2D Virtual Image Generation for Document Security Using Bezier Cubic Splines

Anmar Ali Mohammed
Computer Sciences Department, University of Technology/Baghdad
Email: anmar_aljanabii@yahoo.com

Received on: 29/10/2013 & Accepted on: 5/6/2014

ABSTRACT

This paper presents suggested algorithm to draw or generate a 2D virtual secure image from Arabic text in any document to make it a secure document and avoid any fraud operations which can be done on that document. Bezier curves had been used to generate this image depending on some computed values which is defined and evaluated from the read text in the document. Those defined values have a great effect on secure image generation where those values represent the control points which they are used in Bezier cubic function. Bezier cubic had been used to implement the secure image. Many segments written had been taken as samples to test the suggested algorithm.
Keywords: 2D virtual secure image, Bezier Cubic Splines.

توليد صورة تخيلية ثنائية الابعاد لامنية الوثيقة باستخدام شرائح مكعب بزاير

الخلاصة

هذه البحث يقدم خوارزمية مقترحة لرسم او توليد صورة ثنائية الابعاد من النص المقروء في اي وثيقة لجعلها وثيقة امنه وتجنب اي عملية تزوير او احتيال ممكن ان تكون على هذه الوثيقة. منحنيات بزاير استخدمت في توليد تلك النص المدخل والذي هو جزء من الوثيقة. الصورة بالاعتماد على بعض القيم المحسوبة والتي عرفت وحسبت من مكعب بزاير. لهذه القيم المعرفة تأثير كبير في توليدالصورة التخيلية حيث ان هذه القيم تمثل نقاط السيطرة في معادلة تم استخدام مكعب بزاير لتوليد الصورة التخيلية. تم اخذ العديد من القطع المكتوبة من وثائق مختلفة كعينات لاختبار الخوارزمية المقترحة.

INTRODUCTION

The text of document describes the whole document contents, details. All of those factors can met on the same point and create the abstraction. So if the oil painting can be compiled in words i.e. the viewer can translate the lines and colors into words and those words can construct a story which talks about that oil paint. This thing can be done but in reverse way, this means that the whole document or part of it can be converted into image and this image can describe the content virtually. Of course this process can be implemented in some way. This research focuses on revealing the nature of document image from a cognitive perspective and aims at building an original typology of images in Arabic text at the future.

In this research the document text is implemented as 2D virtual image during a suggested algorithm. This implementation is considered very important in document analysis applications and extracting the document print corresponding to the texts. In the

https://doi.org/10.30684/etj.32.6B.6
2412-0758/University of Technology-Iraq, Baghdad, Iraq

suggested algorithm, the 2D virtual images of some text documents are created corresponding to some factors.

Bezier Cubic Splines

Bezier Cubic Splines are an excellent and preferred method to draw the smooth continuous curves often found in typography, CAD/CAM, and graphics in general. Among their many advantages is a very sparse data set allowing a mere eight values (or four x, y points) to completely define a full and carefully controlled and device independent curve. Figure (1) shows a cubic spline in its graph space [1, 2, and 7]. Here is a cubic spline appears in its equation space:

$$r(u) = (1-u)^3 r_0 + 3u(1-u)^2 r_1 + 3u^2(1-u)r_2 + u^3 r_3 .....(1)$$

Where r_0, r_1, r_2, and r_3 are the control points, u (for time) always goes from zero at initial point to a one at the final point.

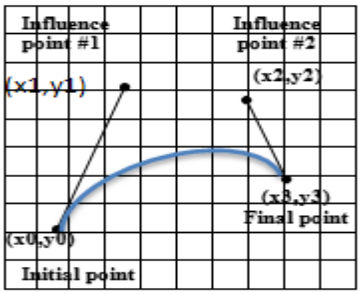

Figure (1) Cubic Spline in its graphic space.

