

Nuevos descriptores para la identificación de personas basados en la simetría del trazo

Verónica I. Aubin (1), Jorge H. Doorn (1,2), Gladys N. Kaplan (1)

(1) Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza, Florencio Varela 1903, San Justo, Argentina.

(2) INTIA, Facultad de Ciencias Exactas
Universidad Nacional del Centro de la Provincia de Buenos Aires,
Paraje Arroyo Seco, Campus Universitario, Tandil, Argentina.
e-mail: vaubin@unlam.edu.ar, jdoorn@exa.unicen.edu.ar

Resumen. La identificación de autores de trazos manuscritos es un área del procesamiento de imágenes en la que se han realizado muchos aportes en los últimos años. Sin embargo se trata de un dominio en el que restan aspectos por resolver. En el presente trabajo se reportan resultados relacionados con la identificación de nuevos parámetros descriptores que contribuyen a la identificación del autor del trazo. Estos parámetros se obtienen de los residuos observables en el papel luego de realizar un escrito.

Keywords: grafología, análisis de trazos, presión del trazo, perfiles residuales.

1 Introducción

Tal como ocurre con la mayoría de los parámetros vitales de un ser humano, los trazos manuscritos realizados por una persona tienen una gran variabilidad dependiendo de numerosos factores. Sin embargo en los trazos manuscritos existe un núcleo de aspectos invariantes que hace viable el reconocimiento manual o automático del autor.

Solange Pellat hijo, fue uno de los principales constructores de la grafonomía, junto con el filósofo André Lalande. Desde 1903 Pellat demostró y dio origen a un análisis fino de los movimientos de la escritura, realizando una investigación sobre las leyes fundamentales de la escritura. Finalmente en 1927 publicó una síntesis de sus trabajos en donde enuncia una serie de leyes que actualmente se utilizan. [1]. Citando a Viñals y Puente “La fiabilidad del sistema era fruto de una nueva visión sobre la escritura: [2] no considerarla como arte, sino como un reflejo fisiológico y psicológico del individuo. Es por ello que se convierte en un elemento identificativo <...>. Pero en la escritura existe una jerarquía de signos, tales como la profundidad, intensidad, presión, rapidez, dirección, continuidad que son muy difíciles de imitar. Esta metodología que se demostró altamente efectiva pues se adentra en la anatomía de la letra, y consiguió el reconocimiento de la Justicia” [2].

En el presente artículo se ha tratado profundizar el abordaje analítico del grafismo, utilizando recursos del procesamiento de imágenes. Procurando encontrar aspectos característicos del trazo, invisibles o muy poco visibles, que tengan la propiedad de ser altamente repetitivos.

El estudio semi-automático o automático de trazos manuscritos es una actividad de importancia en una gran variedad de dominios. En algunos de ellos el objeto de interés es la comprensión del texto escrito, mientras que en otros se procura identificar al autor del mismo [3]. Este artículo se concentra en algunos aspectos relacionados con la identificación del autor de los trazos.

La fuerza ejercida durante la escritura debe medirse durante el acto de la escritura propiamente dicha. Existen dispositivos [4] [5] [6] que permiten conocer esta fuerza y los estudios realizados con los mismos han sido valiosos para estimar las características de las mismas. Sin embargo son poco transportables al problema práctico ya que en general los estudios de identificación del escribiente sólo tienen acceso al resultado de la escritura.

La fuerza ejercida en el momento de la escritura deja algunos residuos tales como el color relativo de cada fragmento del trazo o el ancho del mismo. Este hecho ha sido reconocido y analizado por muchos autores [7] [8] [9] [10].

Los trabajos cuyos resultados se reportan en este artículo comenzaron analizando algunos aspectos de los trazos manuscritos, a través del procesamiento de imágenes. Se relacionó, bajo condiciones controladas, la fuerza ejercida cuando una persona escribe, con el grosor y valor de gris del trazo [11] [12] [13].

Se estableció que el ancho medio y el valor de gris son, dentro de ciertos límites, casi proporcionales al peso, pero una vez que el papel alcanza la máxima deformación condicionada por la base ya no varía significativamente. Se comprobó que no había variación en los resultados anteriores utilizando distintos colores de tinta [11] [12].

