



CENTRO DE INVESTIGACIONES EN OPTICA, A.C.

TESIS DOCTORAL

“EVALUACIÓN DE OBSERVADORES HUMANOS EN TAREAS DE
DETECCIÓN DE SEÑALES USANDO LOGICA DIFUSA Y LA TEORIA ROC ”

PRESENTA:

M.I. JOSE ALFREDO PADILLA MEDINA

ASESOR:

Dr. FRANCISCO JAVIER SÁNCHEZ MARIN

LEON, GTO.

FEBRERO DEL 2003.

INDICE

Introducción.	1
I.- Sistemas Expertos Difusos.	
1.1.- Introducción.	4
1.2.- Conjuntos y Sistemas Difusos.	5
1.3.- Propiedades de los Conjuntos Difusos.	10
1.3.1.- Sistemas Difusos Aditivos.	12
1.4. Partes de un Sistema Experto Difuso.	15
1.4.1.- Diseño de un Sistema Experto Difuso.	18
Bibliografía.	20
II. –Teoría de Detectabilidad de Señales.	
2.1. – Introducción.	21
2.2. – Los Métodos de Selección Forzada.	23
2.2.1. – El Método 2AFC.	24
2.2.2. – El Método MAFC.	25
2.2.3. – Cálculo del Índice de Desempeño o de Detectabilidad.	26
2.2.3.1. – Cálculo de d' en experimentos 2AFC.	26
2.2.3.2. – Cálculo de d' en experimentos MAFC.	28
2.3. – La Curva ROC.	30
2.3.1. – Uso de la curva ROC.	32
2.3.2. – El Método ROC SI-NO.	34
2.3.2.1. – Cambiando el criterio de decisión.	38
2.3.3. – El Método ROC por categorías.	39
2.3.3.1. – La curva ROC como una gráfica binormal.	45
2.4. – Distribución asumida de Señal y Ruido.	47
2.4.1. – Historia del supuesto Gaussiano.	48
2.4.2. – El supuesto Gaussiano y la tarea de simple observación.	49
2.4.2.1. - Caso de igual varianza ($\sigma_s = \sigma_n$).	50
2.4.2.2. – Caso de varianzas desiguales ($\sigma_s \neq \sigma_n$).	51
2.5. – El Supuesto Exponencial.	53
2.5.1. – La distribución exponencial negativa y la razón de verosimilitud.	54
Bibliografía.	58
III. – Implementación de un Sistema Experto Difuso para la Evaluación de Observadores Humanos en la Tarea de Detección de Señales.	



3.1. – Introducción.	59
3.2. – Las Imágenes de Prueba.	60
3.3. – Desarrollo del Sistema Experto Difuso (SED).	61
3.3.1. – Condiciones Teóricas.	62
3.3.2. – Variables de Entrada y Salida.	63
3.3.3. – La Máquina de Inferencia.	65
3.4. – Resultados.	65
3.5. – Conclusiones.	71
Bibliografía.	74

IV. - Sistema Experto Difuso para desarrollar experimentos de percepción visual adaptivos.

4.1. - Introducción.	76
4.2. – Métodos y Procedimientos.	77
4.2.1.- Las imágenes de prueba.	77
4.2.2. – Diseño experimental.	77
4.2.3. – Los observadores humanos.	79
4.2.4.- Variables de entrada y salida del sistema difuso.	79
4.2.5.- La máquina de inferencia.	81
4.3.- Resultados.	82
4.4.- Discusión.	88
Bibliografía.	90



INTRODUCCIÓN.

Nuestros sistemas sensoriales responden a cambios de energía en el entorno. Estos cambios pueden expresarse en forma de estímulos electromagnéticos, mecánicos, químicos o térmicos. El problema de la detección radica en qué tanto cambio de energía, a partir de cero, es necesario para que un sistema sensorial registre dichos estímulos electromagnéticos. Entonces, en términos clásicos, la cantidad mínima de energía detectable se denomina *umbral absoluto*. Es de suponerse que, para un estímulo de energía por debajo del umbral absoluto, una persona sería incapaz de detectarlo; tan pronto como se supere la intensidad de energía por arriba del umbral absoluto, se esperaría que el observador detectará siempre el estímulo físico.

Las mediciones del umbral absoluto y de las capacidades de las personas para detectar un estímulo físico varían con la sensibilidad de los observadores y con las variaciones en las estrategias de decisión de éstos. La teoría de detectabilidad de señales pretende abordar este problema. Se trata de una teoría matemática que supone que cualquier estímulo debe manifestarse, sobre alguna plataforma ruidosa, en nuestros sistemas sensoriales. Por consiguiente, y dada una secuencia de presentaciones, en cada presentación el observador debe decidir si una señal se encuentra presente sobre la plataforma de ruido o si sólo se trata de ruido únicamente. En la teoría de detectabilidad de señales no existe un umbral absoluto; sólo hay una serie de presentaciones, y cada una debe clasificarse como señal presente o ausente. Es posible utilizar una secuencia de presentaciones para determinar qué tan sensible es una persona a una señal determinada, independientemente de cualquier efecto de motivación o expectativa que pudiera sesgar su decisión.

La teoría de detectabilidad de señales requiere un tipo especial de experimentos a para poder establecer un índice de desempeño de observadores. Este tipo de experimentos utilizan 2 tipos de presentaciones de estímulos (de *señal ausente* y de *señal presente*). En una presentación de señal ausente los observadores experimentan eventos sensoriales debido únicamente a la plataforma ruidosa. En una presentación de señal presente los observadores experimentan eventos sensoriales debido a la plataforma de ruido y a un estímulo visual. Las respuestas típicas de un observador en un experimento de percepción visual de señales son *Sí* y *No*. *Sí* indica que el observador decidió que la señal estaba presente en la presentación. *No* indica que el observador decidió que la señal no estaba presente en la presentación. Un “sí” en



una presentación con señal presente es denominado *acierto*, en tanto que un “sí” en una presentación con señal ausente es denominado *falsa alarma*. De igual manera, un “no” en una presentación con señal ausente se denomina *negativa correcta*, y un “no” en una presentación con señal presente se denomina *error*. La probabilidad relacionada con cada uno de estos resultados se debe a dos mediciones independientes de la ejecución del observador, a su *criterio de decisión* y a su capacidad para diferenciar entre los eventos de ruido únicamente (n) y de señal y ruido (sn).

Por lo general la detectabilidad de una señal y el desempeño de un observador se analizan de acuerdo a la proporciones de aciertos y de falsas alarmas, relación que cambia a medida que varía el criterio de decisión utilizado por el observador. Por lo regular la proporción de aciertos se grafica en la ordenada y las falsas alarmas en la abcisa. Las curvas obtenidas se denominan curvas ROC (Receiver Operating Characteristic) y muestran la relación entre la proporción de correctas y la de falsas alarmas. El término ROC implica que la curva mide y describe las características de operación y de desempeño del receptor (observador) para detectar cierto tipo de señales.

Uno de los experimentos más comunes en la teoría de detectabilidad de señales es el ROC por categorías, en los cuales se le presentan estímulos a los observadores para que decidan, seleccionando una de varias respuestas acorde a su criterio de decisión, si ha ocurrido una presentación con ruido únicamente o con señal y ruido. Es decir, el observador utiliza una escala de categorías para expresar su grado de confianza de que la señal se encuentra presente o ausente.

En un análisis realizado, utilizando los supuestos de distribución Gaussiana y exponencial de eventos sensoriales generados por un estímulo visual, de los datos de los experimentos realizados en el desarrollo de este trabajo de investigación fue notario que existen dos factores que afectan a la aplicabilidad de los experimentos ROC por categorías y los resultados que estos arrojan. El primero factor corresponde a que los índices de desempeño tradicionales arrojados por la teoría de detectabilidad de señales no presentan suficiente información para discriminar entre observadores de bajo y alto desempeño. El segundo factor corresponde a que, dado que el número de presentaciones (N) no está claramente definido y N normalmente se selecciona tan grande como sea posible para reducir efectos de variabilidad en los resultados, el tiempo de ejecución de los experimentos de percepción visual es normalmente muy grande lo cual puede causar errores debidos a la



fatiga y pérdida de atención en los observadores y, por consecuencia, afectar la confiabilidad de los resultados obtenidos.

En el presente trabajo se presenta una propuesta para resolver cada uno de los factores anteriormente mencionados utilizando como base a la lógica difusa para representar y manejar la incertidumbre que afecta al proceso de evaluación de observadores humanos cuando se utiliza el método ROC por categorías como experimento de percepción visual de señales. En cada caso fue necesario la implementación de un sistema experto difuso que expresara y manejara la incertidumbre propia de la aplicación y representara el conocimiento de los expertos en el desarrollo de experimentos de percepción visual utilizando el método ROC por categorías.

La estructura del presente trabajo de tesis se ha determinado de tal manera que en el primer capítulo se presenten los conceptos básicos de la lógica difusa y la forma en cómo estos conceptos se utilizan en el diseño de sistemas expertos difusos que representen el conocimiento de un experto humano al momento de analizar datos provenientes (índices de desempeño) de un experimento de percepción visual.

El capítulo II de este trabajo trata acerca de la teoría de detectabilidad de señales y los dos supuestos más comúnmente utilizados para asumir una distribución de eventos sensoriales cuando se experimenta una sensación debida a un estímulo físico. Además, se denota la forma en que se obtienen índices de desempeño bajo cada supuesto.

El capítulo III muestra la metodología utilizada para el diseño de un ROC difuso cuyo objetivo fue encontrar un nuevo índice de desempeño en base a los índices de desempeño obtenidos bajo el supuesto exponencial en cada punto de la curva ROC. Este nuevo índice presentó mejoras, con respecto a los índices tradicionales, acerca de una mejor discriminación entre observadores con índice de desempeño tradicional similares.

Finalmente, el capítulo IV muestra la forma en que utilizamos un sistema experto difuso para el desarrollo de experimentos ROC por categorías adaptivos. El objetivo de este sistema fue diseñar experimentos cortos de percepción visual cuyos resultados no fueran fuertemente afectados por la fatiga y pérdida de atención de los observadores. De esta manera es posible ampliar la aplicación de experimentos ROC por categoría en áreas donde no se cuenta con observadores experimentados.



CAPITULO I.- SISTEMAS EXPERTOS DIFUSOS.

1.1. – INTRODUCCION.

Los sistemas expertos son programas de computadora que simulan el proceso de razonamiento de un experto humano, o que de alguna manera realizan una tarea en forma experta en un dominio en el cual no existe un experto humano[1].

Los Sistemas Expertos(SE) fueron desarrollados como un área de la Inteligencia Artificial a mediados de los 60's, diseñándose en esta época sólo sistemas con fines experimentales. Este período fue dominado por la idea de que una cuantas leyes de razonamiento acoplados a algoritmos eficientes producirán un experto. Uno de los primeros sistemas expertos fue el GPP (General Purpose Problem), éste fue desarrollado por Newell y Simon en un intento por crear una máquina inteligente que simulara el razonamiento humano[2]. El GPP pretendía trabajar con los pasos necesarios para cambiar una cierta situación inicial en una meta deseada.

Sin duda alguna el SE que ocasionó que los SE progresaran en diversas áreas fue MYCIN, éste sistema fue desarrollado por el Dr. Edward H. Shortliffe en la escuela de medicina de Stanford y es considerado el padre de los SE. MYCIN fue desarrollado para realizar diagnósticos de algunas enfermedades de la sangre y asignar el correspondiente tratamiento. La aportación de MYCIN fue haber introducido el factor de certidumbre, concepto ampliamente usado en el desarrollo de SE para el manejo de la imprecisión. Cada regla en MYCIN tiene un peso entre 0 y 1 denominado factor de certidumbre el cual es asignado cuando la regla es evaluada para un conjunto de datos de entrada[1].

PROSPECTOR es otro ejemplo de SE desarrollado en la década de los 70's por Peter Duba, Peter Hard y Rene Regoh, PROSPECTOR es un experto en geología, predice la posibilidad de que ciertos depósitos de minerales se encuentran en una región en particular.

A diferencia de los sistemas expertos tradicionales, los Sistemas Expertos Difusos (SED) utilizan reglas de razonamiento difusas para el manejo de la incertidumbre propia de la aplicación. Los SED han ido sustituyendo a los sistemas de ingeniería convencionales, especialmente en sistemas de control y de reconocimiento[1,4]. Esto se debe a la facilidad con que estos sistemas pueden implementarse y a los resultados tan satisfactorios que estos producen. La base de todo sistema difuso se encuentra en el manejo de “funciones de



membresía” concepto con el cual el sistema es capaz de manejar la incertidumbre y vaguedad en diversas aplicaciones.

La idea de tener un SE en cualquier área (investigación, económico, militar, médico, etc.), que reproduzca el conocimiento tal como un experto humano lo haría y con las características de aprehensivo, comprensivo, explicativo y flexible, ha originado la necesidad de implementar algoritmos de computo poderosos (sistemas inteligentes) teniendo como herramientas principales, entre otras: Sistemas de computo poderosos, múltiples líneas de razonamiento y manejo de información con técnicas difusas.

Así pues, se incluye en este trabajo la teoría manejada en el ámbito de la lógica difusa poniendo especial interés en la metodología de implementación de un SED.

1.2. - CONJUNTOS Y SISTEMAS DIFUSOS.

Desde niños aprendemos rápidamente cómo interpretar e implementar instrucciones difusas (“acuéstate como a las diez”, por ejemplo). Desde niños asimilamos y usamos datos difusos, reglas vagas e información imprecisa bajo las cuales tomamos decisiones en situaciones que parecen estar gobernadas por elementos de suerte o azar[1][6][7]. Los modelos convencionales de los sistemas reales también deben ser capaces de reconocer, representar, manipular y usar incertidumbre tanto difusa como estadística. Los investigadores de la lógica difusa proponen una ontología en donde tiene cabida la vaguedad implicada cuando un evento puede ser “más o menos” cierto.

Los conjuntos difusos fueron introducidos por Lotfi Zadeh en 1965 como una nueva manera de representar, en forma matemática y sencilla, la vaguedad de nuestra vida diaria y en diversas áreas de investigación. Con la lógica difusa es posible representar las vaguedades que surgen a cada momento utilizando “datos difusos” que representan ideas vagas e información imprecisa dentro de modelos computacionales que tienen la capacidad de reconocer, representar, manipular e interpretar la incertidumbre difusa y estadística[1][3][6].

Un conjunto difuso A contiene a un elemento x en algún grado, por ejemplo, el conjunto difuso $(1, 3/4, 2/3, 1/3, 1/4, 0)$ contiene 4 elementos parcialmente, donde 1 corresponde al grado de total pertenencia y 0 corresponde al estado de no pertenencia; los valores



intermedios ($3/4, 2/3, 1/3, 1/4$) representan grados de pertenencia con características inciertas de los elementos de un conjunto. Es posible extender este concepto a casos continuos donde los conjuntos son subconjuntos de una línea real denominada función de membresía[9].

La gran diferencia entre los conjuntos difusos y los convencionales se encuentra en la forma en que cada elemento o regla de un conjunto adquiere un valor numérico. Los elementos de un conjunto convencional solo pueden tener como grado de pertenencia un valor “0” o un valor “1”, mientras que los elementos de un conjunto difuso pueden adquirir infinitud de valores en el intervalo $[0,1]$. Además, es importante diferenciar entre lo difuso y lo probabilístico, la representación difusa señala objetos con propiedades imprecisas y la representación probabilística denota información basada en frecuencias relativas[1,3,6].

En un sentido muy amplio es posible interpretar el modelado difuso como un esquema de modelado cuantitativo el cual podemos describir usando nuestro lenguaje natural. Visto desde el punto de vista matemático, el modelado difuso es la descripción de un sistema usando cantidades difusas, a su vez que las cantidades difusas son números o conjuntos difusos asociados a proposiciones difusas[5].

Zadeh, en sus primeros artículos, pretende dar la idea de que “la mayoría de los problemas reales tienden a ser complejos y, por lo tanto, son algorítmicamente complejos de resolver”[3]. Zadeh sugiere un análisis lingüístico en lugar de uno cuantitativo y recomienda los siguientes puntos para la implementación de un SED:

- 1) Uso de variables lingüísticas en lugar de o en conjunto con variables numéricas normalmente definidas de manera determinística o probabilísticas,
- 2) Caracterización de las relaciones simples entre las variables por medio de declaraciones difusas condicionales,
- 3) Caracterización de relaciones complejas por medio de algoritmos difusos.

Los conjuntos son difusos cuando asemejan propiedades imprecisas o vagas. La mayoría de los problemas son al principio difusos, la solución de éstos inicia con el proceso



de difusificación, resolviéndose el problema en este dominio y posteriormente se vuelve al dominio original con un procedimiento que se denomina desdifusificación.

Para entender mejor el tipo de datos que se utilizan en la solución de problemas con lógica difusa considere el siguiente caso en el cual se maneja incertidumbre:

Sea F el conjunto de todos los números reales *cercanos a 7* definido en un intervalo denominado *alrededor de 7*. La propiedad difusa “cercano a 7” obliga a tener más de una función de membresía (grado en que los objetos satisfacen las propiedades imprecisas del conjunto en cuestión) y el modelador debe elegir la que más se acople a sus necesidades basándose en su intuición. De esta manera, una gráfica que muestre los posibles valores de la función de membresía para el conjunto F se muestra en la Fig. 1.1.

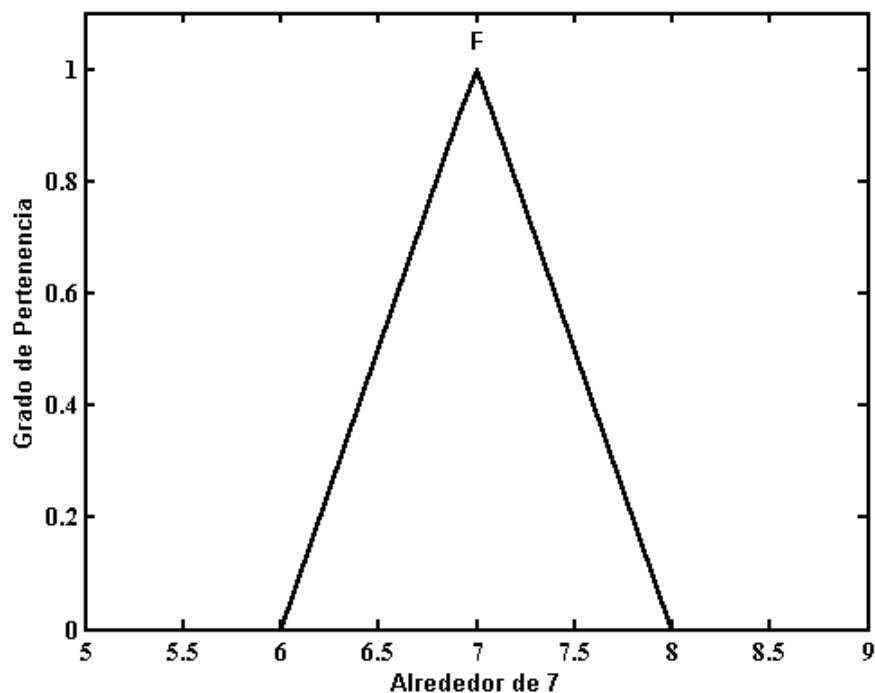


Fig. 1.1. - Ejemplo de Función de Membresía triangular.

en la Fig. 1.1 es posible observar 3 propiedades de las funciones de membresía:

- 1) Normalidad: $m(F) = 1$,
- 2) Monotonicidad: cuanto más cercano es a 7 el número real, mF es más cercano a 1,



- 3) Simetría: números con la misma distancia a la derecha y a la izquierda de 7 tendrán el mismo valor de mF.

Es importante señalar que no solamente se pueden utilizar gráficas de difusificación triangulares, también podemos utilizar formas trapezoidales, gaussianas, senoidales o algún otra forma geométrica.

Con lógica difusa es posible asignar diversos grados de certidumbre a una oración o sentencia. Hay hechos en el mundo que no necesariamente son completamente ciertos o falsos, sino que tienen un cierto grado de verdad (o pertenencia)[8]. Por ejemplo, “Viena es una ciudad grande” podría ser una oración válida en 0.6.

Ejemplo:

Considere el ejemplo de un sistema de aire acondicionado. La temperatura se trata como una variable difusa o variable “lingüística”. El aire puede tomar valores sobre los conjuntos difusos “helado”, “fresco”, “templado”, “tibio” y “caliente”. Podemos definir las curvas para estos conjuntos difusos como se muestra en las Figs. 1.2 y 1.3.

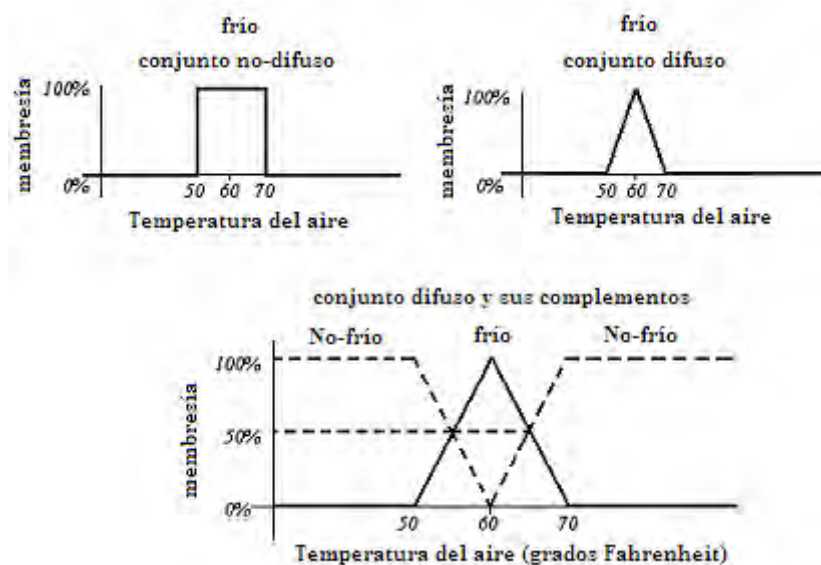


Fig. 1.2. – Conjunto difuso bivalente y conjunto difuso sobre la línea real. Cada curva define una función de membresía.

La Fig. 1.2 compara conjuntos bivalentes con conjuntos multivalentes de grados fahrenheit de aire frío. El conjunto difuso presenta contradicciones parciales que suman la



unidad por cada objeto x . El aire es frío y no frío para algunos grados. Si el aire es 60% frío entonces debe ser 40% no frío. Un conjunto difuso A siempre intersecta a su complemento en algún valor de x , en este caso existe un cambio abrupto de A a $not-A$ y viceversa.