The 2D virtual image creation

In this research, 2D virtual image is created depending on important computed values. Those values are evaluated by using defined rate which is called the letter occurrence rate; this value is computed depending on defined criteria. Three types of this rate are calculated for each letter. The computed values (the three occurrence rates for each letter) play an important role in plotting Bezier curves and as result create 2D virtual image.

The Arabic Letters form and Their Representations

Many samples of texts had been used to implement the corresponding image for each sample. Firstly, the sample of chosen text must be implemented as numerical values, where each Arabic letter must take a unique numerical value. The letters in Arabic language are different from letters in English language where there is a primary letter such as هـ and this letter may be appeared in the middle of the word and is written : هـ also may be appeared in the end of word then it is written: ه such as الله. This case occurs in most of Arabic letters. Table (1) contains all Arabic letters with their writing shape showing numerical values which will be given to the each of the Arabic letters.

Table (1) Arabic letters with its corresponding coordinates.

Arabic Letters	Coordinate Values	
	x-axis	y-axis
ء	0.269566	0.295171
ئ	0.125338	0.375724
ا	0.1998	0.332485
ب	0.349646	0.083538
ت	0.760535	0.466365
ة	0.648455	0.866872
ث	0.157452	0.48062
ج	0.026987	0.540392
ح	0.852965	0.898626
خ	0.856732	0.853703
د	0.102687	0.573137
ذ	0.75896	0.584012
ر	0.937292	0.72803
ز	0.788885	0.679389
س	0.159384	0.249535
ش	0.199504	0.632133
ص	0.5043	0.413481
ض	0.620572	0.842456
ط	0.722385	0.47648
ظ	0.641639	0.48056
ع	0.361509	0.023935
غ	0.527526	0.182922
ف	0.997977	0.012257
ق	0.719512	0.332152
ك	0.752934	0.182612
ل	0.869021	0.791172
م	0.246168	0.84936
ن	0.817777	0.180669
ه	0.505474	0.645856
و	0.622849	0.110385
ؤ	0.478682	0.582033
ي	0.462563	0.00476

After giving each letter its corresponding numerical value. The first value is the absolute appearance of letter in the whole text. Each Arabic letter will take its appearance rate in the given text. The second value of the letter is the rate of the letter appearance together with the next letter in the same word in the text. Finally, the third value which the letter must be taken is the rate of appearance of that letter with the previous letter and this evaluated for the whole text (every word in the text).

The 2D virtual image creation using Bezier Cubic

Bezier curve needs points called control or influence points to plot and control the curve. Bernstein Polynomial of order three is exactly what we need to relate Bezier control points to cubic spline coefficients.

The fourth point is a point which is chosen carefully to make the generated curves tide together and to avoid random generation which can cause generated curve to be out of the determined 2D coordinates. One end of each generated curve should be connected with the other end of curve, i.e. all generated curve should be met together in the same point (start point or end point). The formula for a degree three Bezier curve is [1, 2, 3, and 6]:

$$Q(u) = B_0(u)p_0 + B_1(u)p_1 + B_2(u)p_2 + B_3(u)p_3 .....(2)$$

Where the four functions B_t (u) called blending functions, are scalar-valued. The blending functions B_t (u) are clearly degree polynomials. Indeed, when their definitions are expanded they are equal to;

$$B_0(u) = (1-u)^3, B_1(u) = 3u(1-u)^2, B_2(u) = 3u^2(1-u), B_3(u) = u^3 .....(3)$$

Here, the three defined control points can represent the sense point in the segment. These defined points of appearance rate control the shape of generated curve. One curve is generated for each letter in the given text, and all of those curves represent the image or drawing of the text.

