Integration of Anti-counterfeiting Features into Conventional 2D Barcodes for Mobile Tagging

Stoyan Maleshliyski* and Fernando García*

Keywords: Anti-counterfeiting, mobile tagging, Data Matrix, security, colour

Abstract

The presented technique shows an innovative approach in the field of mobile tagging. A new high density microcode (DataGrid), combined with colour for contrast preservation, is used as a raster pattern for the macro cells of standard Data Matrix codes. As the symbols of the DataGrid microcode are optimized for the usage in security applications, when printed on a substrate, a stochastic error signal (EpiCode) occurs. The EpiCode is a result of the random structure of the substrate surface and the microscopic imperfections of the marking process. In contrast to most security printing solutions, an explicit individualisation of the product packaging is not needed, which makes the technique applicable even if using offset printing.
The observations on experimental data, indicate that the implementation of colour DataGrid into Data Matrix code: i) does not disturb the decoding algorithms of standard mobile tagging applications and ii) the authentication signal (EpiCode) has enough discriminatory information to correctly identify an original out of over 10,000,000,000 samples.
The presented concept could be used as an anti-counterfeiting tool, as the mobile tags are soon to become an inseparable part of most consumer goods packages.

1. Introduction

Mobile tagging is defined as the link between printed media and digital media. The cell phone increasingly takes the role of an input device, such as the mouse of a personal computer. The main idea of mobile tagging is to use one and two dimensional (1D, 2D) barcodes for direct connection to websites, consequently eliminating the need to input long URL names on uncomfortable mobile device keyboards.
Since most of the packages of consumer goods already have 1D and 2D barcodes for product identification and logistical needs, the medium for mobile tags is provided.

* Institute of Digital Signal Processing, Mannheim University of Applied Sciences, Germany, s.maleshliyski@hs-mannheim.de

A typical application for mobile tagging could be a product package supplied with a tag (*Figure 1*), linked to a website with up-to-date product-specific information. But could the buyer verify whether the purchased product is genuine?

Figure 1: An exemplary product package supplied with a mobile tag with integrated anti-counterfeiting features

The presented innovative technique integrates security features into standard two- dimensional codes, commonly used for product identification and mobile tagging, aiming to extend their usage in the anti-counterfeiting world, with no need of explicit package individualisation. This fact makes it applicable even if offset printing is used.
The described research is concentrated on the integration of security features into Data Matrix codes, but theoretically most of the standard 1D and 2D codes are suitable for the proposed approach.

2. Methods

The integration is achieved by a special two dimensional microcode (DataGrid [DG03]), used as a raster for the macro cells of Data Matrix, combined with different colours for contrast preservation (*Figure 2*).
The security features (EpiCode [EC03]) occur during the printing of the high density DataGrid, due to the interaction of the printing medium (ink, printing plate, etc.) and the substrate (paper, carton, etc.). These features can be used for identification as they are unique in their nature.

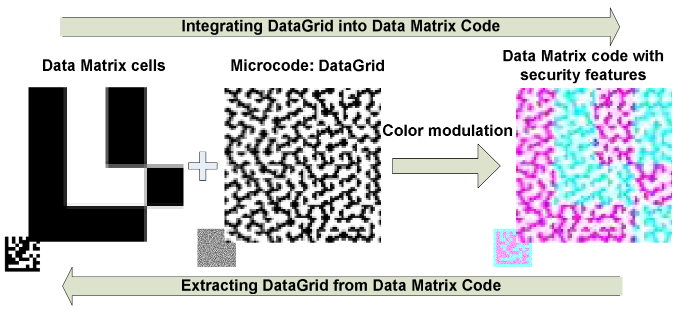

Figure 2: Method for integration of anti-counterfeiting features in a standard Data Matrix code

Data Matrix: Standard Data Matrix consists of dark and bright cells, ordered to form a square or a rectangle. Some regions of the Data Matrix are reserved for control information needed by the decoding algorithms.
For the tests in this experiment a 12x12 cell Data Matrix has been used, with two cells in horizontal and vertical direction for the quiet zone (as specified by the Standard ECC 200 (ISO, 2003)).

