3D CAPTCHA

Captcha based on spatial perspective and human imagination

Author: Juraj Rolko, Žilina, Slovakia (rolko@3dcaptcha.net)

Date: 01/10/2010

ABSTRACT

This document describes a new type of CAPTCHA technology based on spatial perspective and human imagination. The basic idea is rotation of a special 3D model and finding the correct position of rotation. 3D model can be created from any 2D image. First, the picture is divided into several parts, which are then randomly projected into 3D space, subject to certain rules. Parts projected spatially this way generate together 3D models so that the model looks like the original 2D image from a single observation point. The meaning of 3D model observed from any other observation point is incomprehensible for human. The task for a user is to rotate the model to find the right observation point and solve the CAPTCHA successfully. The technology is based on the fact that the meaning of the model is recognizable from the correct observation point only by human. Automated system would not be able to achieve the goal without using a high degree of artificial intelligence. The document describes in detail the 3D CAPTCHA technology principle, method of implementation, examples, comparison of the main advantages and shortcomings and, ultimately, the analysis of safety and security risks.

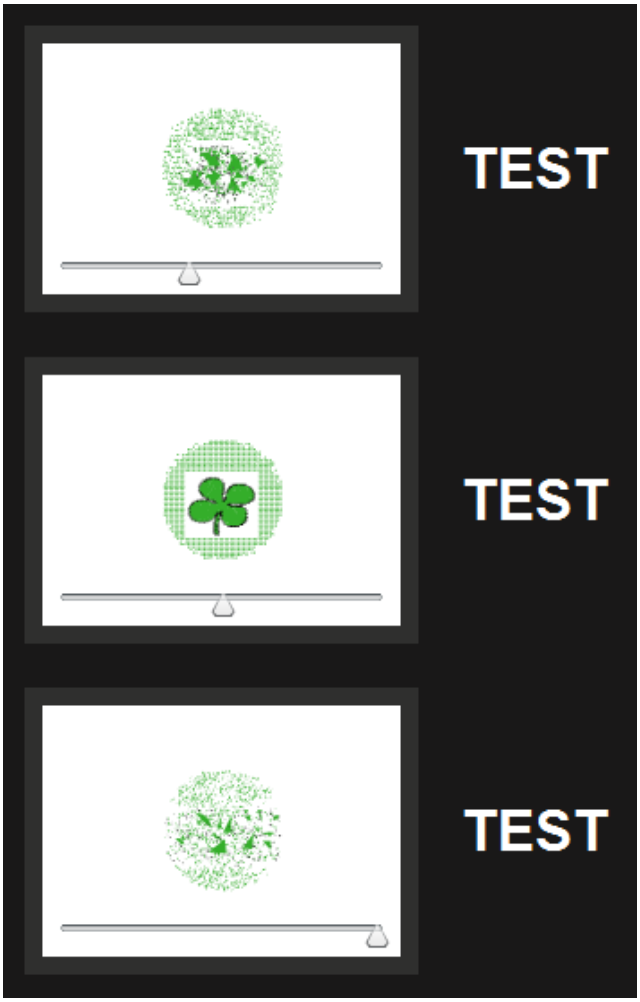


Fig. 01 – example of 3D captcha. The user's goal is to rotate 3D model and find the correct position (position where the meaning is understandable for human)

General terms
Human Factors, Security, Experimentation

Keywords
CAPTCHA, security, 3D models, spatial perspective

1. INTRODUCTION

CAPTCHA (completely automated public Turing test to tell computers and humans apart) is a commonly used security feature of websites that prevents from unwanted access to automated systems (so called bots) to certain parts and functions of Internet applications. Nowadays there are many variations of CAPTCHA system, such as:

- Text-based CAPTCHA
- Text-based CAPTCHA which uses dynamic elements
- Image-based CAPTCHA based on rotation of images (what's up captcha, sketcha [1],[3])
- Video-based CAPTCHA [4]

The one of the most commonly used variations is the text-based CAPTCHA, which uses principle of description of randomly generated alpha-numeric characters. The task of a user, investigator of the test, is to recognize these characters and to rewrite them in input box on website.

Fig. 02 – text based CAPTCHA example 1

Fig. 03 – text based CAPTCHA example 2

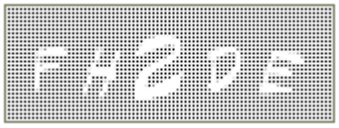


Fig. 04 – text based CAPTCHA example 3

Characters tend to be deformed in general. They also use different contrast backgrounds. The disadvantage of the method is that the pattern recognition techniques are constantly improving and overcoming such method of computer protection systems is getting easier. More and more deformed text strings with more and more complex background is being used as a prevention from breaking through, which makes the text to be unreadable to people. In addition to the text-based CAPTCHA version, there are other alternative modifications, such as different alternatives with dynamic content changes, overlapping letters, and also CAPTCHA based on rotation of images and setting their correct horizontal position.

