UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE CIENCIAS

MÉTRICAS DE DESEMPEÑO EN EL RECONOCIMIENTO DE PATRONES EN IMÁGENES DIGITALES QUE PRESENTAN RUIDO

T E S I S

QUE PARA OBTENER EL TÍTULO DE

LICENCIADO EN FÍSICA

PRESENTA:

CÉSAR ALEJANDRO GUTIÉRREZ CERVANTES

Ensenada, Baja California, México. Mayo de 2016

MÉTRICAS DE DESEMPEÑO EN EL RECONOCIMIENTO DE PATRONES EN IMÁGENES DIGITALES QUE PRESENTAN RUIDO

T E S I S P R O F E S I O N A L

QUE PRESENTA

CÉSAR ALEJANDRO GUTIÉRREZ CERVANTES

Y APROBADA POR EL SIGUIENTE COMITÉ:

DRA. SELENE SOLORZA CALDERÓN

DIRECTOR DEL COMITÉ

DR. JESÚS LERMA ARAGÓN

SECRETARIO

DRA. BRENDA DE LA ROSA NAVARRO

PRIMER VOCAL

DEDICATORIA

A mis padres, por su apoyo incondicional tanto moral como económico durante toda la vida aun estando lejos de casa.

AGRADECIMIENTOS

Al CONACYT por la beca otorgada para la realización de esta tesis bajo el proyecto de Ciencia Básica SEP-CONACYT 2011 con número 169174 y titulado reconocimiento de patrones en imágenes digitales mediante máscaras binarias de anillos concéntricos y firmas unidimensionales invariantes a posición, rotación y escalamiento.

RESUMEN de la Tesis de César Alejandro Gutiérrez Cervantes presentada como requisito parcial para la obtención de la Licenciatura en Física. Ensenada Baja California, México, Mayo de 2016.

MÉTRICAS DE DESEMPEÑO EN EL RECONOCIMIENTO DE PATRONES EN IMÁGENES DIGITALES QUE PRESENTAN RUIDO

En este trabajo se propone medir la similitud de firmas unidimensionales generadas mediante el sistema de reconocimiento de patrones invariante a traslación, escala y rotación de Radon-Fourier-Mellin(RFM), utilizando la distancia de Hausdorff, la distancia euclidiana y la distancia del máximo. Así como usar la distancia de Hausdorff para proponer una métrica de desempeño para imágenes que presentan ruido gaussiano y, sal y pimienta.

Palabras Clave: Reconocimiento de patrones, distancia de Hausdorff, sistema RFM, métrica de desempeño.

Resumen aprobado por:

Dra. Selene Solorza Calderón

Índice general

1. INTRODUCCIÓN **1**

2. Fundamentos **3**

2.1. Representación de imágenes digitales 3

2.2. Métricas 6

2.3. l_p-métricas 7

2.4. Distancia de Hausdorff 10

2.5. Distancia de Hausdorff modificada 12

3. Sistema RFM **15**

3.1. Transformada de Fourier e invariancia a traslación 15

3.2. Transformada analítica de Fourier-Mellin e invariancia a escala 18

3.3. Transformada de Radon e invariancia a rotaciones 20

4. Métricas de desempeño **24**

4.1. Ruido 27

4.2. Coeficiente de similaridad de Hausdorff 30

4.3. Comparación del coeficiente de similaridad para imágenes en blanco y negro 32

4.4. Comparación del coeficiente de similaridad para imágenes en escala de grises 37

5. Conclusiones **41**

Índice de figuras

2.1.1.Representación esquemática de una imagen a color en el espacio de color RGB [12]. 3

2.1.2.Ejemplificación de la generación de una imagen digital. (a) Imagen continua. (b) Barrido de A a B en la imagen continua. (c) Cuantización y muestreo. (d) Línea escaneada digitalizada. (e) Imagen continua proyectada sobre un arreglo. (f) Resultado de la cuantización y muestreo[11]. 4

2.3.1.Comparación de la geometría taxicab con respecto a la distancia euclidiana[7]. 7

2.3.2.Ilustración de la métrica del máximo para un rey en un tablero de ajedrez[3]. 9

2.4.1.Distancia mínima entre dos formas irregulares ilustrada como la distancia entre los puntos rojos. 11

2.4.2.Componentes de la distancia de Hausdorf[4]. 12

2.5.1.Ilustración de la sensibilidad ante ruido de la distancia de Hausdorff. (a) Imagen de un círculo. (b) Imagen del mismo círculo con un punto trasladado. 14

3.1.1.Ejemplos de espectros de amplitud para imágenes trasladadas y rotadas. (a) Imagen I_1. (b) Imagen I_2. (c) Imagen I_3 . (d) $F_1 = \left|\mathcal{F}\{I_1)\}\right|$. (e) $F_2 = \left|\mathcal{F}\{I_2\}\right|$. (f) $F_3 = \left|\mathcal{F}\{I_3\}\right|$. 17

3.2.1.Ejemplos de la respuesta del sistema para imágenes trasladadas, escaladas y rotadas. (a) Imagen B_1. (b) Imagen B_2. (c) Imagen B_3. (d) $G_1 = \left|\mathcal{F}\{I_1\}\right|$. (e) $F_2 = \left|\mathcal{F}\{I_2\}\right|$. (f) $F_3 = \left|\mathcal{F}\{I_3\}\right|$. (g) M_1. (h) M_2. (i) M_3. 20

3.3.1.Representación de las recta L y p[20]. 21

3.3.2.Ejemplos de la respuesta del sistema RFMs. (a) Imagen C_1. (b) Imagen C_2. (c) Imagen C_3 . (d) Gráfica de comparación de las tres firmas 1D con $0 < p < 366$, pero para mejor visualización de la información relevante, se graficó de $50 < p < 320$. 23

4.0.1.Ejemplos de imágenes utilizadas para su comparación y sus firmas. (a) Imagen I_1. (b) S_1, firma 1D de I_1. (c) Imagen I_2. (d) S_2, firma 1D de I_2. (e) Imagen I_3. (f) S_3, firma 1D de I_3 26

4.1.1.Ejemplo de la distribución de ruido gaussiano con $\mu = 0$ y $\sigma = 0.2$. . . 28

4.1.2.Ejemplo de una imagen con ruido aditivo gaussiano. (a) Imagen sin ruido. (b) Imagen con ruido aditivo gaussiano de parámetros $\mu = 0$ y $\sigma^2 = 0.01$. 29

4.1.3.Visualización de una imagen con ruido sal y pimienta. (a) Imagen sin ruido. (b) Imagen con ruido sal y pimienta de parámetro $d = 0.05$. . . . 30

4.3.1.Ejemplos de imágenes en blanco y negro utilizadas para medir el coeficiente de similaridad. (a) Imagen objetivo. (b) S_T: Firma de la imagen objetivo. (c) Imagen problema con ruido aditivo gaussiano de media $\mu = 0$ y varianza $\sigma^2 = 0.3$. (d) S_{TN}: Firma de la imagen problema. (e) Imagen del ruido de fondo. (f) S_{FN}: Firma de la imagen del ruido de fondo. 33

4.3.2.Imagen problema con ruido aditivo gaussiano para diferentes valores de varianza y su gráfica del coeficiente de similaridad. (a) $\sigma^2 = 0$. (b) $\sigma^2 = 0.1$. (c) $\sigma^2 = 0.2$. (d) $\sigma^2 = 0.3$. (e) $\sigma^2 = 0.4$ (f) Gráfica del coeficiente de similaridad. 35

4.3.3.Imagen problema con ruido sal y pimienta para diferentes valores de densidad de píxeles y su gráfica del coeficiente de similaridad. (a) $d = 0$. (b) $d = 0.1$. (c) $d = 0.2$. (d) $d = 0.3$. (e) $d = 0.4$ (f) Gráfica del coeficiente de similaridad. 36

4.4.1.Ejemplos de imágenes en escala de grises utilizadas para medir el coeficiente de similaridad. (a) Imagen objetivo. (b) S_T: Firma de la imagen objetivo. (c) Imagen problema con ruido aditivo gaussiano de media $\mu = 0$ y varianza $\sigma^2 = 0.3$. (d) S_{TN}: Firma de la imagen problema. (e) Imagen del ruido de fondo. (f) S_{FN}: Firma de la imagen del ruido de fondo. 38

4.4.2.Imagen problema en escala de grises con ruido aditivo gaussiano para diferentes valores de varianza y su gráfica del coeficiente de similaridad. (a) $\sigma^2 = 0$. (b) $\sigma^2 = 0.1$. (c) $\sigma^2 = 0.2$. (d) $\sigma^2 = 0.3$. (e) $\sigma^2 = 0.4$ (f) Gráfica del coeficiente de similaridad. 39

4.4.3.Imagen problema en escala de grises con ruido sal y pimienta para diferentes valores de densidad de píxeles y su gráfica del coeficiente de similaridad. (a) $d = 0$. (b) $d = 0.1$. (c) $d = 0.2$. (d) $d = 0.3$. (e) $d = 0.4$ (f) Gráfica del coeficiente de similaridad. 40

1. INTRODUCCIÓN

La facilidad con la que reconocemos un rostro, entendemos las palabras que escuchamos, leemos caracteres escritos o identificamos las llaves del carro por simple tacto, esconden el proceso tan increíblemente complejo que hay detrás de estos actos de reconocimiento de patrones[8]. El reconocimiento de patrones en su forma más general, es la disciplina científica cuyo objetivo es la clasificación de objetos dentro de un número de categorías o clases. Dependiendo de la aplicación, estos objetos pueden ser imágenes, señales o cualquier tipo de medición que se desee clasificar[20].

Es una tendencia común en el ser humano crear tecnologías imitando a la naturaleza o al cuerpo humano con el fin de hacer la vida más sencilla, por lo que no es ninguna sorpresa que existan máquinas capaces de hacer reconocimiento de patrones de manera similar a como nosotros lo hacemos, y algunas veces incluso mejor. El reconocimiento de patrones en imágenes digitales es una de las áreas que más aportes ha generado en ese sentido, algunos ejemplos de estos son el reconocimiento de caracteres alfanuméricos, huellas dactilares, rostros, identificación de diferentes objetos en cintas transportadoras y detección de anomalías en análisis médicos (rayos X, tomografías).

Uno de los problemas centrales en el reconocimiento de patrones es determinar el grado en el cual la forma de un objeto difiere de otra. Operaciones como la correlación y comparación de plantillas son técnicas empleadas para determinar la diferencia

o similitud entre las formas. También se han utilizado métricas de distancia, como la euclidiana o la del máximo para lograr dicho objetivo[13]. En ese sentido, la distancia de Hausdorff podría ser una opción de métrica de discriminación entre objetos, ya que para dos conjuntos A y B mide que tan lejos se encuentran todos los puntos del conjunto A con cualquier punto del conjunto B y viceversa[14].

La tesis comprende cinco capítulos: en el primero se da una breve introducción general al reconocimiento de patrones en imágenes digitales. En el segundo capítulo se presenta la base teórica que sustenta este trabajo de investigación. En el tercer capítulo se detalla la construcción del sistema de reconocimiento de patrones de Radon-Fourier-Mellin. En el cuarto capítulo se explica la metodología utilizada para realizar la comparación de las métricas de desempeño para imágenes digitales con ruido. Finalmente en el quinto capítulo se presentan las conclusiones.

2. Fundamentos

2.1. Representación de imágenes digitales

Una imagen se define como una función bidimensional $f(x, y)$, donde x e y son coordenadas espaciales en el plano, y el valor de f para cualquier par de coordenadas se llama la intensidad de f en ese punto. Las imágenes pueden ser monocromáticas o a color. El término escala de grises es comúnmente utilizado para referirse a la intensidad de las imágenes monocromáticas[12]. Por otro lado, las imágenes a color están formadas por una combinación de imágenes monocromáticas individuales. Por ejemplo, en el sistema RGB una imagen a color está constituida por tres imágenes monocromáticas individuales, a las que nos referimos como las componentes roja (R), verde (G) y azul (B) de la imagen como se aprecia en la Fig. 2.1.1. Por esta razón, muchas de las técnicas

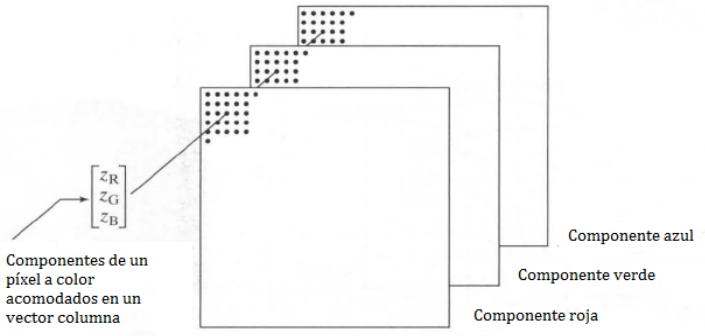


Figura 2.1.1: Representación esquemática de una imagen a color en el espacio de color RGB [12].

desarrolladas para procesamiento de imágenes monocromáticas se pueden extender a imágenes a color procesando cada componente indivualmente[12].