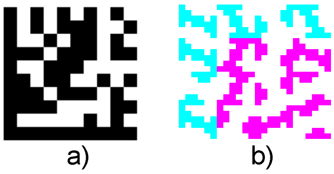

Figure 3 a) Data Matrix code, b) a part of DataGrid code

DataGrid: The version of the microcode, the DataGrid, used in the experiment, is a 2D-code that can store up to 168 bytes of information. DataGrid can be produced by conventional printing, engraving or stamping (Cameron, 2003). DataGrid symbology has eight different symbols, each consisting of 6x6 pixels, which provide means for storing three bits of digital information. The symbols are designed to be optimal for security applications, high data density storage, compensation of poor printing quality and poor scanner optics. The test DataGrid, used in this paper, consists of 36x36 symbols in horizontal and vertical direction.

EpiCode: Whenever the DataGrid is printed onto a substrate, an individual stochastic code is generated during the production process (EpiCode) (Wirnitzer, 2003).
The EpiCode cannot be reproduced; it is a naturally occurring result of the stochastic surface structure of the marked object and the microscopic imperfections of the marking process.
The stochastic microscopic pattern appears during the production of DataGrid with all commercial printing, engraving and stamping methods, in particular also with individual printing technologies like laser printing as well as with mass printing technologies like offset printing.
The EpiCode is robust, for example against distortions, disturbances and calibration as well as environmental influences such as dirt and abrasion. Present tests indicate the applicability of EpiCode under industrial conditions (Cameron, 2003).

Clustercode: DataGrids printed using offset technology show some basic components, which remain relatively constant, during the whole print process. This is because the EpiCode consists in this case of two components: (i) a stochastic component from the medium-substrate interaction e.g. microscopic ink smudging and (ii) a stochastic component from the irregularities of the printing plate. The second component will have a systematic impact on the *EpiCodes* in a cluster. A cluster is a batch of *DataGrids* on consecutive printed sheets which originates from the same master *DataGrid* on the printing plate (Bonev, 2009). So the second stochastic component which presents the printing plate signature is called *ClusterCode* (Maleshliyski, 2007).

Integration: During the process of integration there are two main process parameters which need to be combined so that both the Data Matrix and the DataGrid codes remain readable:

- *The contrast between the macro cells of Data Matrix (C_{DM})*: Recent tests have shown that standard Data Matrix code is readable if there is a contrast of at least 20% between its dark and bright cells.
- *The contrast in the microcells of the DataGrid code (C_{DG})*: Recent tests have shown that DataGrid codes are readable if the contrast between the microcells and the background is at least 30%.

These parameters are negatively correlated, which means that increasing the contrast in the microcells of the DataGrid code (C_{DG}) decreases the contrast in the Data Matrix (C_{DM}) and vice versa.

A colour separation is needed for the identification of the two different types of regions in the macro cells of the Data Matrix and so to extract a homogeneous and "pure" DataGrid code, preserving as much EpiCode information as possible (*Figure 4*). Two methods have been used for this separation: Local Histogram Stretching (Alparslan, 1981) and Principal Component Analysis (PCA) (Vidal, 2003).

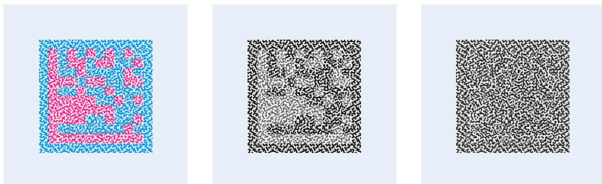

Figure 4: Colour separation and DataGrid extraction problems

- *Local Histogram Stretching:* This method analyses the distribution of the pixel values on the image scale, and adjusts them so that they are distributed over the whole range of the scale (*Figure 5*).

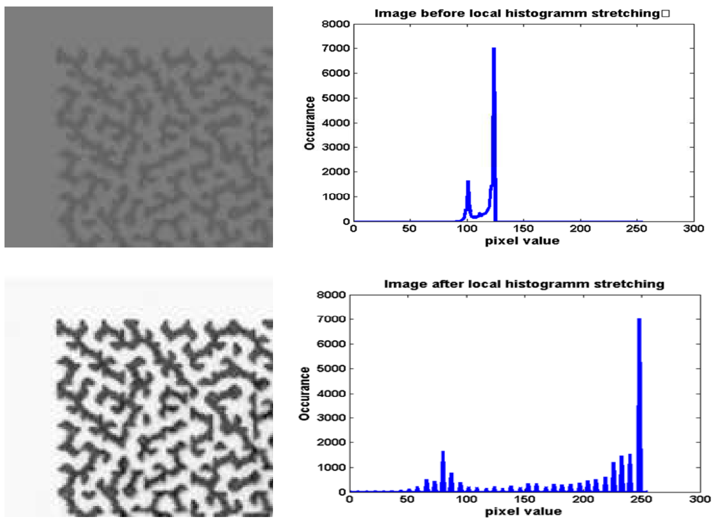

Figure 5 Local Histogram stretching

- *Principal Component Analysis:* PCA is used abundantly in all forms of analysis because it is a simple method for extracting relevant information from a complex data set, reducing it to a lower dimension, and thus revealing the basic structures, which sometimes underlie the complex data set (for more information see Vidal (2003)).