The new 3D CAPTCHA system aims to used the spatial perspective and human imagination and ability to recognize real objects in the context of random graphical information. It is based on a simple principle. First, any separate parts are created from any 2D picture. These are then projected into 3D space. These elements together generate a 3D model in space. Elements are reflected into the space randomly, but with the aim for 3D model to look as an original 2D picture from a certain angle of vision or observation point. The task of a user, solver of the CAPTCHA, is to rotate the model and find the correct observation point. The right solution is basically only guessed by the user. Yet, the process can easily come true thanks to experience, human imagination and intelligence. The basic principle of 3D captcha system closely resembles the techniques of optical illusions, where a group of several objects scattered in space does not make a meaningful object as a whole from many observation points (does not resemble a real object), but it makes sense from one particular observation and people can easily tell that the observation point is an appropriate one.

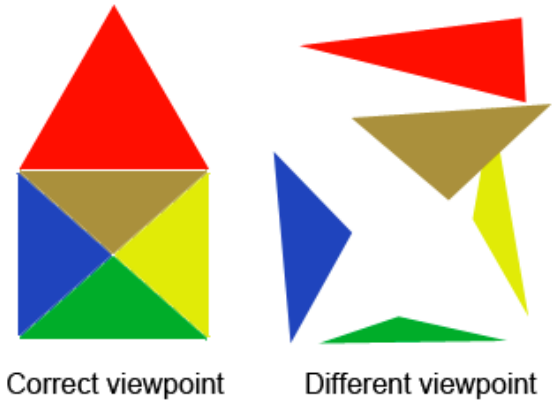


Fig. 05 – spatial optical illusion example 1

Fig. 06 – spatial optical illusion example 2 (penrose triangle)

Identification of a correct object in such case is extremely difficult for the machine, automated computer system, as the objects can be virtually anything. Furthermore, the object may even lack any realistic proportions. It may even be just an abstract reflection of a real object. Any picture used in creation of a model be modified before the transformation to a 3D model additionally, i.e. changing the shape, applying various graphic filters to change a colour, contrast and so on. Variety of modifications can be made from one image. Database of usable files is basically infinite. The only way to solve the 3D captcha test is to use the human imagination. If there was a computer system capable of such a degree of recognition and interpretation of graphic information, any other captcha based on image perception would not have been applicable any more.

The main focus in crating the 3D CAPTCHA technology was put on an innovative, brand new and different approach to solving task aimed at distinguishing human from computer system. A key task was to incorporate human abilities and characteristics, which cannot be emulated by computer systems, into the technology. Moreover, it was obvious to take into account as simple and fast as possible method of solving the task in terms of user, user interface and comfort.

In general, good captcha system must meet three basic properties [1]:

1. Easy for most people to solve
2. Difficult for automated bots to solve
3. Easy to generate and evaluate

3D captcha system meets all the three conditions, which is significant difference compared to other image-based captcha systems. They have great problems meeting the third condition above all. In that case it is very difficult to ensure good enough quality database of source documents, or the creation of the database is very expensive and inefficient.

The existing prototype 3D captcha system along with basic description can be found at:

