PROTECTION OF DOCUMENTS AGAINST SOPHISTICATED PHOTOCOPYING AND SCANNING

Habib Hamam
Faculty of Engineering, University of Moncton, NB, E1A 3E9, Canada
Habib.Hamam@umoncton.ca

Abstract

A novel method, for designing pantographs, is presented. An iterative algorithm is used to design a distribution of bright and dark spots to be embedded in the original document. This distribution should make the word “void”, “invalid”, or “copy” visible on the photocopy or the scanned version of the original document. We used an iterative algorithm to optimize the pantograph. We can start from a random distribution, or from a deterministic distribution based on Dirac pulses properties.

Keywords: *Pantographs, Protection of Documents, Fourier Optics.*

1. Introduction

There is a need to protect original documents in order to distinguish them from forgeries or fraudulent duplications. For example, it is illegal to photocopy prescriptions provided by doctors and present them as original documents to pharmacists. In order for a copy to be different from the original there are two main avenues: 1) the optical system of the photocopying machine or the scanner is not perfect, or 2) the original paper is physically different from the paper on which the photocopy is performed. In this spirit, many technologies have been proposed such as (Ref. [1-9]): void pantographs, latent words, control numbers, scrub, gilding, special inks (brighter metallic ink, blind ink, UV-reactive ink, iridescent ink, optically variable ink, Phosphorescent ink), holograms, reactive tags, RFID tags, dry stamps, pigments and fibers responding to the flash of the photocopier, optically variable brands (Crystagram, Kinegram, Exelgram, Movigram, Pixelgram, Stereogram), tapes (magnetic stripes, optically stripes), micro-line printing, and perforation (needle perforation, laser perforation, micro- perforation).
Void pantographs remain one of the most efficient methods and least expensive technologies. There is, however, a serious challenge. Recent developments in photocopying and scanning technologies have made void pantographs less effective. According to pantograph providers [10], the imaging system of the photocopier is not behaving as a “low-pass filter” [11] anymore. Thus, the word “void”, “invalid”, or “copy” does not appear anymore in the photocopy or scanned version of the document as requested by the standards dictated within the marketplace. To the best of our knowledge and of our industrial partner’s knowledge there today is no solution for this very recent problem.
In reality sophisticated photocopiers and scanners are still behaving as low-pass filters because the optical systems of these machines include optical components of a finite extent. However, the cut-off frequency is higher.
A pantograph is a distribution of spots that does not present readable information. For example, if the word “void”, or a similar version of it, appears on the original documents, the use of the pantograph would not be appropriate, since the original document is useless. When photocopying pantographs, because of the low-pass behavior of the optical system of the photocopier, the suppressed high frequencies make the word “void”, “invalid”, or “copy” appear.

2. Method

In this paper we present an iterative method designing pantographs with variable cut-off frequencies fc. We use the constraint-degrees of freedom approach [12-13]. As freedom degrees, we can use the size and the way "void", or equivalent, is written. The main two constraints are: first, the word "void", or equivalent, should clearly appear in the copy or the scanned version of the paper containing the pantograph; second, the pantograph does not contain any readable information and should not include large dark or white areas otherwise the original document looks tarnished. This leads to the following quantitative constraints:

Dark spots should have a small density (for example, one dark pixel for each 20 pixels in average).

No large white areas. For example, after a set of 30 white pixels we should impose one dark pixel.

Cannot allow long sequences of dark spots. For example, after a set of 3 dark spots we should impose a white spot.

The pantograph should be implementable on a paper in a 2D way (no 3D holograms). A binary pantograph (black and white) is easier to implement.

Frequencies beyond fc are cut in the Fourier plane. After low-pass filtering the word "void" should be seen by the naked eye. A cost function should be defined and minimized during the iteration process.

