Computer-Vision based Pharmaceutical Pill Recognition on Mobile Phones

Andreas Hartl*
Supervised by: Clemens Arth†

Institute for Computer Graphics and Vision
Graz University of Technology
Graz / Austria

Abstract

In this work we present a mobile computer vision system which simplifies the task of identifying pharmaceutical pills. A single input image of pills on a special marker-based target is processed by an efficient method for object segmentation on structured background. Estimators for the object properties size, shape and color deliver parameters that can be used for querying an online database about an unknown pill. A prototype application is constructed using the Studierstube ES framework, which allows to perform pill recognition on off-the-shelf mobile phones. System runtime and retrieval performance with the estimated features is subsequently evaluated on a realistic test set. The retrieval performance on the exemplarily used Identa database confirms that the system can facilitate the task of mobile pill recognition in a realistic scenario.

Keywords: pill recognition, mobile phones, marker, feature estimation, database

1 Introduction

Correct and timely identification of drugs which are encountered without packaging, is a major problem in healthcare. If the exact type of drug taken by a patient is known, a more directed therapy can be applied and unnecessary medical effort can be avoided. The identification of pharmaceutical pills like tablets, capsules and pills (dragées) must often be performed as fast as possible. If pills need to be identified on the way, only a visual inspection is possible. This task can be supported by performing computer vision on mobile devices. The ubiquitous nature of mobile phones, the built-in camera as well as advancements in processing power make up an attractive platform for application development. The main challenges can be seen in the changing environment of operation as well as the limited computational resources of mobile devices.
Some domain research reveals three main methods for pill identification: (1) look-up in a book, (2) application of a special identification scheme, (3) querying a database. From the aforementioned methods, an identification scheme offers best accuracy within a reasonable amount of time. In mobile applications, the availability of an identification method and the necessary skill of the operator are critical. This drawback may be mitigated by using mobile phones for pill identification.
In a visual identification system a set of features must be determined. These could be shape, size, color, scores and imprints, but also transparency or texture. Such features can be estimated directly on the device or on a server which is used to process an image taken by the mobile phone. It is important to note that mobile network coverage is still imperfect, which can render a solution that employs server-side processing entirely useless. In the proposed approach, properties of pills may be instantly estimated without requiring a network connection. Thus, features such as object size can still be used to identify pills by other means. Besides, on-device feature estimation is an essential building block for a truly independent solution in which pill information is stored directly on the device.
In the following we present a system for instant visual pill recognition on mobile phones based on an image taken with the built-in camera. Initially, a set of features is estimated for each individual pill. This information is used to query an online available database. Then, a series of candidates is presented on the mobile phone along with additional information. In this work the features shape, size and color are estimated since they are supported by freely available online databases.
The remainder of this work is structured in the following way: In Section 2 a short review of related work in the area of pill recognition as well as computer vision on mobile phones is presented. Section 3 is devoted to segmentation, as it is a fundamental problem in the application to be developed. Section 4 deals with the feature extraction process from the segmented regions. In Section 5 an overview of a prototype implementation for mobile devices is given, followed by an evaluation of the system on an exemplary database in Section 6. A conclusion and additional remarks are given in Section 7.

*ahartl@student.tugraz.at
†arth@icg.tugraz.at

2 Related Work

To the best of our knowledge there is only a single report of a pill recognition system so far, which is described in a US patent [13]. It deals with an automatic method for the task of verification that the content of a container filled by a dispensing system, corresponds to a given prescription. Little knowledge is available about the inner workings of the system. According to the given source, this is a fixed system which uses a frame grabber as input and employs color, geometry and surface features to identify all pills that a given machine may process. Prior knowledge is used to simplify the problem which gives better results.
Ubiquitous mobile devices are used for various tasks in computer vision. Wagner et al. perform robust 6DOF natural feature tracking using modified SIFT and Ferns as descriptors [11]. In order to allow computation on mobile devices, extensive modifications are carried out to the basic concepts of SIFT and Ferns, followed by an instructive evaluation of system performance. Despite severe limitations in processing speed and memory bandwidth, they achieve real-time performance when using textured planar targets on current-generation phones.
Klein and Murray present a system for parallel tracking and mapping on camera phones [7]. They implement a key-frame based SLAM system that is capable of generating and augmenting small maps. Limited computational resources and problems in image acquisition such as a rolling shutter are specifically accounted for to allow computation on mobile devices (iPhone 3G).
In the system at hand robust estimation of object features from a single input image must be performed. Due to the nature of the problem, it seems justified to run through a separated segmentation, feature extraction and classification step (depending on the feature), in which tasks may be optimized independently. For each of the steps needed in our application a vast amount of literature is available, but self-contained work on how to perform these steps efficiently on mobile devices is not available. Thus, we propose a series of solutions which make the problem computable on current mobile phones in instant time.

