our detection and segmentation stage.

4 Feature Extraction and Recognition

Although the use of natural feature based recognition approaches like SIFT [12] or SURF [2] is popular in wide areas of object recognition, using these methods is not reasonable for our application. Medical pills do not exhibit enough texture, which negatively influences the number of features found by the aforementioned algorithms. Thus, the most dominant features for medical pill identification remain size, shape and color. More evolved features like imprints or transparency are hard to detect automatically, thus we focused on the robust and instant estimation of these features.

Size Estimation Since we got the reference measurements from the marker target we can infer the size of the object with high accuracy. A subset of pixels from the segmented area is chosen as presented in the work of Wijewickrema and Paplinski [18]. This subset is transformed by the homography H, followed by the determination of the covariance matrix. The length and the width of a tablet are defined along the major and minor axis of its projection, given by the minima and maxima values. An illustration of this technique is depicted on the left of Figure 4. The size estimate is tolerant to rotation, translation, scale and a certain degree of perspective distortion.

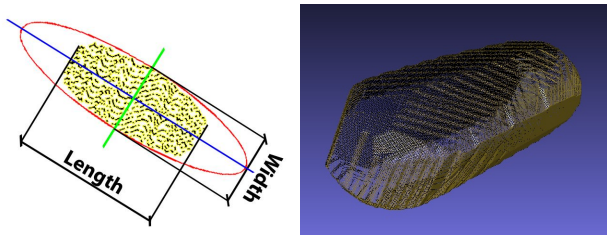

Figure 4. Size estimation for a single medical pill. A subset of pixels is sampled to serve as the basis for covariance matrix calculation (depicted on the left). The silhouette-based reconstruction of a single medical pill (depicted on the right).

Accurate information about the size of a medical pill might serve sufficient for identification. However, we are also interested in getting a more complete representation of the medical compound, if computationally feasible. Since our segmentation algorithm runs in instant time and we rely on onboard image processing, it is possible to acquire an entire set of images from slightly different viewpoints. We can also acquire information about the height of the tablet on the marker target by using a silhouette-based reconstruction approach (see the work of Martin and Aggarwal [13]). By combining the segmentation results over several viewpoints using volumetric reconstruction it is possible to estimate a bounding volume of the tablet. Due to our special setup, the minimum viewing angle is limited, however. So, undesired roof-like structures may be introduced on the upper part of the object. In order to refine those measurements we performed a correction based on the assumption that medical pills are convex symmetric objects. In the first step we cut off the upper part of a pill based on an analysis of the amount of valid voxels on each slice along the height of the object. We got the final result after considering a precomputed systematic measurement error.

On the right of Figure 4 an illustration generated from exemplary volume data is shown, which is obtained from this approach (visualized in Meshlab[2] for simplicity). Note that triangulation is not necessary, rather the resulting point cloud can be used to estimate the height of the medical compound. While not mandatory for our application, it allows us nevertheless to add a third dimension to our size estimation and to gain additional information about the object shape.