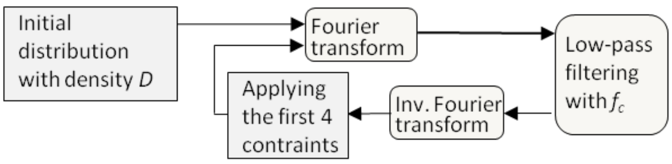

Figure 1. Simulated annealing based algorithm for designing pantographs.

Figure 1 illustrates the principle. We start from an initial distribution, which may be random or determined. We apply the Fourier transform back and forth. The first four aforementioned constraints are applied in the signal plan, whereas the fifth constraint is applied in the Fourier plan. The sixth constraint is applied in the signal plan by calculating the cost function. In this optimization process the cost function is the uniformity of the word "void" or equivalent.

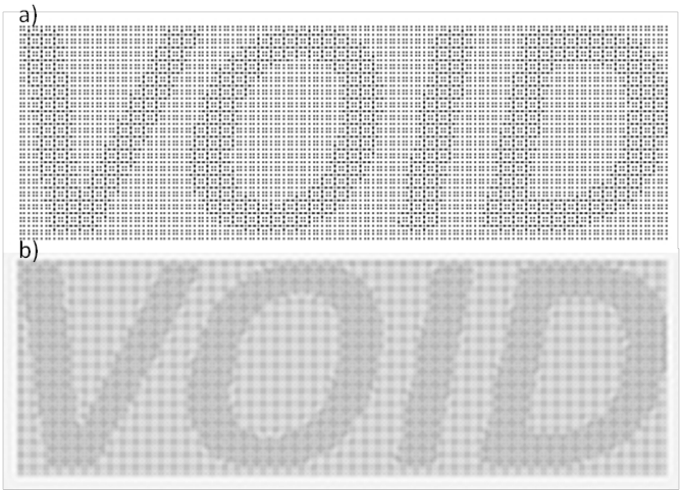

Figure 2. a) The pantograph is the binarized version of the signal h(x,y) which generates H(u,v) in the Fourier plan, respecting Eq (3), with the special frequency $u_o = v_o = 3/8\ N\ \Delta u > f_c = R_c \Delta u$, where R_c=100 and N=512.
b) The expected photocopy or scanned version which corresponds to low-passed filtered of the pantograph.

Let us call the image containing the word "void" f(x,y), and $f_{Low}(x,y)$ the low-pass version of the pantograph. The cost function is as follows:

$$cost = \frac{1}{N_{void}} \int_{void} \left(f(x,y) - \left| f_{Low}(x,y) \right| \right)^2 dxdy \qquad (1)$$

N_{void} is the number of pixels over the word "void". Of course, before calculating the cost function, the function $f_{Low}(x,y)$ should be normalized so that it contains the same amount of energy as f(x,y) over the word "void".

We embed the void image into an NxN image, with N=512, in order to use the fast Fourier transform (FFT). The sampling interval of the pantograph is Δx, which is referred to as the pantograph resolution. Depending on fabrication process it may vary from the 1 nm-range to the 10 µm-range. See for example Table 1 in reference (Ref. [14]).

The sampling interval in the FFT plane is Δu with:

$$\Delta x \quad \Delta u = \frac{1}{N} \qquad (2)$$

The cut-off frequency is $f_c = R_c \Delta u$, where R_c is an integer between 1 and N/2.

Here we considered a random initial distribution. The convergence of the Algorithm depends on this distribution. We could accelerate the convergence by starting from a deterministic distribution. One of the ideas is the following. We can start from the image containing the word "void", f(x,y). In order to hide it, the image can be drowned in another distribution h(x,y). The operation of drowning is merely an algebraic addition. The distribution h(x,y) should disappear after low-pass filtering is performed. This means it should contain only high-frequencies beyond the cut-off frequency f_c. It may contain only one high frequency, which means that the spectrum is a unique shifted Dirac: $H(u,v) = \delta(u-u_o,v-v_o)$, where the frequencies, u_o and v_o, are bigger than the cut-off frequency f_c of the photocopier.

