Image-Based e-Document Authentication

Haiping Lu, Alex C. Kot and Jiancheng Zeng

School of Electrical and Electronic Engineering,
Nanyang Technological University, Singapore 639798

Abstract- **Fingerprint verification is one of the most reliable personal identification methods. With the fast development of the Internet, the usage of e-documents becomes more popular. Nonetheless, the old way of putting fingerprints or stamps on a hardcopy is no longer possible for e-documents. In this paper, a document fingerprinting and de-fingerprinting system is presented. In the document fingerprinting system, a thinned fingerprint image is compressed and embedded in an e-document in the spatial domain. A conversion to image format is needed if the e-document given is in PDF format. The document de-fingerprinting system does the reverse operation. Our experimental results show that the distortion after embedding is almost invisible and the embedding capacity is sufficiently large to embed a compressed thinned fingerprint. The results are encouraging to make it possible in practice for document authentication.**

I. INTRODUCTION

With an increasing emphasis on the information security, biometrics-based verification methods, in particular fingerprint-based techniques, are receiving major attention. With the explosive boom of the Internet in the past decade, the use of the Internet is increasingly integrated into our daily life. However, the old way of putting fingerprints or stamps on documents may no longer be practical. It will be preferable if fingerprints could be put in and integrated with e-documents digitally and securely, and the quality of e-documents could be preserved at the same time.

Research in information hiding has grown explosively recently and a large variety of techniques have been developed [1]. These techniques can be categorized into spatial (image) domain techniques and frequency (transform) domain techniques. Spatial domain techniques include least significant bit (LSB) manipulation, patchwork, word or line shifting and altering of text features [2-5]. Frequency domain techniques include Discrete Cosine Transform (DCT) domain watermarking, secure spread spectrum techniques and wavelet domain watermarking [6-8].

Auguste Kerckhoffs [9] enunciated the first principle of cryptographic engineering, in which he advises that we assume that the method used to encipher data is known to the opponent, so security must lie only in the choice of a key. Without the knowledge of the key, nobody should be able to extract secret information out of the cover. Therefore, if security is a requirement, a secret key has to be used for the embedding and extraction process, which is a common practice in various information hiding techniques.

In this paper, we present a secure document fingerprinting and de-fingerprinting system, which could embed fingerprint information securely into e-documents with low perceptibility while keeping the file size small. The fingerprint can be correctly extracted later for authentication. In this system, we take a spatial-domain approach for embedding and extraction. First of all, a conversion between Portable Document Format (PDF) documents and Portable Network Graphics (PNG) images provides the feasibility of embedding information in e-documents using image processing techniques. Secondly, the spatial-domain embedding algorithm suggested gives low perceptibility even though large amount of information has been embedded. Finally, practical concerns from the view of application, such as security measures and extraction elements required, are considered. In addition, the fingerprint image is compressed before embedding in order to achieve a small document file size with minimum distortion.

II. SYSTEM OVERVIEW

The document fingerprinting system (DFS) consists of a fingerprint embedding system and a fingerprint compression system as shown in Fig.

