

# Watermarking Framework for Authentication and Self-recovery of Tampered Colour Images

J. Molina, V. Ponomaryov, *Member, IEEE*, R. Reyes, *Member, IEEE*, S. Sadovnychiy, *Member, IEEE*,  
and C. Cruz, *Member, IEEE*

**Abstract**—Methods of authentication and self-recovery of tampered information in digital images have been in constant development during the last years. These frameworks are developed by watermarking techniques, which generate an image digest, which is embedded inside the image that must be protected. In this paper, a method of authentication and self-recovery of tampered information in digital images is proposed. Our approach first generates the authentication watermarks, which are based on XOR operations on non-overlapping blocks, subsequently by using a halftoning technique the recovery watermark is generated. To have a higher performance, three copies of each watermark are embedded to have more chance to extract the recovery watermark. To evaluate the quality of the obtained images, the objective criterion of peak signal-to-noise ratio (PSNR) and Structural Similarity Index (SSIM) are used. The experimental results demonstrate the effectiveness of our method in comparisons with other schemes reported in the literature, where the quality of the watermarked images, the quality of the reconstruction images and the recovery rate of each scheme were evaluated.

**Index Terms**—Image Watermarking, Self-recovery, Image Authentication, Halftoning, Image restoration, Image Coding.

## I. INTRODUCCIÓN

EL uso de imágenes digitales ha estado en constante crecimiento en distintas áreas de investigación. Así mismo, han surgido herramientas que permiten editar el contenido visual de las imágenes digitales sin que el usuario final se percate de la presencia de las modificaciones realizadas, las cuales se pueden llevar a cabo con fines maliciosos. Una alternativa para solucionar estos problemas son las técnicas de marcas de agua, las cuales se centran en el ocultamiento de la información en archivos digitales [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12] y son principalmente utilizadas para la protección de derechos de autor y autenticación de contenido digital [1, 2, 3, 4, 5]. Durante los últimos años se han implementado métodos basados en marcas de agua para la reconstrucción a ciegas de información alterada en imágenes digitales [6, 7, 8, 9, 10, 11], los cuales consisten en generar una versión reducida en bits de la imagen original a proteger, estos bits son insertados como marca de agua, posteriormente durante la etapa de reconstrucción esta marca de agua es

extraída para generar una aproximación al contenido original de la imagen y reconstruir las regiones detectadas como alteradas.

En [6] se propone un método para la reconstrucción de información alterada, donde la imagen original se cifra mediante un esquema caótico, posteriormente se divide en bloques de  $2 \times 2$  píxeles no traslapados y para cada bloque se obtiene el valor promedio, el cual es insertado como marca de agua en dos bloques de mapeo mediante el método LSB en el dominio espacial, la principal aportación de este sistema es la aplicación del esquema caótico para cifrar la imagen [12, 13]. En [7] se propone un método para la detección y reconstrucción de alteraciones basado en SVD, este método divide la imagen a proteger en bloques no traslapados de  $2 \times 2$  píxeles, para cada bloque se obtiene el valor promedio y se extraen los 5 MSB, los cuales son utilizados como marca de agua para reconstrucción, por otra parte, se aplica SVD a cada bloque y se generan 3 bits utilizados para autenticación. En [8] se propone un esquema basado en DCT, la imagen a proteger es dividida en bloques no traslapados de  $2 \times 2$  píxeles, posteriormente se aplica la DCT a cada bloque y se cuantizan los coeficientes obtenidos, para cada bloque se generan 10 bits de reconstrucción y dos bits de autenticación. Lin [9] propone un método para la detección y reconstrucción de alteraciones basado en un esquema de marca de agua frágil, este sistema genera 6 bits de reconstrucción y 2 bits de autenticación por cada bloque de  $2 \times 2$  píxeles, para obtener un mejor desempeño durante la detección de alteraciones se emplean distintos niveles jerárquicos de detección. En [10] utilizan una tabla de búsqueda (Look-up table) para seleccionar las regiones de inserción de la marca de agua de reconstrucción, mediante esta tabla se minimiza un problema llamado coincidencia de manipulación. Es importante destacar que ningún método anteriormente citado es capaz de soportar ataques que no modifican el aspecto de la imagen, como son presencia de ruido, escalamiento, ecualización de histograma, etc., debido al método frágil de marca de agua implementado.

Entre los problemas fundamentales de los sistemas de detección y reconstrucción de información alterada en imágenes se encuentra el problema de coincidencia de manipulación, el cual se da cuando una región de la imagen es alterada y la ubicación donde se encuentran los bits insertados para reconstruir dicha región también es alterada, de esta manera es imposible reconstruir la región. Para solucionar este problema, algunos autores insertan como marca de agua dos copias de los bits de reconstrucción [7, 10], para tener así una segunda oportunidad para reconstruir la región

This work was supported by the Instituto Politécnico Nacional of México and the Consejo Nacional de Ciencia y Tecnología (CONACyT, under Grant 220347).

J. Molina, V. Ponomaryov, R. Reyes and C. Cruz are with Instituto Politécnico Nacional (IPN), Mexico City, México (e-mails: jmolina@alumno.ipn.mx, vponomar@ipn.mx, rreyes@ipn.mx and ccruz@ipn.mx).

S. Sadovnychiy is with Instituto Mexicano del Petróleo, Mexico City, México (e-mail: ssadovny@imp.mx).

alterada, así mismo, otros esquemas realizan una fase de post-procesamiento basado en *inpainting* para regenerar las regiones que no se pueden reconstruir [10].

En el presente trabajo se propone un método para la detección de alteraciones y reconstrucción de información visual alterada en imágenes a color, el cual utiliza un esquema de marca de agua frágil, donde se insertan tres copias de la marca de agua de reconstrucción y autenticación, atacando así el problema de coincidencia de manipulación [6, 7, 8, 9, 10], así mismo se presenta un método para la detección de alteraciones mediante niveles jerárquicos como en [9] y se generan tres marcas de agua de autenticación para que en caso de que exista una falsa detección, esta pueda ser corregida mediante las otras dos marcas de agua. La generación de los bits de reconstrucción se realiza mediante el método Halftoning [14] y un procesamiento aplicado a los canales de crominancia. La generación de los bits de autenticación se generan mediante operaciones XOR aplicado a los canales RGB de la imagen empleando bloques de  $n \times n$  no traslapados.