Además, se encontró que un trazo espontáneo de un grafema aparecen zonas donde el ancho medio y el valor de gris son notoriamente diferentes del resto del trazo. Estas zonas son casi invariantes en su ubicación relativa para todas las muestras del mismo grafema realizadas por la misma persona [13].

Las comparaciones realizadas sobre los gráficos característicos del trazo arrojaron resultados muy favorables. Por un lado los valores de grises y los anchos medios del trazo son altamente repetitivos para trazos que representan el mismo grafema realizados por el mismo autor [13].

2 Identificación del autor mediante niveles de grises

El término “pseudo-dinámicas” se usa para distinguir datos reales dinámicos, grabados durante el proceso de escritura, de información, que puede ser reconstruido de la imagen estática. En la literatura se han propuesto diversas características de este tipo, las cuales se revisan a continuación. Una gran cantidad de los trabajos publicados en este campo tratan la autenticación de firmas, la cual es un subproblema de la identificación del autor desde un texto cualquiera.

En [7] se analiza el histograma de la imagen en escala de grises de la firma, y se propone el cómputo de un umbral de presión (Umbral de Alta Presión). En conjunto

con dicho parámetro, se representa la firma mediante el porcentaje de píxeles que superan el umbral, el valor mínimo y máximo de nivel de gris de la imagen, el rango dinámico de la firma, entre otros. Para validar este conjunto de descriptores el trabajo adopta un clasificador basado en distancia. En [8] se propone un método para la verificación de firmas off-line basado en características geométricas y redes neuronales. Las características consideradas son el esqueleto del trazo, el contorno, y regiones de alta presión.

En [14] a partir del esqueleto de la imagen, se detectan los trazos que superen una longitud mínima, y se procede a calcular un Índice de Suavidad (SMI) para cada uno de los trazos. El descriptor corresponde a la razón entre el número de trazos suaves y el número total de trazos de la firma.

En [9] se introduce un enfoque orientado a la detección de falsificaciones elaboradas. Considera información dinámica como la presión que se ejerce con la punta del instrumento de escritura, y pone énfasis en la extracción de los píxeles de baja presión de acuerdo al nivel de gris en la imagen. No se hace referencia en éste trabajo al tipo o gramaje del papel utilizado en las muestras, ni se hace referencia al tipo de instrumento de escritura ni de tinta utilizado, tampoco se menciona la base de apoyo. Los autores utilizan un umbral de decisión basados en una hipótesis de trabajo de que existe una relación entre la presión ejercida en la escritura y valor del nivel de gris en la imagen, pero nunca la probaron.

En [15] el modelo propuesto utiliza un conjunto de características estáticas y pseudo-dinámicas para la verificación. Las características estáticas corresponden básicamente al calibre, el cual describe la relación entre la altura y el ancho, un parámetro que representa la suavidad de la firma, el espaciado y la alineación con respecto a una línea base. Las características pseudo-dinámicas consideradas son la progresión, la forma y la inclinación.

En [16] se propone un método de verificación de firmas Chinas off-line, utilizan tanto características estáticas como dinámicas, y Support Vector Machine para la clasificación. El sistema de verificación propuesto combina cuatro conjuntos de características: Características de momento (la proporción entre la altura y el ancho, el grado de inclinación, el grado de extensión, el grado de excursión horizontal y vertical), distribuciones de dirección, distribución del nivel de gris, y distribución del ancho de trazo. Tampoco los autores hacen referencia al tipo o gramaje del papel utilizado en las muestras ni al tipo de instrumento ni de tinta utilizado, tampoco se menciona la base de apoyo. Para los autores la densidad o presión aparente se describe por la anchura de los trazos, esto es una hipótesis de trabajo que tampoco fue nunca probada.

En [17] se propone un método basado en dos imágenes y su transformación a coordenadas polares. La primera imagen contiene información de los puntos de alta presión, y la segunda corresponde a la versión binarizada de la imagen original. El espacio polar se divide en secciones angulares y radiales, en donde se determina la distribución de los puntos de alta presión. Además, se considera la densidad de dichos puntos respecto del centro geométrico de la firma original. Para poder determinar los puntos de alta presión, se calcula un umbral de alta sobre el histograma de la imagen en escala de grises.