Methodology

This research introduces a method to provide a secure label for graduate students' documents through the use of Bezier curves of third degree. We enter the university of technology statement in our case and the department of interest we want to generate a secure label through choosing a number from 1 to 14 each for a specific department displayed every time program runs. Statements of both university and department are written in Arabic alphabetical letters to be sent to Bezier curve function for additional processing. In order to generate a secure label, a preprocessing is required before applying Bezier. We start loading coordinates of the Arabic alphabetical letters. The x-axis and y-axis coordinates represent each letter between zero as min. value and 1 as max. value. Coordinates generated within Matlab® environment through the use of ***rand*** as random number generator. These coordinates will be fixed every time a program runs.

In order to draw the curves, word segmentation process is done. In other words the coordinate for the words within the university and department statements is done benefiting from spaces between the words. An important step is choosing four control points which considered vital for applying Bezier curves for each word in both university and department statements. This process is done by selecting the first and the last coordinates as the first and last control points, still a middle control points obtained by dividing total number of coordinates by 2 and round it, then the first middle control point selected by (-1) element of coordinate from that divided number, while the second middle control pointby(+1)element of coordinate.Finally, plotting function take place to generate security label by having four control points for each and every word sequentially by applying Bezier formula [4, and 5]:

$$p(t) = \sum_{i=0}^{n} \binom{n}{i} (1-t)^{n-i} t^{i} p_i \ldots\ldots(4)$$

The above equation draw curves representing both university and selected department names. Last but not least, words that have the same letters but in a different appearance or order will create a different curve despite of distinct coordinates for these same letters, meaning that the generated curve is based on the selected control points for that particular order as a basis not the coordinates of these letters which are distinct.

Proposed Algorithm

The proposed algorithm accepts text as input and generates security label representing that same text described as below:

Input:

Enter "d" integer variable number from 1 to 14 refereeing to departments' statements which are pre-inserted along with the university statement to reduce errors of mistyping.

Output:

Secure label represented by Bezier curves of degree three for both university and department statements. The proposed Algorithm implemented in steps as in the following:

Process:

1. A switch statement decides the department of interest based on "d" variable then call Bezier function.
2. Initialization of Arabic alphabetical letters.
3. Load coordinates of control points for each and all Arabic alphabetical letters (alpha_cor) which are pre-generated previously within Matlab® environment through the user of ***rand*** function as random number generator. These coordinates are distinct for each letter.
4. Get control points for both statements of the university and the selected department by comparing with Arabic alphabetical letters.
5. Dictate coordinates of the repeated letters for the university and selected department statements through the use of indexes, ensuring that each repeated letter have the same coordinate of its first appearance within the statement.
6. Get coordinates for each word within university and department statements.
7. Choose only four control points to represent curves for each of the words from step 6.
8. Calculate cubic Bezier curves by using eq. (4) drawing curves using control points from step 7.

$$p(t) = \sum_{i=0}^{n} \binom{n}{i} (1-t)^{n-i} t^{i} p_i \ldots\ldots(4)$$

9. End process.

Results

Many texts of many Arabic documents had been taken as experimental examples as input to the suggested algorithm. Coordinates has been initialized using ***rand*** function within Matlab® environment generating standard uniform distribution values.

Secure image generated by using control points calculated for each of the words that already segmented within the text, forming and affecting the shape of generated curves. Tables (2) and (3) show the four selected control points for both statements university of technology and computer sciences department in a consecutive manner, whilst figure (2) shows 2D virtual image using Bezier curves for those same statements.

Table (2) Control Points for university of technology statement.

الجامعة		التكنولوجية	
x-axis	y-axis	x-axis	y-axis
0.1998	0.332485	0.1998	0.332485
0.026987	0.540392	0.817777	0.180669
0.246168	0.84936	0.869021	0.791172
0.648455	0.866872	0.648455	0.866872

Table (3) Control points for computer sciences department statement.