Los conjuntos difusos forman los bloques base para la implementación de las reglas difusas “si-entonces” tales como “si el aire es frío entonces la velocidad del motor es lenta” o “Si el aire está caliente entonces la velocidad del motor es rápida”. Estas reglas tienen la forma “si X es A entonces Y es B ” donde A y B son conjuntos difusos. Un sistema difuso es un conjunto de reglas difusas que mapea entradas a salidas mediante procedimientos difusos. El acondicionador de aire difuso puede mapear la temperatura de entrada de 68 grados

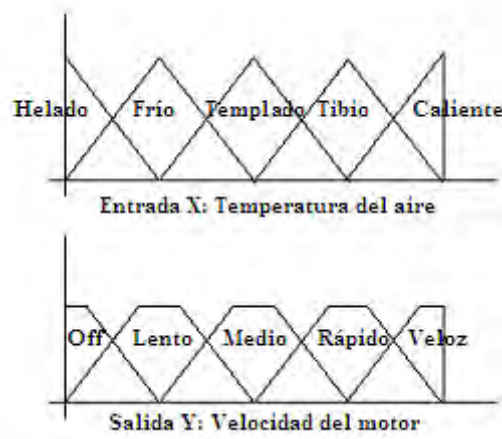


Fig. 1.3. – Variables de entrada y de salida difusa para un control de refrigeración.

fahrenheit a una velocidad de motor de 47 revoluciones por minuto (rpm).

Los sistemas de ingeniería difusa implican tres pasos: el primer paso es seleccionar las variables de entrada y salida X y Y . En nuestro caso la variable de entrada X es la temperatura y la salida Y es un conjunto de valores numéricos de la velocidad del motor. El segundo paso es seleccionar los conjuntos difusos de estas variables. La Fig. 1.3 muestra los conjuntos difusos para las variables de entrada y de salida. El tercer paso es relacionar los conjuntos de salida con los conjuntos de entrada mediante reglas difusas de la forma:

Regla 1: Si el aire esta helado entonces la velocidad del motor es nula(Off).

Regla 2: Si el aire esta frío entonces la velocidad del motor es lenta.



Regla 3: Si el aire esta templado entonces la velocidad del motor es media.

Regla 4: Si el aire esta tibio entonces la velocidad del motor es rápida.

Regla 5: Si el aire esta caliente entonces la velocidad del motor es veloz.

El conjunto de reglas difusas debe ser diseñado por el experto y establecido con base a la forma en que tiene el experto de expresar sus experiencias acerca de la aplicación de interés. Esto normalmente se expresa como un conjunto de reglas FuzzyAnd y/o FuzzyOr.

1.3. – PROPIEDADES DE LOS CONJUNTOS DIFUSOS.

Retomando el concepto de que la función de membresía es la idea básica de la teoría de conjuntos difusos y que su valor mide el grado en que los objetos satisfacen las propiedades definidas imprecisamente. A fin de manipular los conjuntos difusos, es necesario tener operaciones que nos permitan combinar las funciones de membresía de un grupo de variables. Zadeh[3] delineó las operaciones clásicas para los conjuntos difusos, estas operaciones se caracterizan en términos de que una función de variable física difusificada $F(X)$ contiene todos los subconjuntos difusos de X como lo muestra la Fig. 1.4. Es decir $m_A(x) \in F(X)$ y $m_B(x) \in F(X)$ para toda $x \in X$.

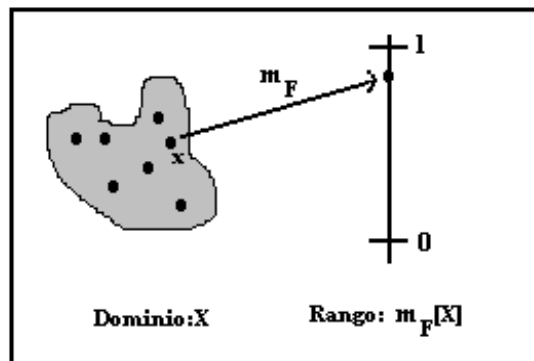


Fig. 1.4.- Mapeo de conjunto X a dominio difuso.

Considerando estas funciones de membresía y que A y B son conjuntos difusos, las operaciones más comúnmente usadas en la implementación de Sistemas Expertos Difusos son:

$$(=) \text{ Igualdad} \quad A = B \leftrightarrow m_A(x) = m_B(x) \quad 1.1$$

$$(\subset) \text{ Inclusión} \quad A \subset B \leftrightarrow m_A(x) \leq m_B(x) \quad 1.2$$



$$(\prime) \text{ Complemento} \quad (m_A(x))' = 1 - m_A(x) \quad 1.3$$

$$(\cap) \text{ Intersección} \quad m_A(x) \cap m_B(x) = \min\{m_A(x), m_B(x)\} \quad 1.4$$

$$(\cup) \text{ Unión} \quad m_A(x) \cup m_B(x) = \max\{m_A(x), m_B(x)\} \quad 1.5$$

Esta aritmética permite el uso de reglas SI - ENTONCES en aquellos algoritmos computacionales que utilicen un razonamiento difuso.

En la lógica difusa una sentencia S es tal que tiene un valor real $t(S)$. Dicho valor define un mapeo de un conjunto de sentencias a un conjunto de valores reales: $t: \{sentencias\} \rightarrow \{valores reales\}$. La teoría clásica de Aristóteles permite únicamente dos valores reales falso y verdadero o 0 y 1, esto es, $t: \{sentencias\} \rightarrow \{0, 1\}$. Una lógica tri-valuada mapea un conjunto de sentencias a tres valores lógicos, esto es, $t: \{sentencias\} \rightarrow \{0, 1/2, 1\}$. Una lógica continua mapea un conjunto de sentencias a un intervalo unitario $[0, 1]$. Esta lógica se denomina lógica difusa e incluye a la teoría clásica como caso especial.

Funciones reales generan sentencias reales en términos de reglas difusas:

$$t(A \text{ AND } B) = \min(t(A), t(B)) \quad 1.6$$

$$t(A \text{ OR } B) = \max(t(A), t(B)) \quad 1.7$$

$$t(\text{not-}A) = 1 - t(A) \quad 1.8$$

Para ejemplificar estas operaciones supóngase: $t(\text{Pasto verde}) = 0.8$ y $t(\text{nieve blanca}) = 0.9$. Entonces $(t(\text{Pasto verde}) \text{ AND } t(\text{nieve blanca})) = 0.8$, $(t(\text{Pasto verde}) \text{ OR } t(\text{nieve blanca})) = 0.9$, $t(\text{No Pasto verde}) = 0.2$. Las implicaciones difusas pueden presentarse de diferentes formas. La forma de Lukasiewicz establece relación difusa entre conjuntos difusos:

$$t_L(A \rightarrow B) = \min(1, 1 - |t(A) - t(B)|) \quad 1.9$$

Las sentencias A y B son equivalentes o $A=B$ si y solo si A implica B y B implica A. Esto nos lleva a la equivalencia de Lukasiewicz que tiene la forma:

$$t_L(A = B) = 1 - |t(A) - t(B)| \quad 1.10$$



Entonces, las sentencias A y B son equivalentes si y solo si tienen los mismos valores reales $t(A) = t(B)$.

La mayoría de los diseñadores de sistemas difusos se enfocan en los sistemas difusos aditivos de forma algebraica simple. Estos son llamados SAM's (Standard Additive Model's). Un sistema difuso SAM, denotado por F, almacena m reglas de la forma "if $X=A_j$ entonces $Y=B_j$ ". Este teorema muestra que un SAM, sin pesos de las reglas, tiene la forma de una suma convexa del conjunto de centroides c_j .

$$F(x) = \frac{\sum_{j=1}^m a_j(x)c_j}{\sum_{j=1}^m a_j(x)} \quad 1.11$$

En donde la función $a_j: R^n \rightarrow [0,1]$ define el conjunto difuso A_j . El vector de entrada $x \in X$ pertenece al conjunto difuso $A_j \subset R^n$ con grado $a_j(x) \in [0,1]$.

1.3.1. – SISTEMAS DIFUSOS ADITIVOS.

Un sistema difuso F almacena m reglas difusas de la forma "Si $X=A_j$ entonces $Y=B_j$ " y computa la salida $F(x)$ como el centroide de los conjuntos difusos B_j generados por la activación de reglas difusas. Cada entrada x activa todas las reglas en algún grado y en forma paralela. De esta manera un sistema difuso aditivo actúa como un procesador asociativo o global de FAM (Fuzzy Associative Memory). La Fig. 1.6. muestra la estructura de la FAM de un sistema aditivo difuso sin retroalimentación.



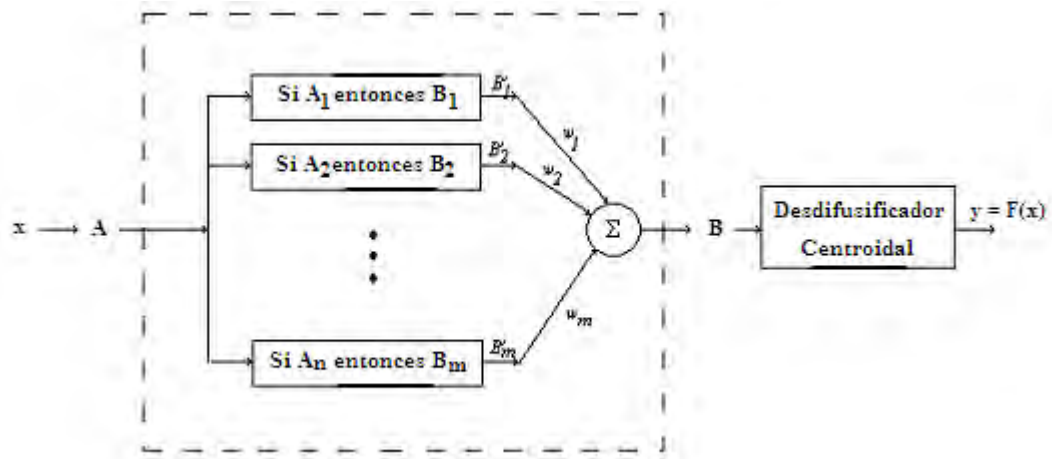


Fig. 1.6. – Arquitectura de un Sistema Difuso Aditivo. Los valores de x activan la parte si-entonces en paralelo y en algún grado.

El modelo aditivo estándar sin retroalimentación es el caso más ampliamente usado. En el caso SAM el conjunto B'_j es el producto $a_j(x)B_j$. La estructura aditiva resulta de la suma de las partes de B activadas.

$$B = \sum_{j=1}^m w_j B'_j = \sum_{j=1}^m w_j a_j(x) B_j \quad 1.12$$

La sumatoria se realiza sobre las m reglas que mapean los subconjuntos difusos $A_j \subset \mathbb{R}^n$ a subconjuntos difusos $B_j \subset \mathbb{R}^p$. Los pesos w_j pueden escalar cada término en la suma para reflejar la credibilidad, frecuencia o “uso” de la regla y poder obtener un término extra para el entrenamiento del sistema de aprendizaje. En la práctica es común ignorar los pesos de las reglas y se toman todas como iguales a la unidad: $w_1 = w_2 = \dots = w_m = 1$.

Podemos establecer la forma más simple del teorema SAM. Este teorema muestra que un sistema difuso mapea entradas x a $F(x)$ como una suma convexa de los m centroides. Estos centroides pueden cambiar con cada valor de x como en un SAM generalizado o como en un SAM adaptivo. El teorema SAM establece el centroide global de B a una suma de centroides locales. Los m coeficientes $p_1(x), \dots, p_m(x)$ son convexos tales que cada término es no negativo, $p_j(x) \geq 0$, y estos suman la unidad.

$$\sum_{j=1}^m p_j(x) = 1 \quad 1.13$$



Así, los pesos definen una función de probabilidad discreta para cada x . El siguiente teorema 1.1 provee un caso más general sobre la base de medias y varianzas estadísticas.

TEOREMA 1.1 (TEOREMA SAM).

Supóngase que un sistema difuso $F:R^n \rightarrow R^p$ es un SAM: $F(x) = \text{Centroide} \left(\sum_{j=1}^m w_j a_j(x) B_j \right)$. Entonces $F(x)$ es una suma convexa de los m centroides.

$$F(x) = \frac{\sum_{j=1}^m w_j a_j(x) V_j c_j}{\sum_{j=1}^m w_j a_j(x) V_j} \quad 1.14$$

$$= \sum_{j=1}^m p_j(x) c_j \quad 1.15$$

Los coeficientes convexos o pesos de probabilidad discreta $p_1(x) \dots p_m(x)$ dependen de las entradas a través de las razones:

$$p_j(x) = \frac{w_j a_j(x) V_j}{\sum_{k=1}^m w_k a_k(x) V_k} \quad 1.16$$

V_j es el volumen positivo finito (o área si $p=1$ en el espacio R^p) y c_j es el centroide del conjunto B_j :

$$V_j = \int_{R^p} b_j(y_1 \dots y_p) dy_1 \dots dy_p > 0 \quad 1.17$$

$$c_j = \frac{\int_{R^p} y b_j(y_1 \dots y_p) dy_1 \dots dy_p}{\int_{R^p} b(y_1 \dots y_p) dy_1 \dots dy_p} \quad 1.18$$



El caso más popular de $p=1$ reduce las ecuaciones 1.17 y 1.18 en:

$$V_j = \int_{-\infty}^{\infty} b_j(y) dy \quad 1.19$$

$$c_j = \frac{\int_{-\infty}^{\infty} y b_j(y) dy}{\int_{-\infty}^{\infty} b_j(y) dy} \quad 1.20$$

1.4. - PARTES DE UN SISTEMA EXPERTO DIFUSO.

Básicamente, un SED se implementa a partir de 3 partes principales: *máquina de inferencia*, *base de conocimientos* y *memoria de trabajo*. La base de conocimientos contiene el conocimiento, proporcionado por el experto humano mediante reglas difusas, necesario para la solución del problema. La memoria de trabajo es usada para almacenar información que el sistema extrae del usuario. La máquina de inferencia utiliza la base de conocimientos y la información adquirida del usuario para poder generar una salida como solución al problema.

Además de las 3 partes principales que contiene un SED, muchos sistemas incorporan un mecanismo de explicación cuya función es la de explicar su razonamiento al usuario. De esta manera el usuario tiene la posibilidad de cuestionar los resultados que el sistema le proporciona y determinar el grado de confiabilidad de las respuestas que el sistema genera.

Las partes que conforman un Sistema Experto se pueden observar en la Fig. 1.7, la estructura puede cambiar para algunos autores, pero las 3 partes fundamentales anteriormente mencionadas siempre estarán presentes[2].

Cada bloque de la Fig. 1.7 tiene una función específica la cual se explica brevemente a continuación.

Base de Conocimientos.- Representa una información con la característica de que cada dato se puede manejar con técnicas de resolución difusas, es decir, cada dato o frase de información deberá tener una función de membresía o un número difuso en el intervalo [0-1].



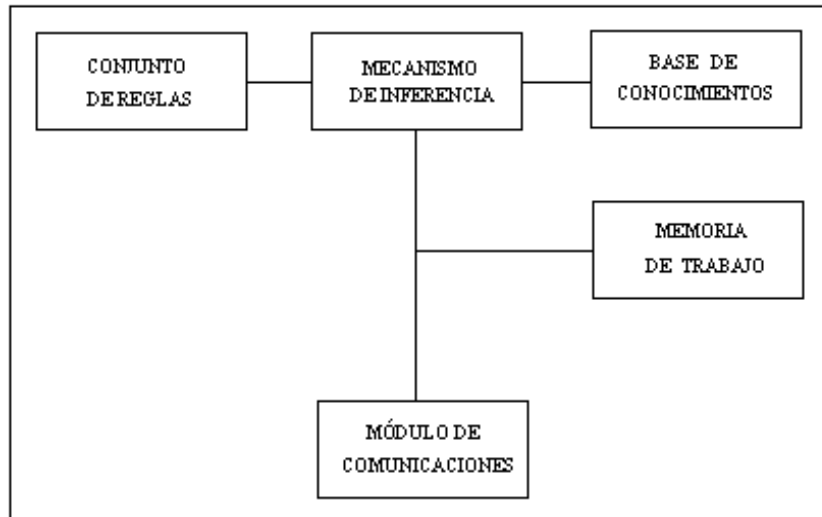


Fig. 1.7. – Diagrama a bloques de un Sistema Experto.

La base de conocimientos consiste de hechos y reglas expresadas de la forma:

El (atributo) de (objeto) es (valor).

o expresada con una terna ordenada tenemos:

(atributo, objeto, valor).

Algunos ejemplos son:

(Edad, Juan, 50) la edad de Juan es 50.

(distancia, (León, Celaya), 150)..... la distancia entre León y Celaya es 150

Pero las reglas pueden no ser siempre tan precisas, por ejemplo:

(Edad, Juan, vieja).... la edad de Juan es vieja.

(Temperatura, Juan, ≈ 40).... la temperatura de Juan es aproximadamente de 40.

donde vieja y ≈ 40 son representadas con funciones de membresía como se muestra en la Fig. 1.8.



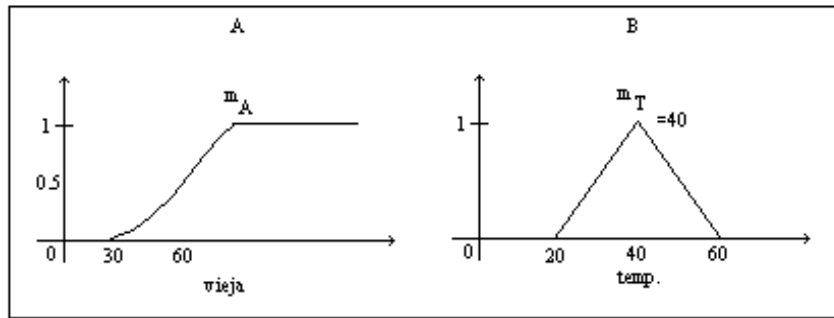


Fig. 1.8. – Representación del conocimiento con funciones de membresía difusas (A-vieja) y (B-temperatura).

Mecanismo de Inferencia.- Es la parte del sistema que utiliza la información de la base de conocimientos, en una presentación muy particular, para llegar a una posible solución; representa el corazón del SED. La máquina de inferencia es responsable de determinar cuándo el usuario realiza una pregunta al sistema o cuándo se trata de incrementar la base de conocimientos. Se encarga de realizar los procesos de difusificación y desdifusificación de las variables de entrada y, en algunos casos, de las variables intermedias.

Conjunto de Reglas.- Son reglas con la característica de difusidad y determinan la forma en que el Mecanismo de Inferencia obtiene una posible solución. Pueden estar almacenadas en una memoria RAM o EEPROM. Este conjunto de reglas está comprendido de un conjunto de reglas IF THEN evaluadas con operaciones de la lógica difusa. Comúnmente operaciones AND y/o OR.

Memoria de Trabajo.- Es la parte del sistema que contiene toda la información inicial que se le introduce al sistema, además, en esta se almacenan los resultados intermedios que se generan en el proceso de decisión del sistema para incluirse en el proceso de explicación y respaldo de los resultados. Físicamente consiste de algún tipo de dispositivo de almacenamiento como memoria RAM, EEPROM, unidades de disco flexibles o de disco duro.

Módulo de Comunicaciones.- Representa la forma en la que le usuario se comunica con el SED, es decir, la forma en que el usuario introduce la información necesaria al SED y



observa los resultados obtenidos. Típicamente, el módulo de comunicaciones esta representado por el teclado y el monitor de un sistema de computo.

1.4.1. – DISEÑO DE UN SISTEMA EXPERTO DIFUSO.

El diseño de un SED implica del manejo de un conjunto de reglas si-entonces y conjuntos difusos que se traslapan. Un vector unidimensional x se presenta como la entrada al sistema. Para proporcionar simplicidad a nuestro análisis, en este parte del texto se utilizan funciones triangulares para representar conjuntos difusos. La Fig. 1.6 muestra la arquitectura de un sistema difuso aditivo, el sistema utiliza un combinador de reglas *min* y como método de desfusificación el método del *centroide*.

Un vector de entrada n -dimensional y la j -ésima regla difusa de la forma $(A^1_j: A^2_j \dots A^n_j: B_j)$ producen un conjunto de resultados B'_j :

$$B'_j = B_j \min(a^1_j(x_1) \cdot a^2_j(x_2) \dots a^n_j(x_n)) \quad 1.21$$

donde $a: X \rightarrow [0,1]$ es la función del conjunto difuso $A \subset X$.

Un sistema con m reglas proporciona m conjuntos difusos B'_j , como se muestra en la Fig. 1.6, para obtener el valor de salida final $F(x)$ comúnmente se utiliza:

$$B = \sum_{j=1}^m B'_j \quad 1.22$$

$$F(x) = \frac{\int_{-\infty}^{\infty} y b(y) dy}{\int_{-\infty}^{\infty} b(y) dy} = \frac{\sum_{j=1}^m a_j(x) c_j}{\sum_{j=1}^m a_j(x)} \quad 1.23$$

Para el modelo SAM el sistema difuso $F: R^n \rightarrow R$ almacena m rutas cartesianas $A_1 \times B_1 \dots A_m \times B_m$. Cada conjunto difuso $A_j \subset R^n$ tiene un conjunto función $a_j: R^n \rightarrow [0,1]$. El conjunto total B tiene un conjunto función $b: R \rightarrow [0,m]$. Cada resultado de las reglas B_j tienen



área o centroide c_j de una región plana o centro de masa para una región sin masa y se considera el punto donde los momentos de equilibrio estático, respecto al sistema de coordenadas, son cero. La localización del centroide se expresa en unidades de longitud a lo largo del sistema de coordenadas.

Para en caso del cálculo del centroide en sistemas difusos solo se considera el centroide en x por ser el eje que corresponde a la variable física de interés. La ecuación 1.14 del centroide de un modelo difuso se puede reducir a:

$$F(x) = \frac{\sum_{j=1}^m a_j(x) P_j}{\sum_{j=1}^m a_j(x)} \quad 1.24$$

donde las modas o “picos” P_j de los conjuntos $B_j \subset R^p$ son iguales a los centroides c_j y los B_j tienen todos el mismo centroide c_j y el peso de las reglas difusas w_j : $P_j=c_j$ y $c_1 = \dots = c_m > 0$. Sugeno y otros ingenieros japoneses han popularizado del modelo de centroide como método de desfusificación, la mayoría de circuitos integrados difusos utilizan este método [9].



BIBLIOGRAFIA.

1. Kandel, Abraham, "Fuzzy Expert Systems", CRC Press, 1991, Tampa Florida, pp. 1-81.
2. Sánchez J. P. Y Beltran, "Los Sistemas Expertos: Una Metodología de Programación", Macrobit Corporation, 1990, México D.F.
3. Zadeh, Lotfi, "Fuzzy Sets, information and control", Academic Press, 1965, no. 3, vol. 8, pp 338-353.
4. Infelise, Nick, "A clear vision of fuzzy logic", Control Engineering, vol. 38, no. 9, pp. 28-30.
5. Sugeno Michio; "A Fuzzy Logic Based Approach to Qualitive Modeling", IEEE trans. On Fuzzy Systems, vol 1, no. 1, feb 1993, pp. 7-31.
6. Bezdek, Jim, "Fuzzy Models –what are they and why?", IEEE trans. On Fuzzy Systems, vol. 1, no.1, feb. 1993. pp. 1-5.
7. Kosko, Bart, "Neural Networks and Fuzzy Systems", Prentice Hall, 1992, U.S.A., pp. 262-295.
8. Russell Stuart and Norving Peter,"Inteligencia Artificial: un enfoque moderno", Prentice Hall, México, D.F., 1996, pp. 439-461.
9. Kosko, Bart, "Fuzzy Engineering", Prentice Hall, 1997,U.S.A.