3 Robust Segmentation

Segmentation is a critical problem in the system at hand, because the results determine the quality of any subsequent feature estimation task. As the setup used for image acquisition is not fixed and lighting conditions may vary, performance with changes in scale, perspective distortion and non-uniform lighting is particularly important. Since it is necessary to estimate the size of arbitrarily colored pills, a marker-based target of known geometry and background is used. For reasons of practicability, the dimensions of this target are chosen to fit into a wallet. Using a checkerboard background (see Figure 1), it is possible to robustly segment objects with reasonable requirements in processing speed. Applicable images may be obtained from an autofocus camera, since fixed focus cameras are unable to give a sharp image of suitable size.

3.1 Reduction of Input Data

The marker-based target can also be used to reduce the amount of input data for segmentation. Besides, feedback about the image geometry is possible before further computations take place. This may be helpful, because rectification of the entire image is not efficient. For image regions, rectification is feasible, however, because the amount of input data is much smaller.
In the problem at hand, a homography between points on the image plane and points lying on a plane defined by the dimensions of the target can be computed. This bijective mapping of coordinates between planes can be represented by a non-singular matrix of size 3x3 with 8 degrees of freedom (see e.g. the work of Hartley and Zisserman [6]). For the computation (estimation) of the 3x3 Matrix $\mathbf{H_{IW}}$ ($\mathbf{H_{WI}} = \mathbf{H_{IW}}^{-1}$) at least $n = 4$ point correspondences must be found between image plane and the corresponding world plane. In consideration of a possibility to realize such a rectification procedure efficiently, the minimum number of 4 points is chosen. They correspond to the corners of the reference rectangle of the marker-based target. Thus, a rectangular region of interest (ROI) can be defined in the image plane using the known location and dimensions of the checkerboard area and the previously computed homography.

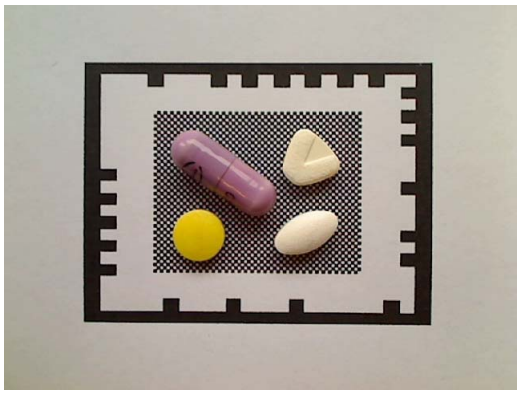

Figure 1: Frame marker with structured background

3.2 Segmentation Algorithm

A segmentation of objects on the proposed card can be obtained by local adaptive thresholding and morphological operations on a grayscale input image. All relevant computations may be carried out on a square neighborhood that extends into both directions w_h pixels, giving a minimum total square length of $2 \cdot w_h + 1$ pixels. The principal

algorithm is as follows:

$$M_{seg} = (\neg(M_{Th_1} \bullet SE_1) \circ SE_2) \bullet SE_1, \qquad (1)$$

where M_{Th_1} denotes a mask obtained by local adaptive thresholding with a neighborhood size of $2 \cdot w_h + 1$, and SE_1 and SE_2 denote structuring elements of length $2 \cdot w_h + 1$ and $2 \cdot w_h + 3$. The symbol $\neg$ is used for mathematical inversion and the symbols $\circ$ and $\bullet$ denote morphological opening and closing. In our implementation, the checkerboard pattern is detected by application of the proposed procedure using an efficient method for local adaptive thresholding by Shafait et al [9].