Las principales contribuciones del método propuesto con respecto a los trabajos mostrados en el estado del arte son:

- ◊ Un mejor desempeño en comparación con otros esquemas durante el proceso de reconstrucción de alteraciones. Esto se debe a dos características importantes implementadas: (I) El método propuesto implementa la fase de autenticación mediante tres marcas de agua, donde cada marca de agua detecta alteraciones en cada canal RGB de la imagen procesada, mejorando así la capacidad de detección de alteraciones. Este proceso aumenta la probabilidad de que la marca de agua de reconstrucción no se elimine ante altas tasas de alteración; (II) Método *inpainting* aplicado a las regiones afectadas por el problema de coincidencia de manipulación, permite mejorar la calidad de la imagen utilizada para reconstrucción.
- ◊ Método de inserción empleado (2-LSB) permite obtener la calidad de las imágenes marcadas arriba de 44 dB en PSNR y 0.9700 en SSIM que son valores de alta calidad, y la marca de agua insertada se mantiene imperceptible al Sistema Visual Humano (HVS, por sus siglas en inglés).

El resto del artículo se encuentra organizado como sigue, en la sección II se presenta el sistema propuesto, en la sección III se presentan los resultados obtenidos durante las pruebas realizadas, finalmente, en la sección IV se muestran las conclusiones generadas.

## II. SISTEMA PROPUESTO

En esta sección se propone un sistema para la detección y reconstrucción de alteraciones en imágenes digitales, el cual puede ser dividido en dos etapas: a) Protección de la imagen, y b) Auto-recuperación. La Fig. 1 muestra el diagrama a bloques del proceso de protección de la imagen. La Fig. 2 muestra el diagrama a bloques del proceso de auto-recuperación.

Cada etapa del sistema propuesto utiliza la siguiente nomenclatura:

$I_h$ : Imagen original a proteger;  $W_{Ib}$ : Marca de agua del canal de luminancia de  $I_h$ ;  $W_{Cb}$ ,  $W_{Cr}$ : Marca de

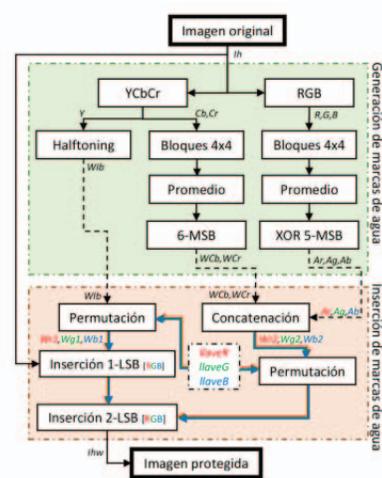


Fig. 1. Diagrama a bloques de la etapa “Protección de la imagen”.

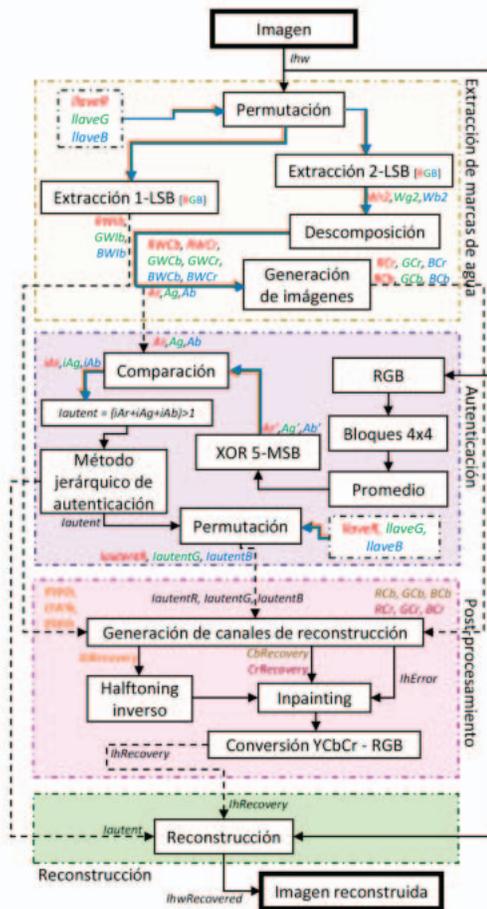


Fig. 2. Diagrama a bloques de la etapa “Auto-recuperación”.

agua de los canales de crominancia de  $I_h$ ;  $Ar$ ,  $Ag$ ,  $Ab$ : Marcas de agua de autenticación para los canales RGB de  $I_h$ ;  $W_{r1}$ ,  $W_{g1}$ ,  $W_{b1}$ : Copias permutadas de la marcas de agua  $W_{Ib}$ ;  $W_{r2}$ ,  $W_{g2}$ ,  $W_{b2}$ : Vectores de marcas de agua permutados, formados por la concatenación de  $\{W_{Cb}$ ,  $W_{Cr}$ ,  $Ar\}$ ,  $\{W_{Cb}$ ,  $W_{Cr}$ ,  $Ag\}$ ,  $\{W_{Cb}$ ,  $W_{Cr}$ ,  $Ab\}$ ;  $I_{hw}$ :

Imagen marcada;  $RWIb$ ,  $GWib$ ,  $BWIb$ : Marcas de agua extraídas de cada canal RGB de  $Ihw$ , aproximación de  $WIb$ ;  $RWCb$ ,  $RWCr$ ,  $GCb$ ,  $GCr$ ,  $BWCb$ ,  $BWCr$ : Vectores de marca de agua para los canales de crominancia, extraídos de los canales RGB de  $Ihw$ ;  $RCb$ ,  $RCr$ ,  $GCb$ ,  $GCr$ ,  $BCb$ ,  $BCr$ : Imágenes de crominancia generados mediante  $RWCb$ ,  $RWCr$ ,  $GCb$ ,  $GCr$ ,  $BWCb$ ,  $BWCr$ ;  $iAr$ ,  $iAg$ ,  $iAb$ : Imágenes de autenticación para cada canal RGB de  $Ihw$ ;  $Iautent$ : Imagen de autenticación general para  $Ihw$ ;  $IautentR$ ,  $IautentG$ ,  $IautentB$ : Errores de extracción para las marcas de agua extraídas de cada canal RGB de  $Ihw$ ;  $IbRecovery$ : Imagen de reconstrucción para el canal de luminancia;  $CbRecovery$ ,  $CrRecovery$ : Imágenes de reconstrucción de los canales de crominancia;  $Iautent$ : Errores de extracción para  $IbRecovery$ ,  $CbRecovery$ ,  $CrRecovery$ ;  $IhwRecovered$ : Imagen reconstruida.