CAPITULO II. TEORÍA DE DETECTABILIDAD DE SEÑALES.

2.1. – INTRODUCCIÓN.

La Teoría clásica de detección de señales inicia en los años 40's con la implementación de experimentos de percepción de señales electromagnéticas en presencia de ruido. Los resultados de estos experimentos se analizaron como un problema de tipo estadístico en donde señales de ruido únicamente son identificadas como las hipótesis nulas, H_0 , mientras que señales con estímulo físico y ruido representaron las hipótesis alternativas, H_1 [1]. Bajo este contexto, los errores de tipo I son denotados como “falsa-alarma” y los errores de tipo II como “aciertos”.

La base de la psicofísica clásica es la mínima energía de la señal que un observador puede detectar en una plataforma ruidosa. De acuerdo con la teoría de detección, el problema de los observadores radica en decidir cuándo un evento sensorial fue causado por una señal o por un proceso aleatorio tal como una plataforma ruidosa [2]. Un ejemplo de este tipo de problemas es el experimento *si-no*, en el cual el observador debe decidir en cada presentación si el estímulo ha ocurrido debido a señal más ruido o si solamente ha ocurrido debido a ruido únicamente. Una variante sobre este tipo de experimentos, que se discutirá ampliamente en el presente trabajo, es el ROC (Receiver Operating Characteristic) por categorías. En este caso, se presentan también estímulos a los observadores para que decida si han ocurrido señales debidas a señal más ruido o a ruido únicamente; ahora los observadores pueden seleccionar una de varias respuestas. Por ejemplo, se puede instruir al observador para manejar una escala de seis puntos siendo el 1 presencia total de señal y el 6 presencia total de ruido únicamente. Una tercera variante de los experimentos de psicofísica son los experimentos de *selección forzada*, en los cuales la señal esta presente en uno de dos o más intervalos (temporales, espaciales) y el observador es instruido para seleccionar el intervalo que él cree estar más seguro de que contiene la señal (o estímulo).

La comparación de los resultados obtenidos a partir de diferentes procedimientos es extremadamente importante debido a que tales comparaciones proveen la principal prueba de validación del análisis de la teoría de decisión. Si el análisis es valido, podemos obtener



medidas de detectabilidad de una señal que son independientes del procedimiento usado para estimar dichas medidas. Entonces, este análisis demuestra la posibilidad de relaciones psicofísicas que son independientes del procedimiento[2].

En los 3 tipos de experimentos citados anteriormente se asume que los observadores son capaces de distinguir una cierta cantidad de eventos sensoriales. Presumiblemente, los observadores desarrollan esta habilidad por experiencia con las señales y el ruido usados en un experimento en particular, por ejemplo, en experimentos que manejan imágenes médicas la señal puede ser un tumor en el cerebro o en cualquier parte del cuerpo y el ruido será el sumado por la modalidad con que se capturan dichas imágenes y las estructuras anatómicas presentes. Si el objetivo es comparar sistemas o algoritmos de procesamiento digital de imágenes el análisis es directo. Si una técnica de PDI (Procesamiento Digital de Imágenes) A tiene una curva ROC superior o el porcentaje de respuestas correctas mas alto se considera que dicha técnica A es mejor que cualquier otra técnica de PDI con una curva ROC inferior o un menor porcentaje de respuestas correctas[4].

Actualmente, experimentos de detección de señales visuales son realizados usando imágenes con ruido únicamente (paciente normal en el caso de imágenes médicas) e imágenes con señal más ruido (paciente anormal en el caso de imágenes médicas). La forma más popular de los experimentos de psicofísica en muchas áreas de aplicación es el método ROC por categorías. En este, una señal específica puede o no estar presente y el observador usa una escala de categorías para expresar su grado de confianza de que la señal esta ausente o presente. El método ROC por categorías es el más ampliamente usado hoy en día y utiliza como medida de detectabilidad el área bajo la curva ROC, A_z , la cual puede ser transformada en el índice de detectabilidad d' . Ambos A_z y d' son medidas de sensibilidad de los observadores respecto a la señal utilizada y bajo algún criterio de evaluación determinado por el experimentador[5].

Otra técnica muy popular para determinar una medida de la desempeño de un observador es el método 2AFC (two Alternative Forced Choice) el cual proporciona una medida directa de la sensibilidad del observador debido a que éste procedimiento experimental produce una medida promedio de todos los criterios de decisión posibles. La generalización del 2AFC es el MAFC (Multiple Alternative Forced Choice), el cual maneja M



imágenes alternativas (o porciones de imágenes). El término “selección forzada” implica que el observador es forzado a seleccionar una de las opciones que se le muestran simultáneamente en cada presentación. La diferencia entre el MAFC y el ROC radica en que en el MAFC la señal esta siempre presente y el observador debe seleccionar una de M imágenes alternativas, en cambio en el ROC se muestra al observador una sola imagen y la señal puede o no estar presente en cada prueba[5].

Estos métodos nos permiten determinar la habilidad del sistema visual humano de distinguir y detectar objetos en un ambiente ruidoso. Muchos experimentos se han realizado utilizando los métodos antes mencionados, Peterson y Birdsall[11] presentaron un modelo matemático para la detección de señales en presencia de ruido. Egan[12], con la orientación teórica de Peterson y Birdsall investigó una gran variedad de tareas auditivas. Tanner y Swets[11] han reportado resultados acerca de la detección de haces de luz superpuestos sobre una plataforma de iluminación constante.

El objetivo de este capítulo es presentar los métodos ROC y MAFC con la teoría y métodos gráficos necesarios para poder utilizarse en la evaluación del desempeño de observadores en tareas de detección de señales o en la evaluación de la calidad de imágenes digitales. Se hace un especial énfasis en el método ROC por categorías por ser el método utilizado en los experimentos base de este trabajo, además de ser el método más ampliamente utilizado hoy en día en diversas áreas de aplicación.

2.2. – LOS MÉTODOS DE SELECCIÓN FORZADA.

Es posible estudiar algunos aspectos de el proceso sensorial utilizando los métodos de selección forzada como procedimientos psicofísicos. La principal característica de estos métodos es que requieren mas de un intervalo de observación antes de obtener la respuesta de los observadores[2]. Las alternativas pueden ser definidas por regiones de espacio, pero es muy común que se definan también por intervalos de tiempo.



2.2.1. – EL METODO 2AFC.

La forma más simple de experimentos de selección forzada son los experimentos de detección de dos alternativas. En este caso se le presentan al observador dos alternativas, visuales o auditivas según sea el caso. En una alternativa se presenta señal más ruido y en la otra únicamente ruido. La señal (estímulo de interés) es aleatoriamente asignada a una o a otra de las alternativas con igual probabilidad de ocurrencia. El observador es entonces cuestionado para indicar cuál de las dos alternativas contiene la señal. A partir de las respuestas del observador, el experimentador determina la proporción de respuestas correctas $P(c)$ para cada alternativa y entonces obtiene el índice de detectabilidad d' .

En caso de que el experimentador este interesado en evaluar la habilidad de un observador para discriminar entre dos señales de diferente amplitud, señales con diferente forma o señales con diferente frecuencia son utilizadas. Entonces, en el caso de los experimentos de discriminación 2AFC, se le presentan al observador dos alternativas, cada una de las cuales contiene una señal asignada aleatoriamente y el observador es cuestionado para seleccionar la alternativa que contiene la señal en cuestión. La medida de desempeño se calcula nuevamente con los porcentajes de correctas para cada alternativa[5].

La Fig. 2.1 muestra una secuencia típica de eventos para un experimento 2AFC. La señal siempre se presenta en alguno de los dos intervalos y el observador es forzado a seleccionar alguno de estos intervalos de observación. Normalmente se instruye al observador para seleccionar el intervalo que le genere un mayor grado de certidumbre de que la señal en cuestión esta presente. El intervalo de retroalimentación normalmente se da como un sonido de error o de acierto para indicar al observador si la respuesta dada ha sido errónea o acertada, esto le permite evaluar cada una de sus respuestas y poder mejorar su porcentaje de correctas.



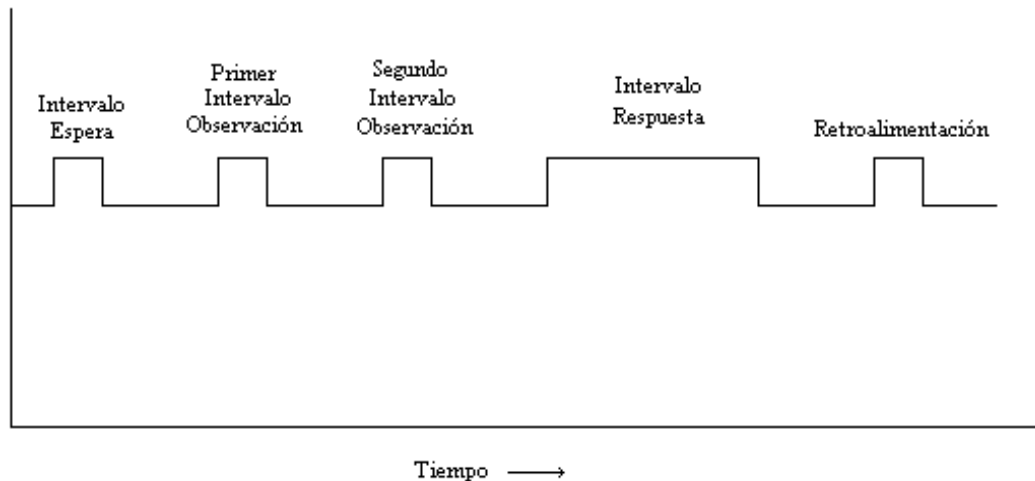


Fig. 2.1. – Intervalos en un experimento 2AFC.

2.2.2. – EL METODO MAFC.

El método MAFC usualmente emplea una señal con M ubicaciones alternativas. La señal está definitivamente presente en una ubicación y las otras $M-1$ ubicaciones contienen únicamente ruido. La ubicación de la señal es asignada aleatoriamente y todas las ubicaciones tienen la misma probabilidad de contener la señal. La secuencia típica en las pruebas MAFC es similar a la del 2AFC, la diferencia es que ahora se presentarán M ($M > 2$) intervalos de observación antes de que el observador de su respuesta por algún medio (normalmente teclado o mouse de una PC en experimentos de percepción de señales visuales).

Los experimentos MAFC se implementan utilizando M señales alternativas de tal forma que todas tengan la misma SNR (Signal Noise Ratio) y sean igualmente detectables por observadores humanos. Existe una ubicación en la cual la señal está definitivamente presente y la tarea del observador es determinar la más probable de todo el menú de opciones que se le presentan. El experimentador registra el porcentaje de correctas, $P(c)$, para cada alternativa y entonces calcula el índice de detectabilidad[5].



2.2.3. – CÁLCULO DEL ÍNDICE DE DESEMPEÑO O DE DETECTABILIDAD.

Asumiendo que los eventos sensoriales que forzan al observador a establecer su criterio de decisión tienen una distribución de probabilidad Gaussiana con desviación estándar unitaria. La relación entre las proporciones de respuestas correctas $P(c)$, y los correspondientes valores normalizados z esta dada por la integral de la distribución normal acumulativa[5].

$$P(z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp(-x^2 / 2) dx \quad 2.1$$

y la relación inversa esta dada por:

$$z(P(c)) = \Phi^{-1}(P(c)) \quad 2.2$$

Estas relaciones pueden describirse por aproximaciones algebraicas polinomiales pero normalmente son presentadas en forma de tablas. Algunas veces estas relaciones son presentadas en términos de la función de error $\text{erf}(z)$ pero es necesario evitar ambigüedades debido a las diferentes definiciones de la función de error. Algunos ejemplos de relaciones $(P(c),z)$ son: (0.3, -0.524), (0.5,0.0), (0.7,0.524), (0.9,1.282), (0.95,1.645). Note que estos pares son simétricos alrededor de $P=0.5$ por medio de:

$$z(P(c)) = -z(1 - P(c)) \quad 2.3$$

2.2.3.1. – CALCULO DE d' EN EXPERIMENTOS 2AFC.

Han sido propuestos diferentes métodos para el cálculo de d' usando experimentos 2AFC. La selección de proporción de respuestas a usar en el análisis 2AFC es arbitraria. McMillan y Creelman definen H como la proporción de respuestas correctas cuando la señal se encuentra en el primer intervalo y F como la proporción de respuestas incorrectas cuando la señal se presenta en el segundo intervalo. También definen a N_L como el número de



presentaciones en las cuales la señal esta presente en el primer intervalo y a N_R como en número de presentaciones en las cuales la señal esta presente en el segundo intervalo.

En el contexto del método 2AFC con igual probabilidad el observador deberá favorecer sistemáticamente un intervalo de observación, ya sea el primero o el segundo intervalo. Se asume entonces que este tipo de observadores tendrán una eficiencia reducida.

Si tenemos un experimento 2AFC en donde $N_H = N_L$, el índice de detectabilidad se obtiene a partir de la expresión 2.4.

$$d'_{zu} = \sqrt{2}[z(P_L) + z(P_R)]/2 \quad 2.4$$

Siendo P_L la proporción de respuestas correctas en el primer intervalo y P_R la proporción de respuestas correctas en el segundo intervalo.

Si N_R es forzada a ser diferente de N_L , es decir $N_R > N_L$ o $N_R < N_L$, será más razonable utilizar pesos promedio de valores z definidos como:

$$w_R = N_R / (N_L + N_R) \quad 2.5$$

$$w_L = N_L / (N_L + N_R) \quad 2.6$$

a partir de estas expresiones podemos obtener el índice de detectabilidad utilizando la expresión 2.7.

$$d'_{zv} = \sqrt{2}[w_L z(P_L) + w_R z(P_R)]/2 \quad 2.7$$

La forma más común es no tomando en cuenta una relación entre N_R y N_L y obtener una proporción promedio de respuestas correctas a través de la expresión 2.8.

$$P_w = (N_{LC} + N_{RC}) / (N_L + N_R) \quad 2.8$$

donde N_{LC} y N_{RC} son las respuestas correctas para el primero y segundo intervalo de observación, el índice de detectabilidad esta dado por:



$$d'_{pw} = \sqrt{2}z(P_w) \quad 2.9$$

hay quienes prefieren utilizar un peso promedio de las proporciones de respuestas correctas de la forma como lo muestra la ecuación 2.10 y obtener el índice de detectabilidad a partir de la expresión 2.11.

$$P_u = (P_R + P_L) / 2 \quad 2.10$$

$$d'_{pu} = \sqrt{2}z(P_u) \quad 2.11$$

Birdsall y Peterson (1954) [10] desarrollaron una tabla similar a la Tabla I, en la cual plasman algunos de los valores de d' que corresponden a un valor de proporción de respuestas correctas. Dicha tabla se aplica para $M=2$ (2AFC) y para valores de M de 4,8,16,32,256,1000. Cada punto de intersección entre $P(c)$ y M representa el valor del índice de detectabilidad d' .

P(c)	M=2	M=4	M=8	M=16	M=32	M=256	M=1000
.51	.04	.84	1.37	1.76	2.12	2.93	3.23
.52	.07	.88	1.40	1.79	2.15	2.96	3.26
.53	.11	.90	1.43	1.82	2.18	2.98	3.28
.54	.14	.94	1.46	1.85	2.20	3.01	3.31
.55	.18	.97	1.50	1.88	2.24	3.04	3.34
.56	.21	1.00	1.52	1.90	2.26	3.07	3.36
.57	.25	1.03	1.56	1.94	2.30	3.10	3.39
.58	.28	1.06	1.58	1.96	2.32	3.12	3.41
.59	.32	1.09	1.61	2.00	2.35	3.15	3.44
.6	.36	1.12	1.64	2.02	2.38	3.18	3.47

Tabla I. – Valores de d' para selección forzada con M alternativas ortogonales.

Elliot[10] también desarrolló unas tablas similares a la Tabla I utilizando la ecuación 2.12 (dada mas adelante) las cuales generalizan los experimentos de selección forzada para M alternativas.

2.2.3.2. – CALCULO DE d' EN EXPERIMENTOS MAFC.

Como se mencionó anteriormente en éste tipo de experimentos una señal puede estar presente en una de M ubicaciones posibles. Esta ubicación es seleccionada de tal manera que las decisiones alternativas sean estadísticamente independientes. En este caso, la metodología



comúnmente utilizada para la obtener el índice de detectabilidad d' se realiza tomando en cuenta los siguientes puntos:

- a) La distribución de probabilidad de las variables (eventos sensoriales) que afectan en proceso de decisión del observador son gaussianas.
- b) Todas las varianzas son iguales.
- c) Cada alternativa tiene la misma probabilidad a priori.
- d) Todas las alternativas son estadísticamente independientes.
- e) Comúnmente los componentes del ruido tienen media cero.

La señal presente en una alternativa en particular incrementará el valor de tal componente y el observador seleccionará la hipótesis correspondiente a la alternativa con más alto valor, es decir, ninguna alternativa debe ser favorecida sobre ninguna de las alternativas restantes.

La siguiente ecuación fue utilizada por Elliot[2] para relacionar la proporción de correctas con un valor de d' para diferentes valores de M.

$$p = 1 - (M - 1) \int_{-\infty}^{\infty} \Phi(t)^{M-2} \Phi(t - d') G(t) dt \quad 2.12$$

Donde:

$$G(t) = \frac{1}{\sqrt{2\pi}} \exp(-t^2 / 2) \quad 2.13$$

Y

$$\Phi(t) = \int_{-\infty}^t G(x) dx \quad 2.14$$

En el análisis de resultados experimentales, la proporción observada de respuestas correctas, $P(c)$, es usada como una estimación de la probabilidad de respuestas correctas[5].



Unos cuantos valores obtenidos a partir de la expresión 2.12 y para diferentes valores de M y $P(c)$ son presentados en la Tabla I.

2.3. – LA CURVA ROC.

La forma de una curva ROC corresponde a un grado particular de detección o de discriminación, mientras que la ubicación de un punto en particular a lo largo de la curva representa un grado específico de desempeño por parte del observador. Esta curva ROC permite localizar la ubicación del mejor criterio de evaluación por parte del experimentador.

La gráfica ROC fue diseñada en el contexto de la teoría de detectabilidad de señales por Peterson, Birdsall y Fon en 1954 para proveer un índice de exactitud consistente con su modelo básico de detección de señales. Se considera que las alternativas de señal-más-ruido (sn) y ruido únicamente (n) pueden ser tratadas como hipótesis estadísticas[1].

La teoría de decisión estadística es la base para un modelo que provee un índice que es independiente de la probabilidad de ocurrencia de las 2 alternativas (s y $1-s$) y de la tendencia discriminatoria a favor de una de las alternativas.

El modelo de teoría de detección (bajo el supuesto gaussiano) se ilustra en la Fig. 2.2. El eje horizontal corresponde a la variable de decisión de la teoría estadística: La variable x es una medida de magnitud de observación, o la magnitud de una muestra estadística. El eje vertical corresponde a la densidad de probabilidad. La distribución de la izquierda es la distribución de probabilidad de x para n , análoga a la hipótesis nula H_0 . La distribución de la derecha es la distribución de probabilidad para sn , análoga a la hipótesis bajo prueba H_1 . El grado de traslape de las dos distribuciones determina una confusión entre las dos alternativas, reflejada en la Fig. 2.2 por la diferencia entre las medias de las distribuciones y denotada por θ .



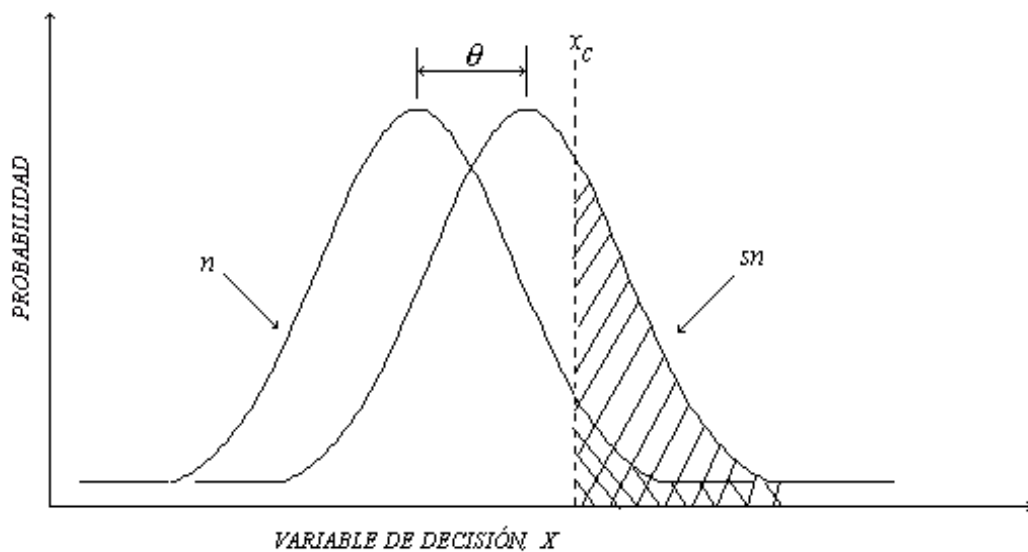


Fig. 2.2. – Modelo Gaussiano de la Teoría de Detección.

Un valor crítico de x , o del criterio de decisión (x_c), separa el valor de x que lleva a seleccionar sn ($x \geq x_c$) de aquellos valores que llevan a seleccionar n ($x < x_c$).

El área bajo la distribución sn a la derecha de x_c es igual a la probabilidad h ($P(S/s)$) (Probabilidad de que el observador dice que existe señal cuando la señal esta presente), y el área bajo la distribución n a la derecha de x_c es la probabilidad f ($P(S/n)$) (Probabilidad de que el observador determina que existe señal cuando se presenta únicamente ruido). La ROC es trazada de izquierda a derecha y de abajo hacia arriba sobre la gráfica h vs. f conforme x_c se mueve de derecha a izquierda como se muestra en la Fig. 2.3.



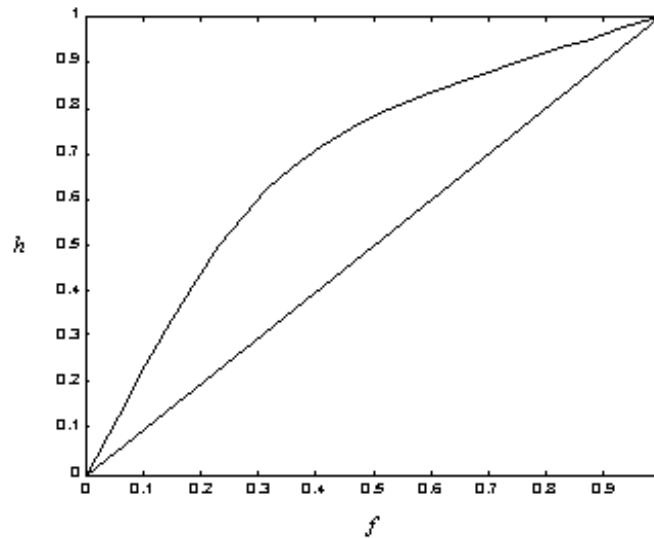


Fig. 2.3. – Una ROC ilustrativa, mostrando probabilidad de aciertos (h) como función de la probabilidad de falsa alarma (f).