www.3dcaptcha.net

2. 3D CAPTCHA

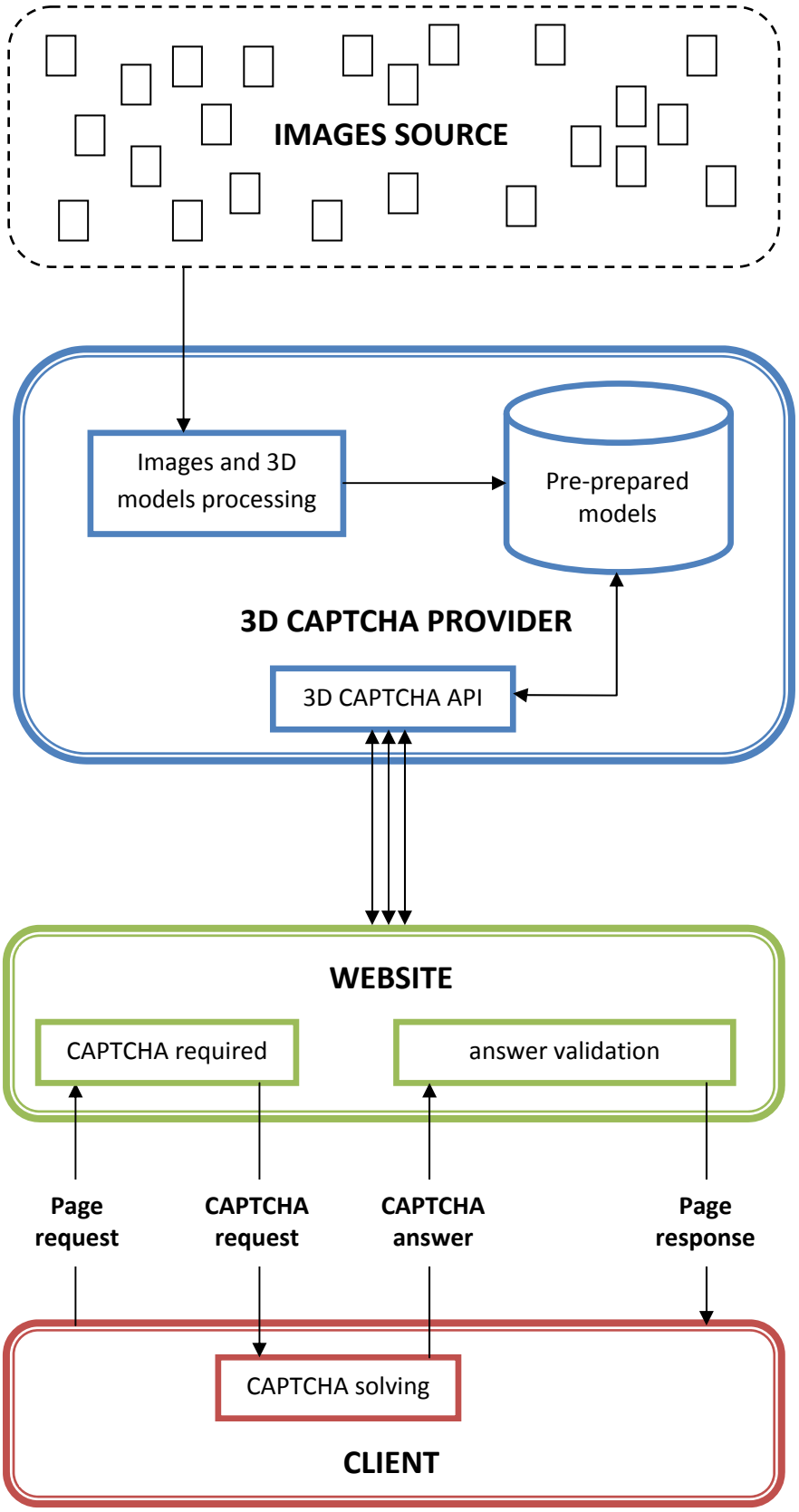
It has been mentioned several times that 3D captcha is based on the space perspective of objects. The main objective is to solve the captcha test by rotation of 3D model in user interface and to find or guess the right position of the 3D model (right rotation angle of 3D model).

Fig. 07 – wrong and correct 3D captcha solution

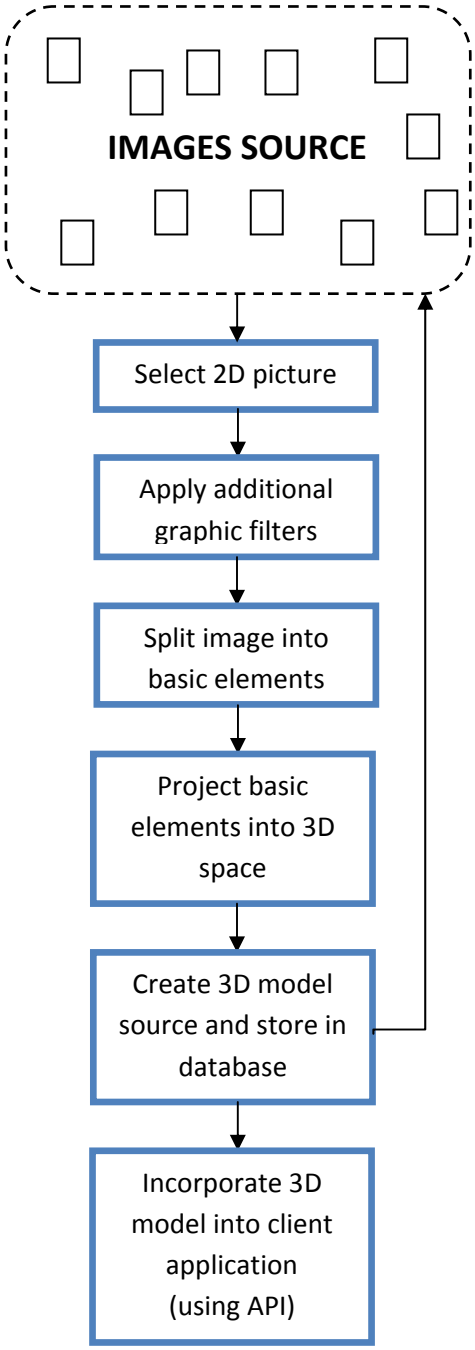
The whole concept of the technology can be divided into two parts. Client section and a captcha provider section. The client part includes solution of the captcha test itself by uses through the interface. Provider ensures and provides a database of prearranged source documents, models and respective interface for programming interface. Web application creators may then integrate the interface itself, client part of the 3D captcha system in their new or existing projects through using of API interface. Thanks to the provider a database of quality and proven source models will be ensured. API interface will be used to provide 3D models, evaluation of the captcha test success and to provide any other functions for the proper functioning of the technology. The very simplified process of creating 3D captcha system, or a kind of prototype solutions can be described by the following six steps:

- **STEP 1: Selection of an appropriate source 2D image**
- **STEP 2: Application of additional filters to 2D image**
- **STEP 3: Division of the image into basic elements**
- **STEP 4: Projection of all basic elements in the space**
- **STEP 5: Creation of 3D model out of the projected elements**
- **STEP 6: Integration of the model into 3D CAPTCHA client application**

2.1 Project design (provider,website,client part)



2.2 3D captcha prototype creations



2.3 Interface

The aim in solving 3D captcha system is to rotate 3D model and to find the point where the model is comprehensible for the user, investigator of the test. It means the point at which the user recognizes an object and evaluates this object as the correct solution. The principle is implemented in the prototype 3D captcha system as a flash application that uses Papervision3D technology [6].External file with 3D model definition is loaded in the flash application. This model is then displayed in the flash animation. The user can then rotate the model using the slider. Through his imagination the user evaluates the most appropriate solution that actually corresponds to a certain degree of rotation of the model. The estimated solution is compared with the correct solution, which is provided for the given model by the 3D captcha system provider. In the case of match the captcha test is evaluated as successful, otherwise it is evaluated as unsuccessful. The principle of operation is basically identical to the principle on which text captcha systems work. The only difference is that in the text captcha systems a text string is sent to the provider for verification, while in 3D captcha system a numeric value / values, which represent the angle of rotation of the model, is sent. In the actual implementation any number of models can naturally be used. As a rule, the more models are used, the less likely possibility of breaking the test computer system exists. However, the user demands for test increase and user tolerance to test decrease at the same time.

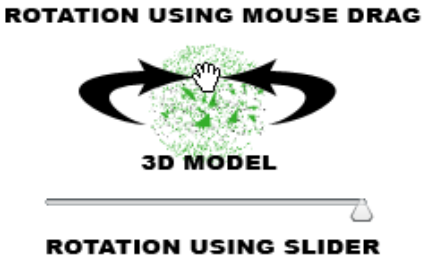


Fig. 08 – 3D captcha client interface principle

2.4 STEP 1 - selection of source 2D image

The first step in the process of creating 3D CAPTCHA technology is the selection of a suitable graphic material for further processing. The big advantage is that it is possible to use any common image, which has some information relevant to people. Of course, that not all types of images can be used for these purposes.

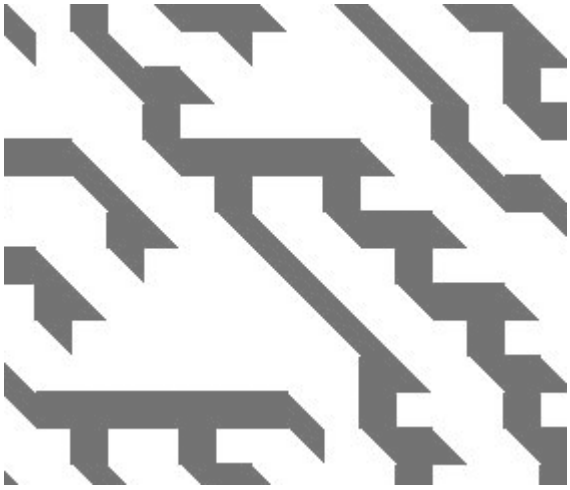


Fig. 09 – unsuitable source picture example

The above image is improperly chosen, since it is the graphic information with very low semantic value. Fortunately, the percentage of these types is very low.

Fig. 10 – suitable source picture example

Unlike previous picture the fig. 10 is suitably chosen type, as its meaning is clear for people immediately. It is picture of a four-leaf clover. People can recognize a real object despite the fact that it is only a simplified, abstract form of a realistic display of an existing object. By the selection of these types of images a database with plenty of source data for 3D CAPTCHA system is created. An important difference compared to other CAPTCHA systems based on the use of images or 3D models is that in this case the source images do not have to undergo such a in-depth scrutiny of the fitness for use. Therefore, the initial creation of the database, filling of the database with appropriate pictures is incomparably faster and more efficient. There is no need for sophisticated image analysers of suitability for use and the amount of manual human intervention in selection is minimised.