#### A. Generación de Marcas de Agua

Sea la imagen a proteger  $Ih$  la cual es de tamaño  $M \times N$  filas y columnas respectivamente, donde se asume que  $M$  y  $N$  son múltiplo de 4, en caso de que el tamaño de las filas o las columnas de la imagen no sea múltiplos de 4, la imagen es recortada de tal forma que cumpla esta condición. Esta imagen es convertida del espacio de color RGB al espacio de color YCbCr. Se aplica el método de Halftoning [14] al canal de luminancia utilizando el filtro propuesto por Floyd [14] obteniendo una imagen de semitono o binaria ( $WIb$ ) de  $M \times N$ , este método es implementado debido a que reduce los bits de profundidad de una imagen pasando de 8 bits a 1 bit, así mismo es posible generar imágenes en escala de grises a partir de su contraparte en semitono obteniéndose una imagen con buena calidad [15, 16]. Los canales de crominancia se dividen en bloques no traslapados de  $4 \times 4$  píxeles, para cada bloque se calcula el valor promedio y se extraen los 6 MSB, los bits generados de cada bloque son concatenados, de esta manera se obtienen dos vectores binarios ( $WCb$  y  $WCr$ ), uno por cada canal de crominancia.

La generación de los bits para detección de alteraciones se realiza de la siguiente manera, cada canal del espacio de color RGB de  $Ih$  es procesado de manera independiente, cada canal es dividido en bloques no traslapados de  $4 \times 4$  píxeles, para cada bloque se hacen cero los 3 LSB de cada pixel y se obtiene el valor promedio, finalmente para el valor obtenido se aplica la operación XOR a los 5 MSB utilizando pares de bits traslapados, obteniendo 4 bits de autenticación por cada bloque. Por ejemplo, si el resultado después de obtener el promedio de un bloque es  $163_{10}$ , esto es  $1010000_2$ , aplicando la operación XOR a los 5 MSB tomando pares traslapados de bits queda  $1110_2$ . Los bits de autenticación obtenidos para cada canal se concatenan, de esta manera se obtienen tres marcas de agua de autenticación ( $Ar$ ,  $Ag$  y  $Ab$ ).

Al finalizar el proceso de generación de marcas de agua se tienen por cada bloque de  $4 \times 4$  píxeles de la imagen RGB la siguiente cantidad de bits: 16 bits de  $WIb$ , 6 bits de  $WCb$ , 6 bits de  $WCr$  y 4 bits para representar  $Ar$ ,  $Ag$  o  $Ab$ . De esta manera, es posible insertar la información de  $WIb$ ,  $WCb$ ,  $WCr$  y  $Ar$  en un solo canal de la imagen a color empleando 2-LSB.

#### B. Inserción de Marcas de Agua

La inserción de las marcas de agua de autenticación y reconstrucción se lleva a cabo mediante el método LSB, realizando la inserción en los dos LSB de cada canal RGB de  $Ih$ . Se generan tres copias de  $WIb$ , denotándose cada copia como  $Wr1$ ,  $Wg1$  y  $Wb1$ , cada copia es dividida en bloques no traslapados de  $4 \times 4$  píxeles, mismos que son permutados utilizando una llave distinta para cada copia ( $llaveR$ ,  $llaveG$  y  $llaveB$ ). Cada copia permutada se inserta como marca de agua en el primer bit LSB de cada canal RGB de  $Ih$ . De esta manera la marca de agua correspondiente al canal de luminancia puede ser recuperada de tres regiones distintas de la imagen. Posteriormente,  $WCb$  y  $WCr$  son divididos en bloques no traslapados de  $1 \times 6$ , por otra parte, cada marca de agua de autenticación ( $Ar$ ,  $Ag$  y  $Ab$ ) es dividida en bloques no traslapados de  $1 \times 4$ . Se realiza una concatenación del  $i$ -ésimo bloque binario de  $WCb$ ,  $WCr$  y  $Ar$  mediante la ecuación:

$$Wr2 = \{WCb_i, WCr_i, Ar_i, \dots, WCb_n, WCr_n, Ar_n\} \quad (1)$$

donde  $i$  representa el  $i$ -ésimo bloque y  $n$  representa el último bloque de cada vector. Se realiza este proceso nuevamente utilizando los vectores  $WCb$ ,  $WCr$ ,  $Ag$  y finalmente utilizando  $WCb$ ,  $WCr$ ,  $Ab$ , mediante las siguientes ecuaciones:

$$Wg2 = \{WCb_i, WCr_i, Ag_i, \dots, WCb_n, WCr_n, Ag_n\} \quad (2)$$

$$Wb2 = \{WCb_i, WCr_i, Ab_i, \dots, WCb_n, WCr_n, Ab_n\} \quad (3)$$

De esta manera se obtienen tres vectores de marca de agua denotados como  $Wr2$ ,  $Wg2$  y  $Wb2$ , finalmente, empleando las llaves  $llaveR$ ,  $llaveG$  y  $llaveB$  cada vector de marca de agua es permutado utilizando bloques no traslapados de  $1 \times 16$ . Finalmente, durante el proceso de inserción  $Ih$  se divide en bloques no traslapados de  $4 \times 4$  píxeles, cada  $i$ -ésimo bloque de cada vector de marca de agua ( $Wr2$ ,  $Wg2$  y  $Wb2$ ) es insertado en el segundo bit LSB del  $i$ -ésimo bloque de  $Ih$ , insertando  $Wr2$ ,  $Wg2$  y  $Wb2$  en los canales R, G y B respectivamente. Al finalizar el proceso de inserción se obtiene la imagen marcada, la cual es denotada en este artículo como  $Ihw$ .