[1] `http://studierstube.icg.tu-graz.ac.at`

[2] `http://meshlab.sourceforge.net/`

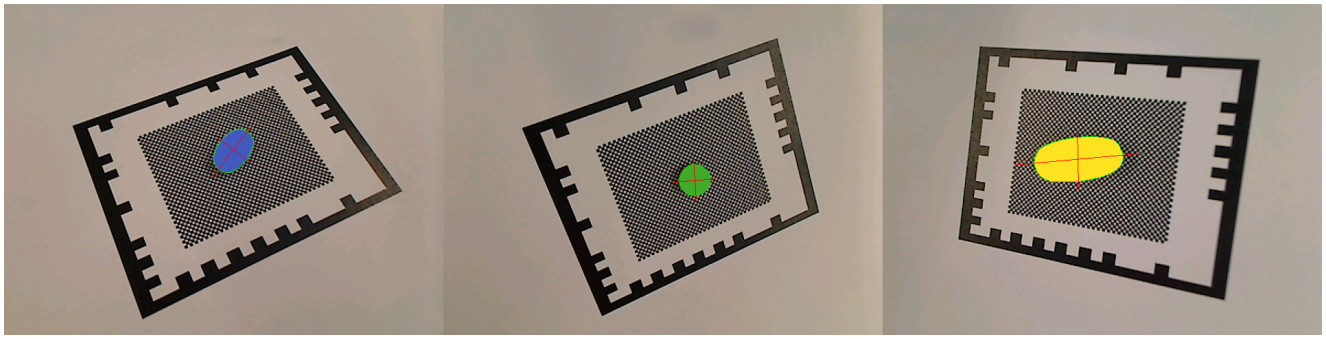

Figure 3. Medical pills on a marker target. The checkerboard background allows a robust segmentation of the individual tablets. The outline of the segmented object is marked, as well as the major and minor axis.

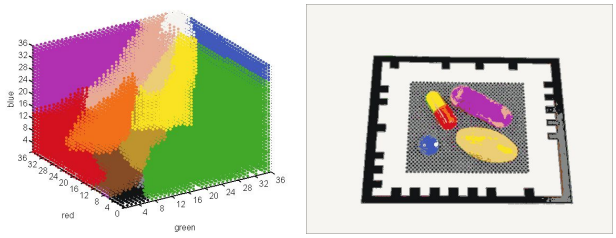

Figure 5. The color clustering in the RGB cube given our lookup table is depicted on the left, a sample per-pixel classification result is shown on the right.

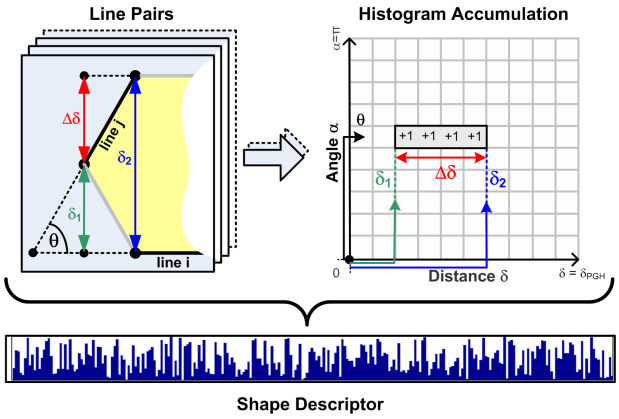

Figure 6. PGH computation from all possible line pairs. After accumulation in a single histogram, the shape descriptor is formed by lining up the columns or rows of the 2D histogram.

Color Estimation Many CV algorithms rely on grayscale imagery. However, for our application the use of color is essential to achieve robust medical pill identification. Since we have a non-stationary setup, the methods used must be able to cope with varying lighting conditions and should deliver results being as close as possible to human perception. The use of mobile phone cameras further complicates the situation, since exposure control or white-balance control is hardly provided through programmatical interfaces.

We performed a local white balance algorithm based on the work of Süsstrunk *et al*. [16] to reduce the influence of varying lighting conditions. The white border of the marker is used to sample reference measurements and to estimate a correction for each individual pixel in our segmented region. Since the use of a very elaborate color classification algorithm is prohibitive due to computational and memory demands, we were using a sRGB lookup table covering the range of possible tablet colors. As there is no global standardization concerning the color of medical pills, we grouped several visually similar hues made of discrete color values into classes by analyzing the contents of the *Identa* [9] database as well as other samples.

The lookup table is then built using all available hues and determining the output color for each entry by evaluating the ΔE_{CIE00} color distance metric in CIE LAB space as advised in the work of Vik [17]. In doing so we are able to generate a histogram of assigned colors and analyze this histogram to get an estimate of the tablet colors[3]. The partitioning of the color cube according to our lookup table, as well as a sample classification result, is depicted in Figure 5.

Shape Estimation In general, medical pills may have an arbitrary shape. Through an analysis of samples as well as the *Identa* [9] database, however, we were able to distinguish 4 major shape classes (circular, oblong, oval, special) for medical pills.