Las imágenes pueden ser discretas o continuas con respecto a sus coordenadas espaciales y con respecto a su amplitud. Una imagen digital es simplemente una imagen $f(x, y)$ que ha sido discretizada tanto en sus coordenadas espaciales como en su amplitud. Digitalizar los valores en las coordenadas es un proceso llamado muestreo, mientras que digitalizar los valores de la amplitud se llama cuantización[11]; en la Fig. 2.1.2 se ejemplifica como se genera una imagen digital; en la Fig 2.1.2(a) se muestra una imagen continua y un barrido sobre ésta de A a B; en la Fig 2.1.2(b) se observa una gráfica de

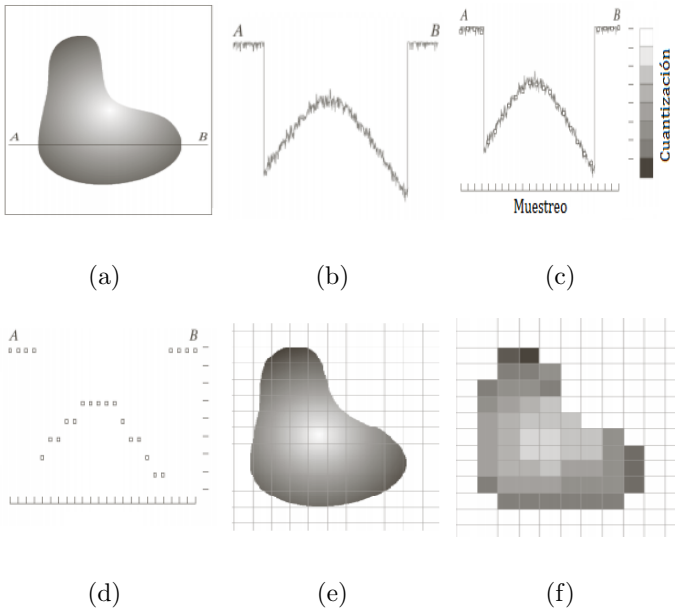


Figura 2.1.2: Ejemplificación de la generación de una imagen digital. (a) Imagen continua. (b) Barrido de A a B en la imagen continua. (c) Cuantización y muestreo. (d) Línea escaneada digitalizada. (e) Imagen continua proyectada sobre un arreglo. (f) Resultado de la cuantización y muestreo[11].

las intensidades del barrido de la imagen continua; para la Fig 2.1.2(c) se tienen dos ejes en la gráfica de intensidades que corresponden al muestreo y la cuantización; en el muestreo se divide espacialmente el barrido de intensidades en partes iguales, entre más grande sea el número de divisiones, mayor será la resolución de la imagen. Por otro lado la cuantización asigna un valor en la escala de grises a todo un intervalo de intensidades. En la Fig 2.1.2(d) se observa la gráfica del barrido discretizado y en la Fig 2.1.2(e) un arreglo bidimensional proyectado sobre la imagen continua, donde la cuadrícula corresponde al número de divisiones realizadas durante el muestreo. Por último se asigna el color correspondiente a la cuantización en cada cuadro con lo que se obtiene una imagen digital.

Una imagen digital se representa por un arreglo bidimensional o matriz de la siguiente manera

$$f(x,y) = \begin{bmatrix} f(1,1) & f(1,2) & ... & f(1,M) \\ f(2,1) & f(2,2) & ... & f(2,M) \\ \vdots & \vdots & \ddots & \vdots \\ f(N,1) & f(N,2) & ... & f(N,M) \end{bmatrix}, \tag{2.1.1}$$

donde cada elemento de la matriz es llamado píxel, derivado de "picture element". Por convención, la coordenada x varía desde 1 hasta M y la coordenada y de 1 hasta N en incrementos enteros[17].

2.2. Métricas

Sea X un conjunto, una función $d : X \times X \to \mathbb{R}$ es una función métrica o distancia en X si para cualesquiera x, y, z elementos de X, se satisfacen las propiedades:

1. no-negatividad: $d(x, y) \geq 0$,

2. identidad de los indiscernibles: $d(x, y) = 0$ si y sólo si $x = y$,

3. simetría: $d(x, y) = d(y, x)$,

4. desigualdad del triángulo: $d(x, y) \leq d(x, z) + d(z, y)$.

Se le conoce como espacio métrico (X, d) a un conjunto X con una métrica d definida entre todos los elementos del conjunto X[6]. Como se mencionó anteriormente, uno de los problemas centrales en el reconocimiento de patrones en imágenes digitales es determinar el grado en el cual la forma de un objeto difiere de la de otro. Una manera de comparar la disimilitud de la forma entre objetos es usando una distancia o métrica, debido a las ventajas que nos ofrece en su definición[14]. Esto es, sean A,B y C conjuntos que representen la forma de tres objetos y $d(A, B)$ una función que mida la disimilitud entre las formas de estos objetos. Si $d(A, B)$ es una métrica, entonces se cumple que:

1. $d(A, B) \geq 0$, ya que no tiene sentido hablar de una disimilitud negativa.

2. $d(A, B) = 0$ si y sólo si $A = B$, ya que esperamos que el objeto se parezca a si mismo.

3. $d(A, B) = d(B, A)$, nos indica que no importa el orden de comparación.

4. $d(A, B) \leq d(A, C) + d(B, C)$.

La desigualdad del triángulo es necesaria, ya que sin ella se podría presentar el caso en el que A es muy similar a B y B muy similar a C, pero que A y C no sean similares, lo cual no sucede en la vida real[1].

2.3. l_p-métricas

Sean $x, y \in \mathbb{R}^n$, la l_p-métrica es una norma métrica en $\mathbb{R}^n$, definida como

$$d_{lp}(x, y) = ||x - y||_p, \tag{2.3.1}$$

donde $1 \leq p \leq \infty$, con la l_p-norma dada por

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}. \tag{2.3.2}$$

En el caso particular en el que $p = 1$, se tiene

$$||x - y||_1 = \sum_{i=1}^{n} |x_i - y_i|, \tag{2.3.3}$$

conocida como geometría taxicab o distancia Manhattan, considerada por primera vez por Hermann Minkowski[6]. En la Fig. 2.3.1 se observan tres curvas arbitrarias en colores rojo, azul y amarillo que representan diferentes caminos entre el punto inicial y final. Estos tres caminos a pesar de ser todos diferentes, su distancia en la geometría taxicab es la misma debido a la forma de medir en esta métrica. El nombre alude al diseño de

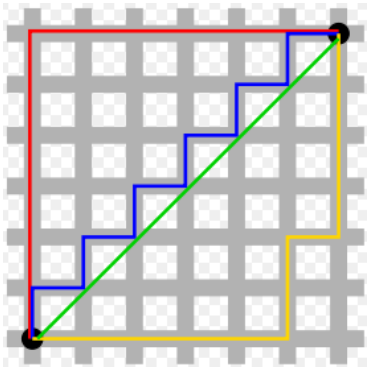


Figura 2.3.1: Comparación de la geometría taxicab con respecto a la distancia euclidiana[7].

cuadrícula de la ciudad de Manhattan, lo que causa que el camino más corto que pueda tomar un taxi entre dos puntos de la ciudad tenga la misma distancia. Podemos observar que los tres caminos mencionados anteriormente recorren 12 cuadras, mientras que en la distancia euclidiana ($p = 2$), dada por

$$||x - y||_2 = \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2}, \tag{2.3.4}$$

el único camino entre el punto inicial y final está representado por la curva verde. Ésta es la distancia ordinaria entre dos puntos en el espacio euclidiano y la que usamos en nuestro día con día. Otra métrica importante es la conocida como distancia de Chebyshev o métrica del máximo y se escribe como

$$D_{Chebyshev}(x, y) = max_{i \geq 1}(x_i - y_i). \tag{2.3.5}$$

A la distancia de Chebyshev también se le conoce como métrica infinito, denotada por $||x - y||_\infty$, y demostraremos a continuación que $\lim_{p \to \infty} ||x||_p = ||x||_\infty$, por lo que la métrica del máximo también es una l_p-métrica. En efecto, sea $x \in \mathbb{R}^n$ y definimos $||x||_\infty = max_{i \geq 1}(x_i)$, tenemos dos casos: para $x = \vec{0}$, se cumple que $||x||_\infty = 0$. Para el caso $x \neq \vec{0}$, $||x||_p$ se reescribe como

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} = ||x||_\infty \left(\sum_{i=1}^{n} \left(\frac{|x_i|}{||x||_\infty} \right)^p \right)^{1/p}. \tag{2.3.6}$$

Notemos que el término $\left(\frac{|x_i|}{||x||_\infty} \right) \leq 1$ para toda i de 1 hasta n, de ahí que

$$||x||_\infty \leq ||x||_p \leq ||x||_\infty n^{1/p}, \tag{2.3.7}$$

pero dado que $n > 0$, entonces $\lim_{p\to\infty} n^{1/p} = 1$, por lo tanto

$$\lim_{p\to\infty} ||x||_p = ||x||_\infty = max_{i\geq 1}(x_i). \tag{2.3.8}$$

y

$$||x - y||_\infty = max_{i\geq 1}(x_i - y_i) = D_{Chebyshev}(x, y). \tag{2.3.9}$$

Otro nombre que recibe esta métrica es el de distancia de ajedrez, debido a que en este juego el número mínimo de movimientos que se requieren para desplazar al rey desde un mismo punto de partida del tablero a otro, equivale a la distancia de Chebyshev o ajedrez entre la posición inicial y la posición final, como se aprecia en la Fig. 2.3.2.

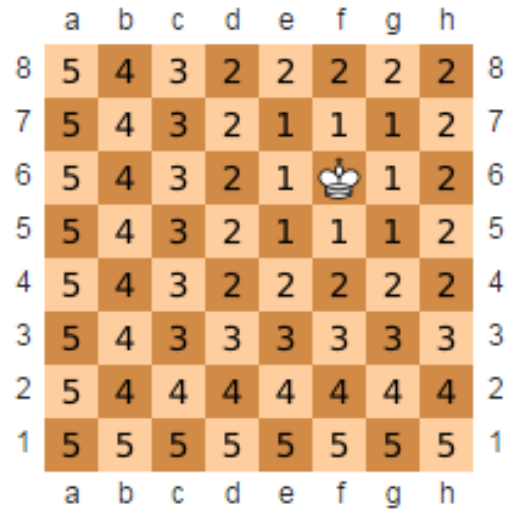


Figura 2.3.2: Ilustración de la métrica del máximo para un rey en un tablero de ajedrez[3].

2.4. Distancia de Hausdorff

Cuando hablamos acerca de distancias, normalmente nos referimos a la distancia más corta, por ejemplo si un punto X se dice que está a una distancia D de la forma de un objeto P, generalmente asumimos que D es la distancia entre X y el punto más cercano de P. La misma lógica aplica para dos formas A y B que están separadas una de otra, donde comúnmente asumimos que la distancia entre ellas es la distancia entre los dos puntos más cercanos de los conjuntos A y B. Formalmente, este tipo de funciones se llaman funciones mínimas, ya que la distancia D entre los dos conjuntos de puntos que representan la forma de A y B está dada por

$$D(A, B) = min_{a \in A} \left(min_{b \in B} ||a - b|| \right), \tag{2.4.1}$$

que se puede leer como sigue: "para cada punto a que pertenece a A, encuentra su distancia más corta a cualquier punto b que pertenece a B"[10]. Esta definición de distancia entre dos formas puede resultar muy insatisfactoria para cierto tipo de aplicaciones. Por ejemplo, para el caso de la Fig. 2.4.1 podríamos decir que las formas están muy cerca una de la otra si consideramos su distancia más cercana, representada por la distancia entre los puntos rojos. Sin embargo, naturalmente se esperaría que una distancia corta entre dos formas signifique que ningún punto entre estas formas se encuentre lejos de la otra, lo cual claramente se aprecia que no es así[10].

