Object Recognition in Smart City

P. Füzi , M. Seebauer
Óbuda University, Alba Regia University Center, Székesfehérvár, Hungary
fuzipeter.oe@gmail.com
seebauer.marta@arek.uni-obuda.hu

***Abstract*—The Intelligent Touristic Project in Smart City helps visitors by the orientation in the case of individual sightseeing. The problem of object localization and recognition can be solved through different ways. The present paper discusses the several methods of object recognition suitable for Intelligent Touristic Project, the application tools for augmented reality, and the practical testing steps.**

I. INTRODUCTION

The city Székesfehérvár, the medieval capital of Hungary is one of the most popular destinations for tourists visiting the Middle Trans-Danubian Region (Fig.1.). The Intelligent Touristic Project is a field of implementation of Smart City concept. In the Smart City all information related to the city is stored and maintained in cloud (Fig.2.). The Intelligent Touristic Project focuses on the improvement touristic services offered to visitors looking for the individual sightseeing using a smart phone.

The smart phone is a tool connecting virtual and real world through linking information and material flow.

During the project's realization the main problem was the orientation of visitors, and the object recognition. The navigation of automatically guided tourists can be solved by using a GPS (Global Positioning System) leading them to the required object. The question was the automatic data capture from the sight to the final user by a smart phone.

Three different methods were examined. The first was QR-code based recognition, the second one was the application of Bluetooth, and the third one was the method of content based image retrieval (CIBR) following to the realization of the augmented reality.

Figure 1. The city of Székesfehérvár is rich on touristic sights

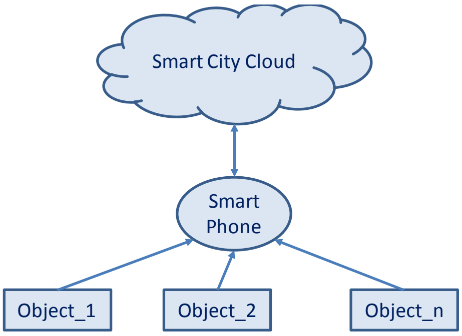

Figure 2. Object identification by using QR-code

II. IDENTIFICATION OF OBJECTS USING QR CODES

The QR (Quick Response) code is a special two dimensional barcode which was created by the Japan Denso-Wave company in 1994 [1]. The QR code (smart phone code) scanning is the most popular data capture method in Japan, but it became more widely used in USA, and also in Europe. The QR-code guarantees a high fault tolerance and therefore this media format became very popular in a relatively short time. In June 2000 on the base of his features was established an international standard ISO/IEC 18804, which was updated several times. [2,3]

The QR code is placed on posters or panels and can be photographed by smart phones (Fig.3.). The code contains a string that can represent an URL or a description. The user scans the QR code by camera, than a special application installed on the smart phone decodes the related string and through a browser application this link will lead him to the required Web-site containing the information about the actual historical sight.

Figure 3. Object identification by using QR-code

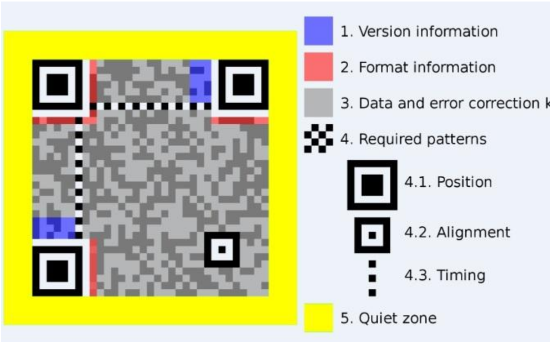

Figure 4. Structure of QR-code [4]

The structure all types of QR code is similar (Fig.4.). The two-dimensional code in form of a quadrate consist of black and white dots representing binary code. Positioning squares are placed in three of four corners that allows the fairly easy orientation by code recognition. It is allowed to take photos from different viewpoints. The reader can correctly recognize and interpret the code in any case. It means that the content can be processed through this method from any angle. Furthermore owing to the smaller positioning square in the forth corner the QR codes are scalable.

40 different type of QR code can be distinguished. The codes consist of units; therefore, the construction of larger codes is based on composition of units as well. As all of optically scanned codes the inaccuracy of the reading device or other optical uncertainties can occur to an incorrect result. This problem can be solved by using error correction methods.