علوم		الحاسوب	
x-axis	y-axis	x-axis	y-axis
0.361509	0.023935	0.1998	0.332485
0.622849	0.110385	0.159384	0.249535
0.361509	0.023935	0.852965	0.898626
0.246168	0.84936	0.349646	0.083538

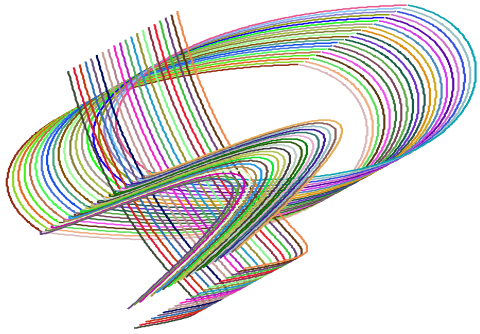

Figure (2) Bezier curves for computer sciences department and university of technology statements.

Tables (4) and (5) show four control points for both statements university and department of machines and equipment engineering followed by figure (3) representing 2D virtual image for the same statements.

Table (4) Control points for university of technology statement.

الجامعة		التكنولوجية	
x-axis	y-axis	x-axis	y-axis
0.1998	0.332485	0.1998	0.332485
0.026987	0.540392	0.817777	0.180669
0.246168	0.84936	0.869021	0.791172
0.648455	0.866872	0.648455	0.866872

Table (5) Control points for department of machines and equipment engineering statement.

هندسة		المكائن		والمعدات	
x-axis	y-axis	x-axis	y-axis	x-axis	y-axis
0.505474	0.645856	0.1998	0.332485	0.622849	0.110385
0.159384	0.249535	0.1998	0.332485	0.361509	0.023935
0.817777	0.180669	0.246168	0.84936	0.869021	0.791172
0.648455	0.866872	0.817777	0.180669	0.760535	0.466365

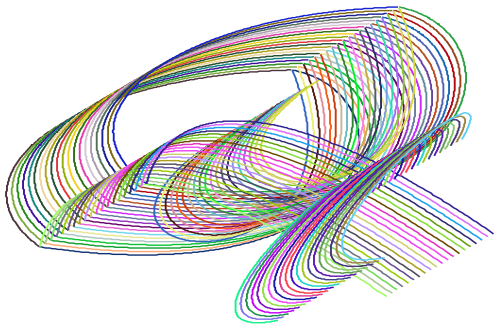

Figure (3) Bezier curves for department of machines and equipment engineering and university of technology statements.

Also, tables (6) and (7) show four selected control points for both university and department of electromechanical engineering statements followed by figure (4) representing 2D virtual image for the same statements.

Table (6) Control points for university of technology statement.

الجامعة		التكنولوجية	
x-axis	y-axis	x-axis	y-axis
0.1998	0.332485	0.1998	0.332485
0.026987	0.540392	0.817777	0.180669
0.246168	0.84936	0.869021	0.791172
0.648455	0.866872	0.648455	0.866872

Table (7) Control points for department of electromechanical engineering statement.

الهندسة		الكهروميكانيكية	
x-axis	y-axis	x-axis	y-axis
0.1998	0.332485	0.1998	0.332485
0.102687	0.573137	0.752934	0.182612
0.505474	0.645856	0.246168	0.84936
0.648455	0.866872	0.648455	0.866872

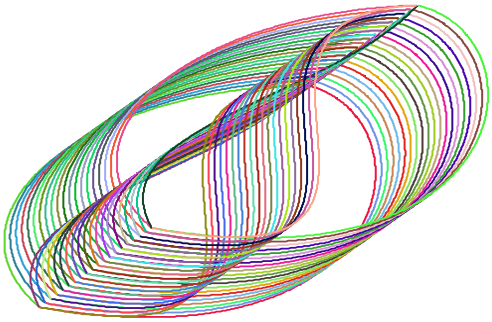

figure (4) Bezier curves for department of electromechanical engineering and university of technology statements.

Finally, tables (8) and (9) show four selected control points for both university and department of chemical engineering statements followed by figure (5) representing 2D virtual image for the same statements.

Table (8) Control points for university of technology statement.

