A framework for fast and secure packaging identification on mobile phones

Maurits Diephuis, Fokko Beekhof, Sviatoslav Voloshynovskiy*, Taras Holotyak,
Nabil Standardo, Bruno Keel[1]
University of Geneva, Department of Computer Science, Stochastic Information Processing Group
7 route de Drize, CH 1227, Geneva, Switzerland
[1] U-NICA Systems AG, Industriestrasse 4, 7208 Malans, Switzerland

ABSTRACT

In this paper, we address the problem of fast and secure packaging identification on mobile phones. It is a well known fact that consumer goods are counterfeited on a massive scale in certain regions of the world, illustrating how existing counter measures fall short or don't exist at all, as can be seen in the local absence of laws pertaining to brand protection. This paper introduces a technological tool that allows the consumer to quickly identify a product or package with a mobile device using a physical non-cloneable features in the form of a surface micro-structure image. This natural occurring identifier allows a producer or brand owner to track and trace all its products and gives the consumer a powerful tool to confirm the authenticity of an offered product.

Notation: We use capital letters to denote scalar random variables X and $\mathbf{X}$ to denote vector random variables, corresponding small letters x and $\mathbf{x}$ to denote the realisations of scalar and vector random variables, respectively. All vectors are assumed to be of the length L. We use $\mathbf{X} \sim p_{\mathbf{X}}(\mathbf{x})$ or simply $\mathbf{X} \sim p(\mathbf{x})$ to indicate that a random variable $\mathbf{X}$ is distributed according to $p_{\mathbf{X}}(\mathbf{x})$. $\mathcal{N}(\mu, \sigma_X^2)$ stands for the Gaussian distribution with mean μ and variance σ_X^2. $\mathcal{B}(L, P_b)$ denotes the binomial distribution with the parameters of length L and the probability of success P_b. $||.||$ denotes Euclidean vector norm and $Q(.)$ stands for the Q-function. $\mathcal{X}$ denotes the set of possible values of x, where $\mathbf{x} = \{x[1], x[2], \ldots, x[N]\}$. The binarized version of $\mathbf{x}$ is represented by $\mathbf{b}_{\mathbf{x}}$. A descriptor with the index k from an image with index m is denoted by $\mathbf{x}^k(m)$ and its individual elements are denoted as $x_i^k(m)$.

1. INTRODUCTION

Physical non cloneable features (PUF's) such as optically acquired micro-structures are an attractive emerging tool in the field of physical object protection. Micro-structures can not be counterfeited and as such can serve as a cheap non-evasive security and authentication identifier for packages, passports, or other physical goods next to being a natural occurring serial number for tracking and tracing applications.

Micro-structures and their process-chain share many aspects with the biometrics field. Examples include the robust extraction of the relevant structure, the selection and deployment of robust features, dimensionality reduction and quantization, and naturally, a need for fast and near error-less identification and authentication on an as large as possible database thus approaching the information-theoretic limits of identification.[15] Naturally, identification capacity, memory-complexity and robustness are conflicting requirements that should be carefully addressed in the design of identification systems.

In previous work, we introduced FAMOS [†],[14] a freely available forensic dataset comprised of micro-structures taken from 5000 unique packages with two different cameras and three acquisition rounds resulting in 30000 images in total. A printed mark was used to ensure that the specific designated region with the micro-structure was captured. Its basic statistical properties were examined and base identification performance for original, dimensionality reduced and quantized data was established.

*The contact author is S. Voloshynovskiy (email: svolos@unige.ch).

[†]`http://sip.unige.ch/famos`

We extended this work by demonstrating that FAMOS micro-structure samples can be identified from raw query images from the entire package without synchronisation based on a printed mark, by deploying robust features and a heuristic matching procedure.[2] The latter approach thus greatly enhances robustness by being partly invariant to geometrical distortions in sacrifice of earlier memory-efficient and fast fingerprinting methods.

In this paper, we will depart from the industrial and purposefully build handheld cameras in favour of a basic consumer mobile phone camera without any special lighting or adaptation. The handheld acquisition can be seen in Figure 1b and the resulting cropped out micro-structure patches in Figures 2e and 2f. Mobile acquired images differ severely from their industrial counterparts in: (a) light, (b) resolution and (c) non linear geometrical distortions and (d) limited control over the digital processing pipeline of the device.

In this work, we will consider three methods for authentication based on: (a) image-based similarity score using design-based synchronisation followed by traditional fingerprinting, (b) feature based geometric matching and (c) a hybrid method that deploys rough synchronisation and accept or rejects features based on geometric consistency according to the pre-defined family of models.