Una métrica que ha mostrado sus capacidades de discriminación entre formas y tiene diversas aplicaciones en el campo del reconocimiento de patrones es la distancia o métrica de Hausdorff[16,9], llamada así en honor a Felix Hausdorff, importante matemático alemán del siglo XIX que fue el primero en describirla. Dados dos conjuntos finitos de puntos $A = \{a_1, a_2, \ldots, a_p\}$ y $B = \{b_1, b_2, \ldots, b_q\}$ en un espacio métrico, la

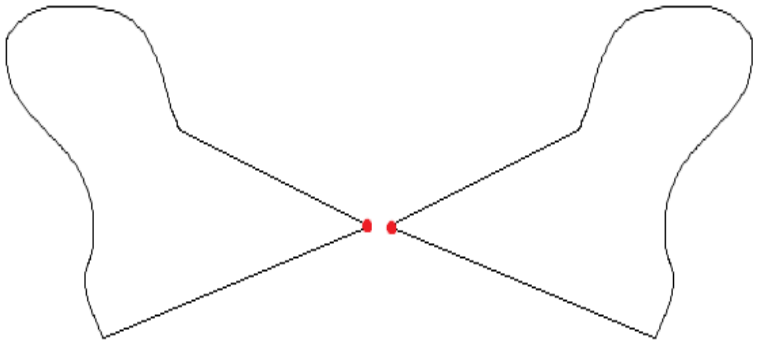


Figura 2.4.1: Distancia mínima entre dos formas irregulares ilustrada como la distancia entre los puntos rojos.

distancia de Hausdorff se define como

$$H(A,B) = max\left(h(A,B), h(B,A)\right), \tag{2.4.2}$$

donde

$$h(A,B) = max_{a\in A}\left(min_{b\in B}||a-b||\right), \tag{2.4.3}$$

con $||\cdot||$ representando una distancia conocida entre los puntos de A y B que es generalmente la distancia euclidiana. A la función $h(A,B)$ se le conoce como distancia de Hausdorff directa. Esta función identifica al punto $a \in A$ más alejado de cualquier punto de B y mide la distancia entre ese a y su vecino más cercano en B[14]. A las funciones $h(A,B)$ y $h(B,A)$ se les conoce como las componentes de la distancia de Hausdorff y están ilustradas en la Fig. 2.4.2, donde podemos observar dos formas diferentes con $h(X,Y)$ representando la componente directa y a $h(Y,X)$ se le conoce como la componente inversa, por lo que $H(X,Y)$ será igual a la más grande de las dos. Se debe notar que en general $h(A,B) \neq h(B,A)$, a este tipo de distancias se les conoce

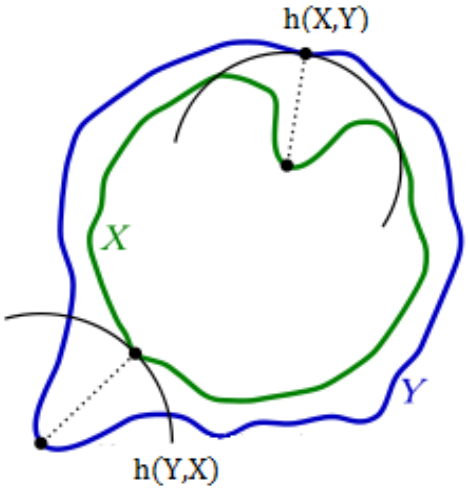


Figura 2.4.2: Componentes de la distancia de Hausdorf[4].

como distancias orientadas o asimétricas.

La distancia de Hausdorff mide el grado con el cual cada punto de un conjunto se encuentra cerca de cualquier punto de otro conjunto. Por esto, la distancia de Hausdorff se utiliza para determinar el grado de similitud entre dos objetos superpuestos uno sobre el otro. Es bien sabido que la distancia de Hausdorff es una métrica sobre el conjunto de todos los conjuntos cerrados y acotados[14], pero nos limitaremos a los conjuntos de puntos finitos, ya que son los que se utilizan en imágenes digitales.

2.5. Distancia de Hausdorff modificada

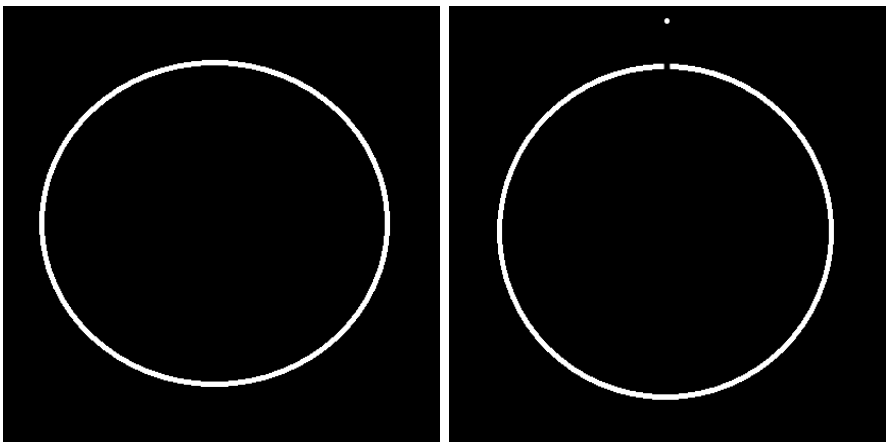
A pesar de lo útil que puede resultar la distancia de Hausdorff para la comparación de formas de dos objetos, ésta resulta ser muy susceptible a los bordes. M. Dubuisson y colaboradores[5] proponen una modificación a la distancia de Hausdorf para evitar esta clase problemas. La distancia de Hausdorff modificada, denotada por H_m está dada por

$$H_m(A, B) = max\left(h_m(A, B), h_m(B, A)\right), \tag{2.5.1}$$

donde

$$h_m(A, B) = \frac{1}{N_a}\sum_{a\in A} min_{b\in B}||a - b||, \tag{2.5.2}$$

de nuevo A y B son un par de conjuntos finitos de puntos de la forma $A = \{a_1, a_2, \ldots, a_p\}$ y $B = \{b_1, b_2, \ldots, b_q\}$, y $||\cdot||$ una distancia conocida que se acostumbra que sea la distancia euclidiana. A diferencia de la distancia de Hausdorff, la distancia de Hausdorff modificada promedia todas las distancias para cada $a \in A$ a su vecino más cercano en $b \in B$. En la Fig. 2.5.1 podemos observar un claro ejemplo de la sensibilidad ante bordes en la distancia de Hausdorff. Si superponemos las imágenes de la Fig. 2.5.1(a) y la Fig. 2.5.1(b), a simple vista podemos decir que son muy similares, incluso casi iguales, pero al calcular la distancia de Hausdorff entre ellas, ésta siempre sería igual a la distancia entre el punto trasladado de la Fig 2.5.1(b) y el punto directamente debajo de él en la Fig. 2.5.1(a), que llamaremos d_T, cosa que no concuerda con lo observado. En cambio si utilizamos la distancia de Hausdorff modificada, el punto trasladado sólo aportaría un factor de d_T/N a la distancia, mientras que el resto de las distancias mínimas serían igual a cero. De ésta manera podemos concluir que si el tamaño del conjunto es grande, la distancia de Hausdorff modificada es mucho menos sensible ante este tipo de ruido.



(a) (b)

Figura 2.5.1: Ilustración de la sensibilidad ante ruido de la distancia de Hausdorff. (a) Imagen de un círculo. (b) Imagen del mismo círculo con un punto trasladado.

3. Sistema RFM

El sistema de reconocimiento de patrones en imágenes digitales Radon-Fourier-Mellin (sistema RFM), es un método invariante a traslación, escala y rotación (TER) basado en tres transformadas integrales. Primeramente la transformada de Fourier, cuyo módulo es invariante a translación, después la transformada analítica de Fourier-Mellin, cuyo módulo normalizado por la componente directa nos proporciona invariancia a escala y por último la transformada de Radon que es invariante ante rotaciones.

3.1. Transformada de Fourier e invariancia a traslación

Llamada así en honor a Joseph Fourier, importante físico y matemático francés quien fue el primero en describirla, la trasformada de Fourier es una transformada integral empleada para transformar señales entre su dominio espacial (o temporal) y el dominio de las frecuencias. La transformada de Fourier en dos dimensiones denotada como $F(u, v)$ de una función $f(x, y)$ está dada por[15]

$$F(u,v) = \mathcal{F}\{f(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-i(ux+vy)}dxdy, \tag{3.1.1}$$

y su inversa

$$f(x,y) = \mathcal{F}^{-1}\{F(u,v)\} = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{i(ux+vy)}dudv. \tag{3.1.2}$$

Si aplicamos un desplazamiento en la Ec.(3.1.1) de la forma $f(x+x_0, y+y_0)$ obtenemos que

$$\mathcal{F}\{f(x+x_0, y+y_0)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x+x_0, y+y_0)e^{-i(ux+vy)}dxdy, \tag{3.1.3}$$

donde $x_0, y_0 \in \mathbb{R}$. Utilizando el cambio de variable $\xi = x + x_0, \eta = y + y_0$, la Ec.(3.1.3) se reescribe como.

$$\begin{aligned}
\mathcal{F}\{f(\xi,\eta)\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta)e^{-i[u(\xi-x_0)+v(\eta-y_0)]}d\xi d\eta \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta)e^{-i(u\xi+v\eta)}e^{i(ux_0+vy_0)}d\xi d\eta \\
&= e^{i(ux_0+vy_0)}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta)e^{-i(u\xi+v\eta)}d\xi d\eta \\
&= e^{i(ux_0+vy_0)}\mathcal{F}\{f(x,y)\},
\end{aligned} \tag{3.1.4}$$

y como

$$\left|e^{i(ux_0+vy_0)}\right| = \sqrt{e^{i(ux_0+vy_0)}e^{-i(ux_0+vy_0)}} = 1, \tag{3.1.5}$$

por lo tanto

$$\begin{aligned}
\left|\mathcal{F}\{f(x+x_0, y+y_0)\}\right| &= \left|e^{i(ux_0+vy_0)}\right|\left|\mathcal{F}\{f(x,y)\}\right| \\
&= \left|\mathcal{F}\{f(x,y)\}\right|,
\end{aligned} \tag{3.1.6}$$

con lo que concluimos que el módulo de la transformada de Fourier bidimensional es invariante ante traslaciones. Al módulo de la transformada de Fourier se le conoce como espectro de amplitud o espectro de Fourier. Utilizando el espectro de amplitud, se construye de manera sencilla un sistema de reconocimiento de patrones en imágenes digitales invariante a traslación, que es el primer paso en la construcción del sistema RFM.

En la Fig. 3.1.1 se muestran tres imágenes en blanco y negro de 257×257 píxeles y sus correspondientes espectros de amplitud. La Fig. 3.1.1(a) contiene a la letra B en

tipografía Arial sin ninguna transformación geométrica, a la que llamaremos I_1; la Fig. 3.1.1(b) tiene a la misma letra B pero ahora trasladada del origen y la denotamos por I_2; la Fig. 3.1.1(c) contiene de nuevo la misma letra B, pero esta vez rotada un ángulo de 36° en el sentido de las manecillas del reloj, y a la que nombramos I_3. Podemos notar que los espectros de amplitud F_1 (Fig. 3.1.1(d)) y F_2 (Fig. 3.1.1(e)) son prácticamente los mismos, salvo pequeñas variaciones debido a los errores de redondeo generados al momento de calcular la transformada rápida de Fourier y que son intrínsecos de las computadoras. Por otro lado, el espectro F_3 (Fig. 3.1.1(f)), es completamente diferente a los dos anteriores, debido a que el módulo de la transformada de Fourier no es invariante ante rotaciones, aunque se asemeja mucho a F_1 y F_2, pero rotados un ángulo de 36°.

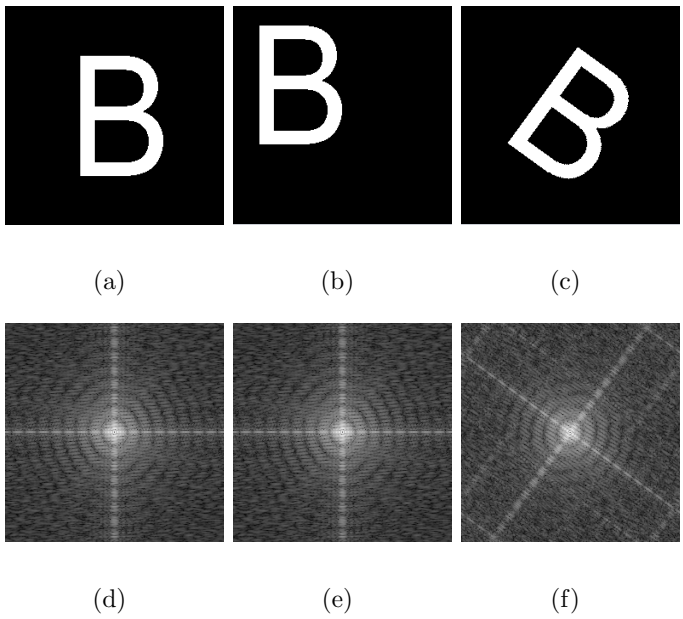


Figura 3.1.1: Ejemplos de espectros de amplitud para imágenes trasladadas y rotadas. (a) Imagen I_1. (b) Imagen I_2. (c) Imagen I_3 . (d) $F_1 = \left|\mathcal{F}\{I_1)\}\right|$. (e) $F_2 = \left|\mathcal{F}\{I_2\}\right|$. (f) $F_3 = \left|\mathcal{F}\{I_3\}\right|$.