2.5 STEP 2 - Additional modifications to the source image

Application of various graphic filters can modify the original image and many other usable combinations can be gained. The change of shape, colour, contrast, resolution and so on can be applied. In general, the smaller resolution of a picture or the bigger simplification of a picture, the more demanding recognition by some automated systems is.

2.6 STEP 3 - Creation of basic elements

A good graphical method creates from the source image a set of basic elements. The image is "cut" into smaller pieces. These are the smallest elements to be projected into space. The basic elements form a shape of triangle. Each triangle is thus defined by its coordinates (x, y) and its colour.

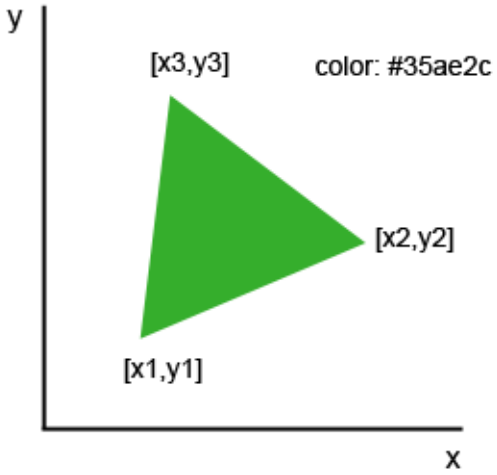


Fig. 11 – basic element

When a triangle is projected in the space randomly, the size and shape of the original basic element cannot be defined. Should rectangles are used instead of triangles the attributes, such are shape and size, of original 2D elements could be estimated out of the group of several projected rectangulars by application of appropriate mathematical methods. Another important condition is to have the basic elements of different shape and size. If all the triangles had the same shape and size, again, an appropriate mathematical method could be used and attributes could be calculated precisely, or estimated precisely enough, out of the original elements.

Fig. 12 – decomposed 2D picture (set of basic elements)

In the figure above shows the decomposition of the image into the basic elements. After dividing the original image into basic elements we obtain a set of elements with known coordinates and colours. These elements are then stored into the database in an appropriate form. Information will then serve us for the construction of the 3D model.

2.7 STEP 4 - Projection of all the basic elements into the space

Now that we have created respective basic elements, we can proceed to construction of the 3D model itself. From that moment on we are beginning to work in 3D space. The first step is the determination of so called original observation point. It is a place from which the 3D model will have the correct shape. It is a point, the initial location of an observer from which the calculations for 3D model construction are derived. It is also the point which has to be found in solving 3D captcha test. Choose of the original observation point is optional. But as it was already mentioned, the resulting 3D model shape depends on its position.

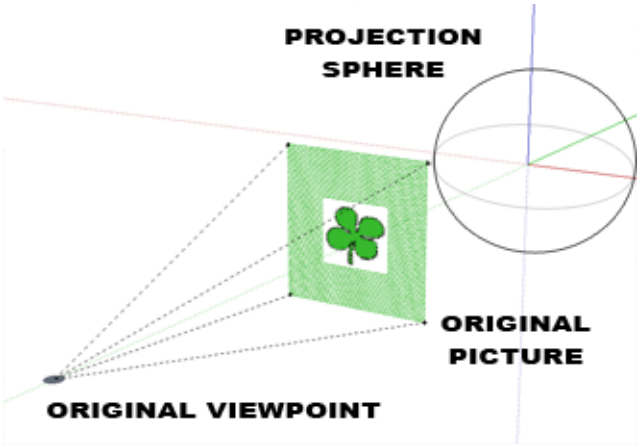


Fig. 13 – original observation point, 2D picture, projection sphere

When we have the original location of the observation point chosen, we shall determine the position and size of so called projection sphere. This is a projection space to which the particular basic elements will be projected on the basis of the coordinates of the original observation point, its own coordinates, the coordinates of the projection sphere and its size.

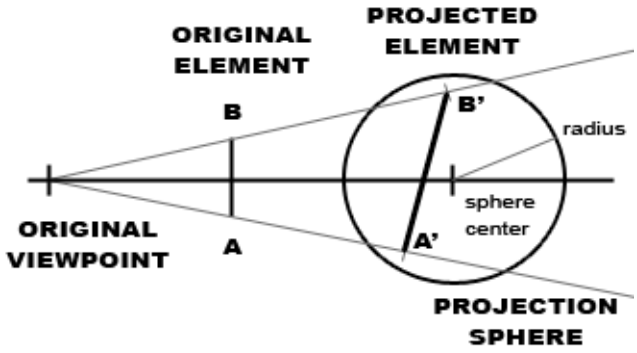


Fig. 14 – example of the projection of a triangle side

Each projected element will have to be inside the projection sphere. Exact projected coordinate is chosen randomly. Random projection of triangles in the space basically ensures that the meaning of the

model from other than the original observation point will not be comprehensible for humans. The projected coordinates must fall within the projection sphere. In case where at least one point of the triangle, the basic element, cannot be projected into the projection sphere, the given element is excluded (will not be projected).The sphere is chosen because its shape from any observation point is always the same and the shape of the projection area is not important for observer in respect of information. Since the original source image is a rectangular and the projection is made into the sphere, it is obvious that some part of graphic information can be lost. Given the previous facts, it is advisable to choose the characteristics of the original observation point and the projection sphere, so that key graphic information from the original 2D image is preserved.