This paper is organised as follows. Section 2 gives a formal problem formulation, after which Sections 3 and 4 give an in-depth overview of all earlier acquired micro-structure databases and the authentication algorithms that were tested and deployed. Section 5 introduces two new datasets, FAMOS-M-L and FAMOS-M-S acquired with a mobile phone, and a new identification algorithm designed to be more robust against lower quality acquisitions and lack of geometric synchronisation. Finally, Section 6 concludes the paper.

2. PACKAGE IDENTIFICATION: PROBLEM FORMULATION

In the identification problem, we assume that the enrolled database contains images acquired from M objects. We suppose that each image of size $N_1 \times N_2$ is lexicographically ordered into a sequence $\mathbf{x}(m)$ of length $N = N_1 \times N_2$ with $1 \leq m \leq M$. The object that is to be tested is represented by its own vector $\mathbf{y}$ which might originate from the observation of some $\mathbf{x}(m)$ contained in the database through the degradation channel $p(\mathbf{y}|\mathbf{x}(m))$ or any randomly generated vector $\mathbf{x}'$ that is not linked with data stored in the database.

The identification problem can be considered as a composite $M+1$ hypothesis testing,[3 15]:

$$\begin{cases} \mathcal{H}_{m'} & : p(\mathbf{y}|\mathcal{H}_{m'}) = p(\mathbf{y}|\mathbf{x}'), \\ \mathcal{H}_{m} & : p(\mathbf{y}|\mathcal{H}_{m}) = p(\mathbf{y}|\mathbf{x}(m)), 1 \leq m \leq M, \end{cases} \tag{1}$$

where the hypothesis $\mathcal{H}_{m'}$ corresponds to the case when some object $\mathbf{x}'$, unrelated to the database, is presented to the system, and hypothesis $\mathcal{H}_m$ denotes a valid case where the object under consideration corresponds to enrolled data item $\mathbf{x}(m)$, $1 \leq m \leq M$. Obviously, geometric synchronization between the related items is explicitly assumed in this formulation. However, this is not the case in practice. As it will be shown below, synchronization plays a crucial role in the identification problem and designing a practical scheme capable of achieving the theoretical limits of the identification problem[15] is a great challenge.

3. IDENTIFICATION BASED ON PACKAGING SURFACE MICRO-STRUCTURES

The packaging identification problem studied in this paper is based on the surface micro-structure images acquired from packaging. To acquire a massive set of micro-structure images under different conditions, an experimental industrial system, shown in Figure 1a, was developed. This dataset, designated as FAMOS[14] consists of 5000 unique packages, photographed three times, with two camera's each, thus giving 30000 acquisitions in total.

Two colour cameras, designated Camera-1 and Camera-2 respectively, are deployed above a conveyor belt that feeds the paper samples through the system. The system can process up to 20000 samples in a single run. Lighting is identical and consists of a white led ring light together with an angled one, approximately 90mm above the surface. Camera-1 has a resolution of 2592 $\times$ 1944 (5Mp) with a sensor size of 5.7 $\times$ 4.4mm and a pixel size of 2.2 μm. It has an optical magnification of $1:0.9$. Camera-2 has a resolution of 1601 $\times$ 1201 (2Mp), a sensor size of 7 $\times$ 5.2mm, a pixel size of 4.4 μm and no optical magnification. This dataset is considered as a

baseline for our further comparisons. Sample images are shown in Figure 2(a-d). The FAMOS dataset can be publicly accessed from `http://sip.unige.ch/FAMOS`.

In this paper, we expand FAMOS with two datasets that were acquired with a consumer mobile phone, as demonstrated in Figure 1b. The deployed mobile device is a Samsung Galaxy S3 phone, which has an onboard camera producing 3264×2448 (8MP) images from the Sony manufactured IMX145 sensor.

The first set consists of 312 packages, acquired twice, designated as FAMOS-M-L. The second set, FAMOS-M-S consists of 19719 unique samples, acquired twice with a handheld mobile phone giving 39438 images in total. No special light or any other modifications were used. Further more, the packages have been acquired in different rooms to verify the impact of variations in the light conditions between the enrolment and identification phases. Sample images are shown in Figure 2(e-f).

4. STATE-OF-THE-ART: FAMOS DATASET

The success of physical object identification critically depends on a proper selection of: (a) informative features extracted from images, (b) synchronisation mechanism and (c) decision rules matched with the statistics of features and distortions.

The selection of features and synchronization mechanism might be decoupled or joint. As an example of a decoupled architecture we consider the *template based identification* where an external template or element of the graphical design is used for the synchronization and consequent extraction of a predetermined micro-structure patch after which the resulting images can be used directly as informative features. Such a design will be considered in Section 4.1.