3.2. Transformada analítica de Fourier-Mellin e invariancia a escala

La transformada de Mellin es una transformada integral presentada primeramente por Bernhard Riemman, importante matemático alemán del siglo XIX que la dio a conocer en sus memorias sobre números primos. Simultáneamente fue desarrollada por Eugene Cahen, de origen francés y Hjalmar Mellin de origen finlandés y de quien toma su nombre. Sea $f(r,\theta)$ una función en coordenadas polares, la transformada de Fourier-Mellin está dada por

$$\mathcal{M}_f(k,c) = \frac{1}{2\pi}\int_0^{\infty}\int_0^{2\pi} f(r,\theta)r^{-ic}e^{-ik\theta}d\theta\frac{dr}{r}, \tag{3.2.1}$$

donde $k, c \in \mathbb{R}$ [21]. La transformada de Fourier-Mellin de una función positiva f existe sólo si es absolutamente integrable, es decir, si se cumple que

$$\int_0^{\infty}\int_0^{2\pi}\left|f(r,\theta)r^{-ic}e^{-ik\theta}\right| d\theta\frac{dr}{r} = \int_0^{\infty}\int_0^{2\pi}\frac{1}{r}f(r,\theta)d\theta dr < \infty. \tag{3.2.2}$$

Debido al factor $\frac{1}{r}$ y considerando a r como la distancia entre el centro de una imagen digital y un píxel sobre la imagen, existen problemas de divergencia para calcular la transformada de Fourier-Mellin cerca del origen. Para solucionar este problema, se sugiere el cálculo de la transformada de Fourier-Mellin de funciones de la forma $f_\sigma(r,\theta) = r^\sigma f(r,\theta)$, donde $\sigma > 0$ es un número real fijo que nos asegura la convergencia de la transformada, esto es

$$\mathcal{M}_{f_\sigma}(k,c) = \frac{1}{2\pi}\int_0^{\infty}\int_0^{2\pi} f(r,\theta)r^{\sigma-ic-1}e^{-ik\theta}drd\theta, \tag{3.2.3}$$

y se le conoce como la transformada analítica de Fourier-Mellin (TAFM)[21].

Al realizar un escalamiento y una rotación de una función f de la forma $g(r,\theta) = f(\alpha r, \theta + \phi)$, con $\alpha \geq 0 \in \mathbb{R}$ y al obtener el módulo de la TAFM de la función g, pero normalizado por el valor de la TAFM en el centro de coordenadas, se obtiene que

$$\left|\frac{\mathcal{M}_{g_\sigma}(k,c)}{\mathcal{M}_{g_\sigma}(0,0)}\right| = \left|\frac{\mathcal{M}_{f_\sigma}(k,c)}{\mathcal{M}_{f_\sigma}(0,0)}\right|, \tag{3.2.4}$$

es decir, el módulo de la trasformada analítica de Fourier-Mellin normalizado por su centro de coordenadas es invariante a escalamiento. Utilizando ese módulo normalizado, podemos construir un sistema de reconocimiento de patrones invariante a traslación y escalamiento, que es el segundo paso en la construcción del sistema RFM.

En la Fig. 3.2.1 se muestran de nuevo tres imágenes en blanco y negro y del mismo tamaño que en la Fig 3.1.1. Las imágenes de la Fig.3.1.2(a) y Fig.3.1.2(c) son las mismas que la Fig. 3.1.1(a) y 3.1.1(c) y las llamaremos B_1 y B_3, respectivamente. La Fig.3.1.2(b) es una letra B en tipografía Arial que está trasladada con respecto al centro y escalada un 30 %, a la que llamaremos B_2. En esta ocasión se observa que a simple vista los espectros de amplitud de la Fig.3.1.2(d), Fig.3.1.2(e) y Fig.3.1.2(f), que corresponden a B_1, B_2 y B_3, respectivamente, son todos diferentes, lo cual tiene sentido, ya que el módulo de la transformada de Fourier no es invariante ni a rotaciones ni escalamiento; llamaremos a estos espectros G_1, G_2 y G_3 en el orden en el que aparecen. Por último a la Fig.3.1.2(g), Fig.3.1.2(h) y Fig.3.1.2(i), que son los módulos de las transformadas analíticas de Fourier-Mellin normalizadas de G_1, G_2 y G_3, los llamaremos M_1, M_2 y M_3. De estas últimas tres imágenes no se puede decir mucho a simple vista, pero salvo pequeñas variaciones debido a los errores de redondeo en cada una de las operaciones efectuadas, M_1 (Fig.3.1.2(g)) y M_2 (Fig.3.1.2(h)) son iguales.

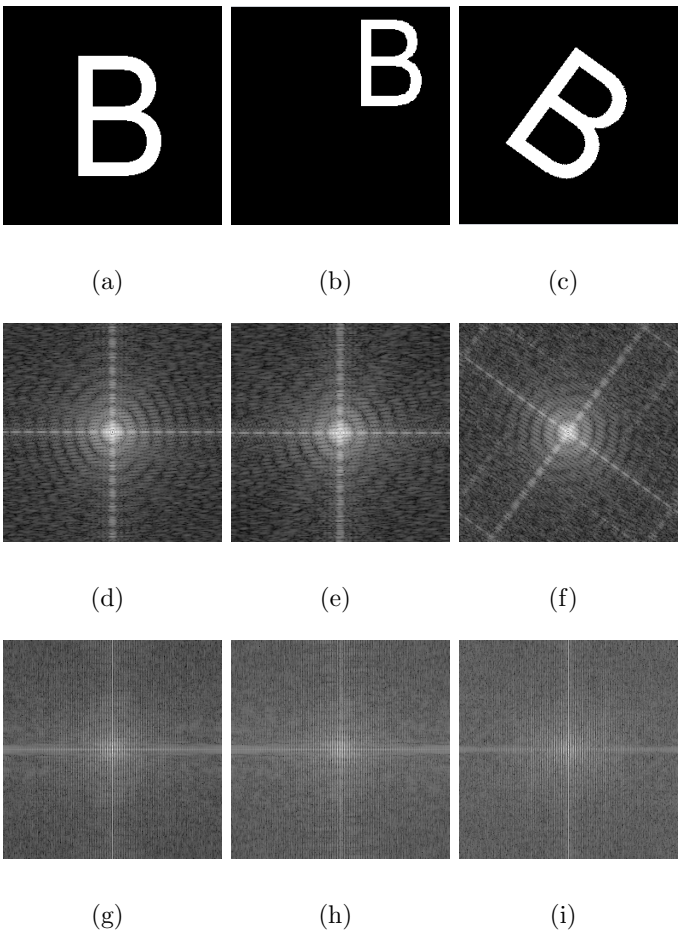


Figura 3.2.1: Ejemplos de la respuesta del sistema para imágenes trasladadas, escaladas y rotadas. (a) Imagen B_1. (b) Imagen B_2. (c) Imagen B_3. (d) $G_1 = \left|\mathcal{F}\{I_1\}\right|$. (e) $F_2 = \left|\mathcal{F}\{I_2\}\right|$. (f) $F_3 = \left|\mathcal{F}\{I_3\}\right|$. (g) M_1. (h) M_2. (i) M_3.

3.3. Transformada de Radon e invariancia a rotaciones

La transformada de Radon bidimensional, llamada así en honor al matemático austríaco Johann Radon, es una trasformada integral que consiste en la integral de una función sobre el conjunto de todas las rectas. Sea p una recta que pasa por el origen y es perpendicular a L, y sea θ el ángulo que forma la recta p con el eje horizontal (Fig. 3.3.1), la transformada de Radon bidimensional está dada por

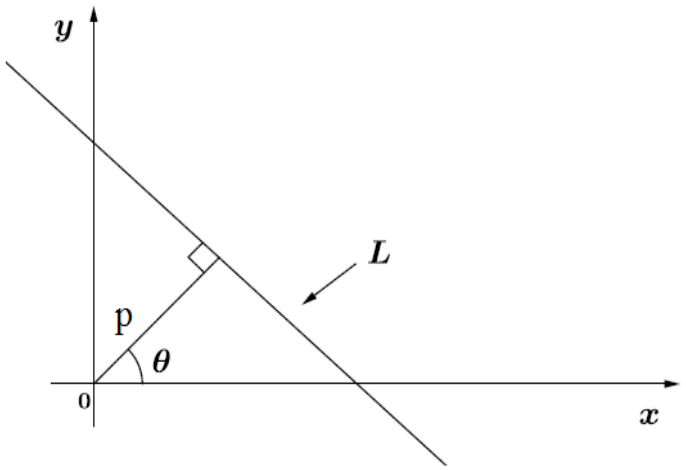


Figura 3.3.1: Representación de las recta L y p[20].

$$R(p,\theta) = \mathcal{R}\{f(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(p - xcos\theta - ysin\theta)dxdy, \tag{3.3.1}$$

donde $0 \leq \theta < 2\pi$. Para un ángulo θ fijo, se dice que la Ec.(3.3.1) es una muestra de la transformada de Radon. Dentro de las múltiples propiedades de la transformada de Radon, la que es de interés para este trabajo es la se presenta al aplicar rotaciones. Para comprender mejor dicha propiedad, la transformada de Radon se escribe en coordenadas polares como

$$R(p,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r,\phi)\delta\left(p - rcos(\theta - \phi)\right)|r|\, drd\phi. \tag{3.3.2}$$

Sea $f(r,\phi)$ una función en coordenadas polares y sea $h(r,\phi) = f(r,\phi+\phi_0)$ la función $f(r,\theta)$ rotada un ángulo θ_0, entonces

$$\begin{aligned}\mathcal{R}\{h(r,\phi)\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(r,\phi)\delta\left(p - rcos(\theta - \phi)\right)|r|\, drd\phi \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r,\phi+\phi_0)\delta\left(p - rcos(\theta - \phi)\right)|r|\, drd\phi.\end{aligned} \tag{3.3.3}$$

Realizando el cambio de variable $\phi' = \phi + \phi_0$ se llega a que

$$\begin{aligned}\mathcal{R}\{f(r,\phi+\phi_0)\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r,\phi')\delta\left(p - rcos(\theta-\phi'+\phi_0)\right)|r|\,drd\phi' \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r,\phi')\delta\left(p - rcos((\theta+\phi_0)-\phi')\right)|r|\,drd\phi' \\ &= R(p,\theta+\phi_0),\end{aligned} \tag{3.3.4}$$

por lo tanto

$$\mathcal{R}\{f(r,\phi+\phi_0)\} = R(p,\theta+\phi_0). \tag{3.3.5}$$

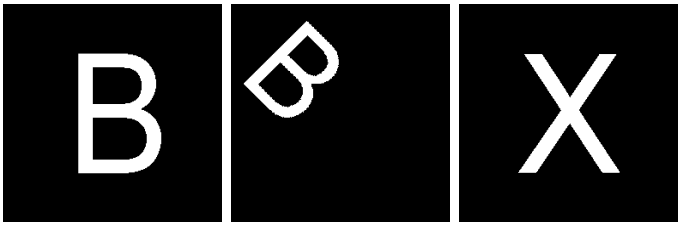
Esta expresión por sí misma no representa una invariancia a rotaciones, pero debido a la posición de los ejes en la transformada de Radon, significa simplemente un desplazamiento horizontal, que a fin de cuentas no afectará al descriptor final debido a la forma de generar las firmas, esto es

$$S(p) = \sum_{\theta=0}^{179} R(p,\theta), \tag{3.3.6}$$

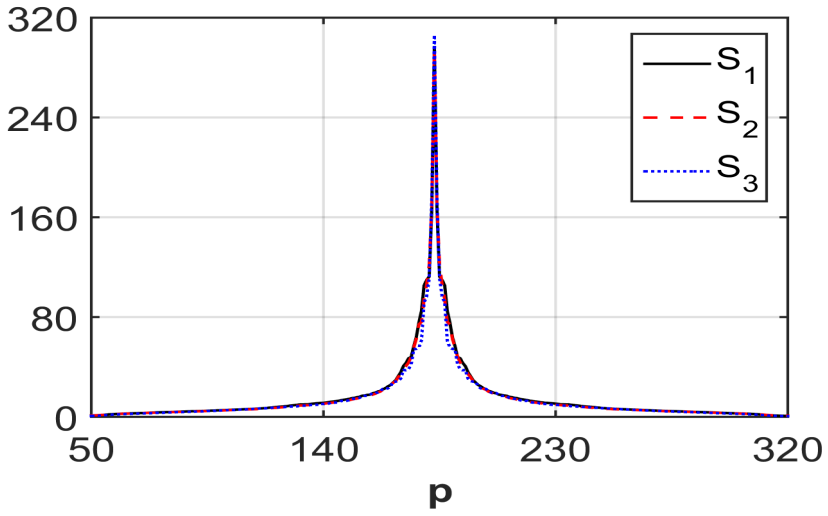
es decir, $S(p)$ es la firma 1D construida a partir de los marginales por renglón de la imagen resultante de la transformada de Radon.