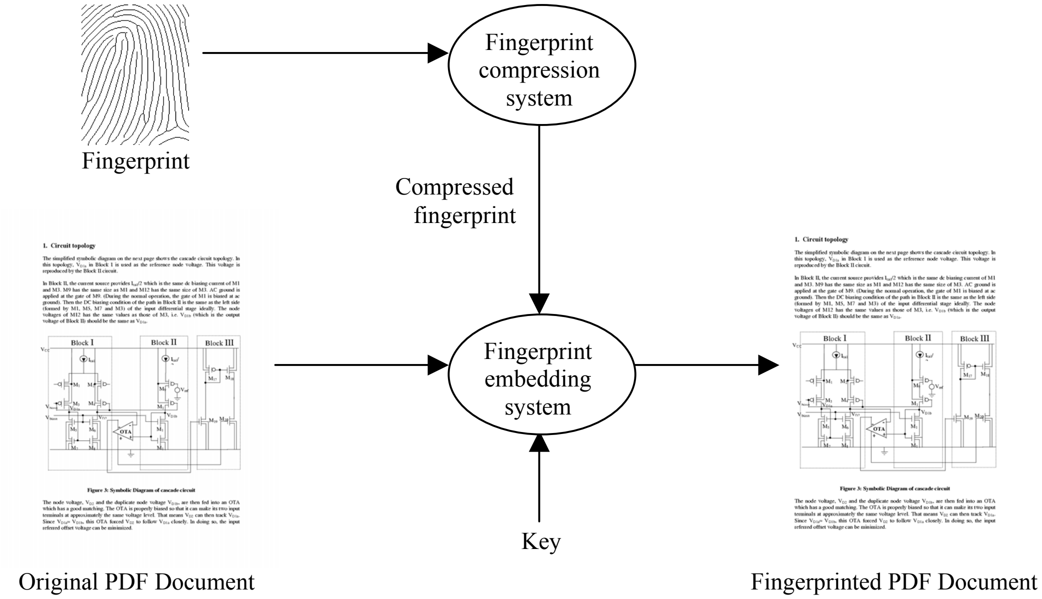

Fingerprint
Fingerprint compression system
Compressed fingerprint
Fingerprint embedding system
Key
Original PDF Document
Fingerprinted PDF Document

Figure 1. Document fingerprinting system

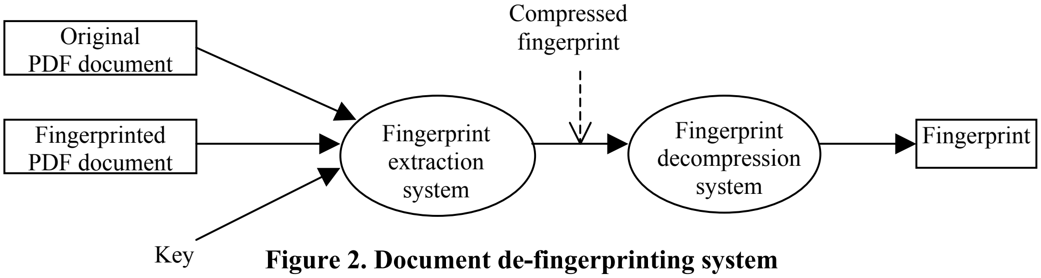

Original PDF document
Fingerprinted PDF document
Key
Fingerprint extraction system
Compressed fingerprint
Fingerprint decompression system
Fingerprint

Figure 2. Document de-fingerprinting system

1. The fingerprint to be embedded is first passed to a fingerprint compression system to be compressed. Then the PDF document to be fingerprinted is passed to the fingerprint embedding system together with the compressed fingerprint. A key is supplied to the fingerprint embedding system and the embedding is done with the fingerprinted PDF document as the output.

The document de-fingerprinting system, which is the reverse of the DFS system, consists of a fingerprint extraction system and a fingerprint decompression system, as shown in Fig. 2.

A. PERFORMANCE CRITERIA

Several performance criteria are used to evaluate the DFS system.

The capacity provided by the embedding system should be large enough to embed the vital information from a compressed fingerprint.

The compression ratio should be high enough to ensure that the size of the compressed fingerprint is not greater than the maximum capacity offered by the embedding system.

The file size of the fingerprinted PDF document should be small for fast Internet access and low storage space.

The quality of the original document should be preserved after the embedding. The distortion introduced in the fingerprinted document image could be measured in Peak Signal to Noise Ratio (PSNR) [1], as defined below:

$$PSNR = 10\log_{10}\frac{X \bullet Y \bullet \max_{x,y} p_{x,y}^2}{\sum_{x,y}(p_{x,y} - \tilde{p}_{x,y})^2}$$

where $p_{x,y}$ represents the gray-level of the pixel at (x, y) of the original undistorted image and $\tilde{p}_{x,y}$ represents that of the pixel at (x, y) of the fingerprinted image. The image is of size X by Y. The PSNR provides an objective measure of the perceptibility. The subjective perceptibility is hard to define. Spatial distribution of noises is an important factor in subjective perception [5].

The embedding and extraction system should be secure by introducing a key system so that a third party could not extract the fingerprint embedded correctly without the knowledge of the key.

The extraction error rate of the extraction system should be kept small.

One of the important performance criteria in the field of watermarking is robustness. In our intended application, the original document is easily available. There will be no intention to remove the embedded fingerprint and therefore robustness is not a concern here.

B. EMBEDDING ALTERNATIVES OF FINGERPRINT INFORMATION

There are several alternatives to embed the fingerprint information into a document. The raw image could be embedded directly into a document. This allows a larger margin of extraction errors since the loss of a small portion of pixels will have little effect on the fingerprint verification process. However, the drawback is that a large capacity is needed. For example, the size of a 300×300 gray (256 levels) image is 90KB. The binarized version would have 11.25KB in size, which is still large in capacity. Hence, direct embedding of a raw fingerprint is not feasible.