Further processing is necessary to extract smooth contours. Thus, region labeling with integrated boundary computation is carried out after the work of Chang et al. [3], in which a linear-time method is described. Subsequently, the convex hull of each boundary is computed using the algorithm by Graham [5]. In the final step, we apply variant of the line drawing algorithm by Bresenham [2] to obtain connected chains (see Figure 2). If necessary, the region may be calculated in a flood fill operation then. It must be noted that the choice of the single parameter w_h is uncritical in the application at hand. However, the length of the checkerboard pattern must be chosen with care, since it determines the quality of segmentation to a large extent. This choice is dependent on the desired resolution, since a sufficiently sharp image of this pattern is required. Currently the length is set to $0.6mm$, favoring images of 640 x 480 pixels, as a compromise between accuracy, runtime and usability.

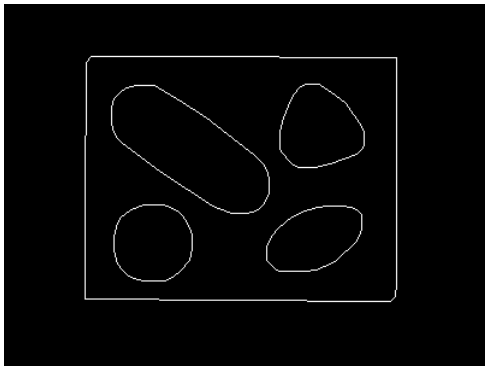

Figure 2: Segmentation borders for Figure 1 (extract)

4 Feature Estimation

In the application at hand we measure objects and perform a coarse estimation of shape and color. The following considerations are based on the assumption that the starting position for property estimation is the boundary of an image region that resembles the silhouette of an object as closely as possible. For each step, an efficient solution is needed to allow for computation on mobile devices.

4.1 Size Estimation

Geometric properties of pharmaceutical pills such as length, width and height are important features for identification (see Section 6). Invariance to perspective distortion is necessary to allow for correct measurements. With a single input image, measurements can be carried out in two dimensions by using an available homography (see Section 3.1). In the current approach for size estimation, distortions in image acquisition are not considered, since the error caused by segmentation is assumed to be dominant.

4.1.1 Measurement of Length and Width

In the following, the length of a pill is defined as the extension of its boundary along the major axis of its perpendicular projection. This suggests that the extension in the perpendicular direction is defined as width. With this definition the majority of pill shapes may be measured correctly without further interaction. Consequently it seems reasonable to determine the direction of maximum variance within a region where a pill is supposed to be. To achieve this goal, we propose a procedure in which a subset of all region points is used for increased robustness. First, a flood fill operation is carried out on the contour and a suitable number of points is selected and rectified using an available homography. Subsequently, the direction of maximum variance $\mathbf{v_1} = [v_1; v_2]$ is determined by analysis of the covariance matrix $\mathbf{C}$ [12]. In order to obtain a more accurate estimate for the length and width of an object, it is necessary to project the boundary that makes up an object using the obtained vector $\mathbf{v_1}$. In the absence of outliers, it is sufficient to project the world points of the convex hull, because they bound the extension of the object. If the vector $\mathbf{m} = [m_x; m_y]$ corresponds to the centroid of all world plane points that make up a region, a projection may be computed as:

$$\begin{bmatrix} xi_{W(p)} \\ yi_{W(p)} \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \\ -v_2 & v_1 \end{bmatrix} \begin{bmatrix} xi_W - m_x \\ yi_W - m_y \end{bmatrix}, \qquad (2)$$

where $[xi_W; yi_W]$ denotes a point in the world plane that belongs to the current region, and $[xi_{W(p)}; yi_{W(p)}]$ denotes its projected counterpart. An estimation of length and width may then be obtained by computation of differences between minimum and maximum values of these coordinates (see Figure 3 for an illustration of this process).