En la Fig 3.3.2 se muestran una vez más, tres imágenes en blanco y negro, del mismo tamaño. La Fig. 3.3.2(a) contiene a la letra B en tipografía Arial sin transformaciones geométricas, a la que llamaremos C_1; la Fig. 3.3.2(b) tiene a la letra B, pero ahora trasladada, rotada 45° en el sentido de las manecillas del reloj y escalada un 25 %, a la que denotamos como C_2; por último tenemos a la Fig. 3.3.2(c) con una letra X en tipografía Arial sin transformaciones geométricas, a la que se nombra C_3. Las tres imágenes se introdujeron al sistema RFM y se generaron sus firmas 1D denotadas por

S_1 para C_1, S_2 para C_2 y S_3 para C_3, las cuales se muestran en la Fig. 3.3.2(d) y donde se aprecia que S_1 y S_2 son iguales y que S_3 difiere de ellas. El hecho de que las curvas S_1 y S_2 no sean completamente iguales se debe principalmente al ruido diente de sierra generado al rotar computacionalmente las imágenes y a los errores de redondeo intrínsecos en las operaciones computacionales.



(a) (b) (c)



(d)

Figura 3.3.2: Ejemplos de la respuesta del sistema RFMs. (a) Imagen C_1. (b) Imagen C_2. (c) Imagen C_3 . (d) Gráfica de comparación de las tres firmas 1D con $0 < p < 366$, pero para mejor visualización de la información relevante, se graficó de $50 < p < 320$.

4. Métricas de desempeño

Las firmas 1D generadas en el Capítulo 3 cumplen con la función de ser descriptores de la imagen de la que proceden. Un descriptor de una imagen debe contener suficiente información como para representar a esa imagen, independientemente de las transformaciones o distorsiones geométricas que presenten los objetos. Sean I_1 e I_2 un par de imágenes representadas por sus descriptores S_1 y S_2 respectivamente, obtenidos mediante el sistema RFM, T.V. Hoang y colaboradores[13] proponen medir la similitud entre las imágenes con la l_2-métrica entre sus descriptores, esto es

$$sim_1(S_1, S_2) = \|S_1 - S_2\|_2 . \tag{4.0.1}$$

El cálculo de $sim_1(S_1, S_2)$ es simple y rápido, lo que nos permite trabajar con problemas que involucran bases de datos grandes. Es posible conservar las ventajas de esta medida de similitud utilizando otras métrica más sofisticadas. En este trabajo se propone modificar la Ec.(4.0.1) utilizando la distancia de Hausdorff modificada, dada por

$$sim_2(S_1, S_2) = H_m(S_1, S_2), \tag{4.0.2}$$

donde H_m está definida en la Ec.(2.5.1). Usando la medida de similaridad de la Ec.(4.0.2) con un umbral tal que $sim_2(S_1, S_2) < \epsilon$, donde ϵ depende del problema a resolver, po-

demos realizar el proceso de clasificación de una nueva manera. Debido a la definición de la distancia modificada de Hausdorff, se pretende reducir la dominancia de algunos términos de las firmas 1D, ya que se toman todos los mínimos entre las firmas y se dividen entre el número de estos, además de ser sencilla y rápida de implementar en 1D.

En la Fig. 4.0.1 se muestran tres imágenes en blanco y negro de tamaño 257×257 píxeles y las firmas correspondientes. La Fig. 4.0.1(a) exhibe la imagen I_1 que contiene una letra D en tipografía Arial sin transformaciones geométricas y su firma 1D, a la que llamaremos S_1, está dada en la Fig. 4.0.1(b). La Fig. 4.0.1(c) es la imagen I_2 con una letra D en tipografía Arial rotada un ángulo de 30°, escalada -25 % y trasladada del centro de la imagen. La Fig. 4.0.1(d) muestra la firma 1D de la imagen I_2 y a la que denotaremos S_2. La Fig. 4.0.1(e) es la imagen I_3 con una letra E en tipografía Arial sin transformaciones geométricas, cuya firma 1D nombraremos S_3 y se muestra en la Fig. 4.0.1(f).

Al comparar las firmas usando la métrica de la Ec.(4.0.2) se obtuvo que $sim_2(S_1, S_1) = 0$, lo que tiene sentido ya que comparamos una firma consigo misma; para $sim_2(S_1, S_2)$ se obtuvo un valor de 0.6627 y para $sim_2(S_1, S_3)$ de 2.2843. Debido a errores de redondeo al rotar y escalar la imagen I_1, se tiene que $sim_2(S_1, S_2) \neq 0$, pero aun el valor es pequeño comparado con la diferencia entre $sim_2(S_1, S_3)$ y $sim_2(S_1, S_2)$ que es más de 3 veces mayor, por lo que podemos fácilmente entrenar al sistema con diferentes rotaciones y escalamientos de I_1 y definir el valor de ϵ para este problema en específico. Para el caso en el que se utiliza la Ec.(4.0.1) los valores obtenidos para las mismas imágenes fueron: $sim_1(S_1, S_1) = 0$; $sim_1(S_1, S_2) = 28.8848$ y $sim_1(S_1, S_3) = 162.0101$.

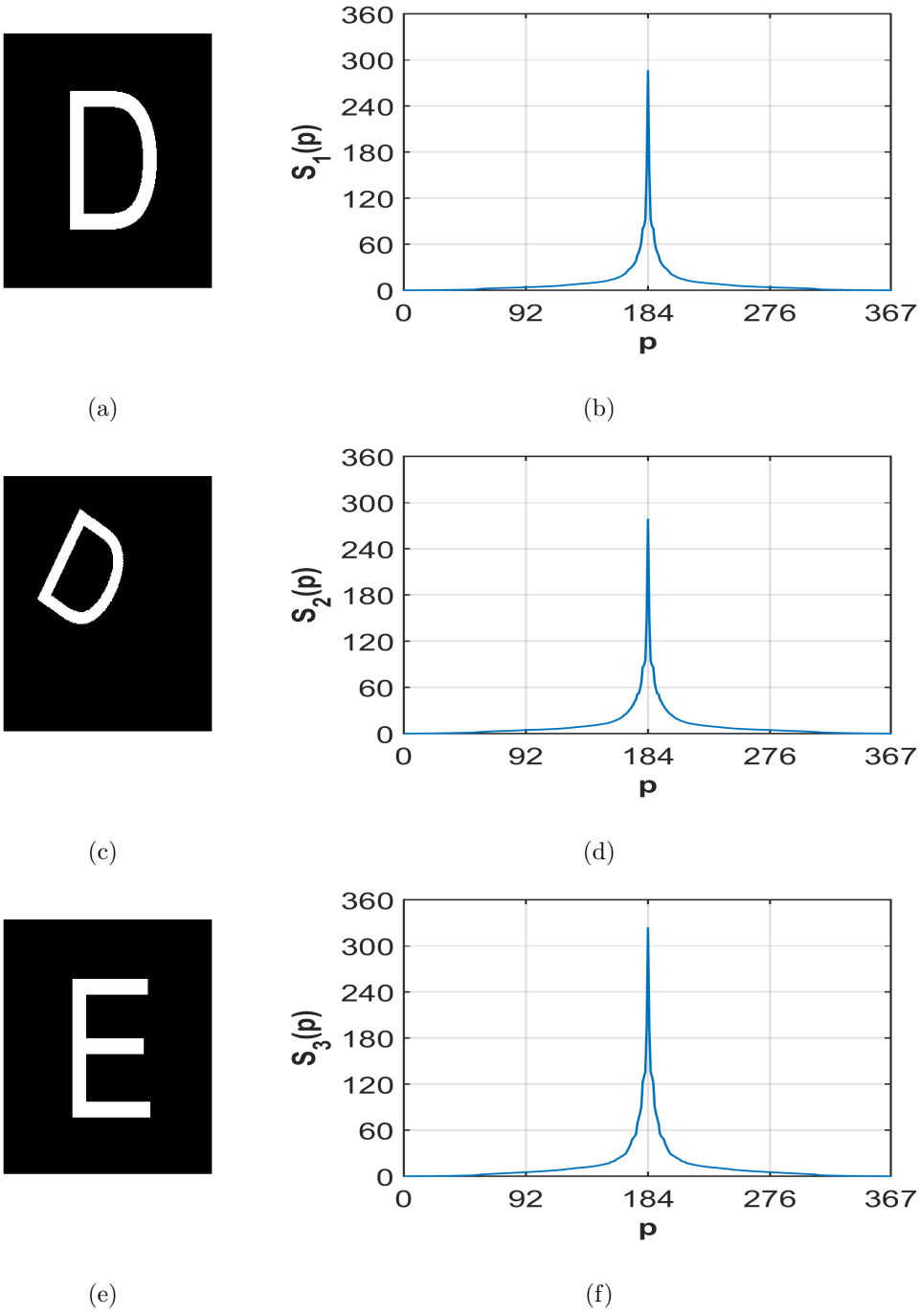


Figura 4.0.1: Ejemplos de imágenes utilizadas para su comparación y sus firmas. (a) Imagen I_1. (b) S_1, firma 1D de I_1. (c) Imagen I_2. (d) S_2, firma 1D de I_2. (e) Imagen I_3. (f) S_3, firma 1D de I_3

Podemos notar que los valores para $sim_1(S_1, S_2)$ y $sim_1(S_1, S_3)$ son bastante grandes comparados con los de $sim_2(S_1, S_2)$ y $sim_2(S_1, S_3)$, pero aún así, $sim_1(S_1, S_3)$ es más de 5 veces mayor que $sim_1(S_1, S_2)$, por lo que su capacidad de discriminación también nos permite definir ϵ con facilidad. Existe una gran variedad de métodos de comparación de descriptores: funciones de correlación, métodos estadísticos, métricas, entre otros, el presente trabajo se enfoca en el estudio del desempeño de distintas métricas para comparar firmas 1D de imágenes digitales con diferentes tipos de ruido. El ruido provoca variaciones aleatorias en las intensidades de una imagen, lo que se refleja también en su descriptor, por lo que es importante saber de qué manera se ven afectadas las imágenes por el ruido y cuál es la mejor manera de comparación bajo tales efectos.

4.1. Ruido

El ruido es una variación aleatoria en las intensidades de las imágenes digitales. El ruido está siempre presente en las imágenes digitales, ya sea generado durante la adquisición, el código, la transmisión y el procesamiento. Se puede considerar al ruido como la información irrelevante de una imagen digital que produce efectos indeseables como bordes poco realistas, líneas invisibles, esquinas, objetos borrosos o perturbación de escenas de fondo[2]. Por ejemplo, el ruido aditivo se modela mediante

$$I_N(x, y) = I(x, y) + N(x, y), \tag{4.1.1}$$

donde I_N es la imagen con ruido, I es la imagen sin ruido y N es el modelo de algún tipo de ruido. A continuación se describen los dos tipos de ruido que serán utilizados en este trabajo de investigación, el ruido aditivo gaussiano y, el ruido sal y pimienta.

El ruido aditivo gaussiano es un tipo de ruido estadístico que perturba los valores

de la escala de grises de una imagen digital en base a la distribución de densidad de probabilidades normal o gaussiana[2]. También es conocido como ruido electrónico, debido a que se genera en amplificadores y en detectores; aunque también puede ser generado por causas naturales como la vibración térmica de los átomos. Para una imagen digital en escala de grises, la distribución del ruido gaussiano P está dada por

$$P(g) = \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\frac{(g-\mu)^2}{2\sigma^2}}, \tag{4.1.2}$$

donde g es la intensidad de la escala de grises de un píxel, σ es la desviación estándar y μ es la media[2]. Generalmente el modelo de ruido gaussiano representa una aproximación acertada a escenarios del mundo real. En la Fig. 4.1.1 se puede apreciar un ejemplo de una distribución de ruido gaussiano con media $\mu = 0$ y desviación estándar $\sigma = 0.2$, donde se puede ver claramente la conocida "campana de Gauss" en la cual aproximadamente el 70 % de los valores de píxeles ruidosos de la imagen degradada

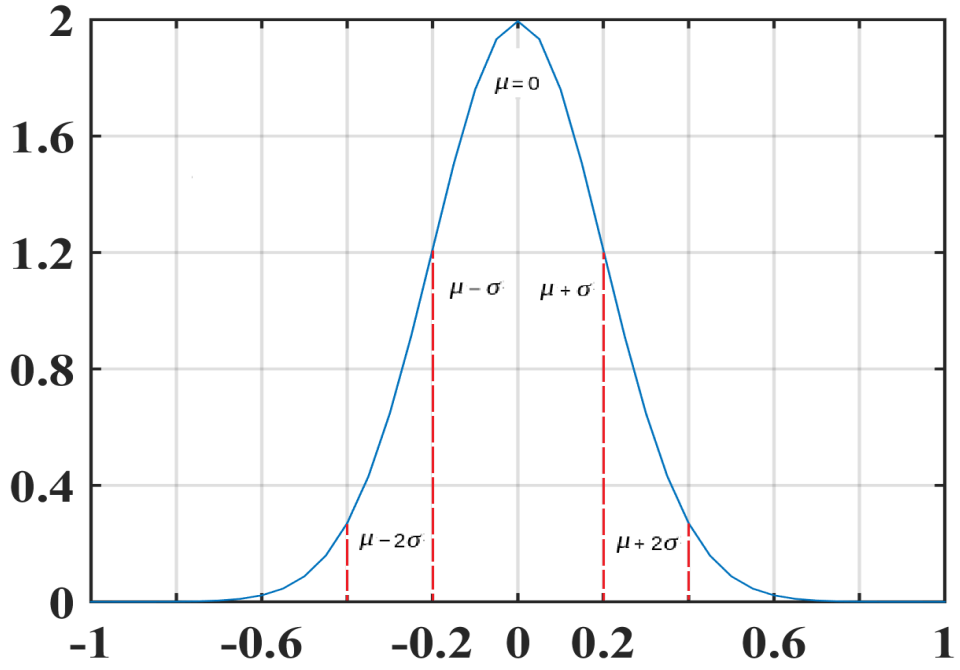


Figura 4.1.1: Ejemplo de la distribución de ruido gaussiano con $\mu = 0$ y $\sigma = 0.2$.