2.3.1. – USO DE LA CURVA ROC.

Suponiendo el ejemplo de un observador que por medio de un experimento obtuvo una probabilidad de 0.1 al responder S (si existe señal) cuando n (únicamente ruido) estuvo presente, y una probabilidad de 0.45 responder S cuando sn (señal mas ruido) estuvo presente. Este punto sobre la gráfica ROC tiene coordenadas 0.1 y 0.45 (punto A de Fig. 2.4). Este punto es obtenido sobre la base de 100 observaciones de n y 100 observaciones de sn .

Supóngase que el observador es inducido a cambiar su criterio de decisión en alguna forma. Por ejemplo, él puede incrementar el número de respuestas “sí” si él está más fuertemente obligado a decir “sí” mediante un mensaje como “Presione “sí” cuando detecte el mínimo indicio de que existe señal”. Presumiblemente ambos valores de las coordenadas incrementarán. Esto es, el observador puede incrementar su probabilidad de acierto ($P(S/s)$) de 0.45 a 0.80 e incrementar su probabilidad de falsa alarma ($P(S/n)$) de 0.10 a 0.30. Este será un nuevo punto sobre la gráfica ROC con coordenadas 0.80 y 0.30 (punto B de Fig. 2.4). Y así sucesivamente cambiando las instrucciones de los experimentos podemos generar un punto sobre la gráfica ROC por cada conjunto de instrucciones. Estos puntos pueden ser conectados por una curva la cual representará los diferentes modos de observación que el observador puede adoptar bajo condiciones de estímulo fijas. La curva entera es generada sin cambiar algún parámetro físico de los estímulos. El nivel de la señal y el ruido son los mismos para todos los puntos, únicamente cambia la forma de selección del observador para decidir



alguna alternativa. Note que las otras proporciones de probabilidad ($P(N/s)$ y $P(N/n)$) no son utilizadas en la curva ROC por ser los complementos de $P(S/s)$ y $P(S/n)$ respectivamente.

A partir de la Fig. 2.4 podemos observar la forma en que se obtiene el área bajo la curva ROC (A_z) calculando geoméricamente el área bajo la curva que une los puntos A y B. Note que el rango de A_z será de 0.5 (peor desempeño) a 1.0 (el mejor desempeño posible). La pendiente en cada punto de la ROC toma diferentes valores a partir de esquina inferior izquierda hasta la esquina superior derecha, usualmente el valor de la pendiente en algún punto dado de la curva se toma como un índice de desempeño de tal punto que indica un criterio de decisión específico.

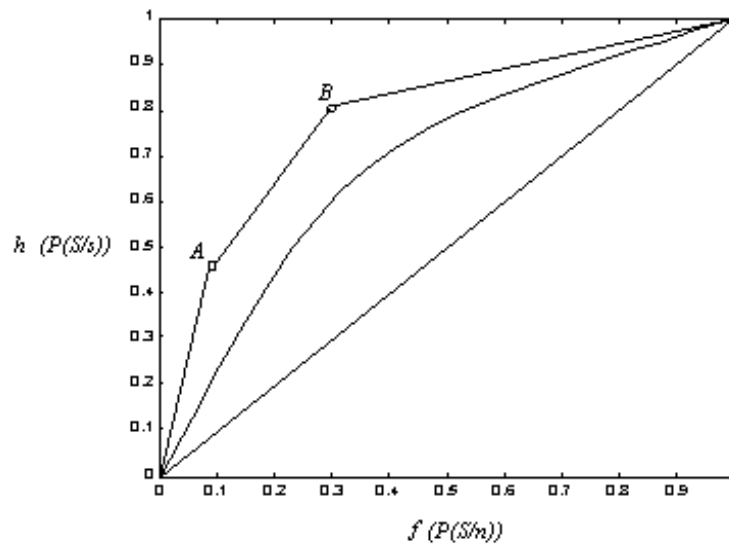


Fig. 2.4. – Ejemplo de gráfica ROC.

La mayoría de los experimentos ROC se realizan utilizando N categorías dentro de las cuales el observador clasifica la forma en que determinados estímulos visuales o auditivos. A partir de estas N categorías se podrán obtener $N-1$ puntos sobre la gráfica ROC y así poder determinar A_z para con la ecuación 2.15 poder determinar el índice de detectabilidad d' .

$$d' = 2 \operatorname{erf} \operatorname{inv}(2A_z - 1) \quad 2.15$$



donde $erfinv()$ es el inverso de la función de error $erf()$, es decir, $x = erfinv(y)$ satisface $y = erf(x)$. La cantidad d' es interpretada como la separación de las dos distribuciones Gaussianas de varianza unitaria correspondientes a los casos de acierto y de falsa alarma.

A continuación se describe la manera en la que operan los métodos que hacen uso de la curva ROC. Se inicia por el método más simple (ROC si-no) indicando cómo se forma la curva ROC con este tipo de experimentos y cómo se obtiene un índice de desempeño o de detectabilidad. Posteriormente, se describe el método utilizado en los experimentos de este trabajo de tesis que es el ROC por categorías, de igual manera se menciona la manera en que se traza la curva ROC y la forma de obtener un índice de detectabilidad a partir de esta curva.

2.3.2. – EL MÉTODO ROC SI-NO.

En un experimento ROC si-no se le presenta al observador una de dos alternativas mutuamente exclusivas. El observador es cuestionado para responder con una de las dos alternativas permisibles. La Fig. 2.5 muestra una secuencia típica de eventos dentro de una tarea de detección binaria. Típicamente, los eventos son marcados por luces en una tarea auditiva y por sonidos en una tarea visual[2].

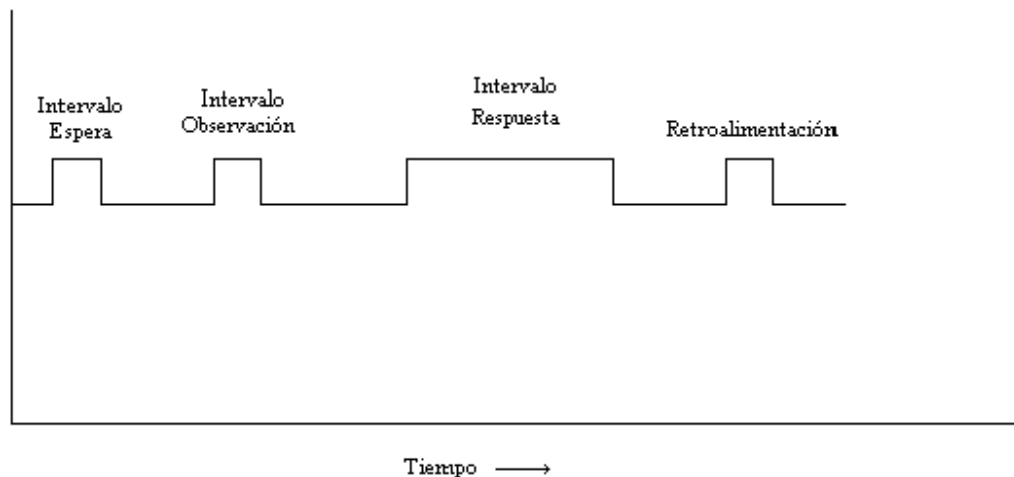


Fig. 2.5. – Intervalos en un experimento ROC si-no.

Un intervalo importante de la secuencia es el *intervalo de observación*, dentro de éste intervalo, el cual es claramente marcado por el observador, uno de los dos posibles estímulos ocurren, esto permite realizar un análisis mucho más simple del método.



En muchos experimentos la plataforma de ruido esta continuamente presente y se le suma la señal para tener el caso de sn . Cuando existe únicamente ruido tenemos el caso n . El observador realiza su respuesta durante un *intervalo de respuesta*. Esta respuesta es la decisión del observador de que una u otra hipótesis es verdadera. Esto es, el observador debe aceptar sn o n . Sus respuestas son denotadas por las letras mayúsculas S y N . Ninguna otra respuesta es permitida en este tipo de experimentos de detección simple. Específicamente, el observador no puede responder “no se”.

Algunas veces, en experimentos de psicofísica, el observador recibe un *intervalo de retroalimentación*; esto es, se le indica (mediante una señal visual o acústica) cuándo su respuesta es correcta o incorrecta. En muy raras ocasiones se le da al observador un *intervalo de espera*, esto es con la finalidad de que él conozca cuándo llegará el intervalo de observación. Debido a lo extenso que son los experimentos de detectabilidad de señales, casi siempre es preferible omitir el intervalo de espera al observador para reducir el tiempo del experimento y así evitar el observador se fatigue y proporcione respuestas irreales[2].

Los experimentos si-no se pueden diseñar de tal forma que el observador varíe su criterio de decisión sobre cada bloque de pruebas en un solo experimento o realizar varios experimentos con diferentes criterios de evaluación para de esta manera poder obtener n puntos sobre la ROC a partir de n experimentos si-no.

MATRIZ DE ESTÍMULO-RESPUESTA.

Debido a que en cada prueba en la tarea de si-no existen dos posibles estímulos y dos posibles respuestas, cada prueba puede ser representada por la intersección de uno de los estímulos alternativos y una de las respuestas alternativas, como se indica en la Fig. 2.6. Algunas de las respuestas S (*existe señal*) o N (*no existe señal*) deberán ser seleccionadas tal que la conjunción de un estímulo y una respuesta podrá ser representada por una de las celdas de la matriz estímulo-respuesta (Fig. 2.6). Note que se ignoran los efectos presentación por presentación y ponemos nuestra atención al promedio obtenido sobre un número determinado de presentaciones. Las presentaciones individuales son promediadas y estimadas para poder ser representadas por medio de cuatro probabilidades.



		Respuestas Alternativas	
		S	N
Estímulos Alternativos	s	$P(S/s)$ Acierto	$P(N/s)$ Error
	n	$P(S/n)$ Falsa Alarma	$P(N/n)$ Rechazo Correcto

$P(S/s) + P(N/s) = 1$
 $P(S/n) + P(N/n) = 1$

Fig. 2.6. – Matriz estímulo-respuesta del procedimiento si-no.

Las entradas para cada celda son probabilidades condicionales. Es decir, $P(S/s)$ es la probabilidad de que un estímulo s (*señal*) invocará una respuesta S , $P(S/n)$ es la probabilidad de que el observador detectó señal cuando no existió señal, de manera análoga se interpretan $P(N/s)$ y $P(N/n)$. Por ejemplo, si el experimentador muestra al observador 100 presentaciones de n y el observador selecciona N 90 veces, el observador debe seleccionar S 10 veces. Entonces, nuestro estimado de $P(N/n)$ es 0.90 y nuestro estimado de $P(S/n)$ es 0.10. Similarmente, si el estímulo s es presentado 100 veces y el observador selecciona la respuesta alternativa N 55 veces, entonces $P(S/s) = 0.45$ y $P(N/s) = 0.55$.

Únicamente dos números pueden introducirse libremente a la matriz, los otros dos números se determinan automáticamente debido a que cada renglón de la matriz debe sumar uno. La matriz tiene entonces dos grados de libertad y no cuatro como el número de celdas lo sugiere.

En los experimentos si-no el observador es cuestionado para variar la proporción de aciertos y de falsa alarma, realizando experimentos con niveles de energía de la señal diferentes. Obviamente su reporte no es el reflejo total de su experiencia sensorial por haber evaluado únicamente un solo nivel de detección. Por lo tanto, al final del primer experimento, la curva ROC consistirá únicamente de un solo punto. Específicamente, la proporción de aciertos $P(S/s)$ debería ser la probabilidad de un estado detectado (D) dada la presentación de



una señal, esto es, $P(D/s)$; y la proporción de falsa alarma $P(S/n)$ debería ser la probabilidad de un estado detectado dada la presentación de ruido únicamente, esto es, $P(D/n)$. El sujeto debe entonces estar adivinando, o cambiando algunos de sus reportes para contradecir su experiencia sensorial y para cumplir las instrucciones experimentales. El punto importante es que *el mecanismo de adivinanza no puede operar sobre alguna base sensorial*. Entonces, asumimos que el mecanismo de adivinanza es independiente del mecanismo *sensorial*.

Partiremos de usar mecanismos geométricos para analizar la curva ROC referente a un solo punto generado por un experimento si-no. Considere las densidades de probabilidad mostradas en la Fig. 2.7a, ambas son uniformes en el intervalo 0 a 1 y también uniformes en el intervalo 1 a 2, específicamente, si la hipótesis n es verdadera:

$$f(x) = 1 - q(n) \quad \text{si} \quad 0 \leq x \leq 1 \quad 2.16$$

$$f(x) = q(n) \quad \text{si} \quad 1 < x \leq 2 \quad 2.17$$

y si la hipótesis s es verdadera:

$$f(x) = 1 - q(s) \quad \text{si} \quad 0 \leq x \leq 1 \quad 2.18$$

$$f(x) = q(s) \quad \text{si} \quad 1 < x \leq 2 \quad 2.19$$

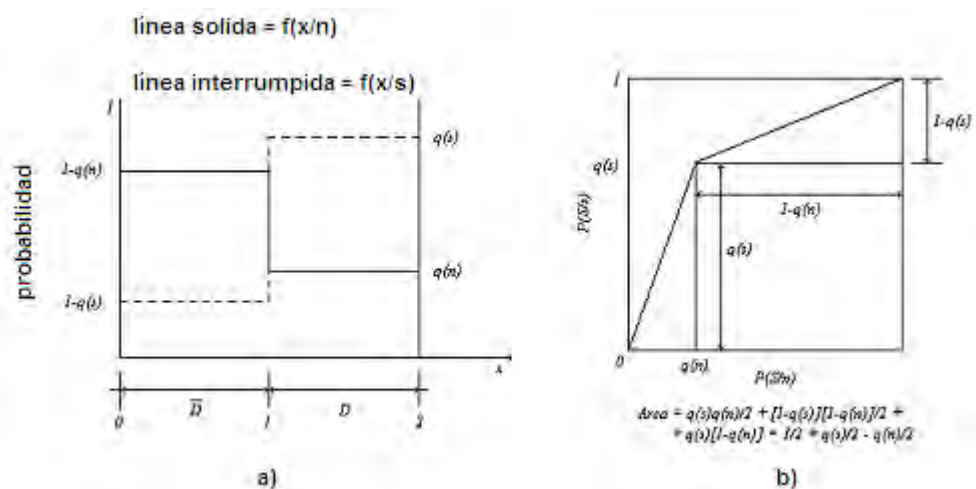


Fig. 2.7. – Interpretación de curva ROC. a) Estados detectados de acuerdo con los valores de x . b) El área bajo la curva ROC es calculada bajo la base triángulos y rectángulos.

Dado algún estímulo, una de las dos razones de posibilidad resultan: ya sea $q(s)/q(n)$ o $[1-q(s)]/[1-q(n)]$. Asociamos con la razón de posibilidad mas alta a D y con la razón de posibilidad más baja a \bar{D} . La variable x puede tomar algún valor sobre el intervalo 0 a 2, pero únicamente dos razones de verosimilitud pueden resultar.

2.3.2.1. – CAMBIANDO EL CRITERIO DE DECISIÓN.

La forma más simple y más directa en el análisis de la teoría de decisión de la prueba si-no es trazar una curva ROC empírica para una determinada señal (y una determinada plataforma de ruido) con una cierta razón señal a ruido.

Para obtener esta curva, el observador debe cambiar su criterio de decisión de un conjunto de pruebas a otro, esto producirá diferentes puntos en el cuadrado unitario teniendo como coordenadas las proporciones de aciertos $P(S/s)$ y las proporciones de falsa alarma $P(S/n)$.

El observador puede ser inducido a cambiar su criterio de decisión de diversas maneras. Simplemente se le instruye para adoptar un criterio “estricto” o “medio” o un criterio “laxo”. Estas instrucciones pueden ser resaltadas aún mas por medio de instrucciones verbales como

“presione la tecla si cuando usted este muy ciertamente seguro de que la señal se ha presentado” para un criterio estricto. Para un criterio “laxo” podrá tener como instrucción lo siguiente *“presione la tecla si cada vez que tenga una ligera indicación de que la señal se ha presentado”*. Dentro de esta serie de instrucciones también podemos considerar criterios como “medio estricto” y “medio laxo” y darles un significado por medio de alguna instrucción verbal[2].



2.3.3. – EL MÉTODO ROC POR CATEGORÍAS.

En experimentos ROC por categorías se requiere que el observador mantenga diferentes criterios de decisión simultáneamente en un mismo experimento. El procedimiento por categorías usa la misma presentación que el procedimiento si-no y presenta los mismo eventos de prueba mostrados en la Fig. 2.5. La secuencia de eventos ocurre de la misma manera que en el experimento si-no, lo que difiere es la naturaleza de las respuestas del observador ya que en el procedimiento por categorías, teóricamente hablando, al menos una respuesta puede estar disponible. Asumiendo que seis categorías de respuesta son empleadas por el observador, la primer categoría representará la total certidumbre de que la señal ha ocurrido y la sexta categoría representará la total certidumbre de que la señal no se ha presentado en el intervalo de observación. Es decir, las categorías entre la primera y la sexta presentan pequeños grados de certidumbre de la ocurrencia o no ocurrencia de s . La Fig. 2.8 muestra la matriz de estímulo-respuesta del procedimiento por categorías (seis categorías) formada por las probabilidades de acierto y falsa alarma ($P(S/s)$ y $P(S/n)$) para cada categoría.

Una forma de tratar los datos de la Fig. 2.8a es considerando el método por categorías como un método si-no. Esto es, alguna prueba sobre la cual el observador uso la primer categoría es tratada como si el observador hubiera dicho “si hay señal”, y todas las otras pruebas sobre las cuales las cinco categorías fueron usadas son tratadas como si el observador hubiera dicho “no hay señal”. Esto permite al experimentador obtener $P(S/s)$ y $P(S/n)$, con esta estimación se define un punto sobre la curva ROC que corresponde a un criterio de evaluación “muy estricto”.

El criterio de evaluación puede cambiar si ahora el experimentador asume que el observador debió decir “si hay señal” para las dos primeras categorías. Esta suposición generará un segundo punto sobre la gráfica con coordenadas $P(S/s)$ y $P(S/n)$. En la Fig. 2.8 se ilustra la reducción de la matriz original asumiendo que se consideraron 6 categorías de respuesta posibles y que en el criterio de evaluación el experimentador consideró que el observador debió considerar que “si hay señal” para las 3 primeras categorías y que "no hay señal" para las últimas 3 categorías.



Categorías

		1	2	3	4	5	6
Estímulos	s	$P(1/s)$	$P(2/s)$	$P(3/s)$	$P(4/s)$	$P(5/s)$	$P(6/s)$
	n	$P(1/n)$	$P(2/n)$	$P(3/n)$	$P(4/n)$	$P(5/n)$	$P(6/n)$

a)

Respuestas si-no asumidas

		S	N
Estímulos	s	$P(S/s) = P(1/s) + P(2/s) + P(3/s)$	
	n	$P(S/n) = P(1/n) + P(2/n) + P(3/n)$	

b)

Fig. 2. 8. – a) Matriz estímulo-respuesta para el procedimiento por categorías. b) Reducción al procedimiento si-no asumiendo las categorías 1,2 y 3 como respuesta “si” y las categorías 4,5 y 6 como respuesta “no”.

Procediendo de esta manera el experimentador puede calcular las probabilidades correspondientes para varios criterios y obtener puntos estimados sobre la gráfica ROC. Note que puntos sucesivos son fuertemente independientes. Además, el número de puntos sobre la ROC que se pueden obtener es el número de categorías menos uno. Si seis categorías son usadas, cinco puntos sobre la gráfica ROC pueden obtenerse. El último punto, la probabilidad de que una de las seis categorías fue usada bajo ciertas hipótesis es la unidad; esto es, el último punto sobre la gráfica ROC tiene coordenadas (1.0,1.0).

Si el observador pretende reducir un experimento ROC por categorías a un experimento ROC si-no a través de la matriz de estímulo-respuesta, entonces la forma de obtener el índice de detectabilidad es como sigue[5].

Con las proporciones $P(S/s)$ y $P(S/n)$ en coordenadas de desviación normal (valores z) el índice de detectabilidad esta dado por la ecuación 2.20.

$$d' = [z(P(S/s)) - z(P(S/n))] \quad 2.20$$



este valor de d' es interpretado como la distancia entre las medias de las curvas de distribución normal para las distribuciones de acierto y falsa alarma y es comúnmente usado como índice de desempeño.

La posición del punto correspondiente al par $(P(S/s), P(S/n))$ depende del criterio de decisión usado por el observador. Este criterio puede ser visto como una respuesta definida usando 2.21.

$$c = -[z(P(S/s)) + z(P(S/n))]/2 \quad 2.21$$

Alternativamente si se presenta el caso en que las distribuciones de probabilidad tengan diferente varianza de acuerdo a la razón $s = \sigma_n/\sigma_s$ para los casos de señal y ruido, respectivamente; el índice de detectabilidad esta dado por la ecuación 2.22.

$$d_a = [2/(1+s^2)]^{1/2} [z(P(S/s)) - sz(P(S/n))] \quad 2.22$$

Alternativamente Swets and Pickett[5] muestran que d_a puede ser calculado usando el estimado A_z del área bajo la curva ROC a través de la ecuación 2.23.

$$d_a = \sqrt{2}z(A_z) \quad 2.23$$

Para ilustrar los pasos necesarios para obtener la curva ROC considere los resultados de un observador (Tabla II) que utilizó seis categorías para determinar su grado de certidumbre de la presencia de una señal.

<i>Probabilidad aposteriori</i>	0.0-0.04	0.05-0.19	0.20-0.39	0.40-0.59	0.60-0.79	0.80-1.0
(1) $f(R_i)$	220	229	170	193	195	181
(2) $f(s \cdot R_i)$	46	57	66	101	154	173
$P(s/R_i)$	0.21	0.25	0.39	0.52	0.78	0.96
$P(R_i/s)$	0.07	0.09	0.11	0.18	0.26	0.29
$P(S/s)$	1.00	0.93	0.84	0.73	0.55	0.29
(3) $f(n \cdot R_i)$	174	172	104	92	41	8
$P(n/R_i)$	0.79	0.75	0.61	0.48	0.22	0.04
$P(R_i/n)$	0.29	0.29	0.18	0.16	0.07	0.01
$P(S/n)$	1.00	0.71	0.42	0.24	0.08	0.01

Tabla II. – Cálculos básicos para el método ROC por categorías. (2) Datos sobre 597 pruebas con señal. (3) Datos sobre 591 pruebas de ruido únicamente. (datos de Swets, Tanner y Birdsall, 1961)



Las columnas de la Tabla II son encabezadas por rangos de probabilidad a posteriori definidas para las seis categorías. La sección (1) contiene como único renglón una lista de frecuencias para las cuales cada categoría R_i fue empleada. La sección (2) muestra cuatro cantidades basadas sobre pruebas que contienen señal; y la sección (3) muestra las cantidades correspondientes basadas en pruebas que contienen únicamente ruido.