Since the segmentation stage provides a boundary of the segmented tablet, we can use an algorithm for estimating the object shape considering the same key aspects as for size estimation: invariance to translation, rotation, scale and some degree of perspective distortion.

In order to achieve this goal, a modified version of the *Pairwise Geometric Histogram* by Evans *et al*. [5] is used.

[3]Note that we are able to distinguish uni- from bi-colored tablets by analyzing the modes in the histogram.

As illustrated in Figure 6, all pairs of line segments are considered and their relative orientation and perpendicular distance is recorded into a common 2D histogram, which forms a descriptor for the shape of the object. The dimensionality of the histogram has to be chosen accordingly.

For forming the PGH, the boundary has to be approximated by line segments, which in our case can be done efficiently by using all contour points which span the convex hull. For non-convex objects an efficient approach (so-called *Critical Point Detector*) by Zhu *et al*. [19] can be used. The boundary is thus approximated by a polygonal representation.

The original form of the PGH descriptor is not scale invariant. However, limited scale-invariance can be achieved by limiting the maximum expected distance in the PGH descriptor to twice the maximum distance of the boundary from the tablet centroid. The histogram bins are then scaled accordingly, so that all descriptors have a common dimensionality. The Euclidean distance can then be used to compare two histograms and calculate a matching value.

5 Prototype Evaluation

Basically, our application is able to run on any modern smartphone. However, our prototype was mainly tested on two different modern smartphones, an *Asus m530w* and a *Samsung Omnia Pro*. In Figure 7 the screen of the latter smartphone is depicted, running the pill recognition task and querying the *Identa* online database [9]. A big problem with mobile phone cameras is the lack of control of white-balance, auto-exposure or auto-focus. Although considerable advances have been made in this area recently (see the work of Adams *et al*. [1]), methods to control these parameters have not widely found their way into customer hardware, yet. Thus, we chose two devices where at least an auto-focus function was available.

Database We evaluated our system with images of 640x480 pixels using a manually annotated set of 95 pills, capsules and tablets, and a total of 225 images (also including multiple views of the same pill for height estimation). This set is a subset of the *Identa* database and is used solely for testing purposes. All objects used for tuning and estimating correction factors were calculated from a separate set of several hundred medical pills retrieved through a pharmacy.

We used 192 discrete hues for the definition of 13 visually similar color classes in the construction of the sRGB lookup table. Through an additional evaluation procedure the best performing spacing in each direction was found to be 7 units in each direction of the color cube.

Through extraction of undistorted contours and manual annotation of representative samples, we collected training data for shape estimation. In total, 27 contours, each corresponding to one of 4 different shape classes are used in the final set. Their pairwise geometric histogram (PGH) representation was stored and used as reference data for nearest neighbor classification. In the following experiments 16 angle bins and 16 distance bins were used for the PGH.

Size Estimation and Correction By evaluating our method on a set of training examples, we can form an estimate of systematic errors occurring during size measurements. As already described in Section 4, the size of the checkerboard pattern used for segmentation has a direct influence on these errors. Due to the inaccuracies in the segmentation procedure, the estimate of the individual dimensions tend to be too high. In Table 1 we list the mean errors and the corresponding standard deviation with and without correction. The systematic error is in the range of about 0.75 times the width/length of the checkerboard pattern for the height estimation and around 0.3 times the width/length of the checkerboard pattern for width and length estimation. All values are given in millimeters $[mm]$.

	without correction		**with correction**	
	mean err.	std. dev.	mean err.	std. dev.
length	0.722	0.409	**0.365**	**0.311**
width	0.525	0.299	**0.285**	**0.268**
height	1.100	0.810	**0.263**	**0.263**

Table 1. Mean error and standard deviation for the size estimation results without and with correction of the systematic errors. The mean errors drop significantly as the systematic error is removed. All measurement values are given in millimeters $[mm]$.

Probabilistic Database Query Compared to our former evaluation (Hartl *et al*. [7]), we now use a probabilistic query strategy, combining the individual feature estimates in a probabilistic framework. However, since the online database is not capable of processing such information, we decided to use our manually annotated database instead. Note that capital letters denote lists of values, while lower case letters describe single values.