In case of QR codes, Reed-Solomon coding is applied. There are four different error correcting levels: L, M, Q, H. The highest security level brings to the less effective data capacity. Level L is used mostly. By applying this level in the case of 7% of lost information the data can be restored well.

The capacity of code type 40 is the largest that can store 7089 numbers or 4296 letters. Two parameters can be defined in it: the level of error correction and the method of masking. The latter fills out the whole area. The data units are placed from right to left in crisscross. (Fig. 5.)

Figure 5 QR-code containing a string

A. Reed-Solomon codes

The Reed-Solomon code [5,6] is a special case of BCH (Bose-Chaudhuri-Hocquenghem) coding. The data storage in this format is safety while the code is wholly redundant.

1.Definition ((n; k) codes).
The coding $A^n \rightarrow A^k$; $(k \geq n+1)$ we call (n, k) code.

2. Definition (RS(n, n+k) code). Under the coding of RS(n, n+k) code we understand a Reed-Solomon code which has a k parity symbol and is created from $GF(2^{\log 2(n+1)})$.

3. Definition (generator polynomial). We call polynomial $Ag(x) \in GF(2^{\omega})[x]$ as the generator polynomial of RS(n; n+k) code if

$$g(x) = \prod_{i=1}^{k} (x - ai) \quad (1)$$

B. Coding

The following message is given:

$$M = \{m0, m1, \ldots, mL_m\};\ (m_i \in GF(2^{\omega})) \quad (2)$$

we define the so called message polynomial

$$m(x) = \sum_{i=0}^{Lm} m_i x^i \quad (3)$$

then the parity polynomial and the created code word is

$$p(x) = x^{n-k} m(x) \bmod g(x) \quad (4)$$
$$c(x) = x^{n-k} m(x) + p(x) \quad (5)$$

C. Decoding

The following received message is given:

$$R = \{r0, r1, \ldots, r_{Lr}\},\ (r_i \in GF(2^{\omega})) \quad (6)$$

we create the message polynomial

$$r(x) = \sum_{i=0}^{L_r} r_i x^i \quad (6)$$

The received message can be described with the help of the e(x) error polynomial

$$r(x) = c(x) + e(x) \quad (7)$$

Statement. The code polynomial can be described as

$$c(x) = q(x)g(x). \quad (8)$$

Proof. By replacing with $x^{n-k} m(x) = q(x)g(x) + p(x)$ we get

$$c(x) = x^{n-k} m(x) + p(x) = q(x)g(x) + p(x) + p(x) = q(x)g(x)$$

Consequence. the radical of the generator polynomial is the radical of the code polynomial as well, that is

$$c(a^i) = 0, \quad (1 \leq i \leq k)$$

4. Definiton (Syndromes). The following defined elements are called syndromes:

$$Si = r(a^i)\ (1 \leq i \leq k) \qquad (9)$$

Statement. The S_i syndromes are equal with the value $e(a^i)$.

Proof. $S_i = r(a^i) = c(a^i) + e(a^i) = m(a^i)g(a^i)+e(a^i) = e(a^i)$

As a consequence of this, if there is no error, that is $e(x) = 0$, then

$$S_i = 0\ (1 \leq i \leq k)$$

D. Determining the position of errors

Let us presume that during the sending of message v number of errors occurred in the $j_1, j_2, \ldots j_v$ indexed symbols.

In this case the e(x) error polynomial can be described as

$$e(x) = e_{j1}x^{j1} + e_{j2}x^{j2} + \ldots + e_{jv}x^{jv} \qquad (10)$$

where the e_{ji} expressed the extent of discrepancy in the i-th position.

Let us define the number of error locator now.

$$\beta_l = a^{jl} \qquad (11)$$

Here by defining the $S_1, S_2, \ldots S_k$ syndromes we get to the following equation system:

$$\begin{cases} S_1 = r(a) = e_{j1}\beta_1+e_{j2}\beta_2+\ldots+e_{jv}\beta_v \\ S_2 = r(a^2) = ej1\beta^2{}_1+e_{j2}\beta^2{}_2+\ldots+e_{jv}\beta^2{}_v \\ \vdots \\ S_k = r(a^k) = ej1\beta^k{}_1+e_{j2}\beta^k{}_2+\ldots+e_{jv}\beta^k{}_v \end{cases} \qquad (12)$$

The equation system consists of k number of equations and $2\frac{k}{2}$ unknown quantities; therefore it has an accurate solution. However, the radical of this equation system cannot be defined with the help of conventional methods, because in its β_i unknown quantities it is non-linear. The kind of method which solves the previously explained equation system is called Reed-Solomon decoding algorithm.