Antes de discutir la construcción de la gráfica ROC, se considera un cálculo que muestra mas directamente la habilidad del observador para usar las categorías que se le presentan. El primer renglón en la sección dos ($f(s \cdot R_i)$), muestra la frecuencia de ocurrencia de la señal por cada categoría R_i . El siguiente renglón ($P(s/R_i)$), muestra la proporción de observaciones colocadas en cada categoría dado que se presentó una señal. Por ejemplo, el dato 0.96, ubicado en la columna mas a la derecha, es obtenido dividiendo la frecuencia de ocurrencia por el total de observaciones en esta categoría. Se puede observar que la cantidad $P(s/R_i)$ se incrementa monotonicamente conforme incrementa la probabilidad a posteriori que define dicha categoría. Este resultado indica que el observador esta usando las categorías como en un principio se habían definido e indica la habilidad para clasificar las categorías de acuerdo a la ocurrencia de la señal. Esto nos deja ver que el observador es capaz de mantener cinco criterios de evaluación simultáneamente.

Los últimos dos renglones de las secciones (2) y (3) de la Tabla II muestran los cálculos para obtener las proporciones de acierto y de falsa alarma. El tercer renglón en cada sección, $P(R_i/s)$ y $P(R_i/n)$, es lo contrario de la probabilidad condicional considerada. Por ejemplo, el dato 0.29 de la columna mas a la derecha del tercer renglón en la sección (2) es obtenido dividiendo el número de observaciones con señal colocadas en cada categoría (173) por el total de observaciones con señal presentadas (597). En el cuarto renglón en cada sección, las proporciones de acierto y de falsa alarma corresponden a cada uno de los cinco criterios de decisión, es una acumulación de las cantidades del tercer renglón, moviéndose de derecha a izquierda sobre la Tabla II. Representa el procedimiento de compresión de datos por categorías a datos si-no considerando exitosamente cada una de las cinco categorías como se definen en el procedimiento si-no.



Con las proporciones de acierto y de falsa alarma se obtiene la curva ROC de la Fig. 2.9. Note que la variabilidad de los puntos respecto de la curva teórica es muy pequeña, este grado pequeño de variabilidad es típico en este tipo curvas. La pequeña variabilidad obtenida es debida a la forma en que fueron calculados los datos; el proceso acumulativo permite el incremento monotonico de la curva resultante. Estas pequeñas diferencias de variabilidad se obtienen siempre y cuando los datos por categorías se obtengan a partir de, por los menos, una cuarta parte de las observaciones requeridas para generar la curva ROC con procedimientos si-no.

INTERPRETACIÓN DE LA PENDIENTE EN UN PUNTO DE LA CURVA ROC.

Los múltiples criterios utilizados por el observador bajo un procedimiento por categorías podrían ser comparados con el criterio especificado por la teoría estadística. En el experimento anteriormente mencionado las probabilidades a priori de la señal y el ruido se pueden considerar iguales [$P(s)=P(n)=0.5$], la relación teórica entre una probabilidad a posteriori y la razón de verosimilitud esta dada por la ecuación 2.24.

$$P(s/x) = \frac{I(x)}{I(x)+1} \quad 2.24$$

Dado que el valor de $I(x)$ es igual a la pendiente de la curva ROC en un punto representando un criterio, el valor de $P(s/x)$ corresponde a cada criterio usado por un observador. Los valores obtenidos pueden ser comparados con los obtenidos en la Tabla II. Estos resultados se muestran en la Figura 2.10. Observamos que los datos obtenidos son consistentes con la relación establecida en la ecuación 2.24.



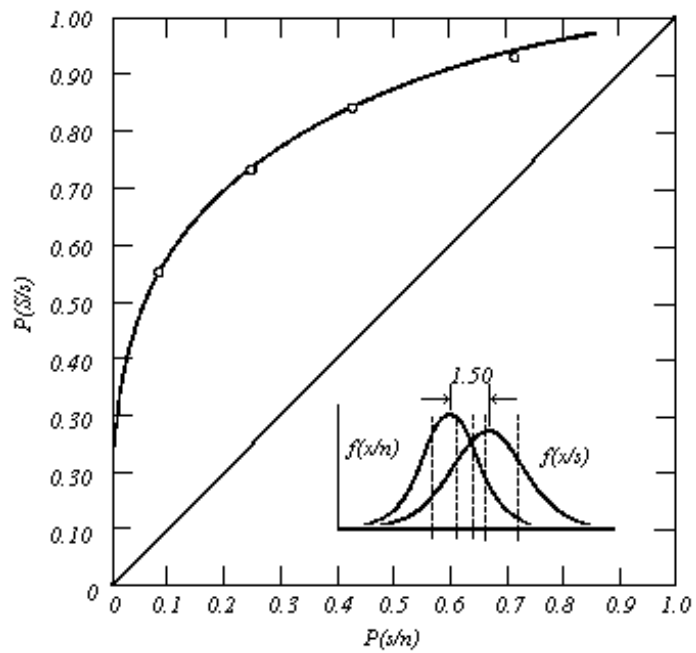


Fig. 2.9. – Ejemplo de curva ROC para un observador. 1188 observaciones fueron realizadas. (Datos de Swetes, Tanner y Birdsall, 1961)



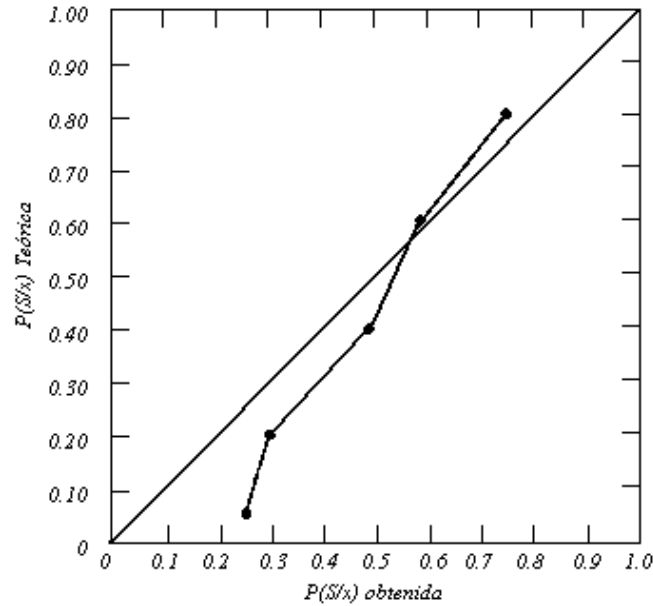


Fig. 2.10. – Comparación de criterio de decisión teórico y obtenido. (Datos de Swets, Tanner y Birdsall, 1961)

2.3.3.1. – LA CURVA ROC COMO UNA GRÁFICA BINORMAL.

La curva ROC se visualiza mejor como una gráfica “binormal” en la cual las coordenadas de probabilidad son escaladas de tal manera que sus correspondientes valores de desviación normal están linealmente espaciados como lo muestra la Fig. 2.11. Sobre cada

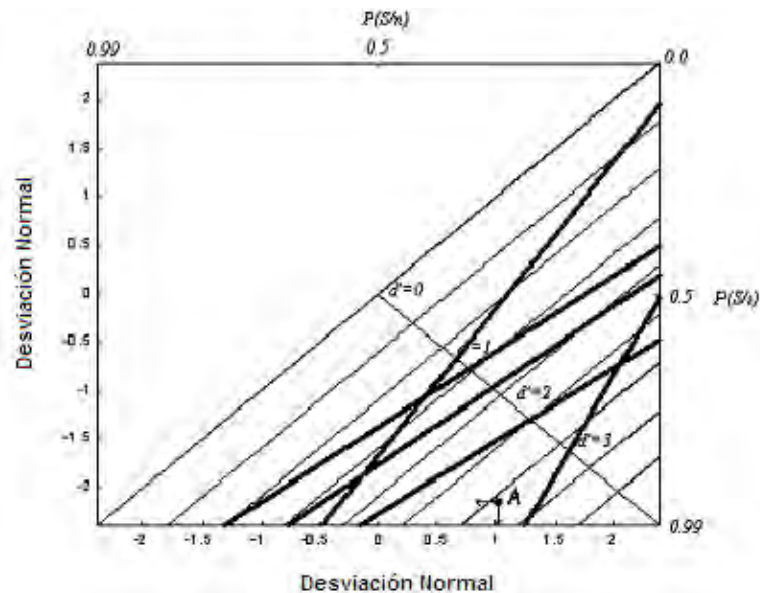


Fig. 2.11. – Curvas ROC teóricas con ejemplos de curvas empíricas obtenidas sobre una gráfica ROC binormal.



gráfica, las curvas ROC empíricas se aproximan a líneas rectas bien definidas con diversas pendientes. La pendiente generalmente oscila en el rango 0.5 a 1.5 como lo indican las líneas continuas de la Fig. 2.11.

La Fig. 2.11 muestra la familia de curvas ROC basada en distribuciones normales de proporciones de aciertos y de falsas alarmas. El análisis ROC nos proporciona una medida de discriminación que es independiente de la ubicación del criterio de decisión y que presumiblemente esta libre de contaminación de procesos como motivación y expectativa[1]. Al mismo tiempo, el análisis ROC provee una medida del efecto neto que influye en la respuesta, específicamente la ubicación del criterio de decisión a un tiempo dado. Esta medida, llamada β en el contexto ROC, es el valor de la razón de verosimilitud a la cual el criterio ha sido establecido para un punto particular de la curva ROC. Como ya se ha mencionado, la medida β es igual a la pendiente a la cual la curva ROC pasa en un punto.

Cuando los datos están presentes para diferentes criterios, ya sea porque el observador varió su criterio de decisión de un grupo de presentaciones a otro o por medio de un experimento por categorías, se puede utilizar la gráfica de la Fig. 2.11 para obtener un estimado de d' . La curva teórica de la Fig. 2.11 contiene líneas rectas de pendiente unitaria tomando en cuenta las coordenadas de desviación normal. Por esta razón el experimentador puede aproximar diferentes puntos a una línea recta con pendiente unitaria y obtener d' con la ecuación 2.20 en el punto en que la recta corta la diagonal.

No todos los datos se aproximan a una recta de pendiente unitaria y esto representa una complicación. En principio, se viola la suposición de normalidad y la pendiente no unitaria viola la suposición de igual varianza. La cuestión es qué hacer cuando se tienen rectas no unitarias, algunas alternativas sugeridas por Swets[1] para atacar este problema son:

- a) Utilizar la diferencia (en coordenadas z) de medias de las dos distribuciones normales superpuestas.
- b) Utilizar la pendiente de la curva ROC.
- c) Utilizar un parámetro que ignore la pendiente de la curva ROC, por ejemplo, el valor de d' en la diagonal negativa o distancia perpendicular desde el centro del



espacio ROC a la curva ROC, cada una de las cuales presenta unidades basadas sobre desviaciones estándar de ambas distribuciones.

- d) Asumir distribuciones que definan pendientes no unitarias, específicamente distribuciones que definan cómo la pendiente varia con la diferencia de medias. El experimentador puede asumir distribuciones de Poisson, gamma, Rayleigh o exponenciales. Este tipo de distribuciones predicen pendientes menores de uno.
- e) En el último de los casos se puede recurrir al parámetro $p(a)$, que es independiente de la distribución utilizada.

2.4. – DISTRIBUCIÓN ASUMIDA DE SEÑAL Y RUIDO.

Hasta ahora se ha usado la teoría de decisión para identificar las variables relevantes de diferentes procedimientos experimentales, y para predecir relaciones entre parámetros del sistema sensorial de los observadores como estimados a partir de los datos recolectados por diferentes procedimientos. Sería de importancia conocer algunos parámetros de la señal como la duración de esta y algunos otros parámetros que afectan la detectabilidad de la misma. Como consecuencia, se asumen formas específicas para la distribución de probabilidad de eventos sensoriales y, que un parámetro de esta distribución tiene alguna relación específica con la variables experimentales. Estos supuestos llevan a predicciones acerca de cómo, cambios en la variable física influirán en la habilidad del observador para detectar la señal.

Para determinar la forma exacta de la distribución se consideran: 1) parámetros físicos de la señal y ruido y 2) el proceso de decisión que lleva a cabo un observador.

El experimentador asume distribuciones acerca del proceso de decisión de los observadores (Gaussiana, Exponencial, Poisson, etc.). La exactitud de estos supuestos se determina comparando las deducciones hechas con los datos obtenidos empíricamente. Específicamente, la distribución normal es analizada en esta parte del trabajo poniendo especial énfasis en cómo las varianzas de las distribuciones de señal y ruido afectan la forma de la gráfica ROC. Posteriormente, se presenta brevemente el modelo Exponencial[1][2].



2.4.1. – HISTORIA DEL SUPUESTO GAUSSIANO.

Thurstone(1927) desarrolló un análisis de dos muy generales métodos de juicio. En el primero (juicio comparativo) el sujeto hace una selección binaria entre dos estímulos. En esta tarea él es cuestionado para establecer si un primer sonido es más fuerte que el segundo, si una coloración es más intensa que otra o si una sentencia representa más la opinión de juicio que otro.

El segundo método (juicio categórico) requiere una respuesta categórica del sujeto ante una presentación de uno de entre varios estímulos. El sujeto es cuestionado para colocar un estímulo en particular en una categoría apropiada. Un ejemplo de juicio categórico es el método si-no.

Thurstone analizó ambos métodos en términos de un concepto que él llamó “dispersión discriminial”. La esencia es que el efecto de estimulación puede representarse por una variable aleatoria. El estímulo tiene un valor medio sobre una escala psicológica relevante pero, por varias razones, su valor exacto fluctúa de momento a momento. A partir de esto se asume que tal estimulación produce una distribución de efectos. Thurstone asumió la distribución Gaussiana, y esta distribución causada por algunos estímulos esta completamente caracterizada por dos parámetros, la media y la varianza.

Existen dos razones para asumir que las distribuciones de los eventos sensoriales son Gaussianas; una teórica y la otra práctica. La razón teórica es el llamado “teorema de límite central”. Un resultado general es que si las variables aleatorias son independientes y todas tienen la misma distribución, sea cual sea dicha distribución, la suma de tales variables se aproxima a una distribución Gaussiana conforme el número de variables crece indefinidamente. Si pensamos que los eventos sensoriales están compuestos de múltiples pequeños eventos, los cuales son independientes, entonces, el teorema del límite central se puede citar para justificar la distribución Gaussiana del efecto neto.

La razón práctica del supuesto de la distribución Gaussiana es que nos permite derivar resultados que son difíciles o imposibles de derivar bajo otros supuestos. El hecho de que los



parámetros y las variables sean fáciles de manipular bajo operaciones lineales ofrece muchas ventajas.

2.4.2. – EL SUPUESTO GAUSSIANO Y LA TAREA DE OBSERVACIÓN SIMPLE.

Cuando se consideran dos hipótesis alternativas, como sucede comúnmente, podemos asumir que los eventos sensoriales pueden ser mapeados sobre una línea específica llamada x . El valor numérico de un evento sensorial observado afecta la confianza del observador acerca de cuál hipótesis es verdadera. No obstante, asumimos que él tiene un criterio k tal que seleccionará n siempre y cuando $x < k$ y seleccionará s , siempre y cuando $x > k$. La variable x puede o no ser la razón de verosimilitud; podemos ahora establecer la distribución Gaussiana en forma precisa: la distribución de la variable aleatoria x es Gaussiana bajo cada hipótesis. Específicamente, bajo la hipótesis n , x tiene una distribución Gaussiana con media m_n y desviación estándar σ_n y, bajo la hipótesis s , x tiene distribución Gaussiana con media m_s y desviación estándar σ_s , matemáticamente:

$$f(x/n) = (2\pi\sigma_n^2)^{-1/2} \exp\left[-\frac{(x-m_n)^2}{2\sigma_n^2}\right] \quad 2.25$$

$$f(x/s) = (2\pi\sigma_s^2)^{-1/2} \exp\left[-\frac{(x-m_s)^2}{2\sigma_s^2}\right] \quad 2.26$$

Las coordenadas de la curva ROC son únicamente determinadas a partir de los supuestos considerados para algún criterio dado k :

$$P(S/s) = \int_k^{\infty} f(x/s) dx = \Phi\left(\frac{k-m_s}{\sigma_s}\right) \quad 2.27$$

y



$$P(S/n) = \int_k^{\infty} f(x/n) dx = \Phi\left(\frac{k - m_n}{\sigma_n}\right) \quad 2.28$$

donde Φ es la integral de la desviación estándar. Las ecuaciones 2.27 y 2.28 contienen cuatro parámetros de las dos distribuciones Gaussianas. En efecto, únicamente dos parámetros son relevantes debido a que las distribuciones base no son directamente observadas, podemos escalar x tal que m_n sea cero y σ_n sea unitaria sobre una nueva escala. Esto es, haciendo y la transformación lineal de x tenemos:

$$y = \frac{1}{\sigma_n}(x - m_n) \quad 2.29$$

La media de la segunda distribución es entonces $\sigma_n/(m_s - m_n)$ y su desviación estándar es entonces σ_s/σ_n . El criterio k puede ser también transformado por estas mismas ecuaciones para obtener dos probabilidades, $P(S/s)$ y $P(S/n)$. Esta transformación equivalente revela que la distancia entre las medias $(m_s - m_n)$ y la razón de desviación estándar σ_s/σ_n son parámetros de interés.

2.4.2.1. - CASO DE VARIANZAS IGUALES ($\sigma_s = \sigma_n$).

Sea el supuesto de que $\sigma_s = \sigma_n$ para derivar la curva ROC para este caso y la razón de verosimilitud. σ_s y σ_n , así como $m_s - m_n$ son medidas en unidades de x . Además, $1/\sigma_n$ es unitaria debido a que $\sigma_s = \sigma_n = 1$. La expresión para la razón de verosimilitud es:

$$l(x) = \exp\left[(m_s - m_n)x - \frac{1}{2}(m_s^2 - m_n^2)\right] \quad 2.30$$

Debido a que el único parámetro es $m_s - m_n$, es conveniente representar $m_s = d/2$ y $m_n = -d/2$; entonces:

$$l(x) = e^{d'x} \quad \text{o} \quad \ln(l(x)) = d'x \quad 2.31$$



La forma de la curva ROC generada por estos supuestos es fácil de determinar considerando que la pendiente de la curva ROC es igual a la razón de verosimilitud sabiendo que la pendiente es unitaria cuando $x=0$. Además, debido a que $e^{d'x}=1/e^{-d'x}$, la pendiente de la curva ROC esta recíprocamente relacionada respecto con la diagonal negativa (ver Fig. 2.11).

Una forma de representar la curva ROC es utilizando papel de doble probabilidad y esta forma se muestra en la Fig. 2.11 la cual se construye calculando valores normalizados z para cada uno de los valores de P de acuerdo a la siguiente ecuación:

$$P(z) = \frac{100}{(2\pi)^{1/2}} \int_{-\infty}^z e^{-x^2/2} dx \quad 2.32$$

La Fig. 2.11 muestra la curva ROC graficada en papel de doble probabilidad. La ventaja de usar este tipo de gráficas ROC es que la curva ROC generada por un experimentador es una línea recta, y la separación entre las medias de las dos distribuciones puede obtenerse simplemente sustrayendo el valor de la coordenada $z(P(S/s))$ a la coordenada $z(P(S/n))$. Por ejemplo, el punto A de la Fig. 2.11 tiene coordenadas $P(S/s)=0.98$ y $P(S/n)=0.16$, este punto esta muy cerca de la línea de $d'=3$. El valor de z asociado con $P(S/n)$ es $+1$; y el valor de z asociado con $P(S/s)$ es -2 ; el valor de d' buscado es entonces la diferencia, en valor absoluto, de ambas coordenadas, es decir, $1-(-2)=+3$.

2.4.2.2. – CASO DE VARIANZAS DESIGUALES ($\sigma_s \neq \sigma_n$).

En este caso la razón de verosimilitud no es monotonica con x . Esto es, las curvas ROC generadas a partir del criterio “decir S si $x > k$ ” no tendrán pendientes que sean monotonicamente decrecientes. Para ver exactamente cómo la razón de las desviaciones estándar afecta la curva ROC es más fácil proceder geoméricamente. La pendiente de la curva ROC para un criterio k fija la razón de distribuciones a un valor de k , esto es:

$$\frac{dP(S/s)}{dP(S/n)} \Big|_x = \frac{f(k/s)}{f(k/n)} \quad 2.33$$



Este resultado es demostrado utilizando las ecuaciones 2.27 y 2.28. La Fig. 2.12 muestra los tres resultados posibles entre las magnitudes de σ_s y σ_n . La razón de las densidades de probabilidad determina la pendiente de la curva ROC que se relaciona con cada supuesto.

En la curva superior $\sigma_s > \sigma_n$, la curva ROC crece rápidamente hasta llegar el punto b , en este punto la razón de verosimilitud y en esencia la pendiente, es unitaria. Entre el punto b y el a la razón de verosimilitud es menor que uno y la curva tiene una pendiente pequeña. A partir del punto a y hasta la esquina superior derecha de la gráfica, la pendiente de la curva crece rápidamente. En la curva del centro $\sigma_s = \sigma_n$, en este caso la curva es simétrica y tiene pendiente unitaria en el punto c , únicamente en éste caso la pendiente de la curva disminuirá monotonamente.

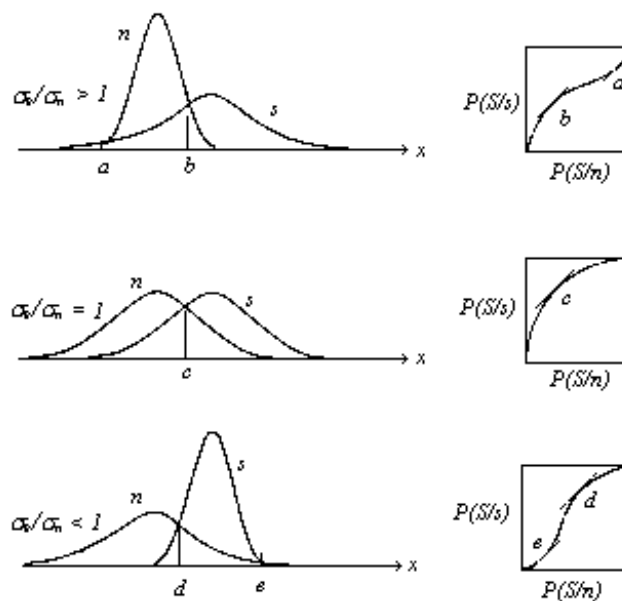


Fig. 2.12. – Curva ROC para diferentes casos de relaciones entre σ_s y σ_n .

En la curva inferior se tiene el caso $\sigma_s < \sigma_n$ y presenta resultados muy similares a la primera curva, esta curva es el reflejo geométrico de la curva superior.

Desafortunadamente para el experimentador la razón σ_s / σ_n normalmente es diferente de uno, lo cual puede detectarse fácilmente en las curvas experimentales ROC. Si por ejemplo, el área bajo la curva ROC es 0.75 y σ_s es el doble de σ_n , entonces la aceleración



rápida de la curva ROC cae cerca de la esquina superior izquierda (punto a) con valores de $P(S/s)$ entre 0.86 y 1.0.