En [10] se propone un método que combina el análisis global y local de la imagen en conjunto con un Support Vector Machine para la clasificación de los patrones. El

análisis global se realiza mediante la estimación de variaciones de nivel de gris de la imagen usando la transformación wavelet, y tratamiento local considera la obtención de información de textura proveniente de la matriz de coocurrencia.

3 Simetría del trazo

Una de las contribuciones principales de este artículo está relacionada con el análisis de la línea que une los puntos más oscuros del trazo. Dado que la línea de mínimos se ubica siempre en la misma posición relativa en las diferentes zonas de un grafema producido por el mismo autor. Estudiando más detalladamente este fenómeno se encontró que la distancia relativa entre la línea de los mínimos y el esqueleto es repetitiva para una persona y varía de individuo a individuo.

El camino que permitió construir la contribución anterior comenzó con la verificación de la hipótesis de trabajo en la que se basan distintos autores, en el sentido que existe una relación entre la presión ejercida en la escritura y valor del nivel de gris en la imagen y el ancho del trazo.

Utilizando un dispositivo diseñado especialmente, se realizaron numerosos trazos rectos con diferentes pesas de manera de realizar trazos con presión constante. Las pesas utilizadas fueron seleccionadas entre 10g y 200g (lo que se corresponde con los 0,1 y 2 Newton).

En la figura 1, se presenta un gráfico de los valores de la escala de grises a lo largo de una línea imaginaria perpendicular al trazo, en un punto determinado del mismo. La zona horizontal representa el papel sin ninguna escritura. Se muestra en la imagen como se tomaron las mediciones de ancho medio del trazo y valor del nivel de gris, para su posterior análisis.

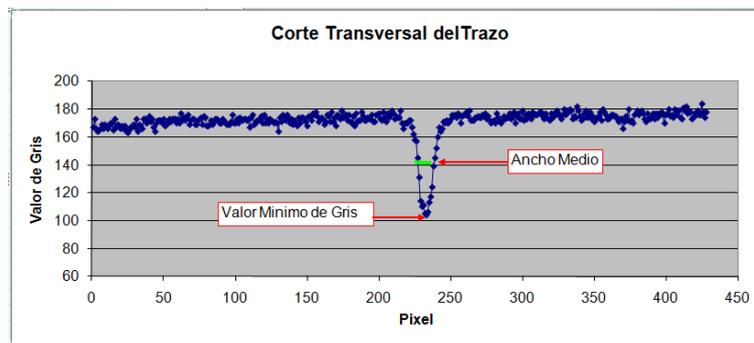
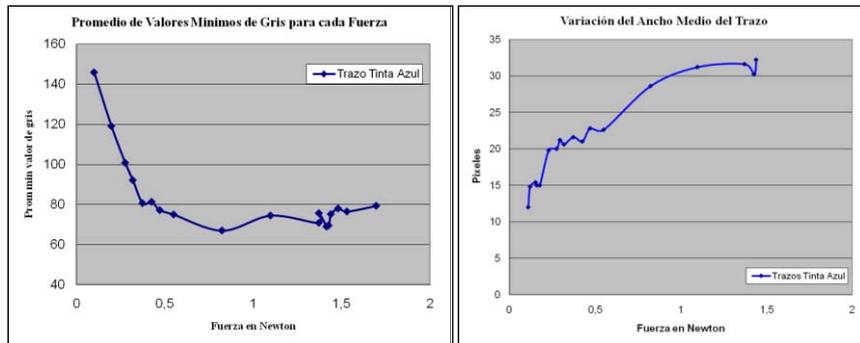


Figura 1 Valor de gris a lo largo de una línea perpendicular al trazo.

Debe notarse que los valores de la ordenada se corresponden con intensidad de tonos de gris, no de fuerza.

La figura 2 muestra lo la influencia en las mediciones de la fuerza aplicada en la escritura. La figura 2 (a) muestra la relación de las distintas fuerzas aplicadas y el valor de gris en el centro del trazo. En la figura 2 (b) se ve la influencia de la fuerza aplicada sobre el ancho medio del trazo.



(a)

(b)

Figura 2. (a) Nivel de gris en el centro del trazo (b) Ancho medio del trazo.