Let us define now the $\Lambda(x)$ error locating polynomial.

$$\Lambda(x) = (1 + \beta_{1x})(1 + \beta_{2x}) : : : (1 + \beta_{vx}) = \qquad (13)$$
$$= 1 + \lambda_1 x + \lambda_2 x^2 + \ldots + \lambda_v x^v$$

In the following stage we describe the $\Lambda(x)$ polynomial that is appropriate for defining the β_i values. We use the Forney algorithm for defining the Berlekamp-Massey and the e_{ji} values.

Note. The radicals of the error locator polynomial are the $1/\beta_i$ values, that is, the inverses of the numbers of the error locator. If the error locator polynomial is not square-free or its radicals do not fall into the $GF(2^\omega)$ field, we cannot correct the received errors.

E. The Berlekamp-Massey iterative algorithm

```
1. Berlekamp-Massey(S, L) [ S – set of syndromes ]
2. k←0,  Λ (x)←1, T(x)←x
3. FOR i ←1 TO |S| DO
4. δ←0
5. FOR j ←1 TO k DO
6. δ←δ + λjSi-j
7. ENDFOR
8. δ←δ- Si
9. IF δ≠0 THEN
10. Λ(x)  ←Λ(x) - δT(x)
11. IF 2k < i THEN
12. T(x)←T(x) + Λ(x)=/δ
13. k ←i - k
14. ENDIF
15. ENDIF
16. T(x)←xT(x)
17. ENDFOR
18. L← Λ(x)
19. END
```

F. Forney algorithm

Let us constitute from the S_i syndromes the following polynomial with infinite scale number

$$S(x) = S_1x + S_2x^2 + . + S_kx^k + S_k+1x^{k+1} + \ldots \quad (14)$$

Let us define the Ω (x) error size defining polynomial

$$\Omega(x) = [1 + S(x)]\ \Lambda(x) \qquad (15)$$

Since we only know the first k number of coefficients of the S(x) polynomial, the decoding problem becomes equal with the definition of the k/2 scaled $\Lambda(x)$ polynomial which satisfies the following equation $\Lambda(x)[1 + S(x)]\ \Omega(x) \bmod x^{k+1}$

Then the coefficients of the error polynomial can be described as

$$e_{ji} = \frac{-\beta_{ji}\, \Omega\, \beta_{ji}^{-1}}{\Lambda'(\beta_{ji}^{-1})} \qquad (16)$$

and the error polynomial as

$$e(x) = \sum_{i=1}^{v} e_{ji} x^{\beta ji} \quad (17)$$

After the definition of the error polynomial, we can get the original message back by the subtraction of the error polynomial

$$m(x) = r(x) - e(x) \qquad (18)$$

The advantage of QR code is that smart phone is able to read it, for example the Kaywa Reader has a client written in Java ME language, which works on most traditional devices, too.

The preference of this method is the low implementation costs but the system is strongly sensitive to the physical influences destroying the contrast of image and causing a failure at recognition. The QR code can be placed on the sights at a not disturbed place. The disadvantage of the system is the barcode's vulnerability. Furthermore, the weather conditions (rain, wind, fog) can influence the scanning process.

TABLE I.
COMMUNICATION DISTANCE OF BLUETOOTH DEVICES

Class	Maximum permitted power		Range
	(mW)	(dBm)	(m)
Class 1	100	20	~100[10]
Class 2	2.5	4	~10[10]
Class 3	1	0	~5[10]

During the research work, the idea of the Bluetooth based identification turned up. At first glimpse, it seems an operating-safer method than the QR code based conception.