#### C. Extracción de Marcas de Agua

En esta etapa, se realiza el proceso de extracción de las marcas de agua de  $Ihw$ , la cual es sospechosa de alteración. Cada canal del espacio de color RGB de  $Ihw$  es dividido en bloques no traslapados de  $4 \times 4$  píxeles, mismos que son permutados utilizando las llaves  $llaveR$ ,  $llaveG$  y  $llaveB$ , respectivamente. Posteriormente, se extrae el primer bit LSB de cada canal de  $Ihw$ , de esta manera se obtienen tres imágenes binarias denotadas como  $RWIb$ ,  $GWib$  y  $BWIb$ . De manera subsecuente, cada canal RGB de  $Ihw$  es dividido en bloques no traslapados de  $4 \times 4$  píxeles, para cada bloque se extrae el segundo LSB obteniendo un total de 16 bits por bloque, los cuales son concatenados hasta formar  $Wr2$  para el canal R,  $Wg2$  para el canal G y  $Wb2$  para el canal B. Tomando

en consideración la generación de  $Wr2$ , el cual es mostrado en el proceso de inserción de marcas de agua, se sabe que cada bloque no traslapado de  $1 \times 16$  de  $Wr2$  está formado por 6 bits correspondientes a  $WCb$ , 6 bits correspondientes a  $WCr$  y 4 bits correspondientes a  $Ar$ , por lo tanto se generan estos vectores definidos como  $RWCb$ ,  $RWCr$  y  $Ar$ , respectivamente. De la misma manera se generan los vectores  $GWCb$ ,  $GWCr$ ,  $Ag$  y  $BWCb$ ,  $BWCr$ ,  $Ab$  de  $Wg2$  y  $Wb2$ , respectivamente.

Finalmente, se generan seis imágenes de tamaño  $M \times N$  denominadas  $RCb$ ,  $RCr$ ,  $GCb$ ,  $GCr$ ,  $BCb$  y  $BCr$ , las cuales son llenadas de la siguiente manera: los vectores  $RWCb$ ,  $RWCr$ ,  $GWCb$ ,  $GWCr$ ,  $BWCb$  y  $BWCr$  son divididos en bloques no traslapados de  $1 \times 6$ , cada bloque es transformado de binario a decimal y el valor obtenido es asignado al  $i$ -ésimo bloque de  $4 \times 4$  pixeles de su correspondiente imagen generada.

#### D. Autenticación

En esta etapa se aplica el proceso de generación de bits de autenticación a la imagen  $Ihw$ , la cual es sospechosa de alteración. El canal rojo de  $Ihw$  es dividido en bloques no traslapados de  $4 \times 4$  pixeles y se aplica el proceso de generación de bits de autenticación, mostrado en la etapa generación de marcas de agua, estos bits generados son comparados con cada  $i$ -ésimo bloque de  $1 \times 4$  de la marca de agua extraída  $Ar$ , en caso de que los bits no sean iguales dicho bloque se considera alterado y cada valor del bloque es seleccionado con un valor 1, en caso contrario es considerado un bloque auténtico y cada valor del bloque se selecciona con un valor 0. Este proceso se realiza para los canales verde y azul y las marcas de agua extraídas  $Ag$  y  $Ab$  respectivamente. Al finalizar este proceso se obtienen tres imágenes de autenticación para cada canal RGB ( $iAr$ ,  $iAg$  e  $iAb$ ). Es importante mencionar que cada una de estas imágenes de autenticación puede tener falsas detecciones debido al problema de coincidencia de manipulación, por lo tanto utilizando  $iAr$ ,  $iAg$  e  $iAb$  se realiza la suma de la  $(i, j)$ -ésima posición de cada pixel de estas imágenes, en caso de que el resultado sea 3 dicho pixel es marcado con 1, lo que significa que el pixel es alterado, en caso contrario el pixel es marcado con 0, la imagen obtenida es denotada como  $Iautent$ . Finalmente se realiza un método jerárquico a  $Iautent$  como en [9] para mejorar la detección de alteraciones.

Por otra parte, se generan tres imágenes, las cuales representan errores de extracción para cada copia de las marcas de agua de luminancia y crominancia, primero se generan tres copias de la imagen  $Iautent$ , las cuales son divididas en bloques no traslapados de  $4 \times 4$  pixeles, empleando una llave por cada copia ( $llaveR$ ,  $llaveG$  y  $llaveB$ ) cada bloque de cada copia es permutado. Las imágenes obtenidas ( $IautentR$ ,  $IautentG$  y  $IautentB$ ) para cada copia muestran los errores de extracción para las marcas de agua extraídas de cada canal RGB.

Este método de autenticación elimina la aparición de Falsos Positivos mediante las imágenes  $iAr$ ,  $iAg$  e  $iAb$ , posteriormente la aparición de Falsos Negativos son minimizados mediante el proceso jerárquico de autenticación.

La Fig. 3 muestra la curva ROC al utilizar este método de autenticación empleando 500 imágenes de prueba y diversos ataques de alteración, donde se tiene un área bajo la curva de 0.8098. Este método de autenticación podría ser mejorado al utilizar bloques más pequeños durante el proceso de autenticación ( $2 \times 2$ ) y una mayor cantidad de bits de marca de agua durante autenticación, sin embargo en la presente propuesta el método de autenticación utilizado fue seleccionado debido a que se requería poca información en bits para esta etapa.

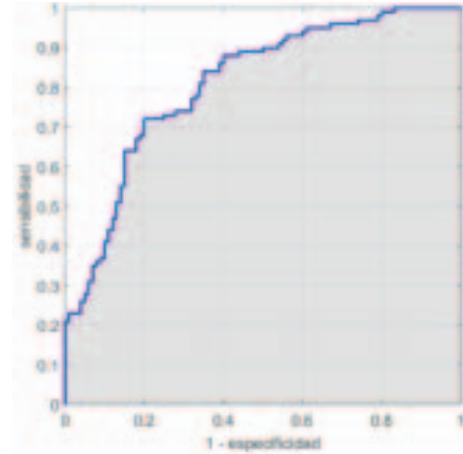


Fig. 3. Curva ROC para el proceso de autenticación.