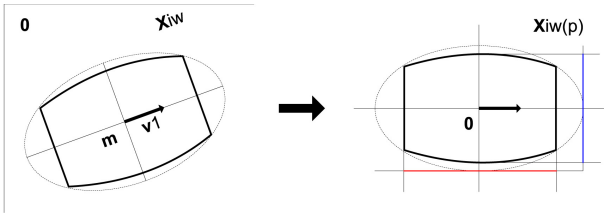

Figure 3: Projecting boundary points (world plane)

4.1.2 Considerations on Accuracy

The square size sq_w of the checkerboard pattern, that limits the accuracy of segmentation, is identified as the major source of error. Under the assumption of a target that entirely fills the viewable area and does not exhibit other kinds of distortion, the expected error is dependent on the square size sq_w. As an upper bound, the error may be assumed to be double the square size then. The minimum usable square size sq_w is dependent on the image resolution. If we assume that pills may be correctly segmented under the previous conditions, the maximum error e_{max} can be estimated as follows:

$$e_{max} = 2 \cdot sq_w \tag{3}$$

With a square size of $0.6mm$, the expected error e_{max} is $1.2mm$. With images of 640 x 480 pixels, the square size may be reduced to $0.5mm$, giving an estimated error e_{max} of about $1mm$. Using an ordinary ruler as measurement tool, it is rather difficult to measure lengths below 1 mm, and the shapes of pills further aggravate measurements. For these reasons, the accuracy of the proposed method can be expected to be better than what can be obtained with a ruler.

4.2 Shape Estimation

For reasons of efficiency, the boundary of an object is chosen to serve as a basis for the estimation of object shape. A shape estimator should be as invariant to changes in translation, rotation and scale as possible. A certain degree of perspective distortion should still be tolerable to account for artifacts. In the following we use a modified pairwise geometric histogram (PGH) as shape descriptor and subsequently perform shape matching.

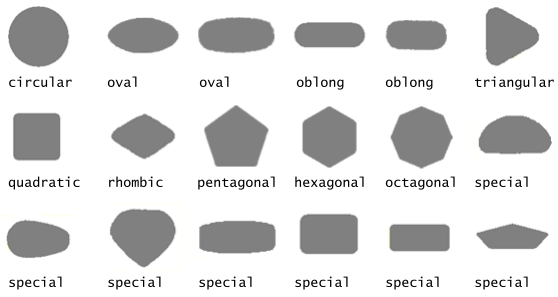

Figure 4: Examples of pill shapes with classes

An analysis of available samples as well as examples from literature led to the decision to categorize each example into one of 9 representative shape classes containing only **convex** instances. A class termed *special* is added to represent shapes that do not fall into any of the other categories (see Figure 4). Shapes of pharmaceutical pills show a pronounced intra-class variance as well as partially low inter-class variance. Thus, an applicable shape descriptor needs to have considerable discriminative power. We propose a modified pairwise geometric histogram (PGH) for this task, which is efficiently computable and sufficiently invariant to changes in scale. This development of a PGH also allows for efficient shape matching.

4.2.1 Modified Pairwise Geometric Histogram

In the PGH descriptor oriented line segments are investigated (see the work of Evans et al. [4]). Their relative orientation and perpendicular distance is analyzed and this information is collected in a 2D histogram of size $D \cdot A$. The set of possible angles and distances is mapped onto the histogram by accumulation of occurrences. During the process, each line is used as a reference line and the angle, as well as the perpendicular distance is computed to all the other lines. In fact, every line is represented as a histogram and the accumulation of all histograms represents the shape of the object. An exemplary histogram is shown in Figure 5. Although mainly designed for polygo-

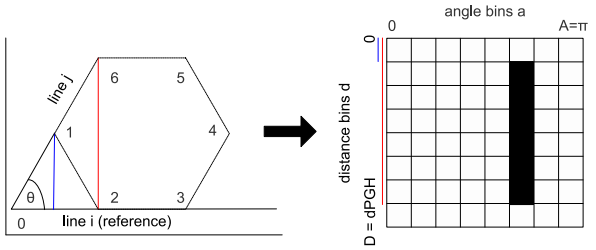

Figure 5: PGH computation of a single line

nal shapes, application to non-polygonal types is possible, if a polygonal approximation is computed as an initial step. Besides, the PGH is not scale invariant in its original form. Limited scale invariance may be achieved by application of a suitable and stable similarity metric[1]. For reasons of efficiency, we perform the following method to achieve invariance for a limited range of scales: We assume that the maximum distance for a PGH can be estimated from the rectified contour. As the computation of the convex hull is part of the segmentation algorithm, no outliers may be expected. So, a measure of scale can be found just by searching for the maximum distance d_{max} in the key from their centroid. The maximum PGH distance may be computed as

$$d_{PGH} = 2 \cdot d_{max} \tag{4}$$

then. As the vast majority of pharmaceutical pills is convex, the PGH is constructed with a reasonable amount of bins using just the points which make up the convex hull. This corresponds to considerable savings in computation.