Los resultados obtenidos representan una realidad que se ajusta a lo que se esperaba, se observa que entre la cota inferior de fuerza y aproximadamente 1 N el ancho medio y el valor de gris son casi proporcional al peso, pero una vez que el papel alcanza la máxima deformación condicionada por la base ya no sigue adelante.

Se repitió la experiencia para estudiar como influía el color de la tinta en las mediciones. Se observó que se mantienen los resultados obtenidos anteriormente independientemente del color de la tinta utilizada.

3.1 Metodología propuesta

El arreglo experimental se basa en la captura de imágenes, usando la luz difusa ambiente; el mismo instrumento de escritura, bolígrafo “bic trazo grueso” de color azul y las características del papel y la base de apoyo en 5 hojas de 75g/m².

El proceso comienza con la umbralización de la imagen y el suavizado de los bordes aplicando los algoritmos de erosión y dilatación [18]. Luego, se esqueletiza el trazo manteniendo la continuidad del mismo [19], y se calcula para cada punto del esqueleto la recta perpendicular al mismo, sobre la cual se mide en la imagen original el valor del mínimo gris. Se procede a rectificar el trazo de manera de poder graficar el valor de gris en un sistema cartesiano.

3.2 Distancia

Como ejemplo, se considera la letra "e", la cual permite observar con facilidad la composición del descriptor. Se definen segmentos sobre el trazo completo del grafema. Estos segmentos procuran contemplar las zonas relativamente homogéneas del trazo. En la figura 3 se muestran los 7 segmentos en que se dividió el grafema para su análisis, los cuales se encuentran etiquetados con letras desde la a hasta la g.

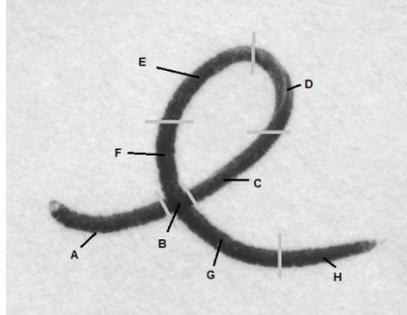


Figura 3 Segmentos del grafema

El proceso de obtención del descriptor es el siguiente, a partir de las líneas perpendiculares de cada uno de los puntos del esqueleto, se identifican las coordenadas y el nivel de valor gris del pixel más oscuro sobre la perpendicular. Los puntos blancos de la figura 4(a) muestra esquemáticamente los pixeles de menor nivel de gris. El cálculo de la distancia entre los puntos del esqueleto y su correspondiente punto más oscuro se realiza sobre la perpendicular la que se señala con los trazos negros en el esquema. Para observar el resultado del computo del esqueleto y la línea de menores niveles de grises en una imagen real, se presenta la figura 4(b) y figura 4(c).

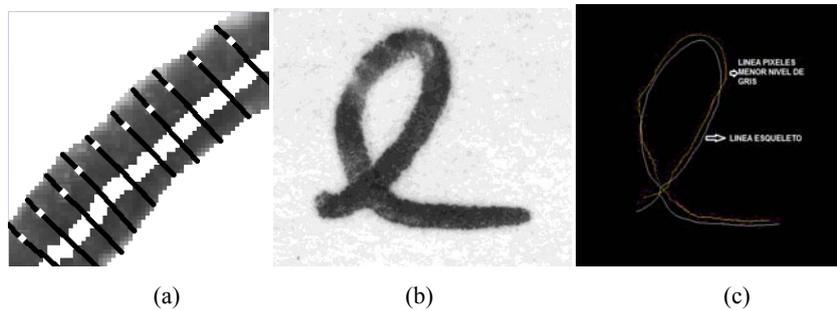


Figura 4 Proceso de Computo del Descriptor

Por último, para cada segmento del trazo se obtiene el promedio de las distancias. A estos promedios por segmento se les denomina ρP . Considerando entonces que para este grafema se divide en 7 segmentos, el descriptor del grafema completo corresponde al vector constituido por los 7 promedios de las distancias normalizadas en cada segmento. Así, el descriptor considerado corresponde al vector $D = (\rho P_1, \rho P_2, \rho P_3, \rho P_4, \rho P_5, \rho P_6, \rho P_7)$.