We first extracted the features of an unknown pill as described in Section 4 including length l and width w. Although a volumetric reconstruction can be a valuable source of information, some tests showed that the obtained height values improved the recognition rates on medical pills only marginally. It is very likely that segmentation accuracy is the limiting factor in this case. In favor of keeping the application accessible for untrained personnel, object height h is not used in the remaining experiments.

Instead of a discrete classification result we now assign probabilities for each known shape class s and color class c. We store in a list P_{shape} the probability for each shape class after Equation 2, where d_s is the minimum Eu-

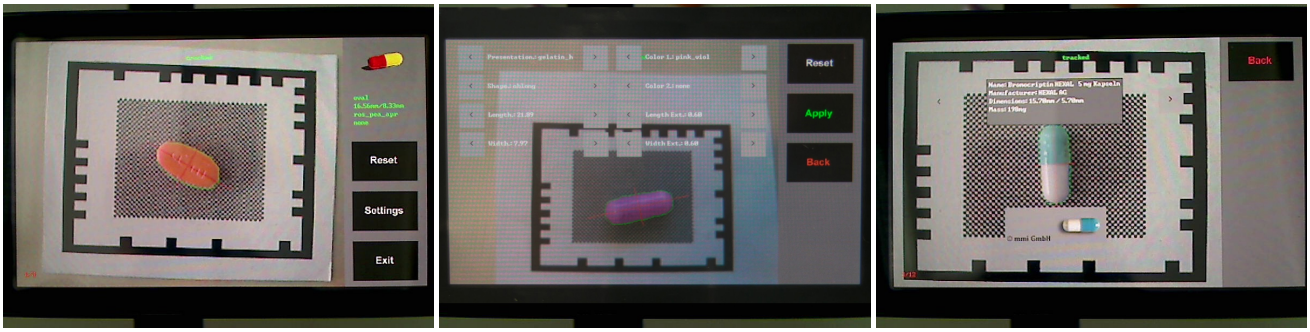

Figure 7. Snapshots from our application running on the *Samsung Omnia Pro* using the *Identa* [9] online database. The user is provided with the initial results of the segmentation and feature extraction step (left image). Additional information can be added, as well as erroneous results from the feature estimation stage can be corrected (middle image). Finally, the information about the actual pill is retrieved from the *Identa* online database and presented to the user (right image).

clidean distance to class s as represented by their PGH vectors and d_{max} is the maximum distance.

$$P_{shape}(s) = 1 - \frac{d_s}{d_{max}} \tag{2}$$

For getting the object color, we analyzed the class histogram h of colors within the region to obtain a probability for each color class c and put it into a list P_{color} (see Equation 3).

$$P_{color}(c) = \frac{h_c}{\sum h} \tag{3}$$

Then we sort the current database content according to the minimum sum of magnitude deviations in l and w, giving a new list L_{lws}, keeping the sorted list of distances D_{lws}. We may then calculate a list of probabilities P_{size}, where each entry is computed after Equation 4.

$$P_{size}(i) = 1 - \frac{D_{lws}(i)}{max(D_{lws}(i))} \tag{4}$$

For each entry in L_{lws} we take its shape- and color-class and use this to select the matching probabilities p_{shape} from P_{shape} and p_{color} from P_{color}. Under the assumption of independent object properties for size, shape and color, we arrived at the final list $P_{size,shape,color}$ of probabilities, whose entries are obtained by computing Equation 5 and sorting the results in descending order.

$$P_{size,shape,color}(i) = P_{size}(i) * p_{shape} * p_{color} \tag{5}$$

The list of candidates can then be directly obtained from the corresponding list of indices $L_{size,shape,color}$.

We chose several combinations to test the influence of different features on the recognition performance. The results of this evaluation are shown in Figure 8.