Alternatively, features extracted from a fingerprint with compact information size of 300-500 bytes can be embedded also. However, visual inspection is not possible in the verification process. In this paper, we use a lossless compression algorithm by employing run-length coding and Huffman coding [10-14] to achieve a two-to-one compression ratio on a thinned binary fingerprint image. Another technique in fingerprint compression can be found in [15]. The only drawback is that small extraction errors could cause significant errors in the decompression process.

III. FINGERPRINT EMBEDDING AND EXTRACTION SYSTEM IN SPATIAL DOMAIN

The fingerprint embedding and extraction systems in Fig. 3 and Fig. 4 consist of a PDF-format to PNG-format convertor, an image embedding and extraction system, a PNG-format to PDF-format convertor, an encryptor and a decryptor. A bit stream is embedded into an image to obtain a fingerprinted image using a spatial domain approach. An error free bit stream is extracted from the fingerprinted image with the help of the original image.

In the fingerprint embedding system in Fig. 3, a PDF document file to be fingerprinted is first converted to a PNG image as an input to the embedding system for preprocessing. A compressed fingerprint is encrypted [12] and embedded within the PNG image, which is then converted to a fingerprinted PDF document through the convertor.

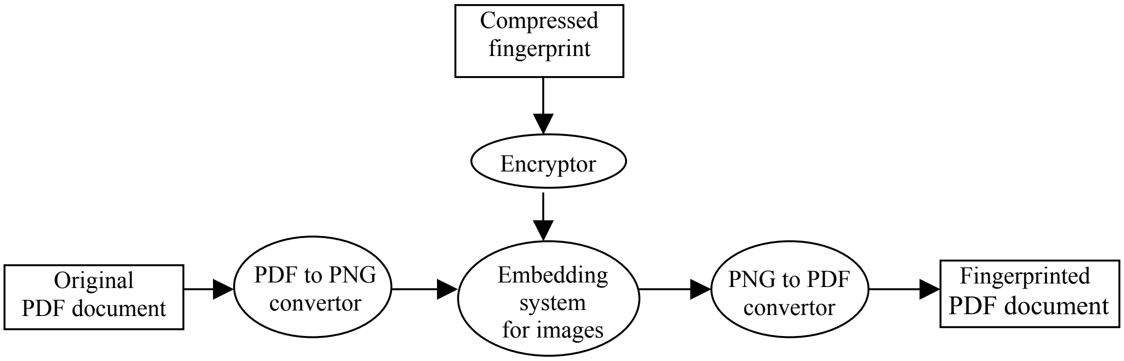


Figure 3. Fingerprint embedding system

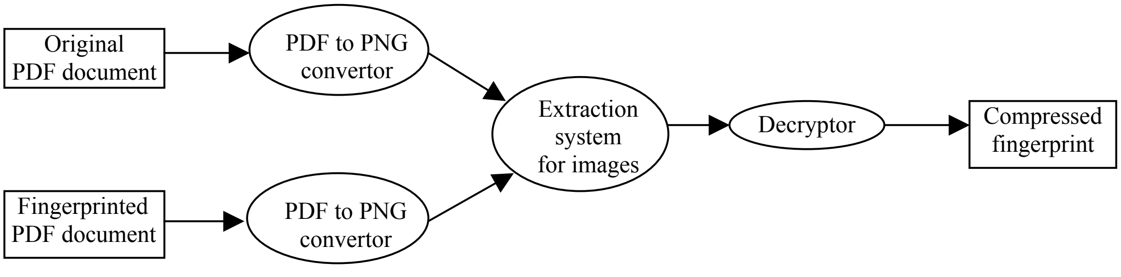


Figure 4. Fingerprint extraction system

A. PREPROCESSING OF INPUT IMAGES

The image directly converted from a PDF file is an 8-bit image with 256 levels. The bit depth of an image, which determines the number of gray levels, is related to its file size. For the same image, the binary representation has the smallest file size, and the file size increases with the bit depth in a nonlinear manner for most document images. Table I shows the file sizes of a text document image with the same resolution (110 dpi) but different bit depths before and after embedding. The parameters used in the embedding were identical for different bit depths.

We observe that a 2-bit (4 levels) image could preserve the quality for most text-dominant black and white documents. Therefore, before embedding, the bit depth of the input image is reduced to 2. It could be reduced to 1 further for a limited file size. Experimentally, the quality of a text document will not be affected with a threshold of 159 in the binarization process.

Table I. Bit depth and file size

Bit depth (bits)	1	2	4	8
Size before embedding (KB)	17.7	28.4	37.6	40.8
Size after embedding (KB)	28.4	37.2	42.9	42.9

B. EMBEDDING OF FINGERPRINT BITSTREAM

We define below three input parameters, α, β and γ, for the embedding process.