La curva ROC binormal puede ser usada cuando las varianzas de las distribuciones son diferentes. Suponga que las ecuaciones de la 2.25 a la 2.28 son ciertas, excepto que $\sigma_s \neq \sigma_n$. Haciendo un cambio de variables, $z_y = (k - m_s) / \sigma_s$, encontramos que $k = \sigma_s z_y + m_s$. Similarmente, un cambio en la otra variable $z_x = (k - m_n) / \sigma_n$ nos llevará a la transformación $k = \sigma_n z_x + m_n$. Las escalas de probabilidad permanecerán linealmente relacionadas, pero ahora un cambio pequeño en k producirá un cambio más pequeño en z_y que en z_x y $\Delta k = \sigma_n \Delta z_x$, la pendiente $\Delta z_y / \Delta z_x = \sigma_n / \sigma_s$. Esto es, conforme la razón σ_s / σ_n incrementa, la pendiente de la línea sobre la gráfica ROC binormal disminuye. Es difícil establecer la pendiente de la recta sobre la gráfica ROC a partir de un grupo de datos específicos[1][2].

2.5. – EL SUPUESTO EXPONENCIAL.

Las familias de curvas ROC bajo el supuesto exponencial (conocidas como curvas power law) se representa mediante la simple ecuación $P(S/s) = [P(S/n)]^k$, donde $0 < k \leq 1$. A partir de esta relación es posible generar una curva ROC que nos de información acerca del desempeño de un observador al detectar un estímulo inmerso en una plataforma ruidosa.

LA FAMILIA EXPONENCIAL NEGATIVA.

La ecuación ke^{-kx} , donde $x \geq 0$ y $k > 0$, representa la distribución exponencial negativa. Algunas de las características de la distribución exponencial aplicadas a la teoría de detección de señales son:

- 1) Un par de distribuciones exponencial negativas genera una power-law ROC.
- 2) Para algún par de distribuciones exponenciales, x es estrictamente monótono con $l(x)$, y en efecto, estas distribuciones representan una familia de razón de posibilidad monótona.



- 3) La función exponencial negativa es una forma estandarizada de la forma chi-cuadrada con dos parámetros de libertad. Esta distribución involucra un observador ideal en tareas de detección.
- 4) El parámetro k (o $1/k$) es considerado como un índice de eficiencia obtenido bajo el supuesto de distribución sensorial Gaussiano.

2.5.1. – LA DISTRIBUCIÓN EXPONENCIAL NEGATIVA Y LA RAZÓN DE VEROSIMILITUD.

Comenzaremos discutiendo algunas de las propiedades de ke^{-kx} . Para esta función de densidad $0 < k < \infty$ y $0 \leq x < \infty$. La función de distribución es:

$$P(X \leq x) = F(x) = \int_0^x ke^{-kt} dt = 1 - e^{-kx} \quad 2.34$$

Entonces, el complemento de la función de distribución es e^{-kx} . Así, la curva de densidad de X (con pendiente $-k^2$) lleva al eje vertical un valor de k . El valor absoluto de la pendiente normalmente decrece a cero y la densidad de X aproxima al eje horizontal asintóticamente cuando x se aproxima a infinito.

La razón de verosimilitud puede obtenerse a partir de un par de exponenciales negativas. Considerando el siguiente par de distribuciones exponenciales:

$$f(x/s) = ke^{-kx} \quad 2.35$$

$$f(x/n) = e^{-x} \quad 2.36$$

donde $0 < k \leq 1$, $0 \leq x < \infty$. La razón de verosimilitud para este par de distribuciones es :

$$l(x) = \frac{f(x/s)}{f(x/n)} = ke^{x(1-k)} \quad 2.37$$



A partir de la ecuación 2.37 es posible observar que $l(x)$ es monotonica con respecto a x . Por lo cual x es la variable de decisión apropiada y el complemento de las dos distribuciones puede ser usado para definir las proporciones de acierto y de falsa alarma y así poder determinar la razón de verosimilitud del observador en cuestión.

La Fig. 2.13 muestra dos funciones exponenciales y su razón de posibilidad asociado.

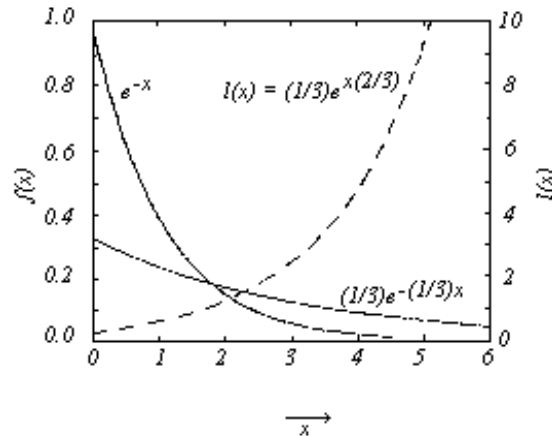


Fig. 2.13. – Funciones exponenciales negativas para $k=1$ y $k=1/3$. La razón de verosimilitud es mostrada con línea interrumpida, note que conforme x se aproxima a cero, $l(x)$ se aproxima a uno.

Las proporciones de acierto y falsa alarma de las dos distribuciones exponenciales son:

$$P(S/s)_c = P(X > c) = \int_0^{\infty} k \exp(-kx) dx = \exp(-kc) \quad 2.38$$

$$P(S/n)_c = P(X > c) = \int_0^{\infty} \exp(-x) dx = \exp(-c) \quad 2.39$$

De lo cual se deduce:

$$P(S/s) = [P(S/n)]^k \quad 2.40$$

La ecuación 2.40 es la ecuación de las ROC bajo el supuesto exponencial, esta ecuación es de forma simple y esta especificada por el parámetro único k . Además, muestra que, para una proporción fija de falsa alarma, la proporción de aciertos se incrementa continuamente conforme k disminuye de uno a cero.



La Fig. 2.14 muestra la familia de curvas ROC power law. Conforme k disminuye, el desempeño del observador aumenta.

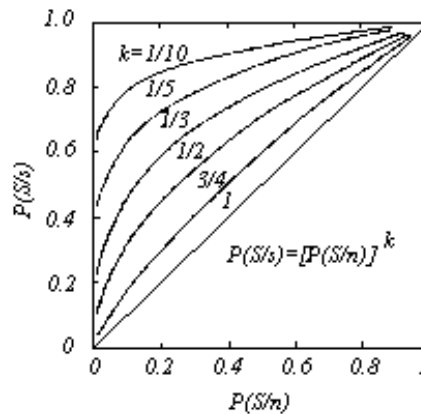


Fig. 2.14. – La familia de curvas ROC bajo el supuesto exponencial.

Considerando las medias de las distribuciones de ruido únicamente y señal y ruido como 1 ($k_n=1$) y $1/k$ ($k_{sn}=1$), respectivamente. Resulta conveniente definir como índice de desempeño el logaritmo de la diferencia de las medias de la distribución exponencial asumida. Esto es:

$$d_k = \log\left(\frac{1}{k_{sn}} - \frac{1}{k_n}\right) = \log\left(\frac{1}{k} - 1\right) \quad 2.41$$

donde k es el índice de desempeño obtenido de la ecuación 2.40. De esta manera, con las ecuaciones 2.40 y 2.41 es posible determinar un índice de desempeño para cada punto de la curva ROC bajo el supuesto exponencial. Si se consideran experimentos ROC 4 categorías para obtener 3 puntos empíricos en cada experimento. La ecuación para calcular los valores de desempeño para cada punto es:

$$d_{kj} = \log\left(\frac{1}{k_j} - 1\right), \quad j = 1, 2, 3 \quad 2.42$$



Con la ecuación 2.42 es posible conocer el desempeño real de los observadores en cada punto de la curva ROC bajo el supuesto exponencial. Dos observadores pueden tener similar desempeño global, pero su desempeño en cada punto de la curva ROC puede ser diferente.



BIBLIOGRAFÍA.

1. J. A. Swets, Signal Detection Theory and ROC Analysis in Psychology and Diagnostics (Lawrence Erlbaum Associates, New Jersey, 1996).
2. D. M. Green and J.A. Swets, Signal Detection Theory and Psychophysics (John Wiley and Sons, New York, 1966).
3. J. A. Swets, Signal Detection and Recognition by Human Observers (John Wiley and Sons, New York, 1964).
4. A. Burgess, "Statistical efficiency of perceptual decisions", SPIE vol. 454, 18-26 (1984).
5. A. Burgess, "Comparison of receiver operating characteristic and forced choice observer performance measurement methods," Med. Phys., 22, 643-655 (1995).
6. D. J. Weintraub and H. W. Hake, "Visual Discrimination and Interpretation in Terms of detectability Theory", JOSA vol. 52, number 10, 1179-1184 (1962).
7. J. Nachmias and E. C. Kocher, "Visual detection and discrimination of luminance Increments," JOSA, 60, 382-389 (1970).
8. A. Burgess, "Image quality, the ideal observer, and human performance of radiologic decision tasks," Acad. Radiol., 2, 522-526 (1995).
9. H.H. Barrett, "Objective assessment of image quality: effects of quantum noise and object variability," JOSA, 7, 1266-1277 (1990).
10. J.A. Swets and R.M. Pickett, Evaluation of Diagnostic Systems (Academic Press, New York, 1982).
11. J.A. Swets, W.P. Tanner, and T.G. Birdsall, "Decision process in perception," Psychol. Rev., 68, 301-340 (1961).
12. J. P. Egan, Signal Detection Theory and ROC Analysis, Academic Press, New York, 1975.



CAPÍTULO III. – SISTEMA EXPERTO DIFUSO PARA EVALUAR OBSERVADORES HUMANOS EN TAREAS DE DETECCIÓN DE SEÑALES.

3.1. – INTRODUCCIÓN.

En este capítulo se presenta el desarrollo de un Sistema Experto Difuso (SED) para la evaluación de observadores humanos en la tarea de detección visual de señales. Dicho SED fue implementado mediante dos subsistemas difusos (cada uno con sus propios métodos de difusificación y de desdifusificación) en cascada para combinar las variables de entrada, obtenidas mediante la teoría del ROC, que consideramos necesarias para el sistema. De esta manera, y con respecto a un sistema difuso de entrada-salida convencional, se redujo el número total de reglas FuzzyAnd necesarias para evaluar los datos. Utilizar 2 subsistemas ayudó a una mejor interpretación de los resultados y a un mejor “ajuste” del sistema. La teoría del ROC se utilizó para definir las variables de entrada y salida de 30 casos empíricos y 23 casos sintéticos. Los datos empíricos se obtuvieron de los experimentos ROC por categorías de 30 observadores humanos, mientras que los 23 casos sintéticos fueron datos ideados con el fin de simular casos extremos que en la práctica se pudieran tener y a los cuales se podría enfrentar el SED.

d_{kd} (índice generado por el SED) mide desempeño visual en una forma muy similar a d_k (índice generado por la teoría ROC) para los casos de desempeño visual promedio. No obstante, existen ventajas importantes cuando el sistema se utiliza para analizar casos de pobre y excelente desempeño aparentemente muy similares. Es decir, el sistema muestra con mucha más claridad las diferencias entre grupos de observadores con pobre y excelente desempeño. Pese a que la metodología del SED fue desarrollada tomando en cuenta experimentos de percepción ROC con cuatro categorías, éste sistema puede ser utilizado para analizar datos generados por experimentos ROC diseñados con cinco o más categorías siguiendo el mismo procedimiento propuesto en este trabajo.

El SED propuesto fácilmente manipula la incertidumbre asociada a la evaluación de observadores humanos en la tarea de detección visual de señales [9,10,17] usando la teoría del ROC [6,18]. Se aplicó esta teoría bajo el supuesto exponencial de distribución de los eventos



sensoriales generando curvas ROC [1,2,11]. Como se mencionó en el capítulo II, la teoría ROC es útil para proveer un índice de detectabilidad en cada punto experimental de la curva ROC que, teóricamente, es independiente de las circunstancias que “forzan” al observador a establecer su criterio de decisión. De esta manera, las entradas del SED son los índices de desempeño en cada punto empírico de la curva ROC generada para cada observador (d_{k1} , d_{k2} y d_{k3}), y el área bajo la curva ROC $p(a)$. La salida final del sistema es un nuevo índice de desempeño d_{kd} . Con este índice es más fácil determinar diferencias entre observadores humanos con un desempeño global similar de acuerdo con los índices de detectabilidad convencionales (i.e. d' , k , $p(a)$) pero con diferencias significativas en el desempeño correspondiente en cada uno de los puntos empíricos de la curva ROC.

3.2.- LAS IMÁGENES DE PRUEBA.

La imágenes utilizadas en nuestros experimentos fueron imágenes de 256 niveles de gris de 384x384 pixeles. Las imágenes con únicamente ruido y con señal y ruido fueron generadas por computadora. Las imágenes con únicamente ruido contenían ruido de distribución Gaussiana con una media de niveles de gris de 128 y una desviación estándar de 32 niveles de gris. Para reducir los efectos de aprendizaje, el ruido Gaussiano fue diferente para cada imagen. Las imágenes con señal y ruido contenían, además del ruido Gaussiano, un objetivo (estímulo) en forma de una onda seno modulada por una exponencial negativa tal que,

$$F(x, y) = 0.03[128 + 127e^{-k_1 r} \cos(k_2 r)],$$

donde $r(x,y)$ es el radio del anillo mas largo en el objetivo, $k_1 = 0.001$ y $k_2 = 0.0001$. El objetivo ocupó un radio de 60 pixeles. Este objetivo, que asemeja una señal de Gabor, fue sumada en el centro de las imágenes de ruido Gaussiano. Las imágenes de señal y ruido presentaron una razón señal a ruido de 2.34. La Figura 3.1 muestra una imagen con señal y ruido (S/N) como las utilizadas en nuestros experimentos pero con una razón señal a ruido de 6% para poder visualizar el objetivo al centro de la imagen.



3.3. – DESARROLLO DEL SISTEMA EXPERTO DIFUSO (SED).

Los sistemas expertos son programas de computadora que simulan el proceso de razonamiento de un experto humano [1]. Los primeros sistemas expertos fueron desarrollados como un área muy particular de la Inteligencia Artificial durante los 60's. Este periodo fue dominado bajo el supuesto que unas cuantas leyes de conocimiento acopladas a algoritmos eficientes de programación producían un experto [1,2,3].

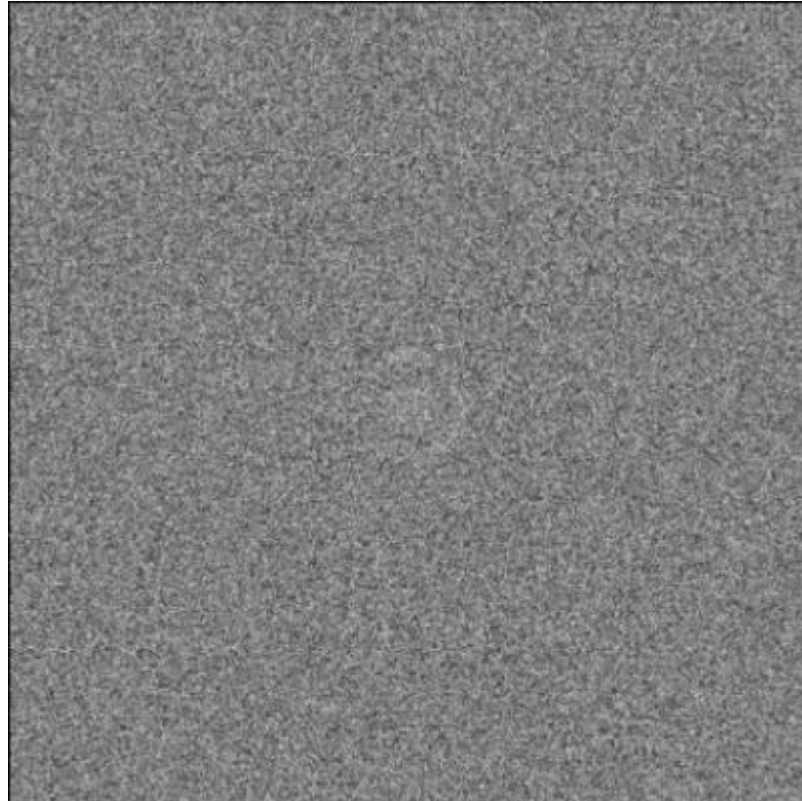


Figura 3.1. – Imagen de señal y ruido similar a las utilizadas en los experimentos de percepción visual (S/N = 6.0).

Los SED utilizan reglas de conocimiento difusas que manejan la incertidumbre propia de la aplicación [4,5]. Gradualmente, los SED han ido sustituyendo a los sistemas de ingeniería convencionales; especialmente a los sistemas de control y de supervisión [1,3,5]. Esto se debe a la fácil implementación de esta clase de sistemas y a los resultados tan satisfactorios que han proporcionado. La manipulación de las “funciones de membresía”



constituye la base de todo sistema difuso, bajo este concepto los SED son capaces de manejar la incertidumbre y la vaguedad asociadas con las diferentes aplicaciones.

Los conjuntos difusos, desarrollados por Lofti Zadeh en 1965, son muy utilizados para el desarrollo de los SED y constituyen una nueva forma de representar, en una forma muy simple, la vaguedad de nuestra vida diaria así como la que encontramos en un buen número de áreas de investigación. La lógica difusa, utilizando datos difusos, representa información imprecisa mediante modelos computacionales que son capaces de reconocer, representar, manipular e interpretar la incertidumbre difusa y estadística [1,7,8].

3.3.1. – CONSIDERACIONES TEÓRICAS.

La teoría del ROC y la base teórica acerca de los SED se han desarrollado mas ampliamente en los capítulos I y II de éste trabajo. No obstante, en este capítulo se considerarán algunos aspectos teóricos relevantes para el desarrollo del sistema descrito. Dichas consideraciones se centran principalmente en la forma en que se manejan y definen las variables de entrada y salida del sistema para la obtención de un SED que proporcione los resultados esperados. Además, se describe la forma en que se implementaron experimentos de percepción visual de señales y su análisis bajo el supuesto exponencial para obtener un índice de desempeño de un observador humano.

Algunas veces resulta difícil o imposible combinar todas las variables de entrada propias de la aplicación en un solo conjunto de reglas difusas. En este caso, una alternativa para resolver este problema es hacer un análisis de todas las variables de entrada y agruparlas de acuerdo a su naturaleza, de esta manera se implementa un conjunto de reglas difusas para cada grupo. En este trabajo, al analizar el conjunto de variables de entrada resultó conveniente implementar un subsistema difuso que proporcionó una salida la cual fue subsecuentemente considerada como una variable de entrada de otro subsistema difuso que combinó el resto de las variables de entrada. Ambos subsistemas constituyeron el sistema difuso total. Esto es, se combinaron 4 variables de entrada usando 2 subsistemas difusos debido a que resultó difícil combinar todas las variables de entrada en un mismo conjunto de reglas difusas. Esto debido a que resultó complicado combinar las variables de entrada (d_{k1} , d_{k2} , d_{k3}) que determinan el



desempeño del observador para cada criterio de decisión (individualmente) con variables de entrada ($p(a)$) que determinan el desempeño global del observador. La Figura 3.2 muestra el diagrama a bloques del SED propuesto.

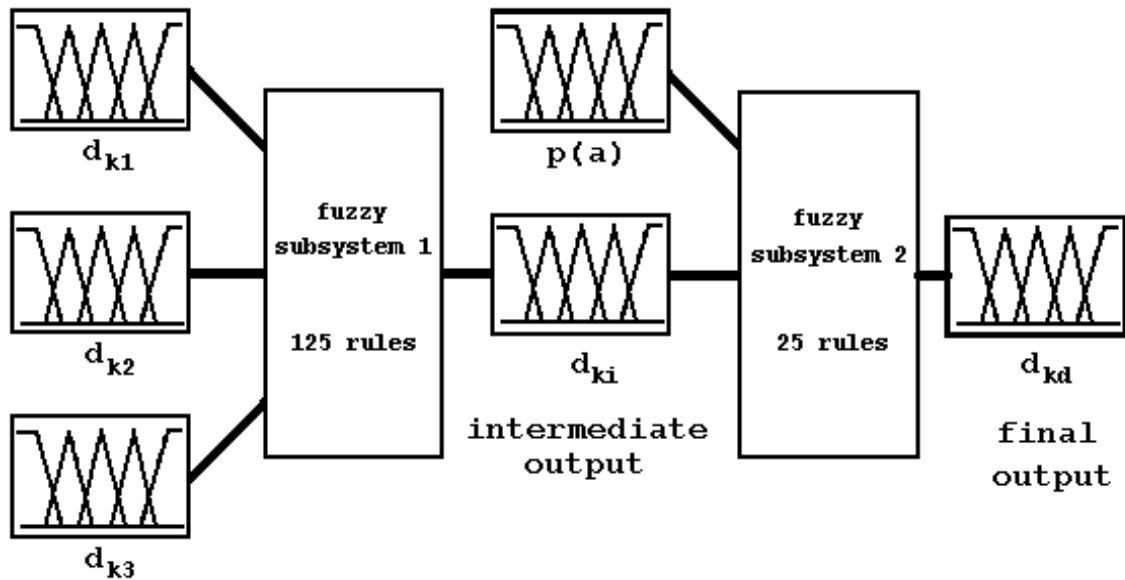


Figura 3.2. – Sistema difuso compuesto por 2 subsistemas difusos.

3.3.2. – VARIABLES DE ENTRADA Y SALIDA.

El SED se probó con datos experimentales obtenidos de 30 observadores que participaron en experimentos de percepción visual usando la técnica ROC por categorías bajo el supuesto de distribución exponencial[20]. Cada uno de los experimentos ROC por categorías proporcionó 3 puntos experimentales sobre la curva ROC (3 criterios de decisión) por cada observador. Antes de probar el sistema con datos reales, se analizaron 23 casos sintéticos que simulaban resultados de experimentos de percepción con entradas perfectamente conocidas que generarían resultados totalmente predecibles.

Las variables de entrada del sistema fueron los valores de los índices de desempeño d_{kj} (Ec. 2.42) para cada uno de los puntos de la curva ROC (experimentales y sintéticos). Esto debido a que se encontró a d_k como el mejor índice para evaluar el desempeño de observadores en cada categoría.

El análisis usando la primera parte del sistema difuso (subsistema 1 en Figura 3.2), resultó en una variable intermedia llamada d_{ki} . El segundo subsistema (subsistema 2 en Figura



3.2) combinó el índice de detectabilidad $p(a)$ con d_{ki} evaluando 9 reglas fuzzyAnd if-then para determinar, en una forma más eficiente, diferencias de desempeño de observadores con similar d_k pero con diferentes valores de d_{k1} , d_{k2} y d_{k3} . d_{kd} es la variable de salida final del sistema. Esta variable d_{kd} tiene el mismo rango de valores de d_k para los puntos experimentales de la curva ROC.