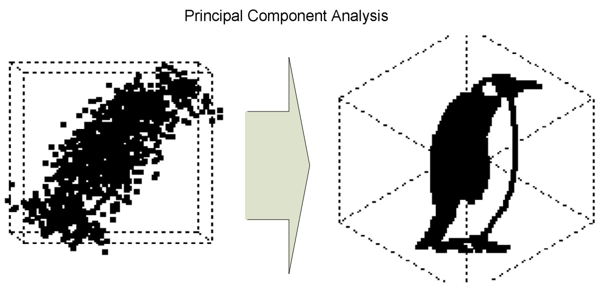

Figure 6 Principal component analysis (PCA) reduces the dimension of a complex data set, thus extracting the relevant information hidden underneath the data cloud

Quality features: The methods used for the verification of the quality of the EpiCode extracted during the tests are the same as those proposed by Bonev (2008), adapted from the biometrical analysis systems (Daugman, 1999):

- *False acceptance rate, false match rate (FAR):* This rate represents the probability that the system would incorrectly declare a successful match between the input pattern and a non-matching pattern in the reference database.
- *False rejection rate, false non-match rate (FRR)*: This is the rate representing the probability that the system would incorrectly declare a failure of match between the input pattern and the matching template in the reference database.
- *Equal error rate (EER)*: This is the error rate of a verification system, at which the operating threshold for the accept/reject decision is selected so that the probabilities of false acceptance and false rejection become equal.

3. Experimental results

The main objectives of the experiment are to explore:

- The impact on standard Data Matrix decoder caused by the DataGrid rastering;
- The influence of multicolour printing on DataGrid decoding and on the discriminatory information in the EpiCode

To understand the experimental concept an example from real life is shown on *Figure 7*. It shows a production process of product packages supplied with a DataGrid for brand protection. There are two stages in the security application:

- ***Enrolment*** (*Figure 7*, left): At this stage the security features (EpiCode) of all packages are extracted, using an intelligent camera, and are saved in a database. This database contains the reference data, which would be used in the next stage of the application.
- ***Authentication*** (*Figure 7*, right): At this stage a package is proved by a customer. For this purpose a cell phone or other handheld device with a camera is used. The device extracts the EpiCode of the package and contacts the database server. The server performs a search in the database for security features that match the ones sent. Depending on the results of the search (correlation) a corresponding message is sent to the client: a genuine product (a match was found in the database) or a counterfeited product (there could not be found any match in the database).

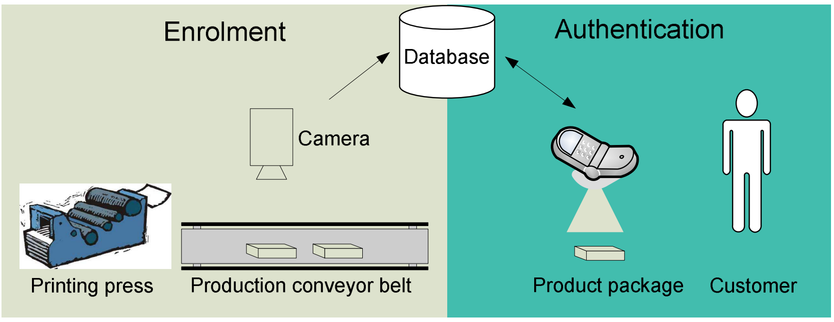

Figure 7 A practical realisation of the experimental concept

The experimental concept used for this paper is presented on *Figure 8*. A single sheet with DataGrids has been created, so that all of the test cases are provided. The sheet consists of N_{DM} Data Matrix codes (*Figure 8*, marked with different patterns, $N_{DM} = 3$) ordered so that in each column the same Data Matrix is provided. The DataGrid codes change every second row. This means that there are N_{DG} pairs of DataGrids in a single column ($N_{DG} = 2$ in the Figure).