An example of joint design is considered in Section 4.2. We refer to this design as *feature based identification*. In this case, local geometrically invariant features are extracted from the micro-structure images and used for the identification. Each local feature is characterised by a descriptor and its coordinates within the image. Object identification is based on establishing a correspondence between the probe features and features stored in the database. The correspondence model is assumed to confirm an affine or projective families of transformations, where matching is performed by one-by-one verification deploying the RANSAC algorithm for example. Leaving the complexity of matching aside, we are interested in establishing the performance limits of identification.

It should be pointed out that to accelerate the speed of matching, the list of the most likely candidates can be deduced based on a bag-of-features (BoF) framework.[4,8] The individual descriptors present within an image are transformed into a fixed-length, so called bag-of-word, vector using most commonly a pooling method, such as *max-pooling*.

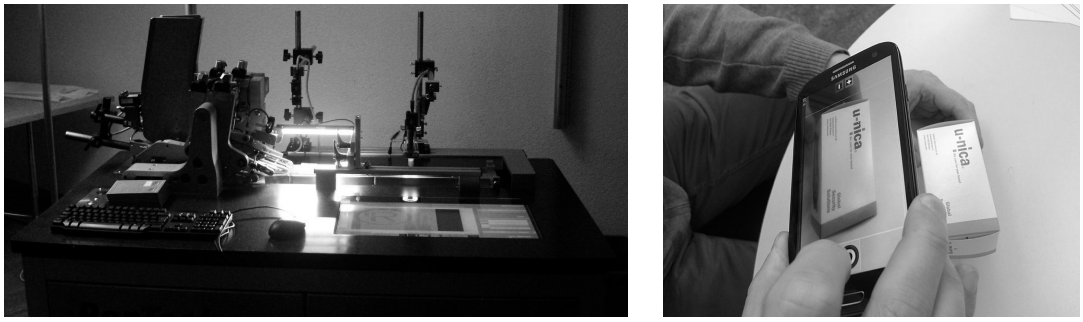
The geometric consistency between features from images from an initial list of retrieved candidates may be verified to attain a final retrieval result. The main difference with image retrieval applications, where BoF methods demonstrate state-of-the-art results, consists in the weakness and low distinctiveness of the used descriptors when they are applied to micro-structure images.

Micro-structure images show a glaring absence of prominent edges and corners where key-points or salient features are traditionally detected and the descriptors are computed. Contrarily, surface micro-structures represent slowly varying random fields that can potentially be modeled using autoregressive or Markov fields.[11]

Micro-structure features tend to be sparse and located in the lower scale-spaces. The latter makes them unreliable. This means that even in a scenario where acquisition conditions, such as the lighting, can be accurately controlled, one can expect a significant number of features that can not, or only poorly, be matched between micro-structure acquisitions.

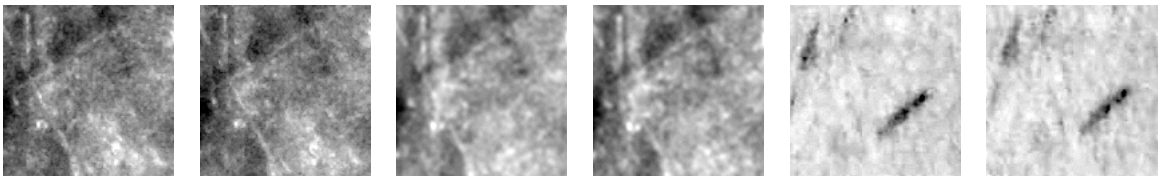
Thus, the geometrical matching of a relatively large list of candidates is of great importance for overall system performance. In fact, we will demonstrate in this paper that the traditional BoF systems fail to produce meaningful results in the considered application.

Finally, the decision rule should form a decision on the identity of the probe by producing either a unique index m or a list of indices $\mathcal{L}(\mathbf{y})$ that contains the most likely candidate indices or reject the probe. Disregarding the particular chosen system design, the decision is based on measuring a similarity between the enrolled samples



(a) (b)

Figure 1: The industrial acquisition device, including the feeder, belt, camera, lighting and integrated computer screen (Figure 1a), and a handheld mobile acquisition (Figure 1b).



(a) (b) (c) (d) (e) (f)

Figure 2: Multiple industrial acquisitions of a single micro-structure sample from Camera-1 (Figure 2a, Figure 2b) and Camera-2 (Figure 2c, Figure 2d). Mobile phone acquisitions without any special equipment or lighting are shown in Figure 2e and 2f. Histogram equalisation was used for visualisation purposes.