The block size during embedding is $2^\alpha \times 2^\alpha$ and 2α bits from the fingerprint bit stream determine the position of the pixel to be modified in a block. E.g., α=5 defines the block size during embedding as $2^5 \times 2^5 = 32 \times 32$ and 10 bits determine a pixel position. The range of α is from 3 to 6. Smaller α determines smaller blocks in embedding to give a large capacity with poorer quality as a trade-off, and vice versa. When the embedding capacity is sufficient for the fingerprint information to be embedded, the largest possible α will be chosen so that distortions appear to be less concentrated and obtrusive.

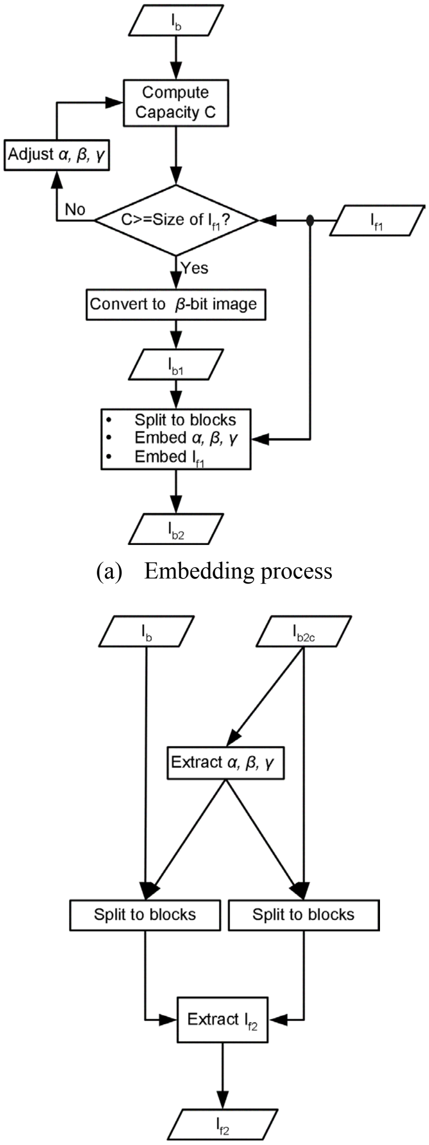
β is the bit depth of the fingerprinted image. E.g., if $\beta=8$, the fingerprinted image will has a bit depth of 8, i.e. 2^8= 256 levels. β is either 4 or 8 so that the modifications are not obtrusive.

γ bits from the fingerprint bit stream are used to modify the pixel intensity in embedding. E.g., if $\gamma=0$, none of bits will be used to modify the intensity of the pixel determined, and the intensity will be modified by a pre-defined level. If $\gamma=2$, 2 bits from the fingerprint bit stream will be used to modify the intensity of the pixel determined. Larger γ provides larger capacity, resulting in poorer quality. $\gamma=0$ always gives a fingerprinted image with the best quality for the same β. The range of γ depends on the bit depth β. For $\beta=4$, γ could be either 0 or 1. For $\beta=8$, γ could range from 0 to 5. Values of γ beyond these ranges will result in obvious distortions in the fingerprinted image.

We use 6 bits to embed these parameters in the first 8×8 block of the input image using the same technique as those for embedding fingerprint bit stream. These 6 bits determine a position in the block and the intensity of the pixel at this position is modified by a default value. However, this modification could be mixed with the modification done using the fingerprint bit stream. Therefore, one additional bit '1' is added to the head of the fingerprint bit stream and the resulted bit stream is treated as the bit stream to be embedded. The additional bit added ensures that the modification done in the first block using the bit stream will only be at the lower half block. For α of value 4, 5 and 6, the first 8×8 block of the entire image will always be in the upper half block of the size 16×16, 32×32 and 64×64, respectively, so that the embedding of parameters will not interfere with the embedding of the fingerprint bit stream. For $\alpha=3$, the 6 bits representing the 3 parameters will start with 00, so the pixel modified using these 6 bits is in the first two rows of the 8×8 block and this modification will not affect the lower half of this block either.

The capacity C, the fingerprinted document file size, the perceptibility *PSNR*, and the extraction error rate *Er* are the four performance measures. C can be computed using $C = \textit{Number of blocks} \times (2 \cdot \alpha + \gamma)$. *PSNR* is given in section II, and *Er* is the ratio of the number of bits wrongly detected versus the total number of bits embedded.