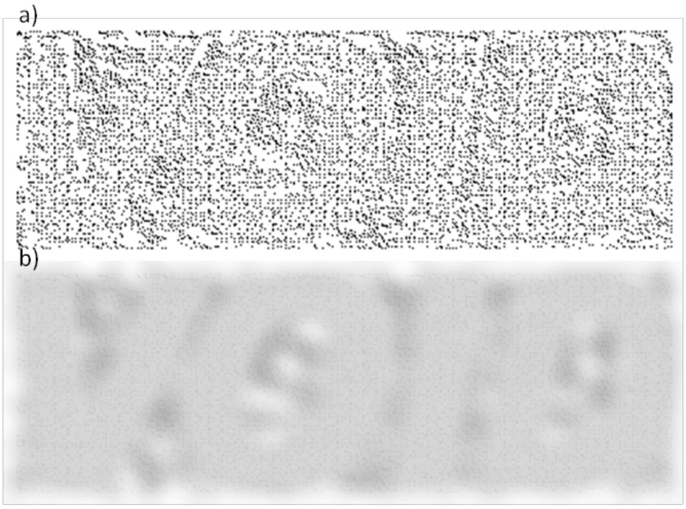

Figure 3. a) The designed pantograph for a cut-off frequency $f_c=R_c\Delta u$ with $R_c = 25$, b) The expected photocopy or scanned version which corresponds to low-passed filtered of the pantograph.

The optical signal h(x,y), generating H(u,v) in the Fourier plan, is complex. In order to obtain a real distribution, we could make H(u,v) an even distribution, such as $H(u,v) = 1/2\ [\delta(u-u_o,v-v_o) + \delta(u+u_o,v+v_o)]$, or:

$$H(u,v)= \frac{1}{4}\left[\delta(u-u_0,v-v_0)+\delta(u-u_0,v+v_0)+ \delta(u+u_0,v-v_0)+\delta(u+u_0,v+v_0)\right] \quad (3)$$

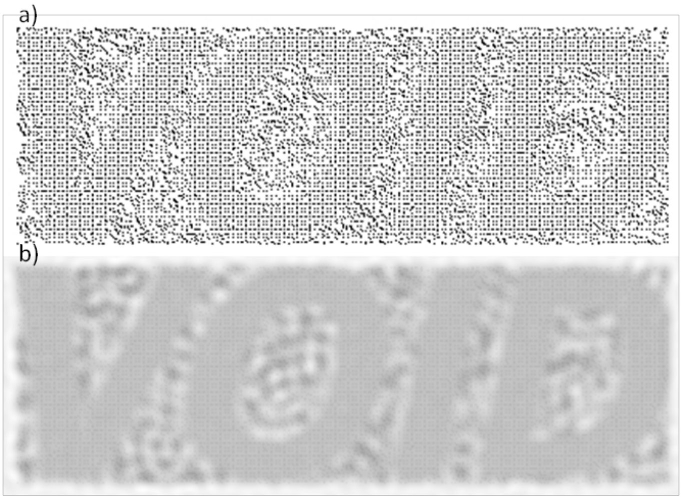

Figure 4. a) The designed pantograph for a cut-off frequency $f_c=R_c\Delta u$ with $R_c = 50$, b) The expected photocopy or scanned version which corresponds to low-passed filtered of the pantograph.

Eq. (3) was our choice for Figure 2, where $u_o = v_o = 3/8\ N\ \Delta u$, and Δu satisfies Eq (2). We clearly see a pattern in Figure 2a, which corresponds to a 2D binary grating. We binarized the grating (constraint 4) by attributing the value 0 if H(u,v)<0, and 1 otherwise. The word "VOID" is visible in the pantograph; it would be less visible if the dark spots are replaced by gray level spots and a logo of the pantograph supplier is put in the background. Preferably, the logo of the pantograph provider should contain significant high frequencies. In Figure 2b we see that the distribution is more uniform and there are no abrupt transitions between white and black spots.