Después de identificar las variables relevantes del sistema así como los rangos de sus valores, se seleccionaron los estados lingüísticos para cada una de las variables y se expresaron por medio de los conjuntos difusos apropiados. Se necesitaron 5 estados lingüísticos para cada una de las variables de entrada y salida, éstos fueron: P – Poor, L – Low, M – Medium, H – High y E – Excellent.

La cuantización difusa de todas las variables fue realizada por medio de gráficas de difusificación triangulares y simétricas. La Figura 3.3 muestra las gráficas de difusificación para las variables d_k (incluyendo d_{k1} , d_{k2} , d_{k3} , d_{ki} y d_{kd}) y $p(a)$.

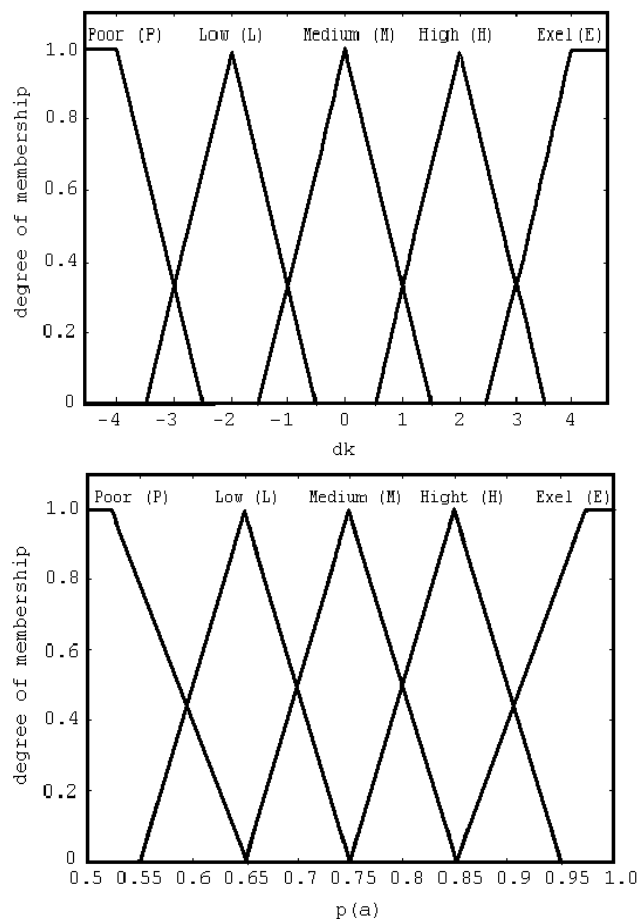


Figura 3.3. – Gráficas de difusificación de las variables d_k y $p(a)$.



3.3.3. – LA MÁQUINA DE INFERENCIA.

El conocimiento adquirido por el sistema fue formulado en términos de reglas de inferencia difusas de tipo fuzzyAnd en la forma de proposiciones condicionales. Los conjuntos de reglas fuzzyAnd para los 2 subsistemas fueron obtenidos a partir de la experiencia de 2 operadores humanos experimentados (los autores). Un ejemplo de una regla fuzzyAnd usada es:

$$\text{if } d_{k1} = H \text{ and } d_{k2} = M \text{ and } d_{k3} = M, \text{ then } d_{ki} = M$$

cada una de las reglas fuzzyAnd fue evaluada usando el mínimo. El propósito de la máquina de inferencia fue combinar cada una de las variables de entrada con reglas de información difusa para poder hacer inferencias acerca de las variables d_{ki} y d_{kd} .

La máquina de inferencia utilizada, a diferencia de las máquinas de inferencia convencionales con un solo conjunto de entradas y salidas, comprendió 2 procesos de difusificación y 2 procesos de desdifusificación (uno para cada subsistema). Finalmente, el sistema difuso transformó cada una de las conclusiones obtenidas por la máquina de inferencia a entidades no difusas utilizando el método de desdifusificación del centroide.

3.4. – RESULTADOS.

En las primeras pruebas con datos sintéticos el sistema difuso contenía 27 reglas difusas para generar la variable intermedia d_{ki} y 9 reglas difusas para generar la variable de salida d_{kd} . En estas pruebas iniciales todas las variables, d_{kj} y $p(a)$, fueron seleccionadas con 3 estados lingüísticos (de los 5 definidos anteriormente) para su representación y cuantificación difusa.

En la etapa de entrenamiento del sistema las reglas difusas y la forma de los conjuntos difusos de todas las variables cambiaron. De esta manera, al final el sistema resultó en 125 reglas difusas para la variable intermedia d_{ki} y 25 reglas difusas para la variable de salida d_{kd} . Las Figuras 3.4a y 3.4b muestran los conjuntos de reglas difusas para las variables d_{ki} y d_{kd} , respectivamente.



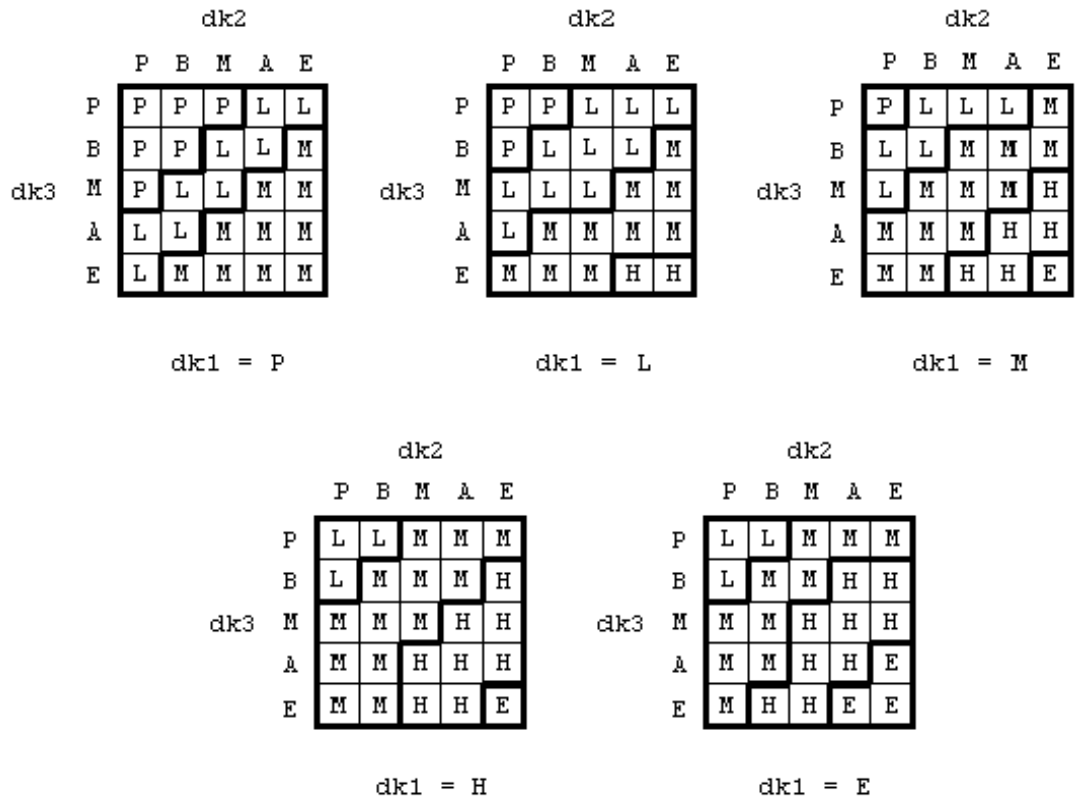


Figura 3.4a. – Conjunto de reglas difusas para la variable d_{ki} .

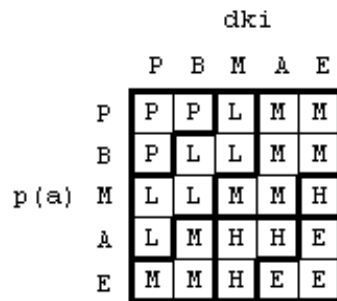


Figura 3.4b. – Conjunto de reglas difusas para la variable de salida d_{kd} .

En la Figura 3.5 son graficados los valores ordenados de d_k y d_{kd} para los datos sintéticos. Es importante notar las diferencias en las pendientes de éstas gráficas especialmente a ambos extremos. Claramente una mayor inclinación de la pendiente es mejor para analizar diferencias entre desempeño de observadores.



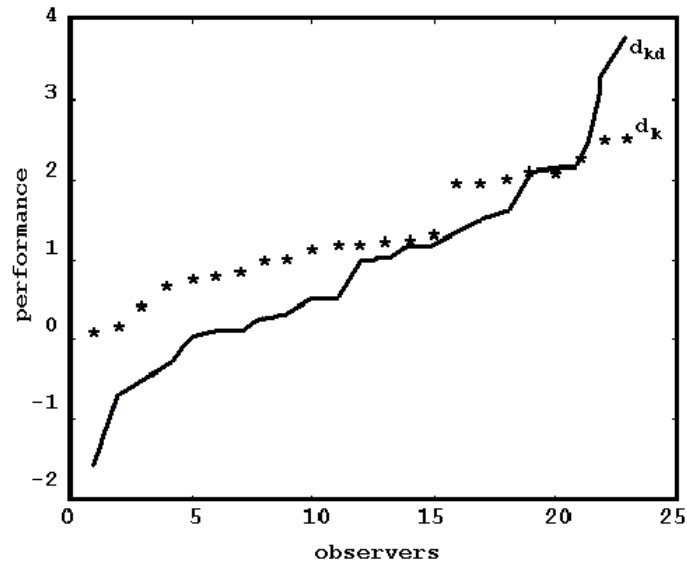


Figura 3.5. – Valores ordenados de d_k y d_{kd} de los datos sintéticos.

Para una mejor ilustración de esta ventaja, considérese el siguiente ejemplo: los valores de d_k (obtenidos como lo indica la teoría del ROC) de 2 observadores con buen desempeño fueron 2.225 y 2.505. El SED reportó valores de d_{kd} de 2.163 y 3.29 para los mismos observadores, respectivamente. En estos datos existe una muy notoria diferencia que facilita determinar cuál observador tiene un mejor desempeño. La diferencia reportada por el SED en los casos mencionados anteriormente es debido al hecho de que los valores de las variables de entrada de los observadores con $d_k = 2.225$ y $d_{kd} = 2.163$, fueron $d_{k1} = -0.934$, $d_{k2} = 1.931$, $d_{k3} = 1.938$, y $p(a) = 0.859$ y para el otro observador, con $d_k = 2.505$ y $d_{kd} = 3.29$, los valores de las variables de entrada fueron $d_{k1} = 1.242$, $d_{k2} = 3.644$, $d_{k3} = 2.865$, y $p(a) = 0.933$. Es muy notoria la diferencia de desempeño en cada una de las categorías de ambos observadores. Además, puede verse fácilmente que el primer observador tiene un pobre desempeño en la primera categoría (d_{k1}) y que existe una clara diferencia entre los índices $p(a)$ de ambos observadores.

En la Figura 3.6 se graficaron los valores ordenados de d_k (índice obtenido como lo indica la teoría del ROC) y de d_{kd} (la salida del SED) de los datos empíricos. Nuevamente existen diferencias importantes entre las pendientes de ambas gráficas.



Considérese ahora 2 observadores con pobre desempeño con valores de $d_k = -0.275$ y $d_{kd} = -1.999$, y otro observador obtuvo valores de $d_k = -0.141$ y $d_{kd} = -1.175$. Estos mismos observadores reportaron valores de $d_{k1} = -0.186$, $d_{k2} = -0.117$, $d_{k3} = 0.034$, y $p(a) = 0.610$; y $d_{k1} = -0.803$, $d_{k2} = 0.654$, $d_{k3} = -0.04$, y $p(a) = 0.686$. El SED “detectó” que existen diferencias en d_{k2} y $p(a)$ y asignó índices de desempeño, como se mencionó anteriormente, de -1.999 y -1.175 .

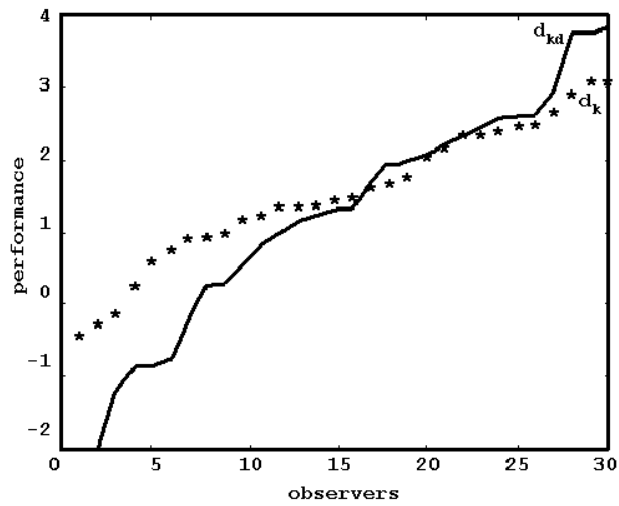


Figura 3.6. – Valores ordenados de d_k y d_{kd} de los datos empíricos.

Para el caso de 2 observadores con buen desempeño por ejemplo, un observador (con $d_k = 3.079$ y $d_{kd} = 2.983$) tuvo variables de entrada de $d_{k1} = 0.745$, $d_{k2} = 3.079$, $d_{k3} = 4.158$ y $p(a) = 0.927$, y otro observador (con $d_k = 3.088$ y $d_{kd} = 3.769$) tuvo valores de entrada de $d_{k1} = 1.636$, $d_{k2} = 3.394$, $d_{k3} = 4.527$ y $p(a) = 0.922$. Nuevamente las diferencias para cada criterio de decisión son notorias. El segundo observador tuvo mejor desempeño en todas las categorías y el SED pudo determinar una clara diferencia entre ambos observadores.

La Figura 3.7a muestra el histograma de d_k para los datos sintéticos, el cual resultó de aplicar teoría del ROC bajo el supuesto de distribución exponencial. En éste histograma no es posible distinguir a los observadores con pobre y excelente índice de desempeño. La Figura 3.7b muestra el histograma de d_{kd} para los mismos datos sintéticos. Podemos ver que en este último histograma es posible distinguir claramente a los observadores con pobre y excelente índice de desempeño.



El histograma obtenido con $p(a)$ como índice de desempeño (Figura 3.8) también separa mejor los pobres y excelentes índices de desempeño. Desafortunadamente el corto rango de valores de $p(a)$ ($[0.5 - 1.0]$) impide detectar fácilmente estas diferencias. A veces estas diferencias resultaron del orden de $1/1000$. Estas razones resultaron suficientes para considerar que d_{kd} es un buen índice de desempeño, útil para detectar ligeras diferencias entre desempeños de observadores.

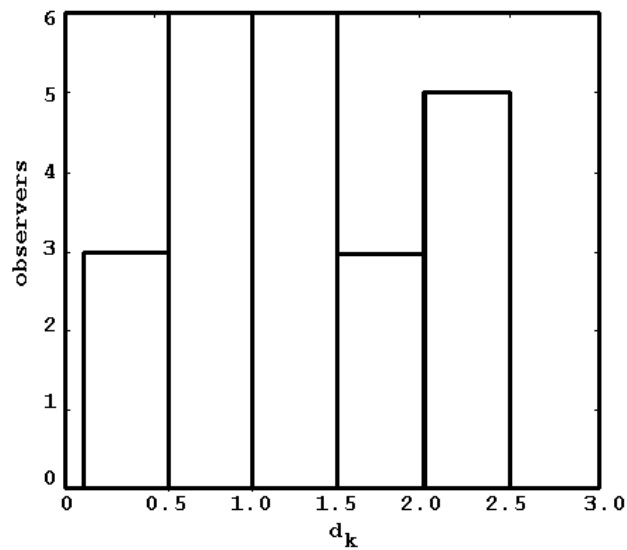


Figura 3.7a. –Histograma de d_k para los datos sintéticos.

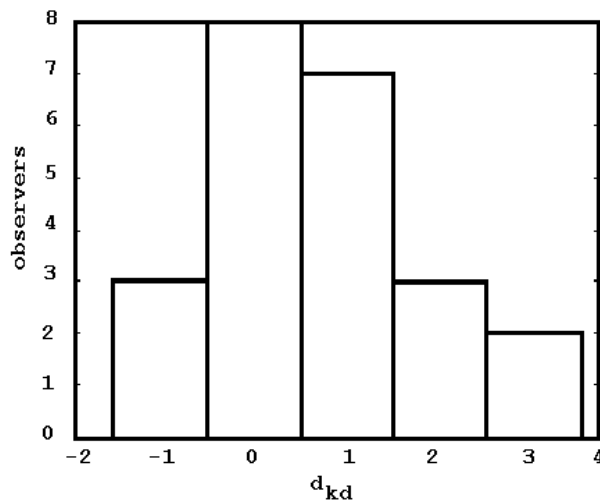


Figura 3.7b. –Histograma de d_{kd} para los datos sintéticos.



Por otra parte, las Figuras 3.9a y 3.9b muestran los histogramas de d_k y d_{kd} respectivamente, de los datos empíricos. En este caso, con d_k es posible detectar diferencias entre observadores con pobre desempeño pero resulta imposible detectar diferencias entre observadores con excelente desempeño. Sin embargo, en el histograma de d_{kd} se percibe que con éste índice es posible determinar diferencias entre observadores con pobre y excelente desempeño.

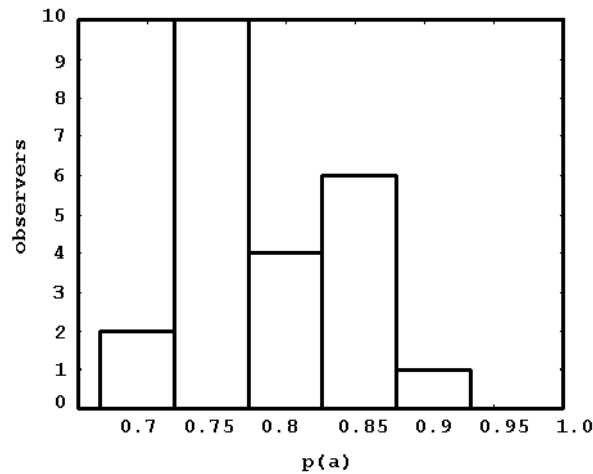


Figura 3.8. – Histograma de $p(a)$ para los datos sintéticos.

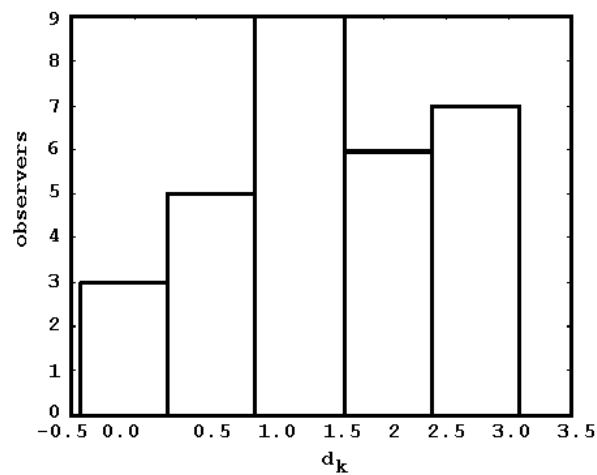


Figura 3.9a. – Histograma de d_k para los datos empíricos.



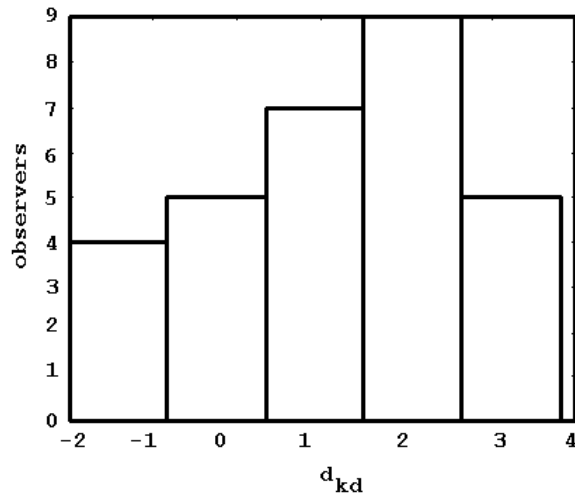


Figura 3.9b. – Histograma de d_{kd} para los datos empíricos.

La Figura 3.10 muestra el histograma de $p(a)$ para los datos empíricos. Como se puede observar en el histograma, con $p(a)$ es fácil detectar diferencias entre observadores con pobre índice de desempeño, pero resulta difícil detectar observadores con excelente índice de desempeño.

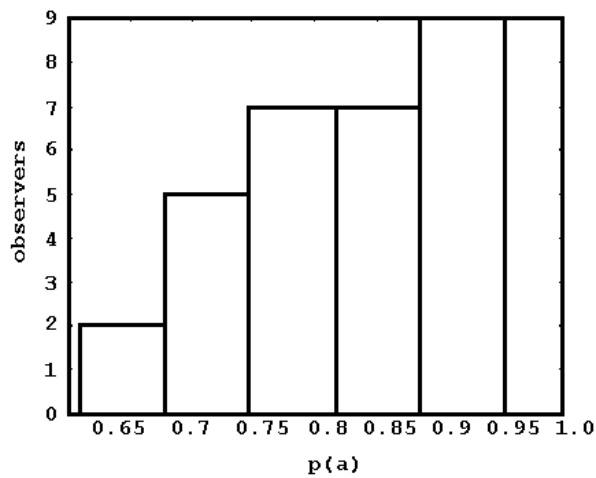


Figura 3.10. – Histograma de $p(a)$ para los datos empíricos.

3.5. – CONCLUSIONES.

Los resultados obtenidos con el SED desarrollado fueron totalmente satisfactorios debido a que fue posible resolver casos para los cuales fue difícil determinar diferencias entre



observadores con índices de desempeño muy similares como fue reportado por la teoría del ROC tradicional.

Se utilizaron los índices de desempeño d_{k1} , d_{k2} y d_{k3} sugeridos por la teoría del ROC bajo el supuesto exponencial. Pero estos índices no fueron suficientes para el diseño de un sistema difuso que pudiera mejorar diferencias y resolver controversias entre índices de desempeño de observadores. Entonces, fue necesario incluir la variable de entrada $p(a)$ combinándose con una variable intermedia d_{ki} a través de un segundo subsistema difuso.

Inicialmente, el sistema fue implementado con 625 reglas fuzzyAnd las cuales fueron creadas combinando las 4 variables de entrada en un solo conjunto de reglas fuzzyAnd. Con esta propuesta inicial de reglas no fue fácil interpretar cada una ellas. Debido a esto se diseñó un sistema difuso conformado por dos subsistemas difusos en cascada, de esta forma se pudo simplificar el diseño reduciendo el número total de reglas fuzzyAnd a 150. Además, la interpretación y evaluación de las reglas fue mucho más fácil.

Al principio se estableció la base de conocimientos usando reglas fuzzyAnd definidas con un criterio “demasiado estricto”. Pero con este conjunto de reglas se obtuvo una pendiente de d_{kd} en su extremo derecho fue muy cercana a cero (Fig. 3.6), esto evitó detectar diferencias entre grupos de buenos observadores. Consecuentemente, se utilizaron criterios de evaluación “estricto” y “no estricto”. Los tres criterios utilizados permitieron detectar fácilmente diferencias entre grupos de malos observadores, pero únicamente los resultados que incrementaron la pendiente de d_{kd} a ambos extremos fueron los obtenidos usando un criterio de evaluación de reglas “no estricto” para los conjuntos de reglas de los dos subsistemas.