In this paper, a spatial domain approach with low perceptibility and high embedding capacity is proposed to embed a large amount of information while keeping the number of pixels modified small. In this approach, only a small number of pixels in the input image are modified. The coordinates of the pixels to be modified are determined by the fingerprint information to be embedded.



(a) Embedding process

(b) Extraction process

Figure 5. Embedding and extraction of the fingerprint bit stream

C. EMBEDDING AND EXTRACTION PROCESS

The embedding and extraction of the fingerprint bit stream are shown in the flowchart in Fig. 5, where I_b is the preprocessed 2-bit image converted from a PDF document D_1, I_f is the compressed fingerprint after encryption, and I_{f1} is obtained by adding a '1' to the head of I_f. If the embedding capacity of I_b with default parameters ($\alpha=5$, $\beta=4$ and $\gamma=0$) is smaller than the size of I_{f1}, these parameters are adjusted until the capacity is greater than or equal to the size of I_{f1}. I_{b1} is an image obtained by converting I_b to an image with bit depth β. I_{b1} is then split into blocks of size $2^{\alpha}\times 2^{\alpha}$ and the values of α, β and γ are embedded into the first 8×8 block of I_{b1}.

The intensity of one pixel in a block blk_i is to be modified. We use every $2\times\alpha$ bits from I_{f1} to determine a pair of coordinates (x_m, y_m), and the next γ bits from I_{f1} are converted to an intensity value Δ_i using gray code [16]. For blk_i, the intensity at position (x_m, y_m) is modified by Δ_i as defined below:

$$\text{for } \beta=4\text{:}\quad blk_i(x_m,y_m)= blk_i(x_m,y_m)+(1+\Delta_i) \quad \text{if } blk_i(x_m,y_m)<2^{\beta-1}$$

$$blk_i(x_m,y_m)= blk_i(x_m,y_m) - (1+\Delta_i) \quad \text{if } blk_i(x_m,y_m)\geq 2^{\beta-1}$$

$$\text{for } \beta=8\text{:}\quad blk_i(x_m,y_m)= blk_i(x_m,y_m)+(2+\Delta_i) \quad \text{if } blk_i(x_m,y_m)<2^{\beta-1}$$

$$blk_i(x_m,y_m)=blk_i(x_m,y_m) - (2+\Delta_i) \quad \text{if } blk_i(x_m,y_m)\geq 2^{\beta-1}$$

where $2^{\beta-1}$ is the middle intensity value. The modified blocks are then combined to produce the fingerprinted image I_{b2}, which is then converted to PDF document D_2.

In extraction, I_{b2c} is the PNG image converted from D_2. Comparing the first 8×8 blocks of I_b and I_{b2c} gives the position of the modified pixel, from which α, β and γ are extracted. The encrypted and compressed fingerprint I_{f2} is then extracted from the position and intensity difference of the modified pixel found by comparing the corresponding blocks in I_b and I_{b2c}.

IV. RESULTS AND DISCUSSIONS

We have tested different types of e-documents with various settings. To demonstrate, we show the results for the 2-bit image from a text circuit document with a resolution of 110 dpi. The image size is 935×1210 and the PNG file size equals to 25.7KB. The original PDF file size is 62.9KB. The size of the information to be embedded is 1163.25 bytes. The results are shown in Table II. The file size in the table is the file size of the PNG image I_{b2c}, which is approximately equal to the size of the PDF file D_2, and the extraction error Er is the error in I_{f2}. Fig. 6 shows the original document image and Fig. 7 shows the fingerprinted document image with $\alpha=4$, $\beta=8$ and $\gamma=0$.

From the table, it can be seen that larger capacity can be achieved for a smaller value of α. For the same values of β and γ, PSNR does not change much for different values of α, which is expected as PSNR is not a good measure for text-dominant document images and it takes only the absolute difference into account.