4.2.2 Classification

Due to the achieved scale invariance, we perform shape matching on a set of predefined samples using the Euclidean distance as a similarity metric. Although the possibly large size of the PGH may still be prohibitive for certain applications, this drawback can be mitigated by application of a suitable data structure.

4.3 Color Estimation

A color estimation procedure should correspond to human perception as good as possible. Besides, the procedure should be able to give reasonable results with varying lighting conditions. In general, pharmaceutical pills may have arbitrary colors (transparent pills are not considered). Research in literature and online databases reveals that most pills have one or two colors, out of a set of 16 basic color tones (black, white and gray are treated like a color). These are black, white, blue, beige, brown, gray, green, ocher, pink, violet, orange, peach, rose, red, cyan and yellow (see Figure 6 for a set of mean colors).

Figure 6: sRGB mean colors for pill color classes

For color estimation, the influence of lighting is reduced by applying a method for local white balance which is based on reference measurements on the marker-based target and subsequent scaling in input space. Due to the small amount of available samples and for reasons of efficiency, color estimation is based on a sRGB lookup table (LUT) which is created by evaluation of the ΔE_{CIE00} color distance metric (see e.g. the work of Vik [10]) in CIE LAB space. Per pixel classification results for a pill can be aggregated in a histogram. An analysis of this information gives the color estimation result.

4.3.1 Look-up Table Computation

Lookup tables allow fast classification at the cost of additional storage and memory requirements. For practical reasons, the LUT size should be comparatively small. In the following, the value s_{LUT} denotes the amount of entries for each channel in a 3D LUT. We propose the following method for the creation of a sRGB LUT that resembles human perception of color, from a small number of samples per class:

- Definition of representative sample colors c_{ij} per class. They need to be visually similar and should show a smooth course of color within a class.
- Assignment of a label l_i to the examples of class i.
- Definition of the desired LUT size s_{LUT} (equal for all dimensions).
- Conversion of all sRGB sample colors into CIE LAB space using the D65 white point (daylight).
- Iteration through all sRGB LUT entries, computation of the corresponding sRGB color in CIE LAB space and the distances d_{nn1} and d_{nn2} to the nearest samples by using the ΔE_{CIE00} color distance metric.
- Assignment of a label l to the current LUT entry after analysis of the distances d_{nn1} and d_{nn2}.

Cases, where a decision is not possible or no suitable color may be found, are handled by suitable thresholds. The proposed approach relies on the assumption that the colors which may be traversed by exhibiting the step size that is defined by s_{LUT} on each channel, show no or negligible perceptual difference.

4.3.2 Color Classification

A classification for up to two colors may be obtained by analysis of the class histogram h_c of per pixel results on white balanced data. Thus, the relative amount of covered pixels in the region as well as the significance of the result, which is based on the distance between the first two entries of a sorted class histogram h_{cd} (descending order), is evaluated. The following metrics are used to decide on the color(s) of a region:

- coverage for a result with one color (label i):

$$c_i = \frac{h_{cd}(0)}{\sum h_c} \tag{5}$$

- coverage for a result with two colors (labels i and j):

$$c_{i,j} = \frac{(h_{cd}(0) + h_{cd}(1))}{\sum h_c} \tag{6}$$

- significance for a result with one color (label i):

$$s_i = \frac{h_{cd}(0) - h_{cd}(1)}{h_{cd}(0)} \tag{7}$$

- significance for a result with two colors (labels i,j):

$$s_{i,j} = \frac{1}{s_i} \tag{8}$$

A decision is possible, if one of these coverages is above a threshold c_{th}, because this assures a more meaningful result. Then, the products $c_i \cdot s_i$ and $c_{i,j} \cdot s_{i,j}$ are used to decide, whether one or more colors are present. If the first product is greater than the second, the result is the corresponding label i of the maximum entry in h_{cd}. Otherwise two colors are assumed and the labels i, j that correspond to the first two entries in h_{cd} are reported.