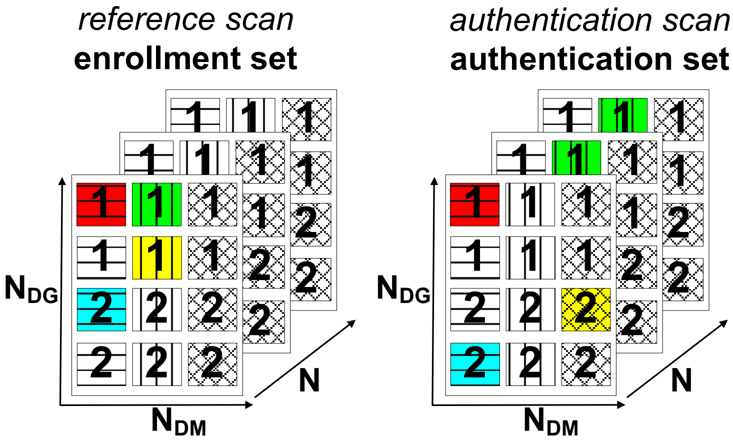

Figure 8: Experiment description: On the left, the enrolment set was used to accumulate the reference EpiCodes, which are then compared with the authentication EpiCodes from the authentication set. There is the same Data Matrix code within a single column (total count of N_{DM}) of the test sheet (marked with different patterns). There are different DataGrids in a single column (total count of N_{DG}). N is the number of test sheets used for the experiment.

Device	Model	Comments
Platesetter	FUJI TCP 36000	2400 dpi physical res.
Printing plate	FUJIFILM Brillia HD LH-PJE	830 nm positive thermal plate
Offset press	MAN ROLAND 506 LV	Six colours offset press
Printing ink	Printcom S102L	CMYK
Micro code	DataGrid	36 x 36 graphic symbols
Macro code	Data Matrix ECC 200	12 x 12 graphic symbols
Flatbed scanner	HP scanjet 8200	2400 dpi scan resolution used, 4800 dpi optical resolution
Paper grade (Substrate)	Chromolux 700	Long grain, 52 x 72 cm, 100 g/m2, 1500 Sheet

Table 1 Printing process information

In order to verify the reproducibility of the concept, multiple prints of the test sheet has been made, using offset printing (details in *Table 1*). In the next step eight different sheets have been selected. On each of the sheets reference and authentication scans were performed thus creating enrolment and authentication sets (Garris, 2006).

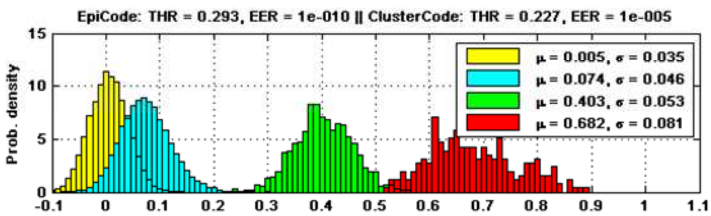

Figure 9 Results for offset printing

The EpiCodes from the enrolment set have been correlated with the EpiCodes from the authentication set. The result of each correlation is a number in the interval [-1.0 : 1.0] (with 0 indicating low correlation, and 1 indicating absolute correlation, which is possible only in an error-free environment). *Figure 9* shows the distribution of the correlation results.

- The red values correspond to correlations between *originals*, this means EpiCodes from *different* scans of the *same* print (*Figure 8*, red).
- The green values correspond to correlations of EpiCodes, extracted from *different* prints of the *same* DataGrid at the *same* position on the test sheet (*Figure 8*, green). The high correlation of this EpiCodes is caused by the *ClusterCode*. (Maleshliyski, 2007).
- Cyan values are produced from the correlation of EpiCodes, extracted from *different* prints of the *same* DataGrid at *different* positions on the

test sheet (*Figure 8*, cyan). In a real life application this would be the group of the possible *counterfeits*.

- The last distribution group, the yellow values, correspond to the correlation of EpiCodes from *different* DataGrids. These are the so called *false matches*.

The distributions shown in *Figure 9* are to be interpreted as follows (according to the biometrical analysis):

a. The "distance" between the red and cyan / yellow values is a property of the system, which gives information on how good the authentication qualities of the EpiCode are, assuming that a counterfeiter has no access to the printing plate properties. At a threshold of 0.293 an equal error rate of 10^{-10} has been achieved. This means that the possibility for the system to falsely qualify a copy as an original is 1:10,000,000,000.

b. The "distance" between the green and cyan / yellow values is a property of the system, which gives information on how good the authentication qualities of the ClusterCode are. At a threshold of 0.227 an equal error rate of 10^{-5} has been achieved. This corresponds to a possibility for the system to falsely qualify a copy as an original to 1: 100,000.