#### E. Post-Procesamiento

En esta etapa se genera la imagen utilizada para reconstruir las alteraciones detectadas en  $Ihw$ . Primero, utilizando las imágenes  $IautentR$ ,  $IautentG$  y  $IautentB$ , y empleando las tres marcas de agua  $RWIb$ ,  $GWIb$  y  $BWIb$ , se genera una copia fiel de la marca de agua original  $Ib$  aplicando los siguientes pasos: primero se generan dos imágenes rellenas con ceros ( $IbRecovery$  y  $IhError$ ) de tamaño  $M \times N$  filas y columnas, respectivamente. Posteriormente se analiza cada  $(i, j)$ -ésimo pixel de  $IautentR$ , si el pixel es auténtico (valor 0) se asigna el valor del  $(i, j)$ -ésimo pixel de la imagen  $RWIb$  al  $(i, j)$ -ésimo pixel de  $IbRecovery$ , en caso de que el pixel no sea auténtico (valor 1) se analiza el  $(i, j)$ -ésimo pixel de  $IautentG$ , en caso de que este pixel sea auténtico se asigna el valor del  $(i, j)$ -ésimo pixel de  $GWIb$  al  $(i, j)$ -ésimo pixel de  $IbRecovery$ , en caso contrario se analiza el  $(i, j)$ -ésimo pixel de  $IautentB$  si este pixel es auténtico se asigna el valor del  $(i, j)$ -ésimo pixel de  $BWIb$  al  $(i, j)$ -ésimo pixel de  $IbRecovery$ , en caso contrario significa que dicho pixel no se puede reconstruir ya que las 3 copias de marca de agua de  $Ih$  en la  $(i, j)$ -ésima posición fueron alteradas, en este caso a la  $(i, j)$ -ésima posición de  $IhError$  se asigna con el valor 1. Este proceso se aplica nuevamente a las imágenes  $RCb$ ,  $GCb$ ,  $BCb$  y a las imágenes  $RCr$ ,  $GCr$ ,  $BCr$ , generando las imágenes  $CbRecovery$  y  $CrRecovery$ . Al finalizar este proceso se obtienen tres imágenes, la primera correspondiente al canal de luminancia, la segunda y tercera correspondiente a los canales de crominancia de la imagen, así mismo se

obtiene una imagen la cual indica las regiones afectadas por el problema de coincidencia de manipulación ( $IhError$ ).

A la imagen  $IbRecovery$ , la cual es una imagen de semitono, se le aplica un proceso de Halftoning inverso el cual es mostrado en [15], este proceso se realiza debido a que este método posee baja complejidad computacional y la calidad de las imágenes generadas es superior a 30 dB, tomando en consideración que esta es una calidad aceptable para los sistemas de Halftoning inverso [15, 16]. Posteriormente, se aplica un proceso de *inpainting*, el cual es mostrado en [17], a las regiones que no se pudieron generar de  $IbRecovery$ ,  $CbRecovery$  y  $CrRecovery$ , estas regiones son mostradas en  $IhError$ . Finalmente las imágenes generadas, las cuales representan los canales del espacio de color YCbCr de  $Ih$ , se pasan al espacio de color RGB, obteniéndose  $IhRecovery$ .

#### F. Reconstrucción de Alteraciones

En esta etapa se reconstruyen las regiones alteradas de  $Ihw$ . Se realiza un análisis de cada  $(i, j)$ -ésimo pixel de  $Iautent$ , en caso de que el pixel sea 1 (alteración) se copia el valor del  $(i, j)$ -ésimo pixel de  $IhRecovery$  al  $(i, j)$ -ésimo pixel de  $Ihw$ . Obteniéndose  $IhwRecovered$ .

### III. RESULTADOS

Durante la evaluación del sistema se utilizaron las imágenes de prueba extraídas de las bases de datos SIPI [18] y Kodak [19], las cuales fueron alteradas utilizando el software Gimp®. Las pruebas se implementaron en cuatro fases: a) calidad de la imagen marcada, b) tasa de reconstrucción de alteraciones, c) precisión durante la detección de alteraciones, d) calidad de las imágenes reconstruidas. Para cada evaluación se utilizaron los criterios de PSNR (dB) y SSIM. Finalmente, se realizaron las comparaciones respectivas con algunos métodos del estado del arte [6, 7, 8, 10]. Para llevar a cabo la presentación gráfica de resultados, se emplearon imágenes con distintas características como regiones texturizadas, regiones lisas e integración de ambas características.

#### A. Calidad de la Imagen Marcada

Durante esta prueba se evalúa la calidad de la imagen marcada o protegida ( $Ihw$ ).

La tabla I muestra los resultados para distintos métodos del estado del arte al evaluar la calidad promedio de las imágenes marcadas. La evaluación de la imagen marcada ( $Ihw$ ) se realizó comparando la imagen protegida ( $Ihw$ ), la cual no presenta alteración, contra la imagen original ( $Ih$ ). Como se puede observar en la tabla I, el esquema propuesto genera imágenes protegidas de mayor calidad a diferencia de las propuestas en [6, 8, 10], esto se debe a que en el sistema propuesto y el esquema en [7] únicamente utilizan dos LSB de  $Ih$  para insertar las marcas de agua generadas, por otra parte, los esquemas en [6, 8, 10] realizan la inserción de las marcas de agua en los tres LSB, esto tiende a alterar de manera significativa la calidad de la imagen. Es importante mencionar que para el sistema propuesto se insertan tres copias de la marca de agua de reconstrucción, en [7, 10] insertan dos copias de la marca de agua de reconstrucción, finalmente en [6, 8] insertan una copia de la marca de agua de reconstrucción.

TABLA I  
CALIDAD PROMEDIO (PSNR Y SSIM) DE LA IMAGEN MARCADA ( $Ihw$ )

Técnica	PSNR	SSIM
Método propuesto	44.12	0.9825
Xiaojun [6]	38.48	0.9353
Sajjad [7]	44.01	0.9780
Durgesh [8]	38.65	0.9553
Lee [10]	40.51	0.9345

#### B. Tasa de Reconstrucción de Alteraciones

La tasa de reconstrucción de alteraciones indica el porcentaje total de alteraciones en una imagen, el cual puede ser reconstruido. Como se mencionó anteriormente, la tasa de reconstrucción de alteraciones está dada principalmente por el número de copias insertadas de la marca de agua, debido a que entre más copias sean insertadas, la marca de agua de reconstrucción puede ser extraída de más regiones distintas de la imagen, solventando así el problema de *coincidencia de manipulación*. Es importante mencionar que la tasa de reconstrucción de alteraciones está dada de manera teórica para el mejor de los casos, debido a que las copias insertadas de la marca de agua pueden ser removidas dependiendo de la ubicación de las alteraciones realizadas, debido a esto, algunos métodos optan por una fase de *inpainting* para reconstruir las regiones cuyas copias de marca de agua de reconstrucción sean destruidas [10]. La Tabla II muestra la tasa de reconstrucción de alteraciones para cada método de acuerdo al número de copias insertadas y al proceso de *inpainting*. Como se puede observar, el sistema propuesto mantiene un mejor rendimiento debido a que se insertan más copias de reconstrucción y al proceso de *inpainting*.