En la figura 5 se compara la distancia promedio de dos personas distintas. Se puede observar en la figura 5(a) que la línea que representa el promedio de las distancias esta cerca de la línea del esqueleto para todos los tramos del trazo. En la figura 5(b) la línea del promedio de las distancias tiene una separación mayor con respecto a la línea central del trazo.

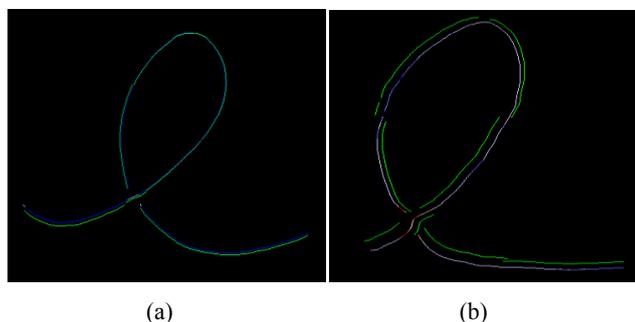


Figura 5. Distancia promedio de dos personas distintas

En los casos que los intervalos de distancia promedio son independientes, se puede utilizar el mismo como característica para clasificar a los autores de las muestras. Evaluando si la distancia de la muestra de la que se quiere probar su autoría se encuentra dentro del intervalo correspondiente.

A modo de ejemplo la figura 6 muestra la distancia \pm el error para los segmentos del trazo inicial, cruce, Trazo Ascendente Inferior, Trazo Ascendente Superior, Trazo Descendente Superior, Trazo Descendente Inferior, Trazo Final 1^{ra} Mitad y Trazo Final 2^{da} Mitad del grafema 'e' realizado por dos autores distintos.

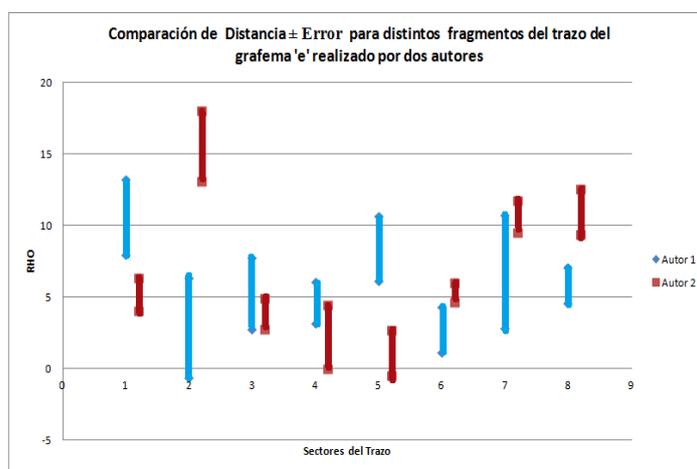


Figura 6 Comparación de la distancia promedio \pm Error, del grafema realizado por dos autores.

3.3 Puntos de Cruces

Se analizaron los cruces de la línea que une los puntos de mínimo valor de gris con la línea del esqueleto del grafema.

En la figura 7 se observan los puntos de cruce y el sentido de los mismos en relación con la línea del esqueleto. Las figuras 7(a), 7(b) y 7(c) corresponden a

muestras realizadas por la misma personal, se visualiza en ellas que las zonas y los sentidos de los cruces se repiten. El cruce en el trazo ascendente es de adentro hacia afuera del trazo. Cruce en el trazo descendente es de afuera hacia adentro del trazo. Para la ubicación del cruce en el grafema se decidió utilizar un porcentaje del total del tramo del trazo en el que se encuentra. Por ejemplo, en la figura 7(a), el cruce 1 se ubica al 66% del tramo ascendente y el cruce 2 a 60% del tramo descendente.