3. Results and Discussions

In our experiment, we embedded the pantograph in a canvas of size NxN with N=512, and considered a cut-off frequency $f_c = R_c \Delta u$ with $R_c = 25, 50, 75, 100, 150, 200$. It worth noting that if $f_c = R_c \Delta u = N/2\ \Delta u$, no frequency is cut. The results are given in images 3 to 8, where, in each figure, a) represents the pantograph and b) represents the low-pass filtered version with the respective cut-off frequency $f_c=R_c\Delta u$ with $R_c = 25, 50, 75, 100, 150, 200$. Figure 3a to Figure 8a are different from the initial distribution of Figure 2a since the pantograph spots are optimized through the iterative process of Figure 1. Energy is normalized in all Figures 2 to 8. To avoid having the algorithm trapped in a local minimum of the cost function, simulated annealing (Ref. [15-16]) is applied. Figure 3b is very blurred, because the low-pass filtering is very selective and most of the frequencies are filtered out since f_c is low. Figure 4b is less blurred because f_c is higher.

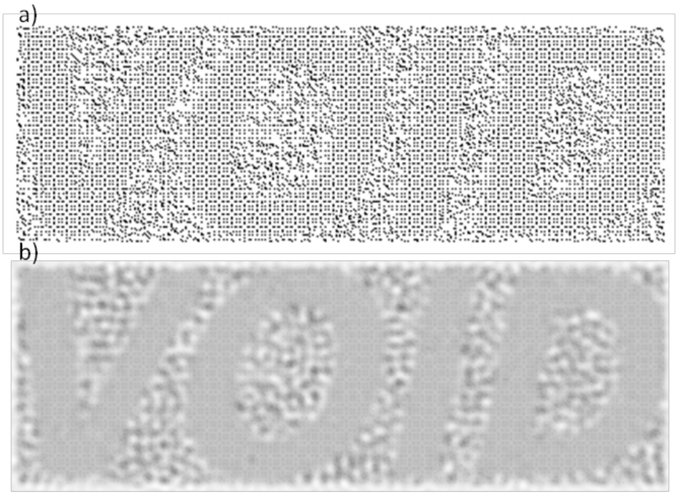

Figure 5. a) The designed pantograph for a cut-off frequency $f_c=R_c\Delta u$ with R_c = 75, b) The expected photocopy or scanned version which corresponds to low-passed filtered of the pantograph.

In Figure 5b to 7b, the "void" pattern is very uniform. In Figure 8, the filtered version, Figure 8b, is not very different with respect to the pantograph, Figure 8a, since only a small part of the frequency domain is filtered out. We notice in Figure 3a to 6a that there are insignificant changes over the "void" area compared to the initial distribution of Figure 2a. However, over the remaining area (between the letters, for example), the iterative process introduced significant changes. In Figure 7a and 8a we see significant changes performed by the iterative algorithm in both the "void" area and the rest of the image. By following the changes through iterations we noticed that the simulated annealing was very active. This means many changes are accepted, although the cost-function is worse, in order for the algorithm not to be trapped in local minima. The reason behind this, lies in the fact that only a small zone of the frequencies is filtered out, making the algorithm very likely to be trapped in local minima. The first minimum corresponds to the initial deterministic distribution itself.