(a) (b)

Figura 4.1.2: Ejemplo de una imagen con ruido aditivo gaussiano. (a) Imagen sin ruido. (b) Imagen con ruido aditivo gaussiano de parámetros $\mu = 0$ y $\sigma^2 = 0.01$.

están contenidos entre $\mu - \sigma$ y $\mu + \sigma$, y alrededor del 95 % de los valores entre $\mu - 2\sigma$ y $\mu + 2\sigma$. En la Fig. 4.1.2 se aprecia visualmente como se ve una imagen antes y después de añadirle ruido aditivo gaussiano.

El otro tipo de ruido utilizado en este trabajo es el conocido como ruido sal y pimienta o ruido impulsional. Este tipo de ruido se observa principalmente en la transmisión de datos y en el mal funcionamiento de sensores de cámaras. En este caso, los valores de los píxeles de la imagen ruidosa son sustituidos por valores mínimos o máximos del píxel, por ejemplo: 0 ó 255 para imágenes del tipo *uint8* y 0 ó 1 para imágenes del tipo doble, lo que se puede interpretar como prender o apagar píxeles aleatorios y que en efecto se asemeja a derramar sal y pimienta sobre una imagen[2]. La función de densidad de probabilidad del ruido sal y pimienta está dada por

$$P(g) = \begin{cases} P(a), & \text{si } g = a, \\ P(b), & \text{si } g = b, \\ 0, & \text{de otro modo,} \end{cases} \tag{4.1.3}$$

donde a es el valor de intensidad mínima con probabilidad $P(a)$ y b es el valor de intensidad máxima con probabilidad $P(b)$. En la Fig. 4.1.3 se muestra una imagen antes y después de añadir ruido sal y pimienta. El valor de d con $0 < d < 1$ es una variable del lenguaje computacional MATLAB llamada densidad de píxeles, que representa aproximadamente el porcentaje de píxeles afectados en la imagen y que correspondería a $P(a) + P(b)$ en la ecuación Ec.(4.1.3).

(a) (b)

Figura 4.1.3: Visualización de una imagen con ruido sal y pimienta. (a) Imagen sin ruido. (b) Imagen con ruido sal y pimienta de parámetro $d = 0.05$.

4.2. Coeficiente de similaridad de Hausdorff

Para poder realizar la comparación de las firmas 1D provenientes de imágenes con ruido, se requiere de un método alternativo de comparación a los métodos descritos en la introducción de este capítulo. Con este fin se definió un coeficiente de similaridad con las

siguientes características: sea S_T la firma de una imagen sin ruido, a la que llamaremos imagen objetivo; S_{TN} la firma de una imagen con ruido, a la que nombraremos imagen con ruido; S_{FN} la firma de la imagen del fondo a la que se le ha añadido ruido y a la cual denotaremos como imagen de ruido de fondo, $\|\cdot\|$ una métrica conocida y CS el coeficiente de similaridad. Necesitamos que el coeficiente de similaridad cumpla que $CS = 1$ cuando se tenga máxima similaridad, o sea el caso en el que la imagen problema no presente ruido y sea igual a la imagen de referencia, y que $CS = 0$ para la mínima similaridad o máximo ruido añadido. Al definir el coeficiente de similaridad de la forma[19]:

$$CS = 1 - \frac{\|S_T - S_{TN}\|}{\|S_T - S_{FN}\|}, \tag{4.2.1}$$

cuando se le añade poco ruido a la imagen se tiene que $S_T \approx S_{TN}$, por lo que $S_T - S_{TN} \approx 0$ y $CS \approx 1$, la igualdad se da cuando el ruido añadido es nulo, que es el caso de máxima similaridad o la igualdad de la imagen de referencia con la imagen problema sin ruido. Al añadir más ruido a la imagen, se observa que ésta se parece cada vez más al ruido de fondo, o dicho de otro modo $S_{TN} \approx S_{FN}$, por lo que $\frac{\|S_T - S_{TN}\|}{\|S_T - S_{FN}\|} \approx 1$ y $CS \approx 0$, observando que la igualdad sólo se daría en caso de poder añadir una cantidad infinita de ruido.

Las métricas que utilizaremos en la Ec.(4.2.1) para realizar la comparación del coeficiente de similaridad son l_2 o distancia euclidiana, dada por la Ec.(2.3.4); l_∞ o distancia del máximo, dada por la Ec.(2.3.9) y la distancia o métrica de Hausdorff modificada, dada por la Ec.(2.5.1). La justificación para elegir estas tres métricas se debe a varios factores, en el campo del reconocimiento de patrones, las l_p-métricas son una clase de métricas de similaridad de imágenes ampliamente utilizadas, cuyos miembros más utilizados son las métricas l_1 y l_2[13,18], por lo que seleccionamos a la distancia euclidiana

para compararla con las otras distancias propuestas en este trabajo. La distancia del máximo o l_∞ es otra métrica perteneciente a la familia de las l_p-métricas, que pese a ser diferente al resto de ellas, en cuanto a que no existe una suma con respecto a la diferencia entre píxeles correspondientes, nos proporciona una noción de que tan alejada está una firma de la otra. La distancia de Hausdorff modificada fue seleccionada debido a su ya conocido poder en la discriminación de formas[9,14,16] y al ser menos sensible a bordes que la distancia de Hausdorff original, además de que no ha sido utilizada anteriormente en la comparación de firmas 1D.

4.3. Comparación del coeficiente de similaridad para imágenes en blanco y negro

En la Fig. 4.3.1 se despliegan tres imágenes en blanco y negro de tamaño 257×257 píxeles y sus respectivas firmas. Éstas son ejemplos del tipo de imagen que utilizaremos para realizar el análisis de la efectividad del coeficiente de similaridad. La Fig. 4.3.1(a) es una letra A en tipografía Arial sin transformaciones geométricas ni ruido añadido; la Fig. 4.3.1(b) muestra su firma 1D invariante a traslación, escala y rotación generada con el sistema RFM, a la que llamamos S_T. La Fig. 4.3.1(c) es la misma letra A de la Fig. 4.3.1(a) pero esta vez se le añadió ruido aditivo gaussiano con $\mu = 0$ y $\sigma^2 = 0.3$; la Fig. 4.3.1(d) exhibe la correspondiente firma 1D, a la que denotaremos S_{TN}. La Fig. 4.3.1(e) es la imagen solamente del fondo de la Fig. 4.3.1(c) a la que se le ha añadido ruido; la Fig. 4.3.1(f) presenta la correspondiente firma 1D, a la que nombraremos S_{FN}. A simple vista podemos apreciar que existe gran diferencia en la forma de las curvas de las tres firmas, por lo que usaremos el coeficiente de similaridad para obtener un valor numérico de qué tan similares o distintas son en función del ruido añadido.

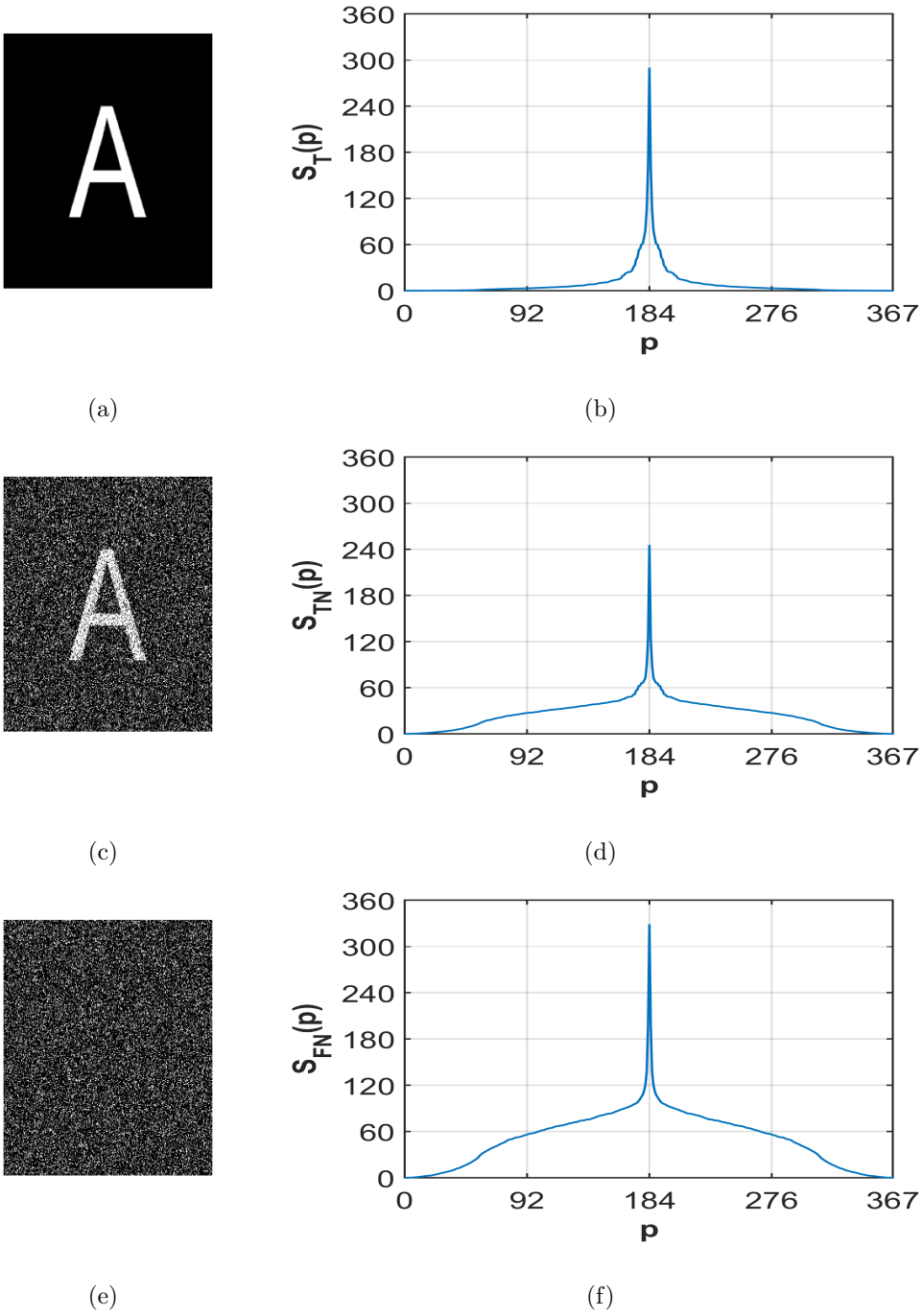
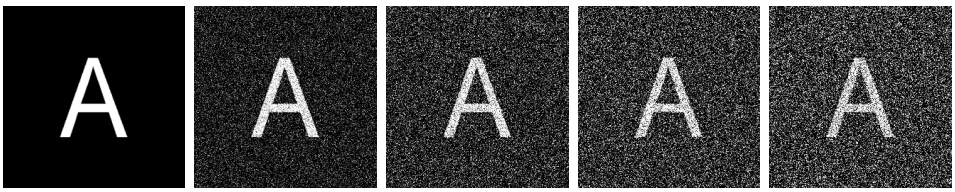


Figura 4.3.1: Ejemplos de imágenes en blanco y negro utilizadas para medir el coeficiente de similaridad. (a) Imagen objetivo. (b) S_T: Firma de la imagen objetivo. (c) Imagen problema con ruido aditivo gaussiano de media $\mu = 0$ y varianza $\sigma^2 = 0.3$. (d) S_{TN}: Firma de la imagen problema. (e) Imagen del ruido de fondo. (f) S_{FN}: Firma de la imagen del ruido de fondo.