Table II. Embedding results for a text circuit document with 110 dpi resolution

α *(bits)*	β *(bits)*	γ *(bits)*	**Capacity (bytes)**	**PSNR (dB)**	**File size (bytes)**	**Extraction error number**	**Extraction Error Rate**
3	4	0	13137	52.11	34399	0	0.00%
3	4	1	15326.5	47.12	34356	0	0.00%
3	4	2	17516	44.03	34336	0	0.00%
3	8	0	13137	70.70	34398	0	0.00%
3	8	1	15326.5	68.11	34354	0	0.00%
3	8	2	17516	66.12	34331	0	0.00%
3	8	3	19705.5	62.97	34256	0	0.00%
3	8	4	21895	58.93	34214	16	0.17%
3	8	5	24084.5	54.26	34226	0	0.00%
3	8	6	26274	48.89	34180	10	0.11%
4	4	0	4350	53.33	34875	0	0.00%
4	4	1	4893.75	48.20	34753	0	0.00%
4	4	2	5437.5	45.05	34675	0	0.00%
4	8	0	4350	71.92	34871	0	0.00%
4	8	1	4893.75	69.19	34739	44	0.47%
4	8	2	5437.5	67.13	34659	3	0.03%
4	8	3	5981.25	63.83	34523	0	0.00%
4	8	4	6525	59.71	34538	18	0.19%
4	8	5	7068.75	54.82	34409	49	0.53%
4	8	6	7612.5	49.66	34424	3	0.03%
5	4	0	1341.25	54.24	35086	0	0.00%
5	4	1	1475.375	49.00	35120	0	0.00%
5	4	2	1609.5	45.81	35054	0	0.00%
5	8	0	1341.25	72.83	35083	0	0.00%
5	8	1	1475.375	69.99	35106	29	0.31%
5	8	2	1609.5	67.87	35047	4	0.04%
5	8	3	1743.625	64.59	34978	7	0.08%
5	8	4	1877.75	60.40	34815	16	0.17%
5	8	5	2011.875	55.45	34795	80	0.86%
5	8	6	2146	50.22	34704	17	0.18%

We can see from the table that despite the difference in parameters used, as long as the amount of information to be embedded is the same, the fingerprinted document file sizes are approximately the same. Also, for $\beta=4$ with all γ, and $\beta=8$ with $\gamma=0$, the extraction is always error free for all test e-documents. There could be some errors for other values of γ when $\beta=8$, and the error rates are below 1% in all the experiments conducted. The error is due to the loss in conversions between PDF files and PNG images.

For a multiple page document, a portion of a fingerprint can be embedded in each page, and the complete fingerprint can be obtained by extracting individual portions.

V. CONCLUSIONS

In this paper, a secure fingerprinting and de-fingerprinting system has been proposed for electronic document authentication. A spatial domain approach is proposed for the fingerprint embedding and extraction system, and a lossless compression algorithm is used for the thinned fingerprint. The performance of the system depends on the embedding parameters chosen. Our experiments show that the fingerprinted documents have good visual quality. The fingerprint information embedded can be extracted without error in most cases. Extraction error can be totally avoided if the conversions between PDF files and PNG images are done perfectly.

REFERENCES

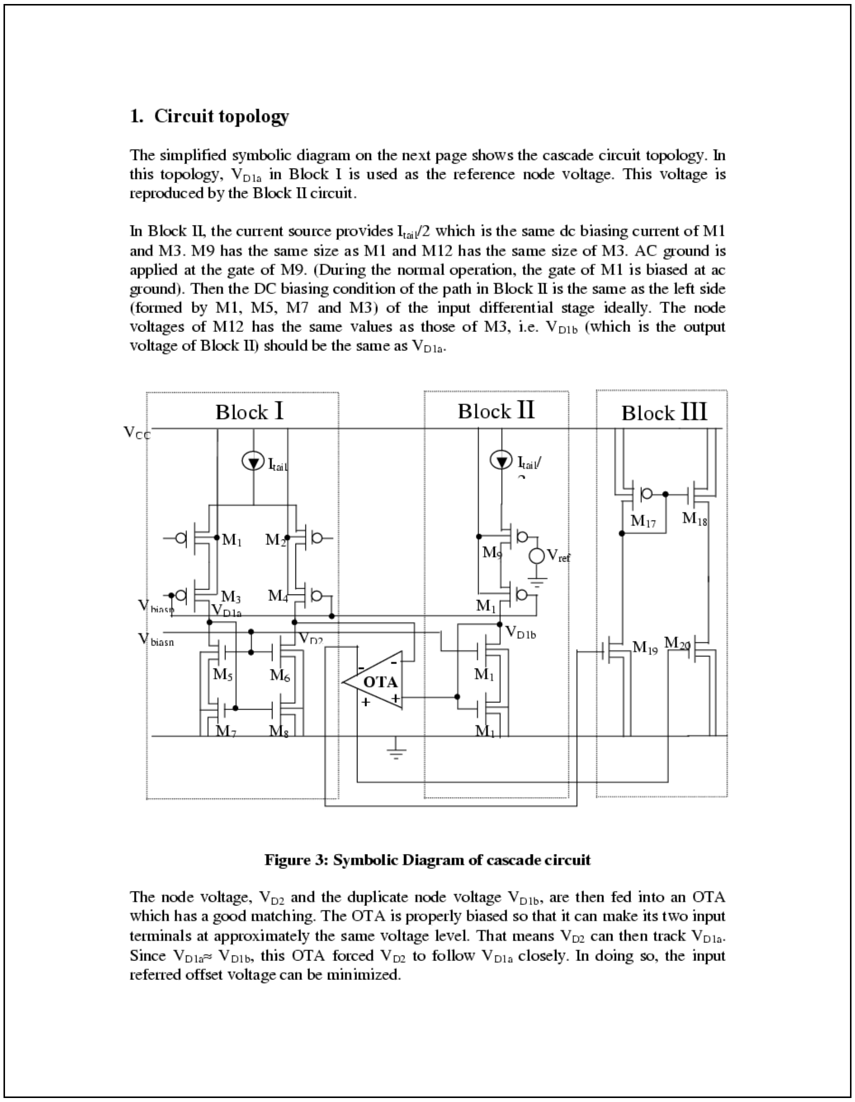
[1] S. Katzenbeisser, and F. A. P. Petitcolas, editors, *Information hiding techniques for steganography and digital watermarking,* Boston: Artech House, 2000.