$\phi(\mathbf{x}(m))$ that are processed via the mapping $\phi : \mathbb{R}^N \to \mathbb{F}^J$ that includes both the synchronization and feature extraction and probe $\psi(\mathbf{y})$ processed by a counterpart mapping $\psi : \mathbb{R}^N \to \mathbb{F}^J$. The feature domain $\mathbb{F}$ can include real values, 8-bit or binary values. Assuming some distance measure $d : \mathbb{F}^J \times \mathbb{F}^J \to \mathbb{R}^+$, which is adjusted accordingly to $\mathbb{F}$, the decision rule is defined as:

$$\mathcal{L}(\mathbf{y}) = \{m \in \{1, \cdots, M\} : d(m) \leq \tau J\}, \tag{2}$$

where $d(m) = d(\phi(\mathbf{x}(m)), \psi(\mathbf{y}))$ and τ is the threshold.

The overall system performance is evaluated by the probability of miss P_M, i.e., the correct m does not appear on the decoder's list under the hypothesis $\mathcal{H}_m$:

$$P_M = \Pr\{D(m) \geq \tau J | \mathcal{H}_m\}, \tag{3}$$

and by the probability of false acceptance P_F:

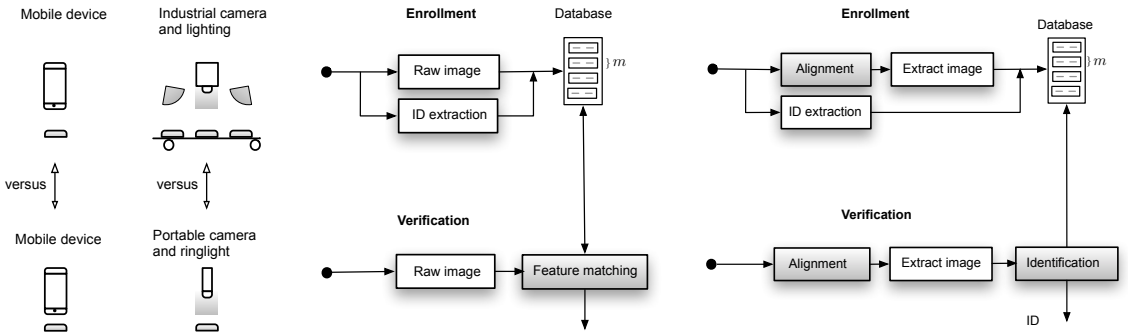
$$P_F = \Pr\{D(m) < \tau J | \mathcal{H}_{m'}\}, \tag{4}$$

i.e., an incorrect m' appears on the decoder's list under the hypothesis $\mathcal{H}_{m'}$. The average list size can be estimated as $\mathbb{E}\{|\mathcal{L}(\mathbf{y})|\} = MP_F$. In the case of unique identification, the list size is 1 that should be ensured by the proper selection of τ, M and J.

4.1. Template based Identification

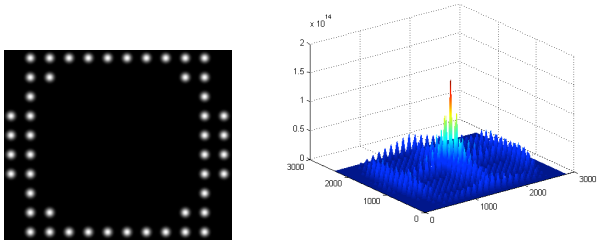
This algorithm leverages a printed mark on the package to ensure that it can extract a micro-structure patch from a predetermined region of interest and undo any geometrical disturbances, i.e., the so called *synchronisation* process. An example of such a printed mark is shown in Figure 4a. In practice many different shapes or marks can be used and have included simple letters or parts of a corporate logo.

The template based synchronisation framework can be seen schematically in Figure 3 and the algorithm in Figure 5a. It executes the following steps:



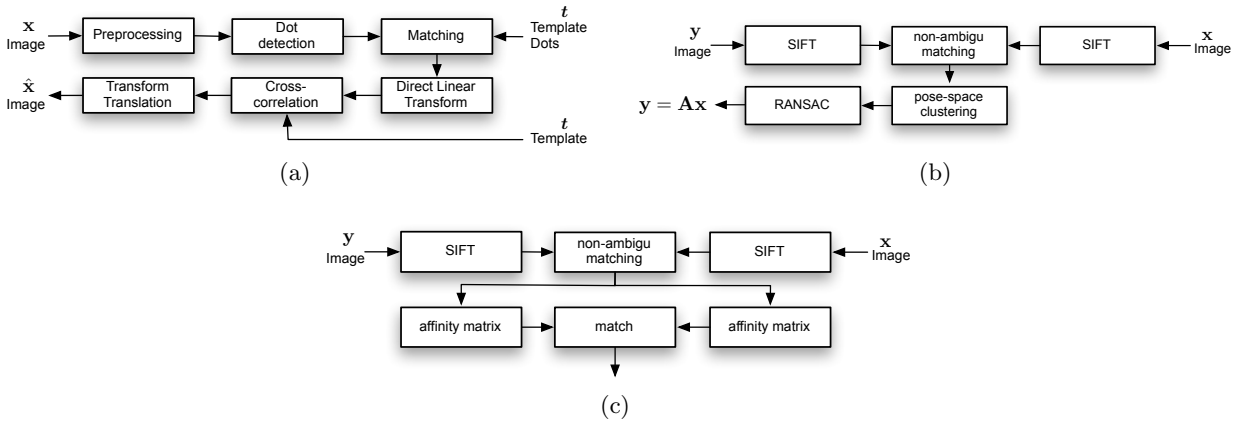
(a) (b) (c)

Figure 3: Acquisition and identification architectures. Acquisition, Figure 3a, is either via an industrial conveyor for the FAMOS dataset, or with a handheld mobile device for FAMOS-M-L and FAMOS-M-S. Figure 3b shows the identification framework when using feature based synchronisation and Figure 3c when the template based synchronisation is used.



(a) (b)

Figure 4: Figure 4a shows example of a set of printed marks to guide template based synchronisation and Figure 4b the resulting autocorrelation.



(a) (b) (c)

Figure 5: Schematics of the major components of the three identification algorithms. Figure 5a shows template based synchronisation, Figure 5b shows feature based identification and Figure 5c shows the hybrid feature based identification algorithm.

- The whole package is photographed in colour, after which the printed mark is isolated using a simple trained Gaussian model to determine the likelihood that a pixel is either part of the background or the printer mark.
- The printed mark is transformed to the auto-correlation domain, as geometrical changes such as scaling and rotation are reflected in this domain. Peaks of auto-correlation function are detected using order filtering. An example result is seen in Figure 4b.
- Autocorrelation peaks originating from the raw image and a template are matched using the Hungarian algorithm,[13] which allows the scaling and rotation between the raw image and the template to be deduced.
- Finally, the translation offset between the raw image and the template is attained via cross correlation.

Obviously, one of the main challenges of developing an optimal identification system, based on micro-structures, is the fact that the distributions $p(\mathbf{y}|\mathbf{x}')$ and $p(\mathbf{y}|\mathbf{x}(m))$ from (1) are hard to obtain even under a perfect synchronisation. That is why we will use the approximation linking the observation model with a metric space via $p(\mathbf{y}|\mathbf{x}(m)) \propto e^{-d(\mathbf{y},\mathbf{x}(m))}$ assuming an exponential family of distortions. Once, however, all micro-structures are synchronised and rid of geometrical disturbances, the hypothesis from (1) can be reformulated using an additive model:

$$\begin{cases} \mathcal{H}_{m'} & : \mathbf{y} = \mathbf{x}' + \mathbf{z}, \\ \mathcal{H}_{m} & : \mathbf{y} = \mathbf{x}(m) + \mathbf{z}. \end{cases} \tag{5}$$

Assuming that the noise is Gaussian, $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \sigma_Z^2 \mathbf{I}_N)$, the optimal decision measure is the Euclidian distance $d(\mathbf{x}(m), \mathbf{y}) = ||\mathbf{x}(m) - \mathbf{y}||^2$. Assuming a proper normalisation of all samples at both enrollment and identification resulting into $||\mathbf{x}(m)||^2 = c$, the Euclidian distance can be reformulated into the equivalent inner product similarity measure, $\rho_{x(m)y} = \mathbf{y}^T \mathbf{x}(m)$, which is a sufficient statistic[9] for the above defined hypothesises (5).

Template based synchronisation and the cross correlation coefficient $\rho_{x(m)y}$ were tested on the FAMOS dataset to ascertain the probability of miss, P_M^T and the probability of false alarm, P_F^T:

$$\begin{aligned} P_M^T &= \Pr[\rho_{x(m)y} \leq \tau J \mid \mathcal{H}_m], \\ P_F^T &= \Pr[\rho_{x(m)y} \geq \tau J \mid \mathcal{H}_{m'}], \end{aligned} \tag{6}$$

where P_M^T and P_F^T are derived with the Neyman-Pearson lemma from the *intra* and *inter* class distances. The *intra*-class is composed of all ρ_{xy} values between acquisitions originating from identical labels, and the *inter*-class from all distances between acquisitions from non identical labels. Results can be seen in Figure 6. Unsurprisingly the best results are obtained when the enrolment and verification camera are identical. Note worthy is also that the lower resolution Camera-2 outperforms the high resolution Camera-1.