TABLA II  
TASA DE RECONSTRUCCIÓN DE ALTERACIONES PARA DISTINTOS ESQUEMAS

Técnica	Copias insertadas	Proceso de <i>inpainting</i>	Tasa de reconstrucción de alteraciones
Método propuesto	3	Si	≈ 77%
Xiaojun [6]	1	No	≈ 50%
Sajjad [7]	2	No	≈ 65%
Durgesh [8]	1	No	≈ 50%
Lee [10]	2	Si	≈ 70%

#### C. Precisión Durante la Detección de Alteraciones

Durante esta etapa se realizaron distintos tipos de alteraciones a las imágenes de prueba utilizadas, posteriormente se ejecutó la fase de autenticación, obteniéndose así la imagen de autenticación ( $Iautent$ ). Las Figs. 4 y 5 muestran los resultados obtenidos para cada alteración realizada y su posterior detección mediante un primer nivel jerárquico de autenticación.

Durante el proceso de detección de alteraciones, debido a que se lleva a cabo mediante marcas de agua ( $Ar$ ,  $Ag$  y  $Ab$ ), el problema de coincidencia de manipulación se mantiene presente, es por eso que existen falsas detecciones durante la

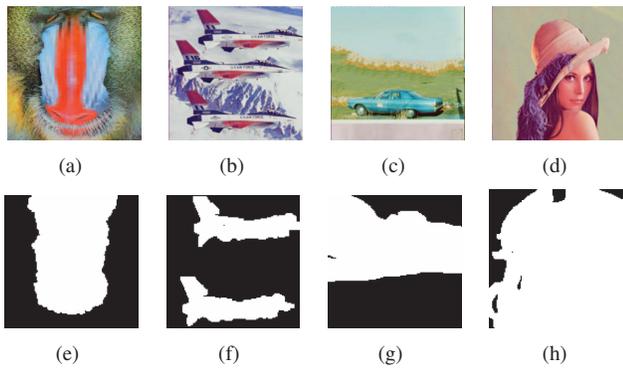


Fig. 4. Precisión durante detección de alteraciones para imágenes obtenidas de SIPI. (a-d) Imágenes alteradas, (e-h) Detección de alteraciones - primer nivel jerárquico de autenticación.

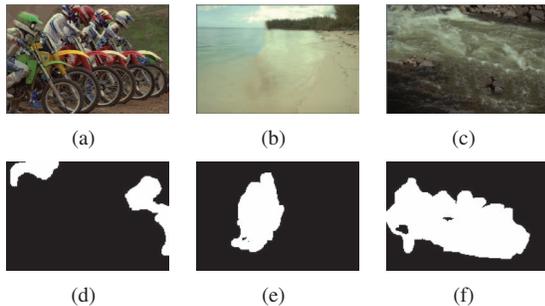


Fig. 5. Precisión durante detección de alteraciones para imágenes obtenidas de Kodak. (a-c) Imágenes alteradas, (d-f) Detección de alteraciones - primer nivel jerárquico de autenticación.

detección de alteraciones. Para solventar esta problemática se aplica un primer nivel jerárquico de autenticación, mostrado en [9], mejorando así la precisión del proceso de detección de alteraciones, como se muestra en Figs. 4 y 5. Se puede observar que el sistema propuesto detecta correctamente los ataques intencionales realizados.

#### D. Calidad de las Imágenes Reconstruidas

En esta etapa se muestran los resultados de las imágenes reconstruidas, el conjunto de imágenes utilizadas son mostradas en Figs. 4 (a) – (d) y Figs. 5 (a) – (c). Fig. 6 y Fig. 7 muestran los resultados obtenidos durante el proceso de reconstrucción.

Como se puede observar, la calidad obtenida en las imágenes es considerablemente alta de acuerdo a las métricas de calidad PSNR y SSIM, a pesar de que se presentan alteraciones que degradan significativamente el contenido visual de cada imagen, el problema de coincidencia de manipulación afecta a las marcas de agua de autenticación y reconstrucción, adicionalmente, se observa que el proceso de *inpainting* elimina significativamente el ruido generado (Figs. 6 (e) – (h) y Figs. 7 (d) – (f)).

#### E. Comparación

En esta etapa se realizó la comparación de la calidad de las imágenes reconstruidas empleando diversas tasas de alteración para diversos esquemas que trabajan con imágenes a color.

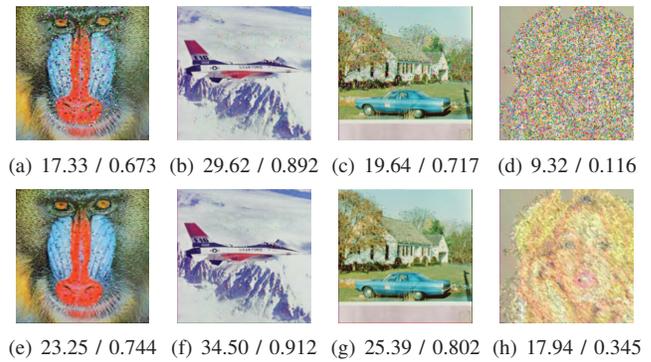


Fig. 6. Calidad (PSNR(dB) / SSIM) de las imágenes reconstruidas para imágenes obtenidas de SIPI. (a-d) Imágenes reconstruidas sin *inpainting*, (e-h) Imágenes reconstruidas con *inpainting*.



Fig. 7. Calidad (PSNR(dB) / SSIM) de las imágenes reconstruidas para imágenes obtenidas de Kodak. (a-c) Imágenes reconstruidas sin *inpainting*, (d-f) Imágenes reconstruidas con *inpainting*.

Fig. 8 muestra las alteraciones realizadas para la imagen Lenna.

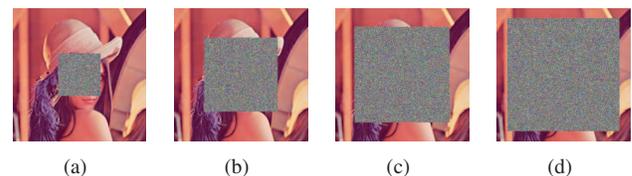


Fig. 8. Imágenes alteradas. a) 10%, b) 30%, c) 50%, d) 70%.

Fig. 9 muestra los resultados obtenidos de la reconstrucción de información con diferentes tasas de alteración (10%, 30%, 50% y 70%) para los esquemas del estado del arte y el método propuesto. Se puede observar que el método propuesto obtiene una mayor calidad en la imagen reconstruida, esto debido a que inserta tres copias de la marca de agua de reconstrucción e implementa una fase de *inpainting* a la imagen generada para el proceso de reconstrucción de alteraciones. Así mismo, los métodos mostrados en [6, 7, 8] únicamente insertan a lo máximo dos copias de la marca de agua de reconstrucción, por lo tanto la calidad de la imagen reconstruida se ve afectada por el problema de *coincidencia de manipulación*.