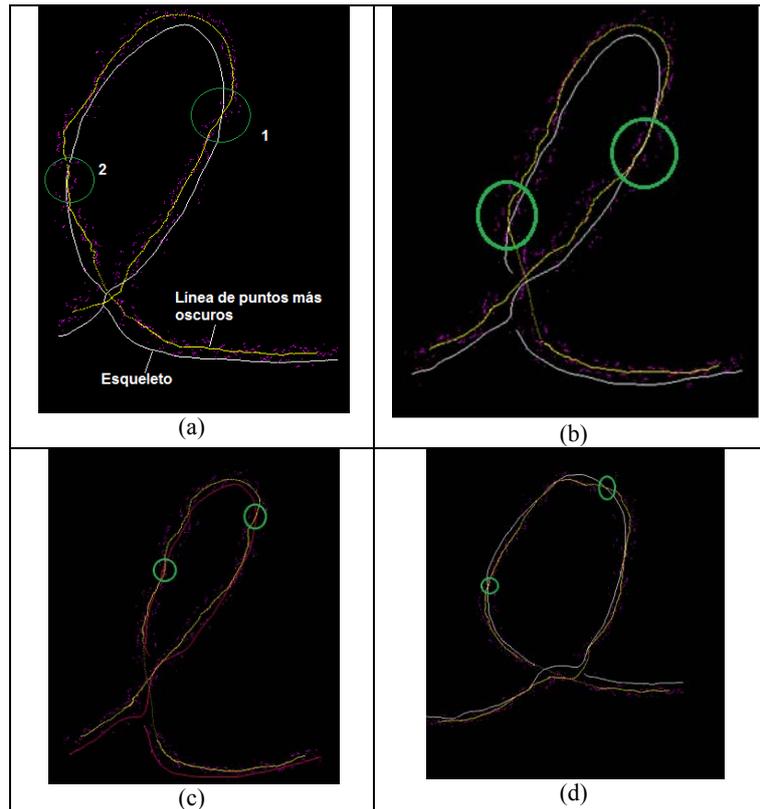


Figura 7 Puntos de Cruces

Para otra persona, se puede observar en la figura 7(d) los cruces. Cruce en el trazo ascendente es de afuera hacia adentro del trazo ubicado al 85% del trazo. El cruce en el trazo descendente es de adentro hacia afuera del trazo al 60% del trazo, como se observa en la Tabla 1.

Tabla 1 Ubicación de los cruces en cada grafema medidos en porcentajes del tramo.

Cruce	Persona1			Persona2
	Imagen 16	Imagen 17	Imagen 18	Imagen 19
Tramo Ascendente	66%	60%	75%	85%
Tramo Descendente	60%	65%	55%	60%

4 Conclusiones

En este artículo se presenta un método no invasivos de bajo costo que permiten extraer características del trazo manuscrito ya producido. Analizando las diminutas deformaciones que la escritura produce sobre el papel y las características del trazo tales como valor de gris del mismo.

El método propuesto no modifica física o químicamente el texto original, lo que posibilita múltiples análisis. Esta característica lo hace muy atractivo para ser utilizado en análisis forenses, ya que permite preservar la muestra original.

Se confirmaron y ampliaron los resultados de otros autores en el sentido que los valores de gris [7] [9] [17] y los anchos medios del trazo [16] son altamente repetitivos para trazos que representan el mismo grafema realizados por el mismo autor.

En este artículo se planteó la extracción de características en función del valor del nivel de gris a lo largo del trazo, en lugar de utilizar histogramas como plantearon otros autores [7] [9] [17]. Un mismo histograma puede responder a trazos con diferentes características, mientras que en el método propuesto el valor de gris está relacionado con su ubicación en el trazo.

Se ha comprobado que la relación de la línea que une los puntos más oscuros del trazo con el esqueleto brinda información muy útil para identificar al autor. Los parámetros que se pueden extraer de esta comparación son:

- Distancia de separación en píxeles.
- Los cruces y su ubicación en el trazo.
- Posiciones relativas con respecto al esqueleto.

Se han confirmado las conclusiones de muchos autores, en el sentido que no hay una única característica del trazo que se pueda considerar suficiente para identificar el autor del trazo. Los resultados de este trabajo permitieron ampliar el conjunto de características pseudo-dinámicas.

Trabajos Futuros

En virtud que todos los estudios fueron realizados con una cantidad exigua de muestras, se hace necesario ampliar la cantidad de muestras en al menos tres dimensiones:

- Más muestras de cada grafema producida por el mismo autor.
- Más autores.
- Más grafemas.