5 Mobile Phone Prototype

We implemented a prototype system suitable for mobile devices using the algorithms described in the previous chapters as well as existing functionality from the StbES framework [8]. It allows to obtain an estimate of recognition performance on mobile devices when querying an online available database. In Figure 7 an overview of the basic components is given. The application frontend may run on a Windows Mobile equipped camera phone as well as on a Windows PC. The database connection is decoupled from the application and may be adapted to arbitrary sources of information.

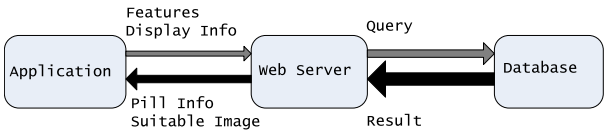

Figure 7: Basic system components

All computer vision related processing takes place on the mobile device (see Figure 8 for an illustration of the required steps). An exemplary pill recognition session is shown in Figure 9 A, B and C. The input image is obtained from live video, using the tracker of the given framework for target detection. The result of segmentation and color estimation as well as the obtained directions of measurement may be instantly verified from overlaid data (see Figure 9 A). In the pill browser, manual correction of query parameters is possible using just the directional stick of the mobile device (see Figure 9 B). Alternatively, input from a touchscreen is possible. In case of the evaluated online database an image for visual verification is provided, along with name, manufacturer, dimensions and mass of the candidate (see Figure 9 C). In order to reduce the initial round-trip time, only textual information is transferred at first and images are cached on the server for on demand retrieval.

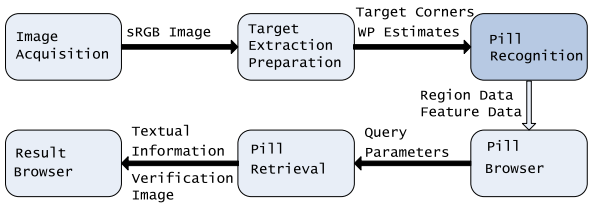

Figure 8: Course of the pill recognition application

Optimizations include excessive preallocation of memory, avoidance of floating point operations or use of fixed-point types. The retrieval of pill information is possible through any of the built-in communication facilities.

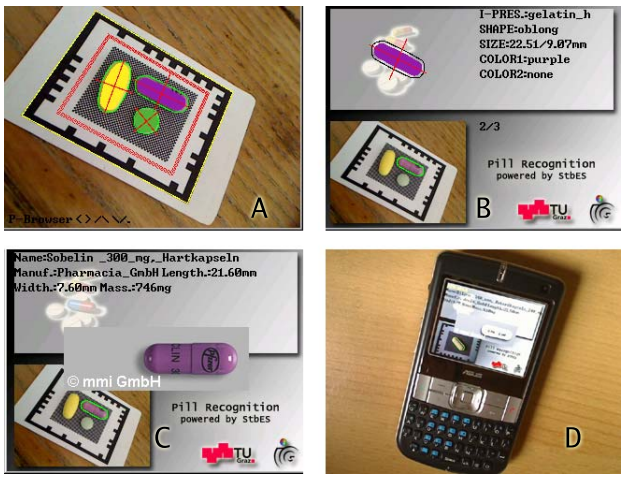

Figure 9: A feature estimation result, B pill browser, C result browser, D test with mobile phone

6 Evaluation

We evaluated the system in terms of estimator performance with suitable test sets that resemble typical operating conditions. Pill identification performance is evaluated with optimal estimator parameters using a set of manually classified examples as well as the Identa[1] online database. The latter is chosen to serve as the backbone for the application, because access is free. Thus, training of feature estimators needs to comply with Identa categories. Only a subset of possible pill shapes are supported as query parameters. The shape categories represent circular, oval, oblong and special shapes (see Figure 10). The created general shape data for pills was subsequently categorized to be in accordance with the shape classes of Identa. In the Identa query form, 14 colors are differentiated by their names as well as a small number of visually similar tones. For reasons of robustness, the tones *rose* and *peach* are merged. The samples for each class are taken from previously collected data on pill color. The Identa database is not fully consistent concerning shape and color, however. Mobile performance is evaluated on an Asus M530w smartphone[2], running Windows Mobile 6 (see Figure 9 D). It features a 2 mega pixel autofocus camera, a 416MHz fixed-point CPU, 64 MB of RAM and 256 MB Flash.