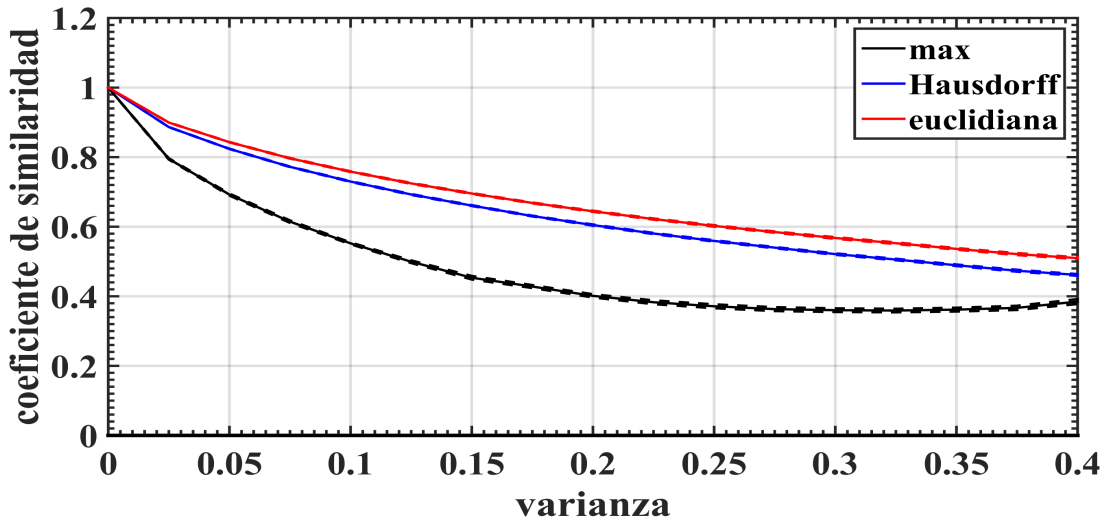
Debido a la aleatoriedad del ruido, se utilizaron 40 imágenes por muestra para cada valor de la varianza, de esta forma se obtiene un valor promedio que representará de una manera confiable el análisis que se está realizando. En el caso del ruido aditivo gaussiano se tiene media $\mu = 0$ y varianza $0 < \sigma^2 < 0.4$, en incrementos de $\sigma^2 = 0.025$. Para el ruido sal y pimienta los valores de la densidad de píxeles fueron de $0 < d < 0.4$, en incrementos de $d = 0.025$.

En la Fig. 4.3.2(f) se observa que las tres curvas comienzan en $CS = 1$, que es cuando $\sigma^2 = 0$, lo que era de esperarse ya que la imagen problema no presenta ruido y $\|S_T - S_{TN}\| = 0$. Después se tiene que las tres curvas decaen lentamente, lo que de nuevo tiene sentido, ya que entre más ruido se le añada a la imagen problema la diferencia $\|S_T - S_{TN}\|$ aumenta, mientras que $\|S_T - S_{FN}\|$ decrece, por lo que el cociente tiene cada vez más peso sobre CS. De la Fig. 4.3.2(a) a la Fig. 4.3.2(e) se exhibe el efecto del ruido aditivo gaussiano en la imagen problema, para todas la media $\mu = 0$ y los valores de la varianza σ^2 van de 0 a 0.4, con tamaño de paso de 0.1. Para la Fig. 4.3.2(c) en la cual $\sigma^2 = 0.2$, le corresponde un valor de $CS \approx 0.6$ bajo la distancia de Hausdorff modificada. Si $CS \approx 0.6$, se tiene que $\frac{\|S_T - S_{TN}\|}{\|S_T - S_{FN}\|} \approx 0.4$, o dicho en palabras, el ruido de fondo pesa 40 % con respecto de la imagen con ruido, por lo que aún se distingue un 60 % de la imagen problema en la imagen con ruido utilizando la distancia de Hausdorff modificada. Entonces, entre mayor es el coeficiente de similaridad significa que la imagen problema con ruido se parece a la imagen de referencia y entre menor sea el coeficiente de similaridad se deduce que la imagen problema con ruido se parece más a la imagen del fondo con ruido.

Para el intervalo $0.2 < \sigma^2 < 0.4$ se observa un comportamiento un poco anti-intuitivo en la curva en color negro, correspondiente al coeficiente de similaridad con la métrica



(a) (b) (c) (d) (e)

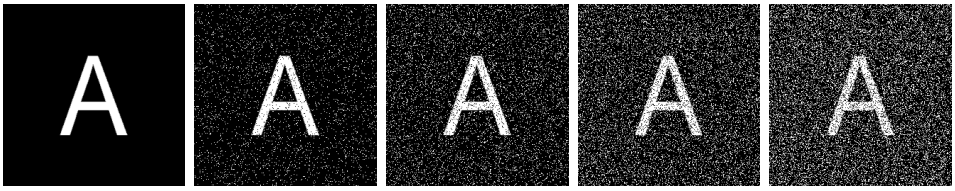


(f)

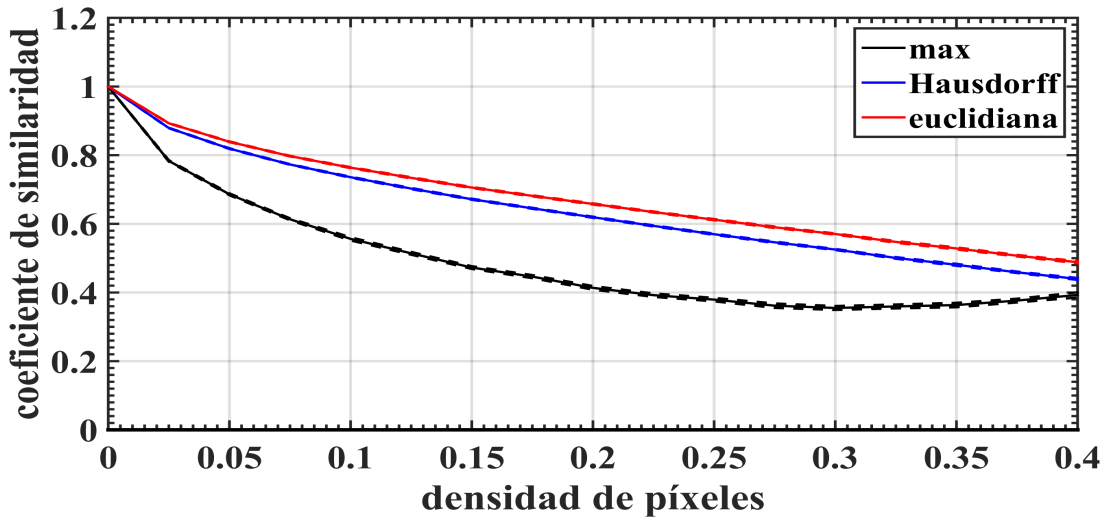
Figura 4.3.2: Imagen problema con ruido aditivo gaussiano para diferentes valores de varianza y su gráfica del coeficiente de similaridad. (a) $\sigma^2 = 0$. (b) $\sigma^2 = 0.1$. (c) $\sigma^2 = 0.2$. (d) $\sigma^2 = 0.3$. (e) $\sigma^2 = 0.4$ (f) Gráfica del coeficiente de similaridad.

del máximo. Esta curva presenta ligeros aumentos en CS para ciertos intervalos, aun cuando se incrementa el ruido, esto se debe a la definición de la distancia en términos de una función máxima y de la aleatoriedad del ruido, tiene sentido pensar que en algunos casos, un poco de aumento de ruido modifique la firma de tal forma que CS crezca pero que de manera global, la tendencia decreciente se mantenga. En este caso se tiene que para el ruido aditivo gaussiano con varianza menor a $\sigma^2 = 0.4$, dominan los coeficientes de similaridad generados por la distancia euclidiana y la distancia de Hausdorff modificada, teniendo ambas un mejor desempeño que la métrica del máximo.

Aunque ambas muestran un buen desempeño muy similar, se recomienda el uso de la distancia euclidiana debido a su ligeramente mejor desempeño, su rapidez y facilidad de implementación. Para el caso del ruido sal y pimienta, en la Fig. 4.3.3 se observa un comportamiento similar al del ruido aditivo gaussiano. Por lo tanto podemos concluir que para ambos tipos de ruido se obtuvo la tendencia que se esperaba para el coeficiente de similaridad y que tanto la distancia euclidiana como la de Hausdorff modificada tienen un mejor desempeño que la distancia del máximo.



(a) (b) (c) (d) (e)

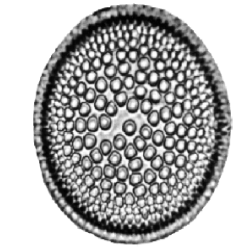


(f)

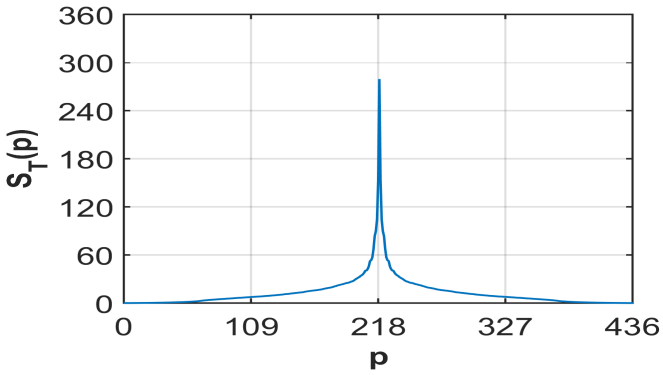
Figura 4.3.3: Imagen problema con ruido sal y pimienta para diferentes valores de densidad de píxeles y su gráfica del coeficiente de similaridad. (a) $d = 0$. (b) $d = 0.1$. (c) $d = 0.2$. (d) $d = 0.3$. (e) $d = 0.4$ (f) Gráfica del coeficiente de similaridad.

4.4. Comparación del coeficiente de similaridad para imágenes en escala de grises

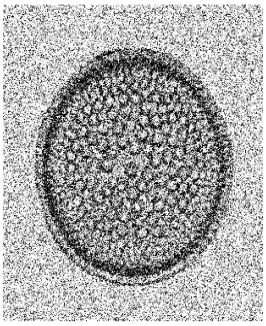
En la Fig. 4.4.1 se observan tres imágenes en escala de grises de tamaño 307×307 píxeles y sus respectivas firmas. Éstas son ejemplos del tipo de imagen que utilizaremos para realizar el análisis de la efectividad del coeficiente de similaridad para imágenes en escala de grises. La Fig. 4.4.1(a) es la imagen de una diatomea sin transformaciones geométricas ni ruido; la Fig. 4.4.1(b) muestra su firma 1D invariante a traslación, escala y rotación generada con el sistema RFM, a la que llamamos S_T. La Fig. 4.4.1(c) es la misma diatomea de la Fig. 4.4.1(a), pero esta vez se le añadió ruido aditivo gaussiano con $\mu = 0$ y $\sigma^2 = 0.3$; la Fig. 4.4.1(d) exhibe la correspondiente firma 1D, a la que denotaremos S_{TN}. La Fig. 4.4.1(e) es la imagen solamente del fondo de la Fig. 4.4.1(c) a la que se le ha añadido ruido; la Fig. 4.2.1(f) presenta la firma 1D, a la que nombraremos S_{FN}. Una vez más, a simple vista se observa una diferencia notoria entre las tres firmas, por lo que se llevará a cabo un análisis del coeficiente de similaridad ahora para imágenes en escala de grises sobre fondos claros. La metodología utilizada es la misma de la sección 4.3. En la Fig. 4.4.2 de nuevo se observa que las tres curvas comienzan en $CS = 1$ para $\sigma^2 = 0$, que es lo esperado debido a la ausencia de ruido en la imagen problema. La tendencia decreciente sigue presente, pero para este caso se mantiene durante todo el intervalo, contrario a lo que sucedía en el caso de imágenes en blanco y negro, donde se observan incrementos repentinos en la curva del coeficiente de similaridad para la distancia del máximo, a pesar de aumentar la varianza.