As expected, using a combination of all estimates gives the best result. In our tests, for any candidate considered, the probability that the correct identification is within the top three candidates retrieved is already $> 90\%$. Since we cannot get hold of the online database content directly

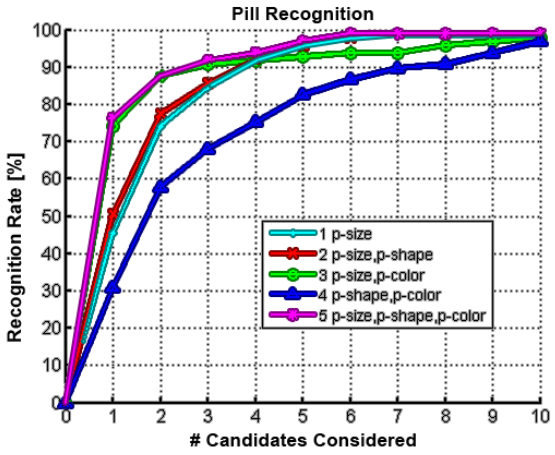

Figure 8. Recognition result for several combinations of features. Note that we counted an identification as correct if it appeared among the first *N* candidates from the database.

or modify the query interface in any way, we left a tight integration of the online database into our system as open issue. However, for a simple *AND* concatenation of the individual feature estimates in an online query, the results from the online database are reasonable but significantly inferior.

Runtime Estimation In order to demonstrate the effectiveness of our computer vision methods, we were recording timings for several runs of our application on the *Samsung Omnia Pro* smartphone.

In Figure 9 the time spent in the individual steps of our system is visualized given two different image resolutions. The time spent for object segmentation clearly dominates. Although the amount of data to be processed in the case of 640x480 pixel images is only four times as high as for 320x240 pixel images, the algorithm takes significantly longer to perform on 640x480 pixel images. A rea-

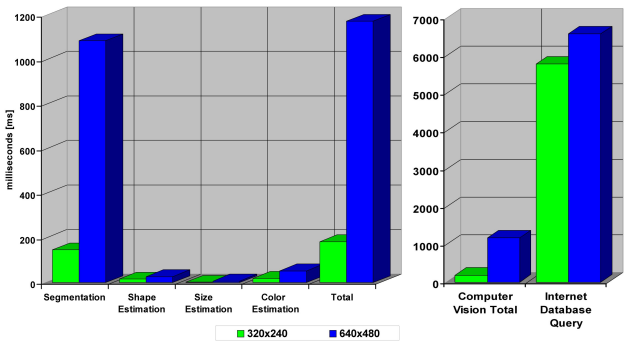

Figure 9. Runtime estimate for our system using two different image resolutions. On the left, the results for running the individual CV tasks is given. On the right, a comparison of the time spent for the CV-based parts of our application is compared to the overall time for performing a simple query on the Internet database over a WiFi connection.

son for this might be a bottleneck in the memory and cache management on the device due to the considerably higher amount of data to handle. By contrast, the feature estimation routines only take several milliseconds, giving a total of 200 milliseconds to 1.2 seconds for the entire CV task.

For the sake of completeness, in Figure 9 also the timing results for querying the online database is given and compared to the time spent in the CV-driven part of our application. We query and retrieve only textual data from the database over a standard WiFi connection in the order of several kilobytes. The latency of the system is compelling, so that the use of a locally cached database on the device is advisable. As a consequence, the time of database querying would become negligible.

Discussion Several parameters influence the performance of our system in terms of runtime and accuracy. With a decreasing size of the checkerboard pattern, the segmentation algorithm becomes slower. This results in an increased overall runtime, the size estimation accuracy becomes significantly better, however. For the estimation of the height of the medical pill, a set of images has to be acquired. The accuracy of the height estimate increases with the number of images and the number of consecutive carving steps respectively.

Using a local copy of the database right on the device reduces the latency for querying significantly. However, since the database content may change several times a year, for a practical application of our system, an appropriate update strategy would be necessary.

6 Summary

In this paper we presented our work on an application for instant medical pill recognition on mobile phones. By using a marker target and Computer Vision methods for robust segmentation and feature extraction, we are acquiring contact-free 3D measurements about the tablets geometry instantaneously without the need for additional equipment like calipers. In our experimental evaluation we showed that our methods are able to estimate the size of the object with high accuracy. By additionally estimating the color and the shape of the tablet we can query existing database resources for accurate identification of the medical compound. This information is presented to the user in an interactive manner with various levels of details.