In general, template based identification can achieve high performance, and as the process rids the micro-structure patches of geometrical deformations it enables the usage of advanced digital fingerprinting methods further down in the processing chain. The algorithm does hinge on the fact that a printed mark of some kind must be present and critically, the printing quality must be high. Fluctuations and noise in the printing influence the synchronisation process which in terms can hurt performance significantly.

4.2. Feature based Identification

Computer vision features such as SIFT[12] or SURF[1] have long been deployed in applications such as the stitching of pictures to form a panorama and CBIR systems. For our specific application, the extraction of a micro-structure patch, we have found such features to be insufficient. Although it is possible to use features to synchronise a template with a query image, the result is not precise enough to extract the relevant micro-structure patch from a predetermined place. It is however possible to build a feature based decision heuristic that can be used for identification while negating the need for explicit synchronisation or the presence of a printed mark.

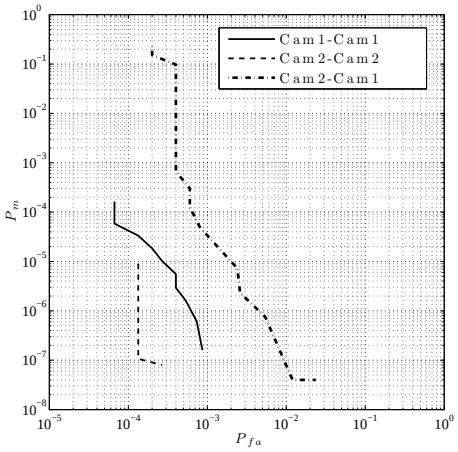


Figure 6: ROC curves for all used camera's in the FAMOS dataset and template based matching.

	Camera 1 vs 1	Camera 2 vs 2	Camera 1 vs 2
P_F^F	0	0	0
P_M^F	**0.040**	0.0773	0.0426

Table 1: The probability of false alarm, P_F^F, and probability of miss P_M^F for feature based authentication on the FAMOS dataset.

The framework can be seen in Figure 3 and the algorithm schematics in Figure 5b.

We assume that the each item in the database is represented by its features and descriptors $\mathbf{x}(m) = \{\mathbf{x}^1(m), \cdots, \mathbf{x}^{J_\mathrm{x}(m)}(m)\}$, $1 \leq m \leq M$, with each descriptor $\mathbf{x}^i(m) \in \mathcal{X}^L$, $1 \leq i \leq J_\mathrm{x}(m)$ and $J_\mathrm{x}(m)$ descriptors per image. In our analysis, we fix the number of descriptors per image to be the same for all enrolled images $J_\mathrm{x}(m) = J_\mathrm{x}$.

The problem is deciding if either a query $\mathbf{y} = \{\mathbf{y}^1, \cdots, \mathbf{y}^{J_y}\}$ is related to some elements of the database or not. In the general case, $J_\mathrm{y} \neq J_\mathrm{x}(m)$. The system should produce a list of indices $\mathcal{L}(\mathbf{y})$ whilst ensuring that the correct index m is always on this list and an empty set, if the probe $\mathbf{y}$ is not related to any item in the database.

The feature based identification algorithm takes the following steps:

- Images are acquired from the packages and stored without any post processing. SIFT features are acquired from each image surface. The SIFT algorithm is tuned to not reject weak features. The database thus contains an image, the SIFT feature vectors and a unique identifier for each enrolled package.
- To identify a package, its image is taken and SIFT feature vectors are attained.
- SIFT descriptors from the probe and the database image are matched using the so called non ambigu matching rule,[12] which only accepts a match when the second best match for that particular point is significantly worse. This factor is usually between 1.4 and 1.6.
- The algorithm coarsely filters matches using Hough pose-space clustering[10, 12] followed by RANSAC[5] and then attempts to attain a relation $\mathbf{y} = \mathbf{A}\mathbf{x}$ using least squares.
- An image for which this affine relation can be established between the probe and the database image is deemed authentic and the index of the corresponding matched features is declared as $\hat{m}$.

Results for the feature based identification on the FAMOS dataset can be seen in Table 1. Although these are obviously very promising, one should make two critical notes. Firstly, the main disadvantage is that this algorithm is computationally expensive. Secondly, as earlier stated, the FAMOS dataset is acquired using an industrial setup and although the photographed samples are misaligned, the lighting and magnification are stable. Samples are also accurately in focus. The following section will therefore cover FAMOS-M-L and FAMOS-M-S, two datasets acquired using a handheld mobile phone.