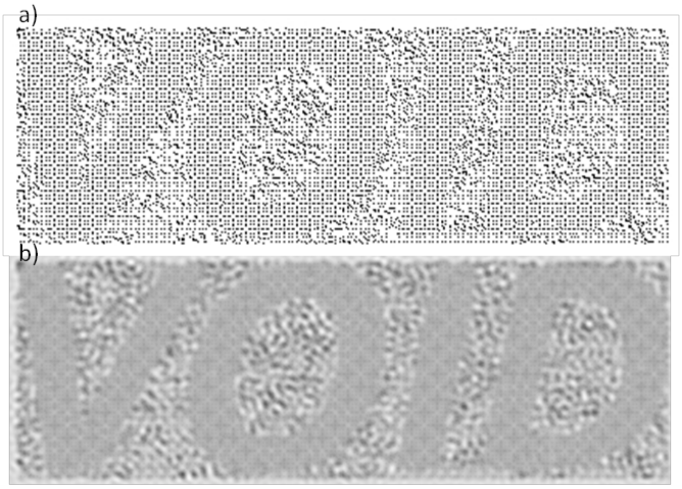

Figure 6. a) The designed pantograph for a cut-off frequency $f_c=R_c\Delta u$ with R_c = 100, b) The expected photocopy or scanned version which corresponds to low-passed filtered of the pantograph.

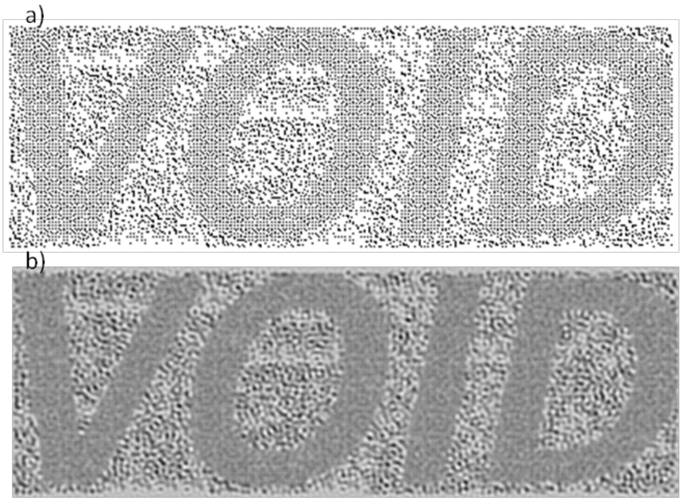

Figure 7. a) The designed pantograph for a cut-off frequency $f_c=R_c\Delta u$ with R_c = 150, b) The expected photocopy or scanned version which corresponds to low-passed filtered of the pantograph.

For all Figures 3a to 8a the simulated annealing based iterative algorithm starts from the signal h(x,y) which generates H(u,v) in the Fourier plan, respecting Eq (3). For Figures 3a to 6a the special frequency $u_o = v_o = 3/8\ N\ \Delta u$, whereas for Figures 7a and 8a $u_o = v_o = 7/16\ N\ \Delta u$, to ensure that $u_o = v_o > f_c=R_c\Delta u$.

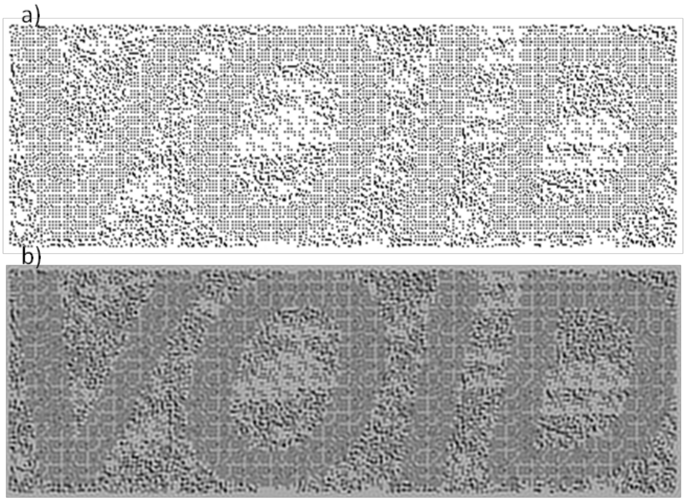

Figure 8. a) The designed pantograph for a cut-off frequency $f_c=R_c\Delta u$ with $R_c = 200$, b) The expected photocopy or scanned version which corresponds to low-passed filtered of the pantograph.