Finalmente, empleando las bases de datos SIPI [18] y Kodak [19] se realizaron diversas pruebas de reconstrucción, donde las Figs. 10 y 11 muestran la calidad promedio (PSNR y SSIM, respectivamente) después de reconstruir regiones alteradas en las imágenes utilizadas, para tasas de alteración que van de

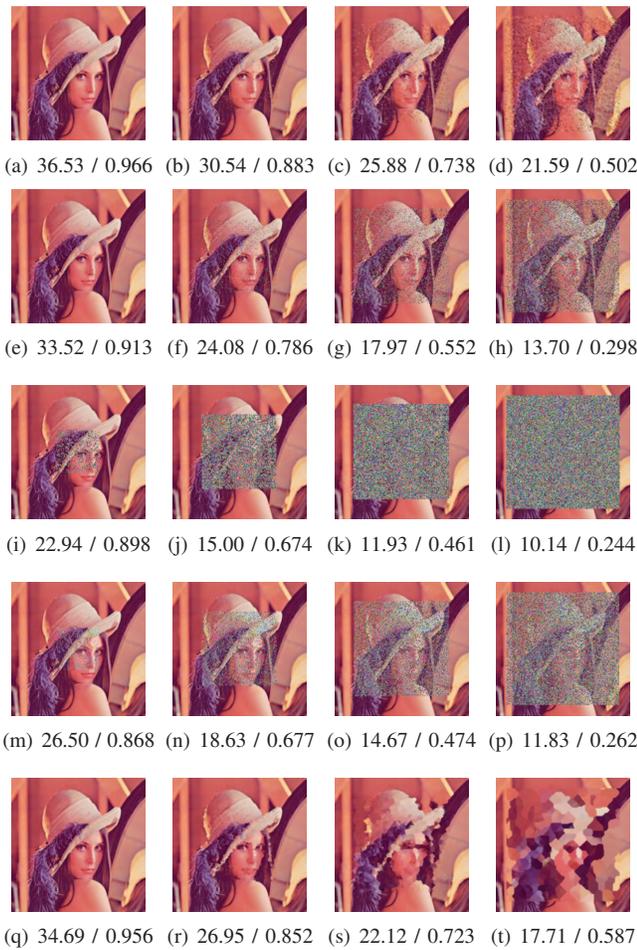


Fig. 9. PSNR(dB) / SSIM promedio para la reconstrucción de la imagen Lenna. (a - d) Método propuesto, (e - h) Xiaojun [6], (i - l) Sajjad [7], (m - p) Durgesh [8], (q - t) Lee [10].

10% hasta 70%.

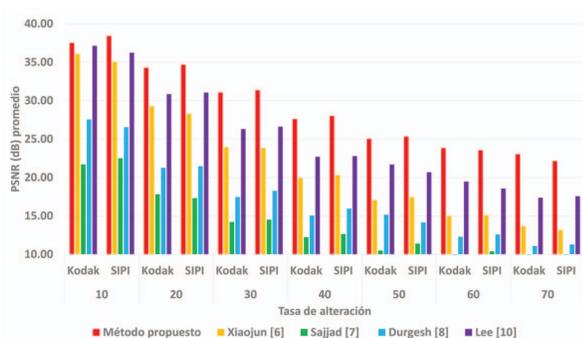


Fig. 10. Calidad promedio (PSNR) para las imágenes reconstruidas ante distintas tasas de alteración.

#### IV. CONCLUSIONES

En el presente trabajo se propone un esquema para la autenticación y reconstrucción de información alterada en imágenes digitales a color, este sistema genera una versión reducida en bits de la imagen a proteger, así mismo se generan bits para la detección de alteraciones, esta información

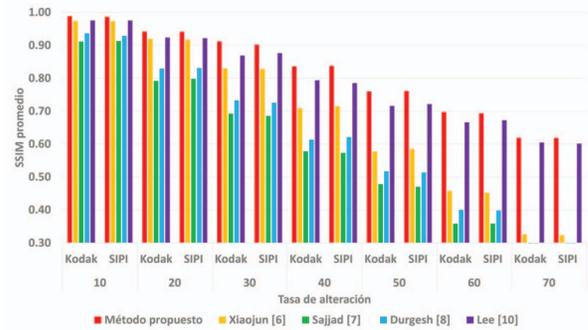


Fig. 11. Calidad promedio (SSIM) para las imágenes reconstruidas ante distintas tasas de alteración.

generada para reconstrucción se inserta tres veces como marca de agua dentro de la imagen a proteger, solventando así el problema de *coincidencia de manipulación*, el cual se encuentra presente en otros esquemas del estado del arte. La generación de la marca de agua de reconstrucción está basada principalmente en *Halftoning* y un procesamiento aplicado a los canales de crominancia de la imagen digital, de esta manera se reduce significativamente la información en bits para representar la imagen a proteger. Finalmente, se implementó un método de *inpainting*, el cual demostró incrementar la calidad objetiva (PSNR y SSIM) de la imagen utilizada para reconstruir alteraciones (Figs. 6 y 7). Los resultados experimentales muestran una excelente calidad para la imagen protegida (Tabla I). Así mismo, el proceso de autenticación presenta resultados precisos en la detección de alteraciones (Figs. 4 y 5). Finalmente, debido a que el método propuesto inserta tres copias de la marca de agua de reconstrucción se puede obtener una tasa de reconstrucción superior a la de otros esquemas del estado del arte, ya que únicamente insertan hasta dos copias de la marca de agua de reconstrucción (Figs. 9, 10 y 11).

#### AGRADECIMIENTOS

Los autores agradecen al Instituto Politécnico Nacional de México, a la Comisión de Operación y Fomento de Actividades Académicas (COFAA) del IPN y al Consejo Nacional de Ciencia y Tecnología (CONACyT, proyecto 220347) por su apoyo otorgado para realizar este trabajo.

#### REFERENCIAS

- [1] Rogelio Reyes, Clara Cruz, Mariko Nakano-Miyatake, Hector Perez-Meana, *Digital Video Watermarking in DWT Domain Using Chaotic Mixtures*, in IEEE Latin America Transactions. Vol. 8, Issue 3, pp. 304-310, 2010.
- [2] Victor Hernandez-Guzman, Clara Cruz-Ramos, Mariko Nakano-Miyatake, Hector Perez-Meana, *Watermarking Algorithm based on the DWT*, in IEEE Latin America Transactions. Vol. 4, Issue 4, pp. 257-267, 2006.
- [3] L. Vargas and E. Vera, *An Implementation of Reversible Watermarking for Still Images*, in IEEE Latin America Transactions, vol. 11, no. 1, pp. 54-59, 2013.