5. HYBRID IDENTIFICATION: A SOLUTION FOR MOBILE IMAGING

This section focusses on package identification based on micro-structure images acquired with a hand held mobile phone without any special equipment or lighting. FAMOS-M-L consists of 312 packages and FAMOS-M-S consists of 19719 unique packages. Both were acquired twice. These packages have a printed mark in the form of four printed crosses, to aid synchronisation and to investigate the performance limits. This synchronization is used as a counterpart of template based methods used in FAMOS.

Contrarily to the FAMOS industrial set, the mobile images suffer from non-linear distortions from the lens, geometrical distortions as the user holds the phone in all different manners while acquiring images and non even and varying lighting. Again, due to the mobile lens, images are not in focus everywhere, specifically they become soft on the borders.

The mobile sets are subjected to the following tests: (a) Template based synchronization (b) Feature based synchronisation and (c) a hybrid combination between the template approach and a modified feature based approach.

Figure 8 shows the identification results of template based synchronisation for both mobile datasets. The extracted and synchronized micro-structure patch is 1350×700 for the FAMOS-M-L dataset and 100×150 pixels for FAMOS-M-S. Performance is clearly significantly less than for the FAMOS set (Section 4.1, Figure 6), not to mention, insufficient for security applications. Contributing factors are next to unfavourable acquisitions conditions the fact that our template algorithm can not recover non linear or projective geometrical distortions.

The feature based algorithm doesn't fair better. The mobile sets suffer from a high number of false accepts when the initial features from a query image and a whole package image are matched. The algorithm is not geared to deal with sets in which the vast majority of the initial matches are outliers. An example of this phenomenon can be seen in Figure 7a.

The situation can be amended by deploying two extra steps. The first is to use the printed marks to attain a rough synchronisation. The semi-synchronized images are then stored, and used for feature based identification. Secondly, the feature based algorithm is modified to deal with a large number of outliers, by dropping the Hough pose-space clustering and RANSAC steps. Instead the algorithm compares relative distances between neighbouring points in the query image against the distances between neighbours for matching points in the database, a step reminiscent of.[6,7] And example of this type of matching can be seen in Figure 7b.

In detail the algorithm, as schematically seen in Figure 5c, takes the following steps:

- Images are all roughly synchronized using the Template based algorithm from Section 4.1.
- SIFT features and descriptors are determined and stored.
- SIFT features from a query and database image are matched with the non ambigu matching rule[12] and a relatively loose threshold of 1.2. All points without valid matches are discarded.
- For all points in the query image, the geometric distance to all other points is determined. The result is a Euclidean distance matrix, as if the points form a fully connected graph, of which the upper triangle is stored. The same is done for the matching points in the database image.
- Matching points pairs are accepted to the final set, only if their distances to neighbors in the query image is consistently identical to those in the database image with a certain scaling factor.
- A query image is deemed authentic if is has three matching point-pairs in the final set between which a relation $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be established.

Results can be seen in Table 2 demonstrating that package identification is possible using mobile phone images whilst using relatively tiny patches, be it at a computational price.

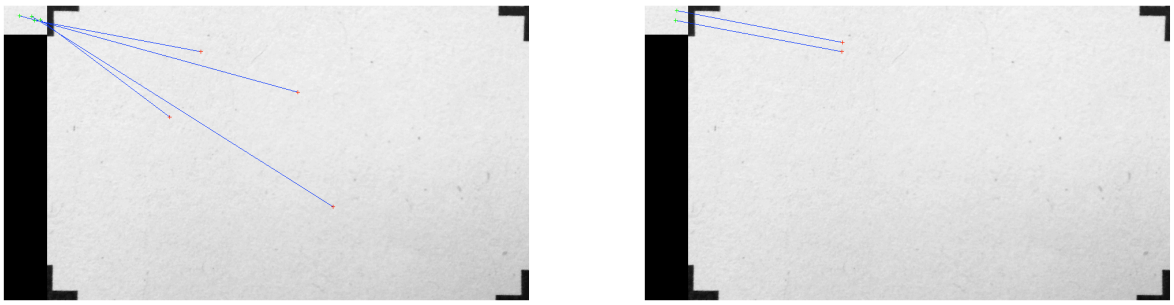
6. CONCLUSION

In this paper we have shown that it is feasible to uniquely identify a set of packages based on micro-structures that were acquired using an unmodified handheld consumer mobile phone.