[2] L. M. Marvel, C. G. Boncelet, and C. T. Retter, "Reliable Blind Information Hiding for Images", in *Proc. of Second Workshop on Information Hiding*, pp. 48-61, 1998.

[3] W. Bender and D. Gruhl, "Techniques for Data Hiding", *IBM Systems Journal*, vol. 35, no. 3 &4, pp. 373-336, 1996.

[4] J. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman, "Electronic marking and identification techniques to discourage document copying", in *Proc. of Infocom'94*, pp. 1278-1287, 1994.

[5] H. Lu, J. Wang, A. C. Kot, and Y. Q. Shi, "An objective distortion measure for binary document images based on human visual perception," in *Proc. of ICPR*, vol. 4, pp. 239-242, 2002.



1. Circuit topology

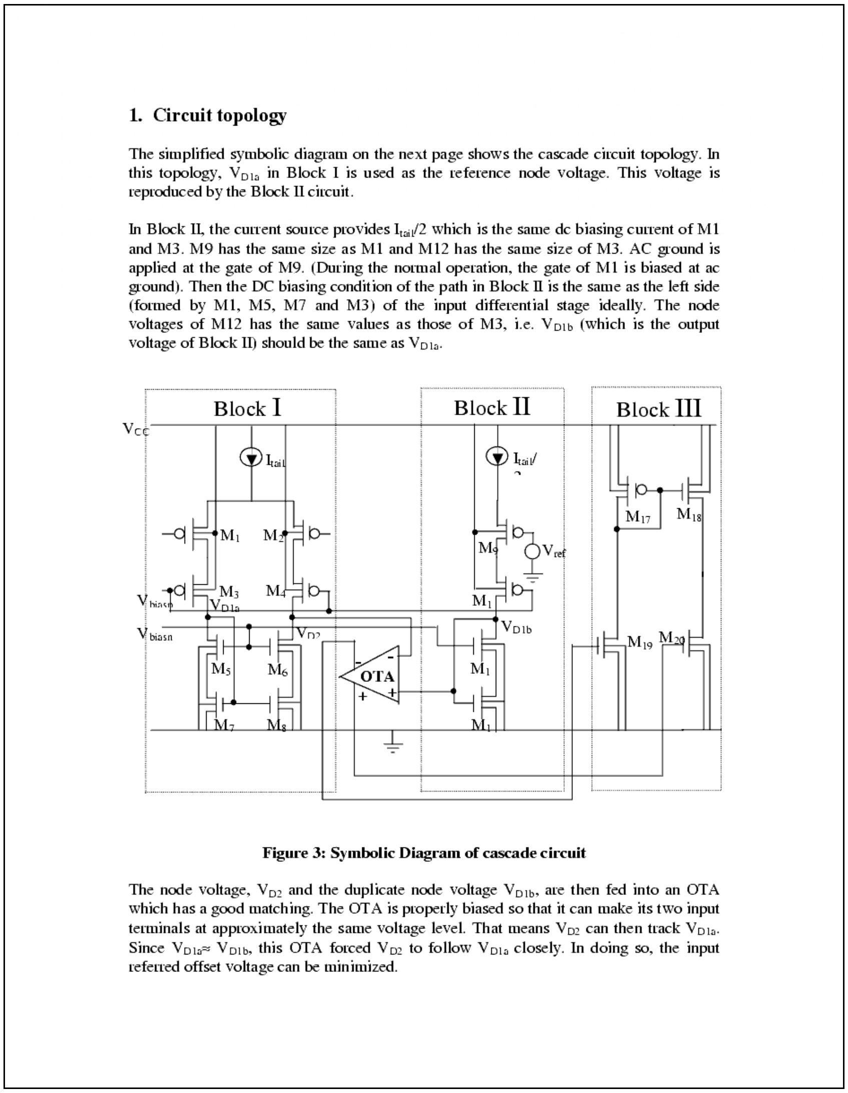
The simplified symbolic diagram on the next page shows the cascade circuit topology. In this topology, V_{D1a} in Block I is used as the reference node voltage. This voltage is reproduced by the Block II circuit.