4. Conclusion

The analysis of the results shows that the following objectives have been fulfilled:

a. The false acceptance (FAR) and false rejection rates (FRR) of the EpiCode remain at an acceptable level, which was proven by the values in the previous section (EER of 10^{-10}).

b. The resulting codes were tested with the mobile tagging applications listed in Mobile Tagging (2008). The tests verify that all of the test codes are readable for the decoders used.

c. The cyan-magenta (CM) is suitable for decoding of both DataGrid and Data Matrix without additional pre-processing steps. For the combination: yellow-black (YK) a special signal processing algorithm (PCA, see Vidal, 2003) for the extraction of DataGrid is needed.

d. The cost / performance ratio of the presented technique is very good, because of the high security level provided by the EpiCode and the low cost of the technology, as it doesn't rely on any special colours or papers.

5. Future work

The future work in the project is concentrated in the following two directions:

a. Development of the concept: The evaluation of the microcode (DataGrid) gives excellent security for the recognition of counterfeited products (EER of 10^{-10}). Further research is done in the error signals produced by factors, other than the mentioned before: e.g. the naturally occurring displacement of the different colours (for example because of the multi-stage offset printing process), the use of non-base colour mixtures and lower resolutions.

A physical aspect which is to be explored is the effect of the substrate nature (paper, metal, plastics) on the qualities of the DataGrid and the EpiCode.

b. Applications of the concept: A programming challenge is the implementation of the DataGrid decoding respectively EpiCode extracting software on a mobile phone device, thus combining the mobile tagging and individualities extraction software in one. The practical problems with this aspect are: the bad optical qualities of most cell phone cameras and the lack of optical macro function (focusing problems on objects at a small distance).

Figure 10: A standard one dimensional barcode

The concept in this paper brightens the horizon for new applications. An extension of the proposed idea could be the integration of security features into the standard 1D barcode (***Figure 10***), which as well as the Data Matrix code is widely spread in the consumer goods marking. Integrations in other standard codes could also be implemented using the same technique.

6. Acknowledgments

The authors wish to thank the BMBF (Federal Ministry of Education and Research) of Germany for the financial support (O-PUR Project FKZ 02PU1153) and manroland AG for providing and printing the materials for the experiments. A grateful acknowledgment to Prof. Dr. rer. nat. Bernhard Wirnitzer for the stimulating, inspiring and always interesting discussions.

Literature

Alparslan, E. and Ince, Fuat

1981, "Image Enhancement by Local Histogram Stretching", IEEE Transactions on systems, man and cybernetics, Vol. SMC-11, no. 5, pp. 376-385

Bonev, S. and Wirnitzer B.

2005, Fachhefte Grafische Industrie, no. 3, p.30-33

2008, "Security printing for product packaging in industrial printing applications", IARIGAI 2008, Valencia, Spain

Bonev, S.

2009, "High volume security printing using sheet-fed offset press", IARIGAI 2009, Stockholm, Sweden

Cameron, M.

2003, Epyxs GmbH, Mannheim, Germany, http://www.epyxs.com

Daugman, J.

2004, *IEEE Trans. CSVT* 14(1), pp. 21 - 30.

Garris M. D., Tabassi E. and Wilson C. L.

2006, *NIST Fingerprint Evaluations and Development, IEEE Proceedings, 94, 11, 1915-1926*

Vidal R., Ma, Y. and Sastry, S.

2003, "Generalized Principal Components Analysis", Proceedings of the 2003 IEEE CVPR, 1063-6919/03, USA

Wirnitzer, B.

2003, Germany, DE10345669, "Datenträger mit Kopierschutz und Verfahren zur Kodierung solcher Datenstreifen"

Wirnitzer, B. and Maleshliyski S.

2007, Germany, DE102007059022, " Verfahren zur Erzeugung eines Sicherungscodes für einen Rasterdruckdatenspeicher und Gegenstand mit Farbrasterdruckdatenspeicher"

Mobile Tagging

Upcode, http://www.upcode.fi,
Nokia, http://www.nokia.com

ISO

2006, Information technology -- International symbology specification -- Data Matrix ISO/IEC 16022:2000