To be able to cope with the large variability in image quality due to for example lighting fluctuations or the distance and angle the phone was held from the package, our authentication algorithm sacrifices speed in favor of robustness. It uses a printed mark on the package to get a rough estimate of the predetermined location of the micro-structure patch, after which it deploys a number of heuristics to match SIFT features between a presented query and the package images it has in the database. Identification hinges on the quality of this potential match.

7. ACKNOWLEDGEMENTS

This work is supported by SNF-grant 200020-146379 and the CTI project between the University of Geneva and U-NICA systems. The authors are also thankful to Christophe Gushiot for his help with the acquisition and adjustment of the FAMOS-M sets.



(a) (b)

Figure 7: Two examples showing a feature based authentication attempt between a large query image and a small enrolled micro-structure patch. Figure 7a shows the failure of the feature based algorithm (Section 4.2) due to the large number of outliers. Figure 7b shows the results for the modified hybrid approach, as detailed in Section 5.

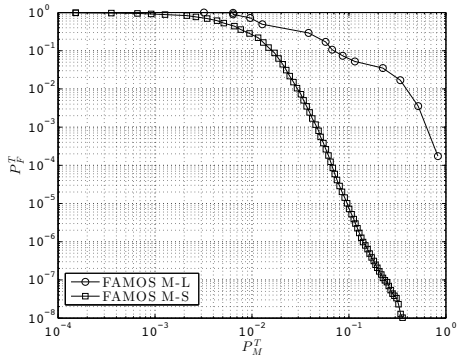


Figure 8: ROC curves for template based matching applied to FAMOS-M-L and FAMOS-M-S.

	FAMOS-M-L	FAMOS-M-S
P_F^H	0.0287	0.016
P_M^H	0.0025	0

Table 2: FAMOS-M-L and FAMOS-M-S results for hybrid identification.

REFERENCES

1. H. Bay, T .Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. In *In ECCV*, pages 404–417, 2006.
2. M. Diephuis and S. Voloshynovskiy. Physical object identification based on famos microstructure fingerprinting: comparison of templates versus invariant features. In *Proceedings of 8th International Symposium on Image and Signal Processing and Analysis (ISPA 2013)*, Trieste, Italy, September 4–6 2013.
3. Farzad Farhadzadeh, Sviatoslav Voloshynovskiy, Oleksiy Koval, and Fokko Beekhof. Information-theoretic analysis of content based identification for correlated data. In *IEEE Information Theory Workshop (ITW)*, pages 205–209, Paraty, Brazil, 2011.
4. B. Girod, V. Chandrasekhar, D. Chen, N.-M. Cheung, R. Grzeszczuk, Y. Reznik, G. Takacs, S. Tsai, and R. Vedantham. Mobile visual search. *Signal Processing Magazine, IEEE*, 28(4):61–76, 2011.
5. R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
6. Geoffrey E Hinton and Sam T Roweis. Stochastic neighbor embedding. In *Advances in neural information processing systems*, pages 833–840, 2002.
7. JHM Janssens, Ferenc Huszár, EO Postma, and HJ van den Herik. Stochastic outlier selection. Technical report, Technical report TiCC TR 2012-001, Tilburg University, Tilburg Center for Cognition and Communication, Tilburg, The Netherlands, 2012.
8. H. Jegou, M. Douze, and C. Schmid. Product quantization for nearest neighbor search. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(1):117–128, 2011.
9. S. M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall Signal Processing Series, 1993.
10. P. D. Kovesi. MATLAB and Octave functions for computer vision and image processing. Centre for Exploration Targeting, School of Earth and Environment, The University of Western Australia. Available from: <http://www.csse.uwa.edu.au/∼pk/research/matlabfns/>.
11. Dirk P Kroese, Thomas Taimre, and Zdravko I Botev. *Handbook of Monte Carlo Methods*, volume 706. John Wiley & Sons, 2011.
12. D. Lowe. Distinctive image features from scale-invariant keypoints. In *International Journal of Computer Vision*, volume 20, pages 91–110, 2003.
13. James Munkres. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial & Applied Mathematics*, 5(1):32–38, 1957.
14. S. Voloshynovskiy, M. Diephuis, F. Beekhof, O. Koval, and B. Keel. Towards reproducible results in authentication based on physical non-cloneable functions: The forensic authentication microstructure optical set (famos). In *Proceedings of IEEE International Workshop on Information Forensics and Security*, Tenerife, Spain, December 2–5 2012.
15. S. Voloshynovskiy, O. Koval, F. Beekhof, F. Farhadzadeh, and T. Holotyak. Information-theoretical analysis of private content identification. In *IEEE Information Theory Workshop, ITW2010*, Dublin, Ireland, Aug.30-Sep.3 2010.