The major number of methods employed in this application is not limited to a special object category. In principle, our application can also be used to segment and classify items within other object categories that meet roughly the same size dimensions as medical pills. This allows us to further develop our algorithms also into other areas, such as media entertainment or Augmented Reality, in the future.

Acknowledgements

This research work was funded by the Christian Doppler Research Association (CDG) and supported through a grant from the Deutsche Bundesdruckerei GmbH.

References

[1] A. Adams, E.-V. Talvala, S. H. Park, D. E. Jacobs, B. Ajdin, N. Gelfand, J. Dolson, D. Vaquero, J. Baek, M. Tico, H. P. A. Lensch, W. Matusik, K. Pulli, M. Horowitz, and M. Levoy. The Frankencamera: an Experimental Platform for Computational Photography. In *SIGGRAPH*, pages 29:1–29:12, 2010.

[2] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding (CVIU)*, 110:346–359, June 2008.

[3] F. Chang, C.-J. Chen, and C.-J. Lu. A Linear-Time Component-Labeling Algorithm using Contour Tracing Technique. *Computer Vision and Image Understanding (CVIU)*, 93(2):206–220, 2004.

[4] C. M. Christoudias, B. Georgescu, and P. Meer. Synergism in Low Level Vision. In *Int. Conference on Pattern Recognition (ICPR)*, pages 150–156, 2002.

[5] A. Evans, N. Thacker, and J. Mayhew. Pairwise Represenations of Shape. In *Int. Conference on Pattern Recognition (ICPR)*, pages 133–136, 1992.

[6] P. Felzenszwalb and D. Huttenlocher. Efficient Graph-Based Image Segmentation. In *Int. Journal of Computer Vision (IJCV)*, volume 59(2), 2004.

[7] A. Hartl, C. Arth, and D. Schmalstieg. Instant Segmentation and Feature Extraction for Recognition of Simple Objects on Mobile Phones. In *Int. Workshop on Mobile Vision (in conjunction with CVPR)*, pages 17–24, 2010.

[8] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2008.

[9] Identa Database. `http://www.gelbe-liste.de/pharmindex/identa`, 2011.

[10] H. Kato and M. Billinghurst. Marker Tracking and HMD Calibration for a Video-Based AR Conferencing System. *Int. Workshop on AR*, 0:85, 1999.

[11] A. Koschan and M. Abidi. Detection and Classification of Edges in Color Images. In *Signal Processing Magazine (Special Issue on Color Image Processing)*, volume 22(1), pages 64–73, 2005.

[12] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Int. Journal of Computer Vision (IJCV)*, 60:91–110, November 2004.

[13] W. N. Martin and J. K. Aggarwal. Volumetric Description of Objects from Multiple Views. *PAMI*, pages 150–158, 1983.

[14] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust Wide Baseline Stereo from Maximally Stable Extremal Regions. In *British Machine Vision Conference (BMVC)*, volume 1, pages 384–393, 2002.

[15] F. Shafait, D. Keysers, and T. M. Breuel. Efficient Implementation of Local Adaptive Thresholding Techniques using Integral Images. In *Document Recognition and Retrieval (DRR)*, volume 6815/1, page 681510. SPIE, 2008.

[16] S. Süsstrunk, J. Holm, and G. D. Finlayson. Chromatic Adaptation Performance of Different RGB Sensors. *Proc. of SPIE*, 4300:172–183, 2001.

[17] M. Vik. Industrial Colour Difference Evaluation: LCAM Textile Data. In *Proc. of the AIC Color and Paints*, pages 138–142, 2004.

[18] S. Wijewickrema and A. P. Paplinski. Principal Component Analysis for the Approximation of a Fruit as an Ellipse. *Int. Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, 2005.

[19] P. Zhu and P. M. Chirlian. On Critical Point Detection of Digital Shapes. *PAMI*, 17(8):737–748, 1995.