6.1 Datasets

In the proposed evaluation procedure, reference data for the captured pills is necessary. For this purpose, each sample was manually classified into the appropriate categories

[1] http://www.gelbe-liste.de/pharmindex/identa
[2] http://www.asus.com

Figure 10: Examples for Identa shape classes: A special class; B oval class; C oblong class; D circular class

for shape and color. The length and width of each sample were measured using a nonius. All collected information about the samples was subsequently stored for later use (see Table 1 for a description of contents). Despite its small size, the set contains pharmaceutical pills with the most current shapes, colors and dimensions. In Table 2 a detailed listing of the distribution of colors for single-colored pills is given. Test images for each example in

Shapes	*circular*	*oval*	*oblong*	*special*
Instances	41	26	33	8
Colors	*single*	*multi*		
Instances	98	10		
Sizes [mm]	*min. length*	*min width*	*max length*	*max width*
	5.68	5.68	18.07	18.07

Table 1: Contents of the reference set: shape, colors, size

the global reference set were captured at 640 x 480 pixels with differing lighting conditions. A labeled example is obtained by storing each filename at its corresponding entry in the global reference database. This means that one representative image for each sample in the global reference set (108 samples) is included for set D (daylight) and set F (fluorescent lighting).

Color	*black*	*white*	*blue*	*beige*	*brown*
Instances	1	29	6	11	4
Color	*gray*	*green*	*ocher*	*pink/violet*	*orange*
Instances	1	5	6	3	6
Color	*rose/peach*	*red*	*yellow*	-	-
Instances	11	5	10	-	-

Table 2: Contents of the reference set: single-colored pills

6.2 Results

Initially shape and color estimators were extensively evaluated to obtain optimal parameters (shape PGH: D=12, A=12; color LUT: $s_{LUT} = 36$). In this step, the best recognition rate for shape is 0.89 and that for color is 0.84, with little dependence on lighting. In size estimation the average deviation is 0.43*mm* (maximum deviation: 1.37*mm*). Runtime on the mobile device was evaluated for a pill of average size and for the largest available pill (see Table 3 for a detailed listing). For this purpose, the input image is read from a file. Segmentation consumes the largest portion of runtime, but the problem may still be computed on the mobile test device in a reasonable amount of time.

Task	**RT [ms] (largest)**	**RT [ms](average)**
overall	**1205**	**1114**
preparation	20	18
segmentation	698	714
shape features	204	213
shape class.	6	4
size estim.	44	29
color estim.	233	136

Table 3: Runtime evaluation on specific pills (640 x 480 pixels): results for largest pill in set DF as well as an average pill (optimum settings; rounded), RT...runtime

6.2.1 Pill Recognition

Due to the small amount of samples, pill recognition performance is evaluated in two different ways. First, the created global reference database is used as the source of pill information. In the definition of a query range for length and width, a deviation of 0.7mm is assumed. In the second part, pill retrieval performance is evaluated with Identa using a filtered set of 27 pills, that can be positively assigned to instances within Identa. In this case a correction of parameters by a human operator is assumed.

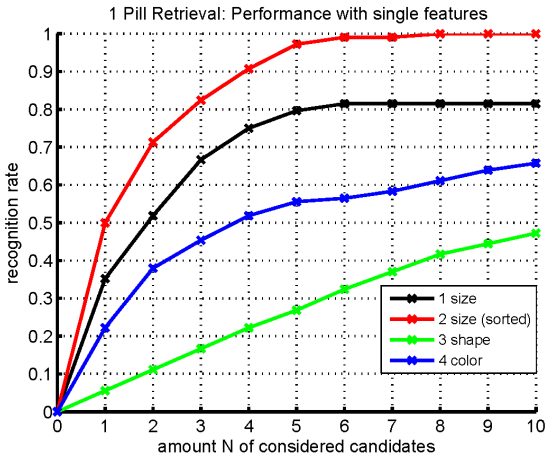

Figure 11: Pill retrieval on the global reference set (single features): results for *set D*