4. Conclusion

In conclusion, a new method for designing pantographs has been proposed. The idea surrounds an iterative process where constraints are applied in both the pantograph and its Fourier plan. Among the constraints, the pantograph should not contain visible information on the word “void”, “invalid”, or “copy”. However, this word should disappear after low-pass filtering with a high cut-off frequency.

We did not carry a real experiment since the production of a pantograph requires a very expensive process. A film should be developed for massive production of high-resolution pantographs. We started a collaboration with an industrial that is specialized in pantograph production, and we hope to soon be able to perform real tests.

In a future work, we also intend to consider different cut-off frequencies in the same pantograph. For example, we will seek to consider two cut-off frequencies f_{c1} and f_{c2}, where f_{c2} is higher than f_{c1}. Whether the sophisticated photocopier or scanner cuts at f_{c1} or at f_{c2}, the pantograph should work.

References

1. *A. Hodgson and S. Sims “Challenges in security printing”, International Conference On Digital Printing Technologies. 148-152. (2013)*
2. *K. Solanki “Print and Scan Resilient Data Hiding in Images”, IEEE Transactions on Info Forensics and Security 1, 464-478 (2006)*
3. *A. Aronoff, S. Simske and M. Sturgill “Automated Optimization of Void Pantograph Settings”, International Conference on Digital Printing Technologies. 690-693, (2011).*
4. *S. Huang and J. Wu “Optical Watermarking for Printed Document Authentication”, IEEE Trans. on Info Forensics & Secur. 2, 164-173 (2007)*
5. *A. Yetisen, H. Butt, T. Mikulchyk, R. Ahmed, Y. Montelongo, M. Humar, N. Jiang, S. Martin, I. Naydenova and S. Yun “Color-Selective 2.5D Holograms on Large-Area Flexible Substrates for Sensing and Multilevel Security”. Advanced Optical Materials. doi: 10.1002/adom.201600162 (2016).*
6. *V. Kolyuchkin, A. Zherdev; E. Zlokazov, D. Lushnikov, S. Odinokov and A. Smirnov “Correlation method for quality control of master matrix used for embossing security holograms”, Proc. SPIE 8776, Holography: Advances and Modern Trends III, 87760A; doi:10.1117/12.2016992. (2013)*
7. *K. Thongkor and T. Amornraksa “Improved watermark extraction for printed and scanned watermarked document”, ISPACS, doi: 10.1109/ISPACS.2011.6146097 (2011).*
8. *S. Ibrahim, M. Afrakhteh, M. Salleh “Adaptive watermarking for printed document authentication” ICCIT,doi: 10.1109/ICCIT.2010.5711127 (2010).*
9. *Yuan-Liang Tang; Chia-Jung Yang, "Print-and-Scan Resilient Watermarking Based on Modulating the Averages of DCT Coefficients," in Biometrics and Security Technologies (ISBAST), doi: 10.1109/ISBAST.2012.31, pp113-117 (2012).*
10. *Rx-Security, https://www.rxsecurity.com/*
11. *J.J. Singleton, First – Pantographic Security Papers, https://www.aceofcoins.com/2016/10/.*
12. *H. Hamam “Digital Holography based Steganography”, Optics Letters 35, 4175-4177 (2010).*
13. *H. Hamam, “Intensity based self-imaging”, Applied Optics - Information Processing 49, 2519-2528 (2010).*
14. *S. Singh, A. Chebolu, S. Mandal, Nagahanumaiah “Development of a pantograph based micro-machine for nano-scratching", Production Engineering 7, 517-525 (2013).*
15. *S. Kirkpatrick, C. Gelatt, M. Vecchi “Optimization by Simulated Annealing". Science 220, 671–680 (1983).*
16. *S. He, N. Belacel, H. Hamam and Y. Bouslimani “A Hybrid Artificial Fish Swarm Simulated Annealing*
17. *Optimization Algorithm for Automatic Identification of Clusters”, Intern. Journal of Information Technology & Decision Making, DOI: 10.1142/S0219622016500267 (2)*