Fig. 15 –possible loss of graphic information

When converting from 2D to 3D space the basic trigonometric functions are used. Any suitable basic element with coordinates x1, y1 will be projected according to the set rules in the space and there will be some resulting coordinates x1 ', y1', z1 '.

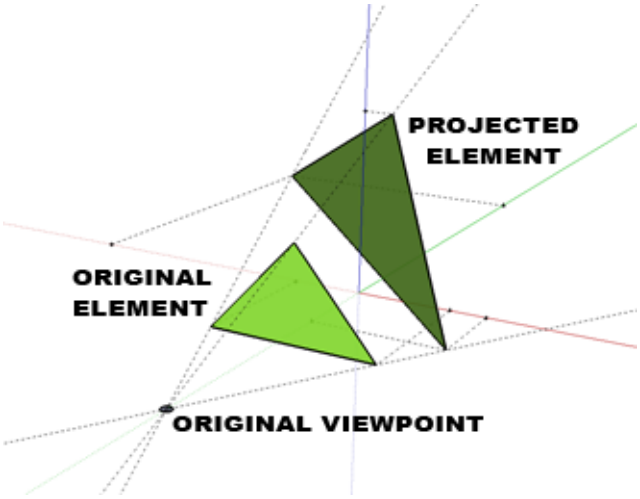


Fig. 16 – element projection

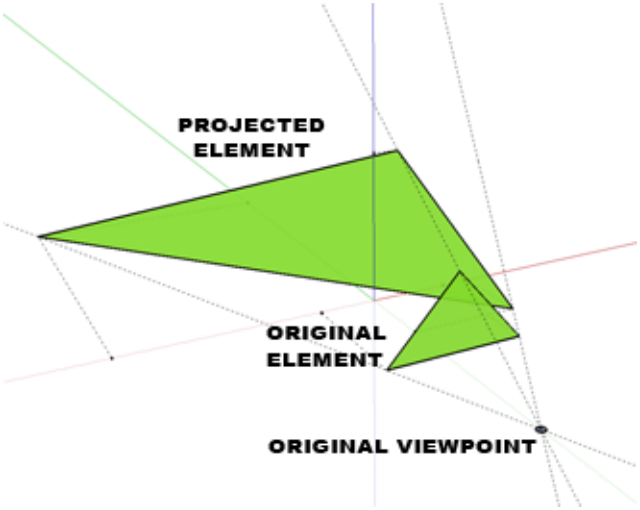


Fig. 17 – element projection - different view

As can be seen from the previous two images 3D perspective has a major role in the captcha system. The projected element is completely different from different angle of vision then from the original observation point. And as the link to a basic element is lost after the projection, it is impossible to determine its original shape from the projected element itself.

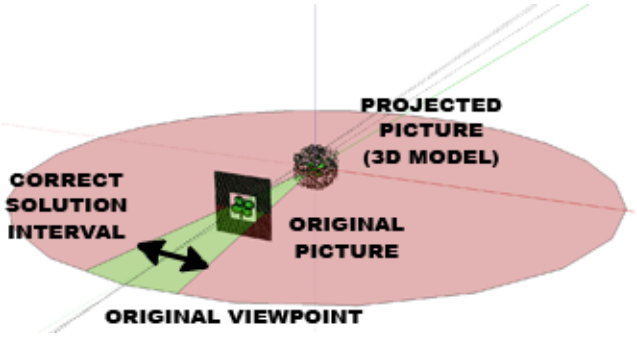


Figure 18 – all elements projection

The above procedure provides calculation of coordinates for all the basic elements. That way we will get a set of data, which are subsequently used to generate a set of 3D models (eg DAE, KMZ, and others).