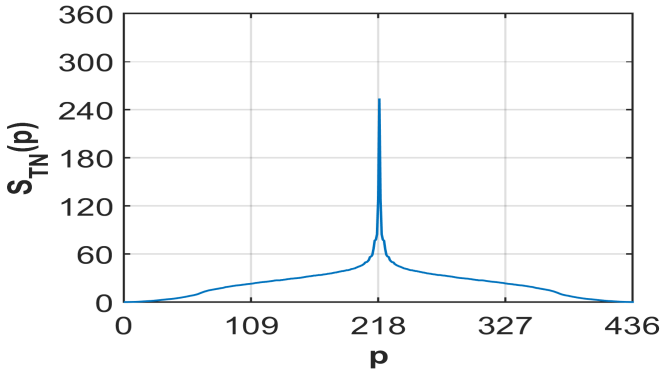
(a)



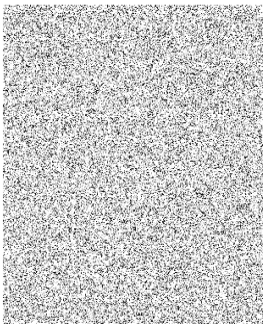
(b)



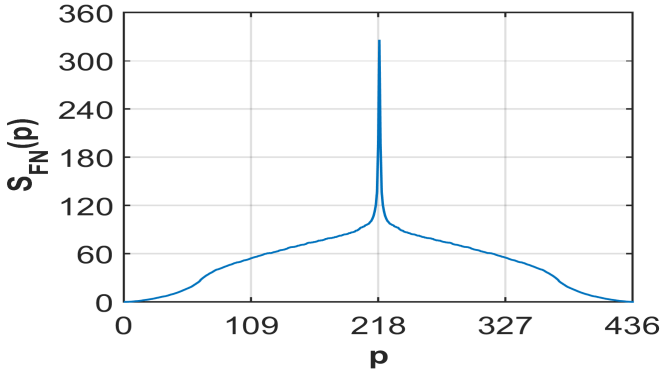
(c)



(d)

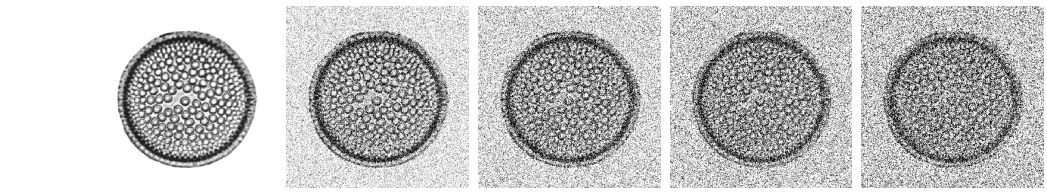


(e)

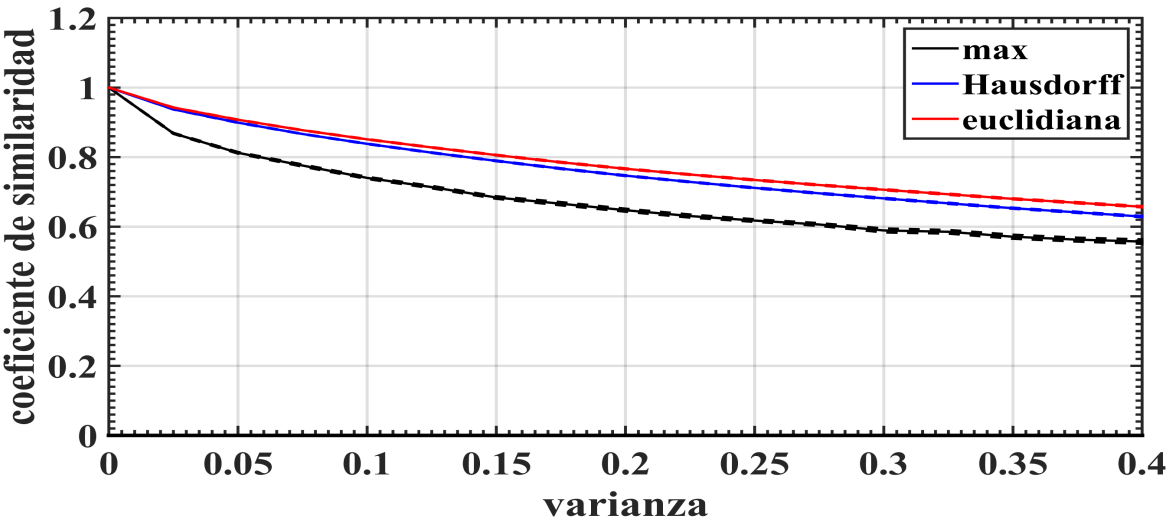


(f)

Figura 4.4.1: Ejemplos de imágenes en escala de grises utilizadas para medir el coeficiente de similaridad. (a) Imagen objetivo. (b) S_T: Firma de la imagen objetivo. (c) Imagen problema con ruido aditivo gaussiano de media $\mu = 0$ y varianza $\sigma^2 = 0.3$. (d) S_{TN}: Firma de la imagen problema. (e) Imagen del ruido de fondo. (f) S_{FN}: Firma de la imagen del ruido de fondo.



(a) (b) (c) (d) (e)



(f)

Figura 4.4.2: Imagen problema en escala de grises con ruido aditivo gaussiano para diferentes valores de varianza y su gráfica del coeficiente de similaridad. (a) $\sigma^2 = 0$. (b) $\sigma^2 = 0.1$. (c) $\sigma^2 = 0.2$. (d) $\sigma^2 = 0.3$. (e) $\sigma^2 = 0.4$ (f) Gráfica del coeficiente de similaridad.

Para el caso de imágenes en escala de grises, se tiene una vez más que para ruido aditivo gaussiano con varianza menor a $\sigma^2 = 0.4$, dominan los coeficientes de similaridad generados por la distancia euclidiana y la distancia de Hausdorff modificada, teniendo ambas un mejor desempeño que la métrica del máximo.

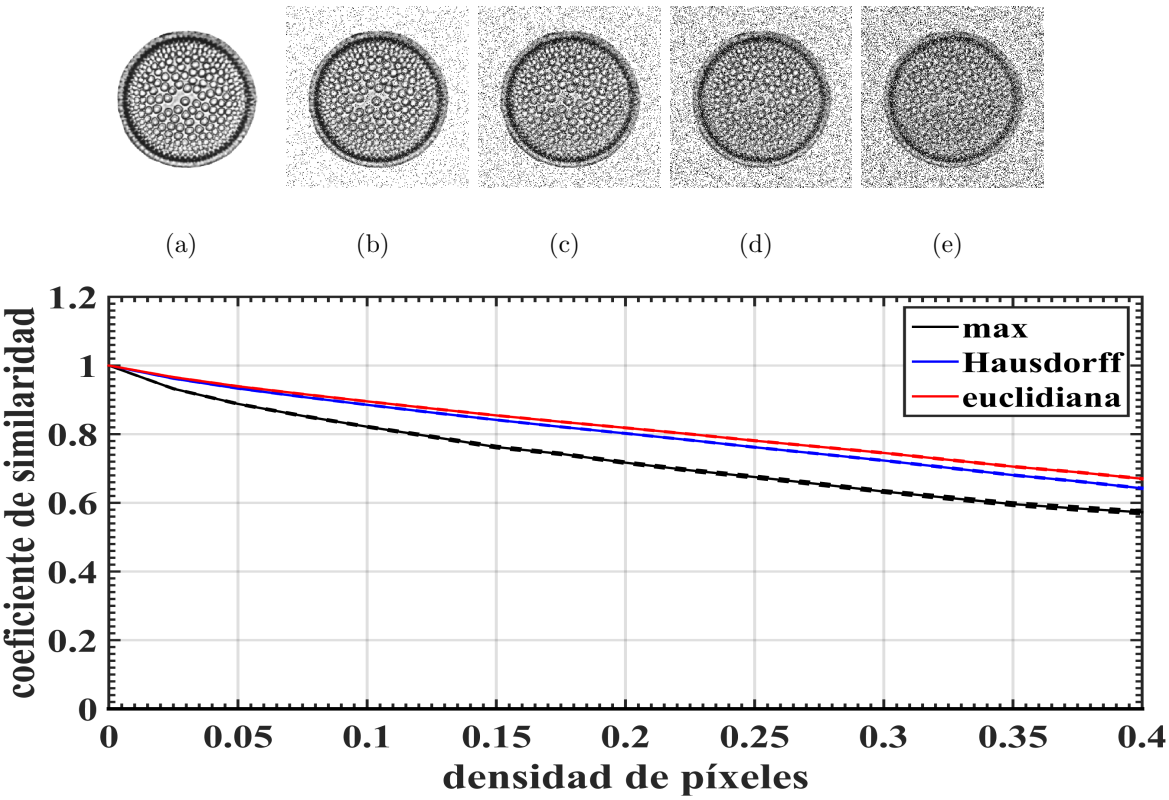


Figura 4.4.3: Imagen problema en escala de grises con ruido sal y pimienta para diferentes valores de densidad de píxeles y su gráfica del coeficiente de similaridad. (a) $d = 0$. (b) $d = 0.1$. (c) $d = 0.2$. (d) $d = 0.3$. (e) $d = 0.4$ (f) Gráfica del coeficiente de similaridad.

Para el caso del ruido sal y pimienta en la Fig. 4.4.3, el comportamiento es idéntico al del ruido aditivo gaussiano, por lo que se concluye que obtuvimos un patrón en el comportamiento del coeficiente similaridad, por lo tanto, como para las imágenes en blanco y negro, y en escala de grises el coeficiente de similaridad con la distancia euclidiana mostró tener siempre los valores más grandes, se recomienda su uso sobre la distancia de Hausdorff modificada y la distancia del máximo.

5. Conclusiones

Utilizando la métrica del coeficiente de similaridad, se analizó el desempeño de las distancias euclidiana, del máximo y de Hausdorff modificada, para la comparación de las firmas 1D generadas por el sistema de reconocimiento de patrones en imágenes digitales de Radon-Fourier-Mellin invariante a traslación, escala y rotación. La comparación se realizó para imágenes en blanco y negro e imágenes en escala de grises.

En este trabajo se propuso utilizar la distancia de Hausdorff modificada como métrica de desempeño para la comparación de las firmas 1D generadas por el sistema RFM debido a sus ya conocidas capacidades de discriminación de formas, pero las pruebas realizadas durante la investigación mostraron que la distancia euclidiana produce mejores resultados tanto para imágenes en blanco y negro como para imágenes en escala de grises bajo la influencia de ruido aditivo gaussiano y sal y pimienta.

Bibliografía

[1] E.M. Arkin, L.P. Chew, D.P. Huttenlocher, K. Kedem, J.S.B. Mitchell. *An Efficiently Computable Metric for Comparing Polygonal Shapes*. Cornell University, Ithaca, NY 14853. 1989.

[2] A.K. Boyat, B.K. Joshi. *A Review Paper: Noise Models In Digital Image Processing*. Signal & Image Processing: An International Journal (SIPIJ) Vol.6, No.2. 2015.

[3] Creative Commons. *Distancia Chessboard*. Recuperado de `https://en.wikipedia.org/wiki/Chebyshev_distance#/media/File:Chess_x5t45.svg`. 2010.

[4] Creative Commons *Distancia de Hausdorff*. Recuperado de `https://en.wikipedia.org/wiki/Hausdorff_distance#/media/File:Hausdorff_distance_sample.svg`. 2007.

[5] M. Dubuisson, A.K. Jain. *A Modified Hausdorff Distance for Object Matching*. Proc. International Conference on Pattern Recognition, Jerusalem, Israel, pp 566–568. 1994.

[6] M.M. Deza, E. Deza. *Encyclopedia of Distances*. Springer, Berlin, Alemania, 1era edición. 2009.

[7] Dominio Publico. *Distancia Manhattan*. Recuperado de `https://en.wikipedia.org/wiki/Taxicab_geometry#/media/File:Manhattan_distance.svg`. 2006.

[8] R.O. Duda, P.E. Hart, D.G. Stork. *Pattern Classification*. Wiley Interscience, NY, USA, 2da edición. 2000.

[9] Y. Gao. *Efficiently comparing face images using a modified Hausdorff distance*. IEEE. 2003.

[10] N. Grégoire, M. Bouillot. *Hausdorff Distance Between Convex Polygons* Apuntes de clase McGill University. Recuperado de `http://cgm.cs.mcgill.ca/~godfried/teaching/cg-projects/98/normand/main.html`. 1998.

[11] R.C.González, R.E. Woods. *Digital Image Processing* Prentice Hall. NY, USA, 2nda edición. 2002.

[12] R.C.González, R.E. Woods, S.L. Eddins. *Digital Image Processing Using Matlab*. Tata McGraw Hill Education, NewDelhi, 2da edición. 2010.

[13] T.V. Hoang, S. Tabbone. *Invariant Pattern Recognition Using the RFM Descriptor*. Patter Recogn. 45, 271-284. 2012.

[14] D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge. *Comparing Images Using the Hausdorff Distance*. Transactions of Pattern Analysys and Machine Inteligence, Vol. 15, No.9. 1999.

[15] H.P. Hsu. *Análisis de Fourier*. Prentice Hall, Edo. de México, México, 1era edición. 1998.

[16] Y. Hu, Z. Wang. *A Similarity Measure Based on Hausdorff Distance for Human Face Recognition*. The 18th International Conference on Pattern Recognition (ICPR'06), IEEE. 2006.

[17] M. Petrou, C. Petrou. *Image Processing: The Fundamentals*. Wiley & Sons Ltd, UK, 1era edición. 2010.

[18] P. Sinha, R. Russell. *A perceptually based comparison of image similarity metrics*. Perception, volumen 40, páginas 1269-1281. 2011.

[19] S. Solorza, J. Verdugo. *A RFM Pattern Recognition System Invariant to Rotation, Scale and Translation*. Proc. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. 2015.

[20] S. Theodoridis, K. Koutroumbas. *Pattern Recognition*. Academic Press, California, USA, 2da edición. 2009.

[21] J. Verdugo. *Reconocimiento De Patrones En Imágenes Digitales Usando La Transformada De Radon*. Tesis de Licenciatura, Universidad Autónoma de Baja California. 2015.