- [4] D. Renza, D. M. Ballesteros and H. D. Ortiz, *Text Hiding in Images Based on QIM and OVSF*, in *IEEE Latin America Transactions*, vol. 14, no. 3, pp. 1206-1212, 2016.
- [5] A. Santos Brandao and D. Calhau Jorge, *Artificial Neural Networks Applied to Image Steganography*, in *IEEE Latin America Transactions*, vol. 14, no. 3, pp. 1361-1366, 2016.
- [6] X. Tong, Y. Liu, M. Zhang, Y. Chen, *A novel chaos – based fragile watermarking for image tampering detection and self – recovery*, in *Signal Processing: Image Communication*, Vol. 28, Issue 3, pp. 301 – 308, 2013.
- [7] S. Dadkhah, A. Abd Manaf, Y. Hori, A. Ella Hassanien, S. Sadeghi, *An effective SVD – based image tampering detection and self – recovery using active watermarking*, in *Signal Processing: Image Communication*, Vol. 29 Issue 10, pp. 1197 – 1210, 2014.
- [8] Durgesh Singh, Sanjay K. Singh, *Effective self – embedding watermarking scheme for image tampered detection and localization with recovery capability*, in *Journal of Visual Communication and Image Representation*, Vol 38, pp. 775 – 789, 2016.
- [9] Phen-Lan Lin, Chung-Kai Hsieh, Po-Whei Huang, *A hierarchical digital watermarking method for image tamper detection and recovery*, in *Pattern Recognition*, Vol. 38 Issue 12 pp. 2519–2529, 2005.
- [10] Tien-You Lee, Shinfeng D. Lin, *Dual watermark for image tamper detection and recovery*, in *Pattern Recognition*, Vol. 41, pp. 3497-3506, 2008.
- [11] L. Rosales Roldan, M. Cedillo Hernandez, J. Chao, M. Nakano Miyatake and H. Perez Meana, *Watermarking-based Color Image Authentication with Detection and Recovery Capability*, in *IEEE Latin America Trans.*, vol. 14, no. 2, pp. 1050-1057, 2016.
- [12] L.Wang, Q. Ye, Y. Q. Xiao, Y. X. Zou, B. Zhang, *An image encryption scheme based on cross chaotic map*, in *International Congress on Image and Signal Processing*, Vol.3, pp. 22–26, 2008.
- [13] X. L.Huang, *A novel improved cross chaotic based image encryption algorithm*, in *Imaging Science Journal* Vol. 9, pp. 1078–1083, 2012.
- [14] Robert W. Floyd, Louis Steinberg, *An adaptive algorithm for spatial greyscale*, in *Journal of the Society for Information Display*, Vol. 17, no. 2, pp. 75–77, 1976.
- [15] N. Damera-Venkata, T. D. Kite, M. Venkataraman, and B. L. Evans, *Fast Blind Inverse Halftoning*, *Proc. IEEE International Conference on Image Processing*, vol. 2, pp. 64-68, 1998.
- [16] F. Pelcastre-Jimenez, M. Nakano-Miyatake, K. Toscano-Medina, G. Sanchez-Perez, H. Perez-Meana, *An Inverse Halftoning Algorithms Based on Neural Networks and Atomic Functions*, *IEEE Latin America Transactions*, Vol. 15, Issue 3, pp. 488 – 495, 2017.
- [17] Idan Ram, Michael Elad, Israel Cohen, *Image Processing Using Smooth Ordering of its Patches*, *IEEE Transactions on Image Processing*, Vol. 22, Issue 7, pp. 2764 – 2774, 2013.
- [18] SIPI Image Database, Volume 3: Miscellaneous [Online] Available: <http://sipi.usc.edu/database/database.php?volume=misc>

(Accesado 23 Junio 2019).

- [19] Kodak Lossless True Color Image Suite [Online] Available: <http://r0k.us/graphics/kodak/> (Accesado 23 Junio 2019).



**Javier Molina García** Recibió el título de Ingeniero en Computación en 2014 y el grado de Maestro en Ingeniería en Seguridad y Tecnologías de la Información en 2016 por Instituto Politécnico Nacional (IPN), Ciudad de México (CM). Actualmente es estudiante de Doctorado en Comunicaciones y Electrónica (IPN), donde realiza investigación en el campo de procesamiento digital de señales, cómputo en paralelo (CPU y GPU) y marcas de agua.



**Volodymyr Ponomaryov** (Ph.D.' 1974, Dr. Sc.' 1981, Full Professor' 1984). Trabaja como profesor investigador en el IPN CM. Es Miembro del SNI (México), Nivel III. Sus principales actividades de investigación se basan en: procesamiento de imágenes y señales, filtrado en tiempo real, sensores médicos, etc. Es autor de más de 200 artículos científicos internacionales, 350 artículos en extenso en congresos internacionales, además de 23 patentes en ex-USSR, Rusia y México, y cinco libros científicos en editoriales internacionales.



**Rogelio Reyes Reyes** Recibió el título de Ingeniero en Comunicaciones y Electrónica, el grado de M. en C. de Ingeniería en Microelectrónica, y el grado de Doctor en Comunicaciones y Electrónica en el IPN, en 1999, 2003 y 2009, respectivamente. Actualmente es profesor investigador en el IPN. Sus áreas de interés son la esteganografía, las marcas de agua y el procesamiento digital de imágenes.



**Sergiy N. Sadovnychiy** Recibió el grado de Doctor en el Instituto de Aviación de Kharkov (ahora Universidad Nacional Aeroespacial), Ucrania, en 1991. Actualmente, se encuentra en el Instituto Mexicano del Petróleo, México. Sus intereses de investigación incluyen procesamiento digital de señales e imágenes, comunicaciones y diagnóstico acústico.



**Clara Cruz Ramos** Recibió el título de Ingeniera en Comunicaciones y Electrónica, el grado de M. en C. de Ingeniería en Microelectrónica, y el grado de Doctor en Comunicaciones y Electrónica en el IPN, en 1999, 2003 y 2009 respectivamente. Actualmente es profesora investigadora en el IPN. Sus áreas de investigación son reconocimiento de patrones, procesamiento de imágenes y marcas de agua.