2.8 STEP 5 - Creation of 3D model out of the projected elements

We now have all the basic elements projected into the space. Each element has thus its coordinates x, y, z defined. At the same time, each element is assigned a corresponding colour. This colour matches the colour of the corresponding basic element. These data are fully satisfactory and based on the elements defined this way, we can create a particular set (egCOLLADA DAE formate, KMZ, etc.).

```
<?xml version="1.0" encoding="utf-8"?><COLLADA
xmlns="http://www.collada.org/2005/11/COLLADASchema" version="1.4.1">
    <asset>
        <contributor>
            <authoring_tool>Tool</authoring_tool>
        </contributor>
        <created>2010-05-08T21:22:47Z</created>
        <modified>2010-05-08T21:22:47Z</modified>
        <unit name="inches" meter="0.0254"/>
        <up_axis>Z_UP</up_axis>
    </asset>
    <library_materials>
    <material id="material_0_8ID" name="material_0_8">
        <instance_effect url="#material_0_8-effect"/>
    </material>
    <material id="material_1_8ID" name="material_1_8">
        <instance_effect url="#material_1_8-effect"/>
    </material>
    ...
    <library_geometries>
        <geometry id="mesh1-geometry" name="mesh1-geometry">
            <mesh>
                <source id="mesh1-geometry-position">
```

Fig. 15 – DAE file example

This file is then used directly in the client application to display the CAPTCHA. These formats are supported for example by Papervision3D, and others. Hence, the choosing a file format for 3D model is optional. It depends on what technology will be used for creation of interface, client part of 3D captcha system. The said Papervision3D technology supports many file types for 3D models, such as DAE or KMZ format.

2.9 STEP 6 - Inclusion of model into the client application to solve CAPTCHA test

After all the previous steps is now possible to create very functional 3D CAPTCHA. Client application can be created by various technologies. In our prototype solution flash in combination with Papervision3D were used. However, with the launch of HTML5 technology, it will be easy to implement this system using other tools too. The whole operation of 3D CAPTCHA system presupposes the existence of provider who would provide a database of prepared 3D models and also presence API interface for communication PROVIDER <-> WEB APPLICATION. Owner, creator of the web application would use simple flash application on your web site and would communicate with captcha providers through API.Verification of the correct solution would again took place in cooperation with 3D captcha providers in cooperation with web application.

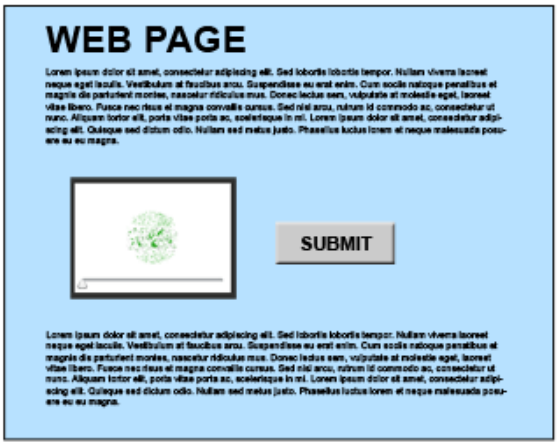


Fig. 16 – web page integration example

3. SECURITY ANALYSIS

Security and resistance to unwanted attacks are essential requirements to be met by quality captcha systems. 3D captcha is a system that uses human imagination and ability to distinguish between different types of objects. In the process of constructing 3D captcha interruption in the connection between the original source 2D image and 3D model occurs, which can be seen by the user in solving the captcha test. The right solution is therefore impossible to obtain by any calculation methods. The right solution is basically just estimated solution. Correct estimate, with high accuracy level, can only be made by human thanks to its advanced intelligence compared to nowadays computer systems. It is likely that if there were computer system with the artificial intelligence with a comparatively high level, so any captcha based on recognition of graphic information would not be applicable any more.

According to standard of captcha systems, in case of possible random estimation of at least 1:10000, the captcha system is considered acceptable. For 3D captcha the model can rotate in the range of 360° and therefore in the case of using one model there is a probability of 1/360 that automated system that can determine the correct solution by random estimate. In practical deployment of technology it is therefore necessary to use several models simultaneously within a single captcha test. In view of the user's comfort, it is also necessary to set certain tolerance level applicable to the solution. When using 3 models, each one with the tolerance level of 5° , there is a likelihood of accidental guessing the correct solution $(5/360)^{3,}$ which is approximately 1:373248. When using 2 models with a tolerance level of 3° , there is a likelihood of accidental guessing the correct solution $(3/360)^{2,}$ which is approximately 1:14400. Both these probabilities are sufficient for such a captcha test to be acceptable. The 3D technology captcha concept naturally allows using horizontal and vertical rotation of the model simultaneously in one model. This method of implementation is however difficult to solve even for humans.

Given that within the 3D captcha system almost any image as a source can be used and any additional graphic filters can be applied to these images, the database of source materials is virtually unlimited. Any possible unauthorized attack based on technique of estimation on the basis of samples from the database of source materials is virtually useless.

3D captcha system displays graphic information dynamically. Consequently, the attacker must analyse every single image fragment corresponding to respective position of a model separately in case of the attack using technique of image analysis. This is a significant increase in time and energy demands in respect of the sources of attacker compared to other captcha system based on the use of graphic

information. In such cases, the attacker only analyzes one or few images.