الجامعة		التكنولوجية	
x-axis	y-axis	x-axis	y-axis
0.1998	0.332485	0.1998	0.332485
0.026987	0.540392	0.817777	0.180669
0.246168	0.84936	0.869021	0.791172
0.648455	0.866872	0.648455	0.866872

Table (9) Control points for department of chemical engineering statement.

الهندسة		الكيمياوية	
x-axis	y-axis	x-axis	y-axis
0.1998	0.332485	0.1998	0.332485
0.102687	0.573137	0.462563	0.00476
0.505474	0.645856	0.462563	0.00476
0.648455	0.866872	0.648455	0.866872

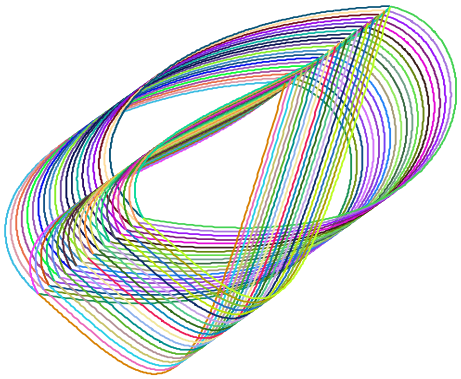

Figure (5) Bezier curves for department of chemical engineering and university of technology statements.

Conclusion and future work

From the experimental results, 2D image can be generated from any text, to draw an image contains secure parameters or embedded secure file which contains some of the image details implicitly. 2D virtual image is generated presenting a secure label for each department along with the university name avoiding fraud operations that can be done on that document. The Arabic text has a wonderful structure between their words implicitly, and the reader can feel this structure when he reads Arabic text carefully. Graphical primitives can be used in the suggested algorithm to generate a virtual secure image or document print and as a result this generated image can describes the details of document text or what is called a document soul. Many 2D images for different document are generated by the suggested algorithm and then those images can be used to create the document soul using specific criteria. Also, B-Spline curves can be used to generate the document secure file but some important points must be taken into account such as how the knots of B-Spline should be chosen and other considerations.

References

[1] J. Connors, G. Elkaim, "Analysis of a Spline Based, Obstacle , Avoiding Path Planning Algorithm," IEEE Vehicle Technology, Conference, IEEE VTC 2007, Dublin, Ireland, Apr. 22-25, 2007Published online 2004 November 24. Doi: 10.1073/pnas.0406398101.

[2] Wu, X., X. Han, and S. Luo. ,2007, Quadratic Trigonometric Spline Curves with Multiple Shape Parameters. The Proceedings of 10th IEEE International Conference on Computer Aided Design and Computer Graphics, Beijing, Oct 15-18: 413-416.

[3] Liu, H., L. Li, and D. Zhang, 2011, Study on a Class of TC-Bézier Curve with Shape Parameters. Journal of Information & Computational Science, 8(7): 1217-1223.

[4] Demidov E., "An interactive introduction to splines", http://www.ibiblio.org/e-notes/splines/intro.htm, 2004.

[5] Uzma Bashir, Muhammad Abbas, The Quadratic Trigonometric Bézier Curve with Single Shape Parameter, ISSN 2090-4304 Journal of Basic and Applied Scientific Research ,© 2012, TextRoad Publication.

[6] Wei X. Xu, L. Qiang Wang, Xu Min Liu, 2011, Quadratic TC-Bézier Curves with Shape Parameter, Advanced Materials Research (Volumes 179 - 180): 1187-1192, 2546.

[7] Aleksas Riškus, "APPROXIMATION OF A CUBIC BEZIER CURVE BY CIRCULAR ARCS AND VICE VERSA", Department of Multimedia Engineering, Kaunas University of Technology Studentų St. 50, LT−51368 Kaunas, Lithuania ISSN 1392 – 124X INFORMATION TECHNOLOGY AND CONTROL, 2006, Vol.35, No.4.