III. THE DEFINITION OF THE OBJECT BY MEANS OF BLUETOOTH

The Bluetooth [7] is a short range, open, wireless standard method for automatic data exchange between low power mobile digital devices (computers, peripherals, phones, etc.) within 2,4 GHz frequency area. It operates by frequency jumping transposition with wide spectrum. Devices depending on their classes are able to communicate within the several distances (Table 1.)

Applying 1st class (100 mW) Bluetooth, the range is 100 meters, which can be reduced due to unfavorable environmental effects. This solution is more convenient, as the QR code in point of view that the communication between the devices can be established without optical link between them.

The advantage of this type of communication between the smart city cloud and a smart phone is the widely used standard communication interface. The weakness of this method is, however, the high establishment and operating costs of transmitters installed at the sights.

IV. THE OBJECT RECOGNITION THROUGH CONTENT BASED IMAGE RETRIEVAL

The content based image retrieval (CBIR) was developed in the Óbuda University as a new method of image recognition [8]. The basic idea of this method is the search in the collection of images by the distinctive features, as color, texture or shape basically in Internet.

In Intelligent Touristic Project we inspected this method for searching in database of sights. The database contains

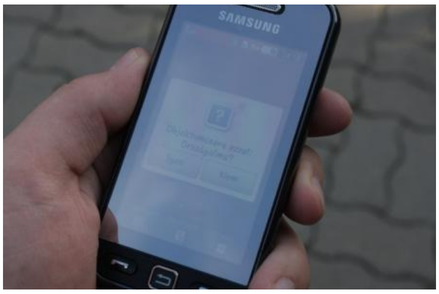

Figure 6. Object identification using Bluetooth

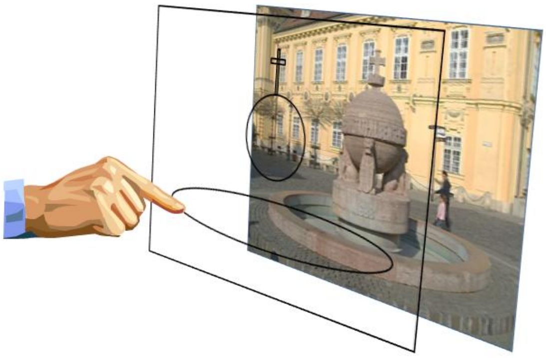

Figure 7. Object identification using CBIR

the shape specific information about the object as a key information.

Fig. 7. presents the idea of this method. The user takes a picture from the related object, then he draws its typical shape on the touch screen. This shape will be used as the searching key in database of sights. By searching process the best fits result a list of the hits in order of best correlation.

This method of recognition will follow to the very perspective applications in augmented reality. The efficiency of this method is satisfying, but regarding its complicatedness, any of the previously mentioned two solutions would be easier to apply.

V. CONCLUSIONS

The material considerations and experience acquired in Intelligent Touristic Project allowed us to draw the conclusion that most efficient solution is the QR code implementation. During the execution of the project, the designer team in first order will apply this method.

REFERENCES

[1] http://www.denso-wave.com/en/adcd/index.html, Retrieved October 28, 2012

[2] http://archive.is/20120915/http://www.qrcode.com/en/qrfeature.html, Retrieved October 28, 2012

[3] BS ISO/IEC 18004:2006. Information technology. Automatic identification and data capture techniques. QR Code 2005 bar code symbology specification. London: BSI. 2007. p.126. ISBN 978-0-580-67368-9.

[4] http://www.lexiconnect.co.uk/qr-codes.html Retrieved October 28, 2012

[5] Robert G. Gallager "Information Theory and Reliable Communication". ISBN: 978-0-471-29048-3, Wiley, January 1968.

[6] Irving S. Reed, Gustave Solomon, "Polynomial codes over certain finite fields". Journal of the Society for Industrial and Applied Mathematics. [SIAM J.], 8, 1960, ISSN 0036-1399, pp. 300–304.

[7] http://www.bluetooth.com/Pages/Bluetooth-Home.aspx Retrieved October 28, 2012

[8] B. Szántó, P. Pozsegovics, Z. Vámossy, Sz. Sergyán, "Sketch4Match – Content-based Image Retrieval System Using Sketches" 9th IEEE International Symposium on Applied Machine Intelligence and Informatics, SAMI 2011, Smolenice, Slovakia, ISBN 978-1-4244-7428-8pp. 1–6, January, 2011.