To increase the efficiency of 3D captcha system the source 2D images can be simplified as much aspossible. The larger abstraction rate of images limits options to be analysed by automated systems but, in appropriate use, it can also make the procedure of solving more simple for the user.

4. ADVANTAGES AND DISADVANTAGES

Like the most technologies, the 3D captcha has its advantages and its shortcomings.

4.1 Advantages

The principle is based on the human imagination. This is the biggest advantage of 3D captcha system. Successful solution requires understanding of the meaning of displayed graphic information. While no computer system can cope with the type of task required by 3D captcha.

Full automation of the system. It is very easy to fully automate the whole process of creating and using 3D captcha system. This greatly reduces the resources that provider needs to spend on system maintenance. Unlike other captcha systems using images or 3D models, it is not necessary that the source materials are strenuously manually checked or additionally modified (different assigning of tags, labels, descriptions, etc.)

Virtually unlimited database of source images. This technology practically enables using of any image as the source image. This greatly eliminates possibility of attack using method of comparison of samples from the database of used images.

Modifications to the advertising system. With the option to use any graphic materials as the source images, the support for advertising can be implemented in the system. Advertiser delivers graphic materials (e.g. logo), provider shall incorporate them in 3D captcha system and owner of the website shall integrate it into his application. Web application can therefore receive additional income in form of share in advertising. This principle is widely used by many websites in case of displaying simple text ads. Another important advantage is that any possible modification of ads is more user-oriented than a traditional text advertisements. The user is "forced" to cautiousness when solving the captcha test. Type of displayed graphic information in 3D captcha system can also be customized according to content of web site in which it is used.

User friendliness of the system. The technology was created to be as friendly to the user as possible. User just needs to use just mouse or other positioning device. The user does not need to use the keyboard to enter text inputs as in text-based captcha systems. Moreover, with increasing experience in using 3D captcha system the time necessary for solution of test reduces. Significant advantage is the independence of the technology in respect of language and suitability for users who encounter issues in solving classic text-based captcha systems.

4.2 Disadvantages

The prototype technology uses flash. To create a prototype 3D captcha system Flash runtime environment has been used. However, this is not supported on all devices. Fortunately, with the progress and mass expansion of HTML5 technology it will be possible to modify the system to be available for more users.

The risk of breaking this security method. As with all security applications, in 3D captcha there is a risk that automated systems (bots) would break this security method. However, the use of the principle of human imagination gives a high level of resistance against attacks. Yet, some abuse of an unknown system imperfections, which is not directly connected with basic principle of protection cannot be excluded.

Relatively high demands on connectivity. Given that 3D captcha uses flash and external files of 3D models, the demands for the connectivity of user are higher than in text-based captcha systems. A prototype of client part of the 3D captcha system is about 500KB in size. This will cause longer loading time of a web site for users with slower Internet connection.

5. CONCLUSION AND FUTURE WORK

This paper provided information on a new technology named 3D captcha. This captcha system is based on human imagination and spatial perspective. Test solution requires from user to find the correct observation point. It is a point which makes 3D image displayed in the captcha meaningful. The user is looking for a proper solution by rotating the 3D model in a simple and well arranged user environment. When creating the technology the emphasis was put on the use of innovative and untraditional methods. Many users find the most widespread text-based captcha systems unfriendly, so in the case of 3D captcha the authors tried creating a simple user interface and providing users with the best possible comfort in solving captcha test.

3D captcha has many advantages. It is easy to solve for man, but difficult to deal with for computer systems. Its development can be fully automated. Database based on source images used in 3D captcha system is huge. Moreover, compared to other image-based captcha systems, no complex process of definition of suitability of images to be used is necessary. We believe that 3D captcha technology is more resistant to cyber attacks than other image-based captcha systems.

Further development of 3D technology requires existence of captcha provider. Provider will provide a database of source images and API, which will enable implementing 3D captcha system under real-world conditions.

6. REFERENCES

1) Rich Gossweiler, Maryam Kamvar, Shumeet Baluja, “What’s Up CAPTCHA?
A CAPTCHA Based On Image Orientation”
http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/sk//pubs/archive/35157.pdf

2) Taylor Hayward, „3D images as a CAPTCHA“
http://taylorhayward.posterous.com/3d-images-as-a-captcha

3) Steven A. Ross, J. Alex Halderman, Adam Finkelstein, „Sketcha: A Captcha Based on Line Drawings of 3D Models“
http://www.cs.princeton.edu/gfx/pubs/Ross_2010_SAC/index.php

4) Kurt Alfred Kluever, Dr. Richard Zanibbi, „Video CAPTCHAs“
http://sudbury.cs.rit.edu/about

5) Michael G. Kaplan, “The 3-D CAPTCHA”
http://spamfizzle.com/CAPTCHA.aspx

6) Papervision3D
http://www.papervision3d.org