In Block II, the current source provides $I_{tail}/2$ which is the same dc biasing current of M1 and M3. M9 has the same size as M1 and M12 has the same size of M3. AC ground is applied at the gate of M9. (During the normal operation, the gate of M1 is biased at ac ground). Then the DC biasing condition of the path in Block II is the same as the left side (formed by M1, M5, M7 and M3) of the input differential stage ideally. The node voltages of M12 has the same values as those of M3, i.e. V_{D1b} (which is the output voltage of Block II) should be the same as V_{D1a}.

Figure 3: Symbolic Diagram of cascade circuit

The node voltage, V_{D2} and the duplicate node voltage V_{D1b}, are then fed into an OTA which has a good matching. The OTA is properly biased so that it can make its two input terminals at approximately the same voltage level. That means V_{D2} can then track V_{D1a}. Since $V_{D1a} \approx V_{D1b}$, this OTA forced V_{D2} to follow V_{D1a} closely. In doing so, the input referred offset voltage can be minimized.

Figure 6. Original text circuit document with 110 dpi



1. **Circuit topology**

The simplified symbolic diagram on the next page shows the cascade circuit topology. In this topology, V_{D1a} in Block I is used as the reference node voltage. This voltage is reproduced by the Block II circuit.

In Block II, the current source provides $I_{tail}/2$ which is the same dc biasing current of M1 and M3. M9 has the same size as M1 and M12 has the same size of M3. AC ground is applied at the gate of M9. (During the normal operation, the gate of M1 is biased at ac ground). Then the DC biasing condition of the path in Block II is the same as the left side (formed by M1, M5, M7 and M3) of the input differential stage ideally. The node voltages of M12 has the same values as those of M3, i.e. V_{D1b} (which is the output voltage of Block II) should be the same as V_{D1a}.

Figure 3: Symbolic Diagram of cascade circuit

The node voltage, V_{D2} and the duplicate node voltage V_{D1b}, are then fed into an OTA which has a good matching. The OTA is properly biased so that it can make its two input terminals at approximately the same voltage level. That means V_{D2} can then track V_{D1a}. Since $V_{D1a} \approx V_{D1b}$, this OTA forced V_{D2} to follow V_{D1a} closely. In doing so, the input referred offset voltage can be minimized.

Figure 7. Fingerprinted text circuit document with *α=4, β=8* and *γ=0*

[6] E. Koch, J. Rindfrey, and J. Zhao, "Copyright protection for multimedia data", *Digital Media and Electronic Publishing*, Academic Press, pp. 203-213, 1996.

[7] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia", *IEEE Trans. on Image Processing*, vol. 6, no. 12, pp. 1673-1687, 1997.

[8] D. Kundur, and D. Hatzinakos, "A Robust Digital Image Watermarking Method Using Wavelet-based fusion," in *Proc. of ICIP*, vol. 1, pp. 544-547, 1997.

[9] A. Kerckhoffs, "La Cryptographie Militaire", *Journal des Science Militaires*, pp. 5-38, 1883.

[10] B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3rd ed., Oxford University Press, Inc., 1998.

[11] R. Hoffman, *Data Compression In Digital Systems*, Chapman & Hall, 1997.

[12] W. Stallings, *Data and Computer Communications*, 6th ed. Prentice Hall Int., Inc., 2000.

[13] G. Held, and T. R. Marahall, *Data and Image Compression: Tools and Techniques*, 4th ed., John Wiley & Sons Ltd., 1996.

[14] D. Salomon, *Data Compression: The Complete Reference*, 2nd ed., Springer-Verlag New York, Inc., 2000.

[15] J. N. Bradley, and C. M. Brislawn, "The wavelet/scalar quantization compression standard for digital fingerprint images", in *Proc. of ISCAS*, vol. 3, pp. 205 –208, 1994.

[16] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1993.