Queda también pendiente evaluar el impacto de la inclinación del instrumento de escritura sobre la simetría de los valores de gris en la dirección perpendicular al esqueleto del trazo. Potencialmente puede ser necesario regresar a la realización de trazos controlados pero con el instrumento de escritura en diferentes ángulos respecto del papel.

Los estudios de grosor y el análisis de la presencia de “minucias” serán reportados en futuros artículos.

5 Referencias

1. Manuel José Moreno Ferrero "Grafología Forense: La Pericia Caligráfica Judicial" www.grafoanalis.com/moreno_forense.pdf (última consulta marzo 2011)
2. F. Viñals and M. Puente. "Pericia Caligráfica Judicial: Práctica, casos y modelos". Ed. Herder, Barcelona. 2001.
3. Plamondon R. y Srihari S. N "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey". IEEE Transactions on Pattern Analysis and Machine Intelligence 22(1), 63–84 (Jan. 2000).
4. D. Sakamoto, T. Ohishi, Y. Komiya, H. Morita and T. Matsumoto, "On-line Signature Verification Algorithm Incorporating Pen Position, Pen Pressure and Pen Inclination Trajectories", Proc. IEEE ICASSP 2001, Vol. 2, pp. 993-996, 2001.
5. W. S. Wijesoma, K. W. Yue, K. L. Chien, and T. K. Chow, "Online handwritten signature verification for electronic commerce over the internet," WI 2001. N. Zhong et al. Eds. Berlin, Germany: Springer-Verlag, 2001, pp. 227–236.
6. Nelson, W. and E. Kishon. Use of Dynamic Features for Signature Verification. IEEE Transactions on Systems, Man and Cybernetics. 201-205. 1991
7. Ammar, M., Yoshida, Y., Fukumura, T.: "A New Effective Approach for Off-line Verification of Signatures by Using Pressure Features", Proc. ICPR (1986) 566-569
8. Huang, K., Yan, H.: "Off-line signature verification based on geometric feature extraction and neural network classification", Pattern Recognition 30(1997) 9-17
9. A. Mitra, P. Kumar Banerjee and C. Ardil. "Automatic Authentication of Handwritten Documents via Low Density Pixel Measurements". International Journal of Information and Mathematical Sciences 2:4 2006.
10. Vargas, J.F., Ferrer, M.A., Travieso, C.M., Alonso, J.B.: "Off-line signature verification based on grey level information using Texture features". 2010.
11. V. I. Aubin, R. S. Wainschenker, J.H. Doorn.: "Determinación de Propiedades de Trazos Manuscritos por Medios Interferométricos". WICC-2005. 2005 ISBN:950-665-337-2 pág 134-137.
12. V. I. Aubin, R. S. Wainschenker, J.H. Doorn.: "Perfilometría Virtual en Trazos Manuscritos Residuales". WICC 2010.
13. V. I. Aubin, R. S. Wainschenker, J. H. Doorn.: "Aspectos Invariantes en Trazos Manuscritos". WICC 2011.
14. Fang, B., Wang, Y.Y., Leung, CH., Tang, Y.Y., Tse K.W., Kwok, P.C.K. and Wong, Y.K.: "A smoothness index based approach for off-line signature verification". In Proceedings of the Fifth International Conference on Document Analysis and Recognition, pages 785–787, 1999.
15. Oliveira L.S., Justino, E., Freitas, C. and Sabourin, R.: "The graphology applied to signature verification". In 12th Conference of the International Graphonomics Society, pages 286–290, 2005.
16. Lv, H., Wang, W., Wang, C. and Zhuo, Q. "Off-line Chinese Signature Verification Based on Support Vector Machine". Pattern Recognition Letters, Elsevier, 26:2390–2399, 2005.
17. Vargas, J.F., Ferrer, M.A., Travieso, C.M. and Alonso, J.B.: "Off-line signature verification based on high pressure polar distribution". In Proceedings International Conference on Frontiers in Handwriting Recognition 2008. Montreal., August 2008.
18. Baxes G. A. "Digital Image Processing" John Wiley & Sons Inc.(1994).
19. Zhang T. Y. and Suen C. Y. "A fast parallel algorithm for thinning digital patterns". In Communications of the ACM, volume 27, pages 236-239, 1984