Retrieval performance for single features gives information about the descriptive power of a feature (see Figure 11). This means that *size* is the most powerful feature, followed by *color* and *shape* because higher recognition rates are achieved with a smaller amount of considered candidates. *Size* and *color* have a steeper slope and give

better results faster than the feature *shape*. Performance with the feature *size* reaches a recognition rate of 0.814 ($N = 6$ candidates), when querying using a range defined by the measurements and the expected accuracy in size. Performance of feature *size* improves considerably, if results are ordered by the minimum sum of deviations in length and width. Peak performance may drop when using several features.The reason may be mutual reactions of errors from the individual feature estimators, because their results were not corrected. We have included the results for a very conservative query strategy, using no prior information about the set and reporting the result of sequential querying for features. In Table 4 a summary of results is given.
When querying the Identa database using the filtered set, the final retrieval performance is 85.19 percent. On average, $N = 8$ candidates must be inspected to obtain the correct match (Identa contents 12/2009).

Size	Size (sorted)	Shape	Color	RR_{max}	#
x				0.8148	6
	x			**1.0000**	8
		x		0.4722	10
			x	0.6574	10
x		x		**0.7315**	6
x			x	0.6389	3
x		x	x	0.5556	3
	x	x		**0.8889**	5
	x		x	0.8148	2
	x	x	x	0.7037	2

Table 4: Pill retrieval on the global reference set (*set D*, $N = 10$): best results, RR...recognition rate

7 Conclusions

In this work we present a mobile pill recognition system for conventional smartphones. An algorithm for segmentation of pills from a business card sized marker is presented. Hereafter, relevant properties, such as size, shape and color are chosen to be determined by means of CV. A domain analysis is carried out to identify the most relevant shapes and colors of pills. The estimated features can be used for identification of medical pills within a database then. An example of this application interfacing an online-available database for medical matters is presented, followed by an extensive evaluation of the system in terms of algorithm speed and pill retrieval performance.
The obtained results show that the approach is able to work in instant time using commonly available smartphone hardware. Furthermore, the retrieval performance of the system on the exemplarily used Identa database confirms that the system can facilitate the task of visual pill recognition in a realistic scenario.
For the verification of practical use an extensive user study will be necessary. The incorporation of additional features such as brick lines or text could be the topic of future work. In the context of a human operator and higher runtime, the benefit must be evaluated. If the pill database is stored on the mobile device, a fully independent solution is possible.

References

[1] A.P. Ashbrook, N.A. Thacker, P.I. Rockett, and C.I. Brown. Robust Recognition of Scaled Shapes using Pairwise Geometric Histograms. In *Proceedings of BMVC*, pages 503–512, 1995.

[2] J. E. Bresenham. Algorithm for Computer Control of a Digital Plotter. *IBM Systems Journal*, 4(1):25–30, 1 1965.

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

[4] A.C. Evans, N.A. Thacker, and J.E.W. Mayhew. Pairwise Represenations of Shape. In *Proceedings of the 11th ICPR*, pages 133–136, 1992.

[5] R.L. Graham. An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set. *Information Processing Letters*, 1(4):132–133, 1972.

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

[7] G. Klein and D. Murray. Parallel Tracking and Mapping on a Camera Phone. In *Proceedings of ISMAR'09*, pages 83–86, 2009.

[8] D. Schmalstieg and D. Wagner. Experiences with Handheld Augmented Reality. In *Proceedings of ISMAR'07*, pages 1–13, 2007.

[9] F. Shafait, D. Keysers, and T. Breuel. Efficient Implementation of Local Adaptive Thresholding Techniques Using Integral Images. In *Proceedings of SPIE*, 2008.

[10] M. VIK. Industrial Colour Difference Evaluation: LCAM Textile Data. In *Proceedings of AIC'04*, pages 138–142, 2004.

[11] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Pose Tracking from Natural Features on Mobile Phones. In *Proceedings of ISMAR'08*, pages 125–134, 2008.

[12] S. Wijewickrema and A. P. Paplinski. Principal Component Analysis for the Approximation of a Fruit as an Ellipse, 2004.

[13] J. R. Wootton, V. V. Reznack, and G. Hobson. US Patent 6535637 - Pharmaceutical Pill Recognition and Verification system, 2003.