El número de reglas fuzzyAnd en el primer subsistema puede incrementarse o reducirse exponencialmente de acuerdo con el número de categorías usadas en los experimentos ROC. No obstante, el número de reglas para el segundo subsistema permanecerá igual (25) para cualquier número de categorías utilizado en los experimentos ROC.



Se optó por utilizar reglas fuzzyAnd debido a que su interpretación para ambos subsistemas fue fácil y clara dada la naturaleza de la aplicación. No obstante, los resultados deberían ser similares si se utilizan reglas fuzzyOr, pero esta opción implicaría complicaciones innecesarias al diseño del SED.



BIBLIOGRAFIA.

1. Kandel, Abraham, Fuzzy Expert Systems, CRC Press, 1991, Tampa Florida, pp. 1 -81 (1991).
2. Sánchez J. P. Y Beltran, Los Sistemas Expertos: Una Metodología de Programación, Macrobit Corporation, 1990, México D.F.
3. Kosko Bart, Fuzzy Engineering, Prentice Hall, 1997,U.S.A.
4. Infelise, Nick, A clear vision of fuzzy logic, Control Engineering, vol. 38, no. 9, pp. 28-30.
5. Sugeno Michio; A Fuzzy Logic Based Approach to Qualitive Modeling, IEEE trans. On Fuzzy Systems, vol 1, no. 1, feb 1993, pp. 7-31.
6. H. H. Barrett, Objective assessment of image quality: effects of quantum noise and object variability, JOSA, 7, 1266-1277 (1990).
7. Zadeh, Lotfi, Fuzzy Sets, information and control, Academic Press, 1965, no. 3, vol. 8, pp 338-353.
8. Bezdek Jim, Fuzzy Models –what are they and why?, IEEE trans. On Fuzzy Systems, vol. 1, no.1, feb. 1993. pp. 1-5.
9. D. M. Green and J.A. Swets, Signal Detection Theory and Psychophysics (John Wiley and Sons, New York, 1966).
10. A. Burgess, Comparison of receiver operating characteristic and forced choice observer performance measurement methods, Med. Phys., 22, 643-655 (1995).
11. J. Markowitz, and J.A. Swets, Factors affecting the slope of empirical ROC curves: comparison of binary and rating responses, Percept. Psychoph., 2, 91-100, (1967).
12. J. P. Egan, Signal Detection Theory and ROC Analysis (Academic Press, New York, 1975).
13. C.E. Metz, ROC methodology in radiological imaging, Invest. Radiol., 21, 720 -733 (1986).
14. J. P. Egan, A.I. Schulman, and G.Z. Greenberg, Operating characteristics by binary decisions and by ratings, J. Acoust. Soc. Am., 31, 768-773 (1959).
15. Kosko, Bart, Neural Networks and Fuzzy Systems, Prentice Hall, 1992, U.S.A., pp. 262-295.
16. A. Burgess, Image quality, the ideal observer, and human performance of radiologic decision tasks, Acad. Radiol., 2, 522-526 (1995).



17. J. A. Swets, Signal Detection and Recognition by Human Observers (John Wiley and Sons, New York, 1964).
18. J. A. Swets, Signal Detection Theory and ROC Analysis in Psychology and Diagnostics (Lawrence Erlbaum Associates, New Jersey, 1996).
19. Russell Stuart and Norving Peter, Inteligencia Artificial: un enfoque moderno, Prentice Hall, México, D.F., 1996, pp. 439-461.
20. Padilla Medina Jose Alfredo and Sanchez Marin Francisco Javier, Alternative Application of ROC theory Under the Gaussian and Exponential Assumptions to Analyze Experimental Visual Detection Task, in paper submitted to Journal of Visual Communication and Image Representation on April of 2000.



CAPITULO IV. - SISTEMA EXPERTO DIFUSO PARA DESARROLLAR EXPERIMENTOS DE PERCEPCIÓN VISUAL ADAPTIVOS.

4.1. - INTRODUCCION.

Evaluar la habilidad de un observador humano para detectar señales en un ambiente dado incluye un modelo matemático de experimentos de psicofísica que implican tareas de detección visual de señales en las cuales un estímulo se encuentra presente o no dentro de una plataforma ruidosa[10,24,27].

En la Teoría del ROC existen dos tipos de experimentos de percepción visual comúnmente utilizados: los experimentos de si-no y los experimentos por categorías. Los experimentos ROC por categorías son preferibles en la mayoría de los casos debido a que son significativamente más eficientes. Con los experimentos por categorías es posible obtener varios puntos sobre la curva ROC (al menos se requieren 3 puntos) en un solo experimento, mientras que con los experimentos si-no cada experimento nos lleva a un solo punto sobre la gráfica ROC. No obstante, cuando los observadores no cuentan con mucha experiencia en la realización de experimentos de percepción visual, no es apropiado utilizar experimentos por categorías por largos períodos de tiempo. Además, normalmente se requiere un gran número de presentaciones para cada experimento por categorías. Estas dos inconveniencias incrementan considerablemente la duración de éste tipo de experimentos de tal forma que la fatiga y pérdida de atención en los observadores pueden afectar considerablemente el desempeño de los observadores así como los resultados obtenidos. Esto, en consecuencia, limita el uso de los experimentos por categorías. Para resolver estos problemas algunas soluciones adaptivas se han sugerido [11,16,21,25,30] pero ninguna de estas propuestas ha hecho uso de la Lógica Difusa.

En éste capítulo se presenta un Sistema Experto Difuso (SED), cuya arquitectura es semejante a la de un Controlador Difuso [10,30]. Este sistema se desarrolló con la finalidad de resolver las inconveniencias anteriormente mencionadas. Con éste SED se reduce la duración de los experimentos de detección visual a aproximadamente 1/4 de la duración de un experimento típico. Aunado a esto, los resultados obtenidos son equivalentes a los obtenidos con experimentos largos. Las entradas del sistema son los índices de desempeño obtenidos a lo largo de los experimentos y el incremento de dichos índices. La salida del sistema es el



valor "controlado" de la energía de la señal, la cual varía de manera adaptiva en términos del desempeño del observador en turno.

4.2.- MÉTODOS Y PROCEDIMIENTOS.

El algoritmo final fue optimizado mediante experimentos de percepción visual con 3 observadores experimentados. Los aspectos del sistema que se modificaron durante la etapa de optimización fueron la energía inicial de la señal, el rango de variación de la salida del sistema difuso y la matriz de conocimiento difuso expresada en términos de reglas FuzzyAnd. Luego, se probó el Sistema Difuso con datos derivados de 5 observadores humanos no experimentados los cuales participaron en experimentos de percepción visual utilizando el método ROC con 4 categorías[8,10,17]. Estos experimentos generaron 3 puntos sobre la curva ROC para cada observador.

4.2.1.- LAS IMÁGENES DE PRUEBA.

Las imágenes utilizadas en nuestros experimentos fueron imágenes de 256 niveles de gris de 384x384 píxeles. Las imágenes con únicamente ruido y con señal y ruido fueron generadas por computadora. Las imágenes con únicamente ruido contenían ruido de distribución Gaussiana con una media de niveles de gris de 128 y una desviación estándar de 32 niveles de gris. Para reducir los efectos de aprendizaje, el ruido Gaussiano fue diferente para cada imagen. Las imágenes con señal y ruido contenían, además del ruido Gaussiano, un objetivo (o estímulo) en forma de un círculo de 64 píxeles de diámetro y cuya energía fue adaptivamente variada de acuerdo al desempeño del observador en turno. Este objetivo fue sumado al ruido exactamente en el centro de cada imagen con señal y ruido. La Figura 3.1 muestra una imagen de señal y ruido pero en este caso el estímulo fue diferente.

4.2.2.- DISEÑO EXPERIMENTAL.

Cada experimento de Percepción Visual se inició con una energía de la señal equivalente a 4% de la energía de un círculo totalmente blanco (nivel de gris de 255). La energía del objetivo se incrementó y decrementó dependiendo del desempeño del observador



en turno hasta que los experimentos arrojaron un porcentaje de correctas del 80% (aproximadamente).

El programa de computadora que exhibió las imágenes y registró las respuestas de los observadores fue diseñado de tal manera que cada 20 presentaciones calculara un índice de desempeño parcial (d'), así como su correspondiente incremento (o decremento) con respecto al índice anterior. El programa detiene el experimento cuando se han realizado 200 presentaciones o cuando se obtienen 3 valores consecutivos del índice de desempeño en el rango de 1.3 a 1.8 (lo cual corresponde al rango de 74% a 82% de respuestas correctas). La secuencia que se presentó en cada sesión consistió de 200 imágenes conteniendo señal y ruido y ruido únicamente, estas imágenes fueron presentadas aleatoriamente y cada una con la misma probabilidad de presentación.

Los experimentos fueron hechos en un cuarto totalmente oscuro (con paredes negras). Los observadores fueron ubicados a una distancia de un metro de un monitor estándar que exhibió las imágenes de prueba de tal manera que el estímulo a identificar en las imágenes subtendió un ángulo visual de 1.146 grados. La duración en promedio de los experimentos fue de 15 minutos.

La secuencia de las imágenes de prueba y el registro de las respuestas de los observadores fueron controladas por la misma computadora que se utilizó para exhibir las imágenes. Al inicio de cada presentación, únicamente las imágenes de prueba aparecieron sobre el monitor de la computadora. Una vez que el observador en turno decidía dar su respuesta, presionando el botón derecho de un “mouse” estándar apareció una barra menú con las siguientes 4 posibilidades de respuesta:

- 1) Señal definitivamente presente.
- 2) Señal probablemente presente.
- 3) Señal probablemente no presente.
- 4) Señal definitivamente no presente.

A los observadores se les dieron instrucciones para que seleccionaran la opción que mejor describiera lo que veían. De esta manera, cuando una respuesta correcta fue dada por los observadores, la computadora emitió un sonido para fines de retroalimentación a los



observadores. En cuanto a los tiempo de exhibición y respuesta por parte de los observadores, éstos fueron ilimitados.

Como forma de entrenamiento para los observadores, se les aplicó una sesión corta de prueba a cada uno de ellos. Durante esta sesión se les presentaron imágenes con únicamente ruido e imágenes con señal y ruido a diferentes razones de señal a ruido, de tal manera que, en algunas imágenes, los observadores pudieron percibir claramente la señal. Continuando con la etapa de entrenamiento, los observadores realizaron un experimento corto de prueba de 60 imágenes bajo un ambiente similar al de los experimentos finales.

4.2.3.- LOS OBSERVADORES HUMANOS.

La mayoría de los observadores que participaron en los experimentos fueron estudiantes (hombres y mujeres) de licenciatura, maestría o doctorado. Solamente aquellos que presentaron una agudeza visual 20/20, natural o corregida, de acuerdo con las cartas de letras de Snellen, fueron aceptados para realizar los experimentos. Este fue la única restricción para participar. Excepto por uno de los autores, todos los observadores ignoraban la hipótesis de nuestro trabajo.

4.2.4.- VARIABLES DE ENTRADA Y SALIDA DEL SISTEMA DIFUSO.

Como se ha mencionado anteriormente, las variables de entrada del sistema fueron los índices de desempeño d'_i ($i = 1,2,3,\dots,9$) para cada 20 presentaciones y los correspondientes incrementos de dichos índices ($\Delta d'_i$). Estas dos variables proporcionan suficiente información acerca de las capacidades de percepción de los observadores en turno. De esta manera, el sistema difuso pudo controlar la energía de la señal a niveles que fueron apropiados para las capacidades de cada observador. Esto es, el sistema no proporciona valores predeterminados de la energía de la señal, sino que infiere niveles de energía apropiados en términos de lo que ha hecho el observador para obtener un 80% de respuestas correctas. Un diagrama a bloques del sistema se muestra en la Figura 4.1.





Figura 4.1. – Entradas y salidas del Sistema Difuso (E = Energía de la señal, d' = índice de desempeño).

La salida final a analizar son los valores de d' junto con sus correspondientes valores de energía que los originaron. Este conjunto de números dan información amplia sobre el desempeño de los observadores.

Una vez identificadas las variables de entrada y salida del sistema, así como los rangos para cada variable, se seleccionaron los apropiados estados lingüísticos de cada variable para expresarlas en términos de conjuntos difusos. Únicamente se requirieron 4 estados lingüísticos para las variables de entrada y 5 estados lingüísticos para la variable de salida, estos estados se muestran a continuación:

Entradas:

$\Delta d'$: NA(NH) NB(NL) PB(PL) PA(PH)

(NA = Negativo Alto, NB = Negativo Bajo, PB = Positivo Bajo, PA = Positivo Alto)

d' : MB(VL) B(L) M(M) A(H)

(MB = Muy Bajo, B = Bajo, M = Medio, A = Alto)

Salida:

E : MB(VH) B(L) M(M) A(H) MA(VH)

(MB = Muy Bajo, B = Bajo, M = Medio, A = Alto, MA = Muy Alto)



La cuantización difusa para todas las variables se realizó con gráficas de difusificación triangulares y trapezoidales.

4.2.5.- LA MÁQUINA DE INFERENCIA.

El conocimiento incluido a nuestro sistema fue formulado en términos de reglas de inferencia difusa FuzzyAnd de tipo if-then. Para una mejor definición del conocimiento del sistema difuso se analizaron las respuestas de tres observadores experimentados. Este análisis ayudó a definir mejor: a) el conjunto de reglas FuzzyAnd, b) los conjuntos difusos de las variables, y c) el apropiado rango de cada una de las variables de entrada y salida.

El conjunto de reglas difusas se muestra en la Figura 4.2.

$\Delta d'$ d'	VH	NL	PL	PH
VL	VH	VH	VH	H
L	M	M	L	L
M	M	L	L	VL
H	L	L	VL	VL

Figura 4.2. – Conjunto de reglas FuzzyAnd del sistema difuso.

Un ejemplo de una de las reglas FuzzyAnd es:

If $\Delta d'$ es NH and d' es M then E es L

Cada regla difusa fue evaluada utilizando el mínimo. El propósito de la máquina de inferencia fue combinar las variables de entrada por medio de reglas FuzzyAnd para poder hacer una inferencia acerca de los cambios de energía cada 20 presentaciones de tal manera que la magnitud de dichos cambios resultara en términos del desempeño del observador en turno durante las 20 presentaciones previas.

Finalmente, el sistema utilizó como método de defusificación el método del centroide explicado en capítulos previos.



4.3.- RESULTADOS.

Para la etapa de entrenamiento del sistema se realizaron experimentos de percepción con tres observadores. Con los resultados de estos tres experimentos se notó que el sistema estaba generando niveles de energía por debajo del umbral de detección de los observadores. Esto ocasionó que los observadores perdieran el control sobre su criterio de decisión en el resto del experimento, es decir, generó confusión en los observadores y los resultados obtenidos fueron erróneos. Para resolver esto se cambió el rango de salida de la variable E y se diseñó una nueva base de conocimientos que evitara cambios fuertes en la salida y que estuvieran por debajo del umbral de detectabilidad de los observadores.

Después de la etapa de entrenamiento del sistema, éste resultó en 16 reglas FuzzyAnd (ver Figura 4.2.) y gráficas de difusificación para las variables de entrada y salida como se muestran en las Figuras 4.3 y 4.4.

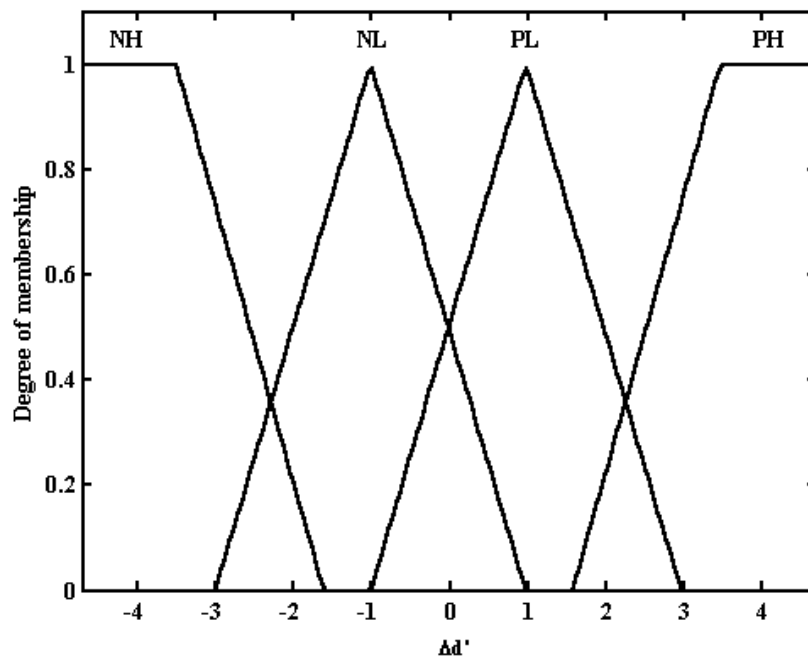


Figura 4.3a. – Gráfica de fusificación para las variables de entrada $\Delta d'$.



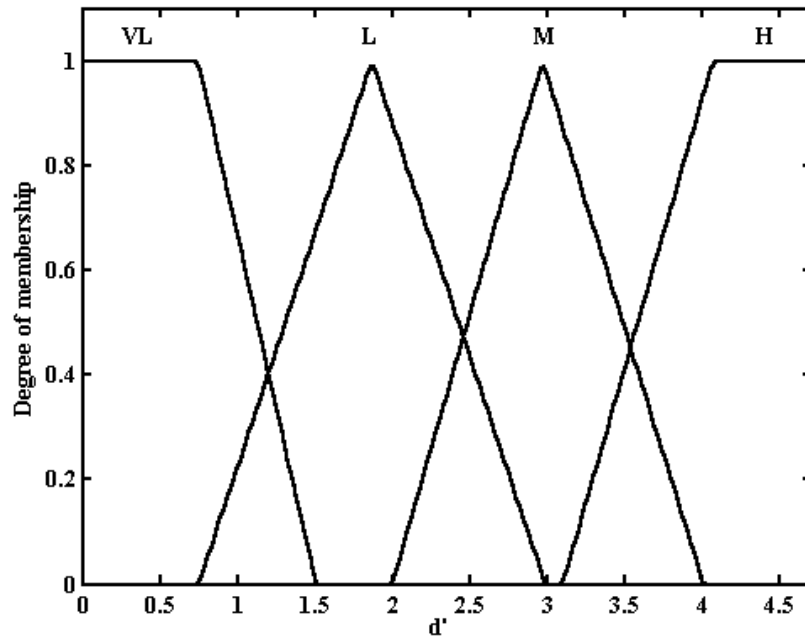


Figura 4.3b. – Gráfica de fusificación para las variables de entrada d' .

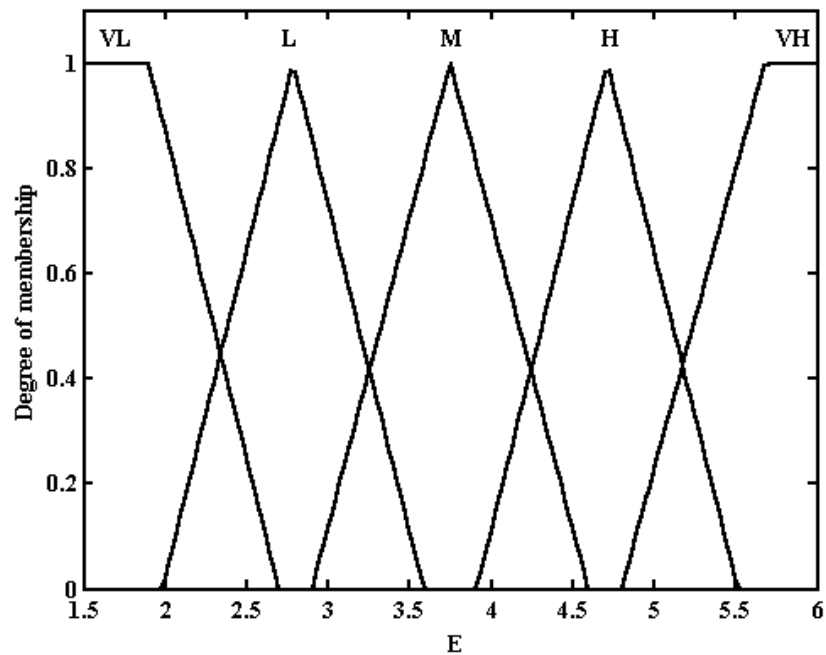


Figura 4.4. - Gráfica de fusificación para la variable de salida E .

En la Figura 4.5 es posible notar cómo la variable de salida del sistema cambia como una función de las dos variables de entrada. Como se han mencionado, la meta del sistema fue



generar cambios en la energía de la señal de tal manera que el observador en turno pudiera obtener entre un 74 y 82% de respuestas correctas.

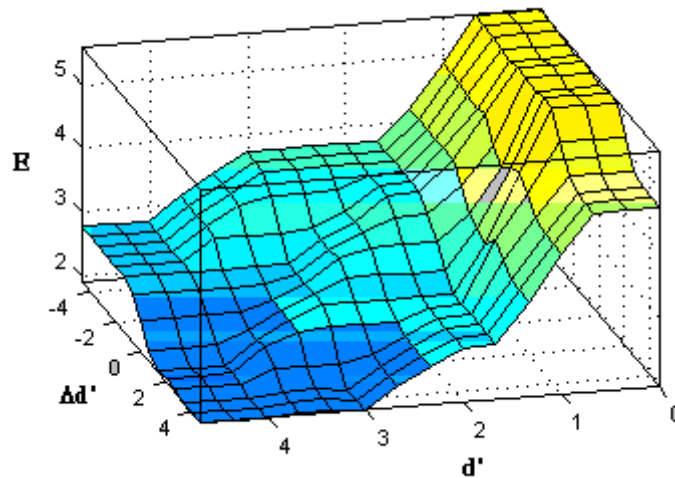


Figura 4.5. – Relación Entrada-Salida del sistema difuso.

En la Figura 4.6 se notan las variaciones de la energía de la señal, así como las variaciones de los índices de desempeño (d'), para dos de los observadores participantes. Es importante notar que las variaciones de la energía de la señal durante las primeras presentaciones fueron significativamente más grandes que las variaciones ocurridas al final del experimento. Además, de la misma Figura 4.6 se puede apreciar que el sistema difuso busca el apropiado nivel de energía para el observador en turno para obtener entre un 80 y 90% de respuestas correctas (que corresponden a una d' de 1.68 y 2.55, respectivamente).

El observador 1 (un observador muy consistente y experimentado) obtuvo una $d'=2.55$; esto significa que con los valores asignados a los parámetros del sistema no es posible mantener a éste observador en el rango de 80 a 90% de respuestas correctas. Este punto será discutido más adelante. No obstante, para observadores menos experimentados (como el observador 2) el sistema si pudo ubicarlo en el rango de porcentaje de correctas para el cual fue diseñado el sistema.



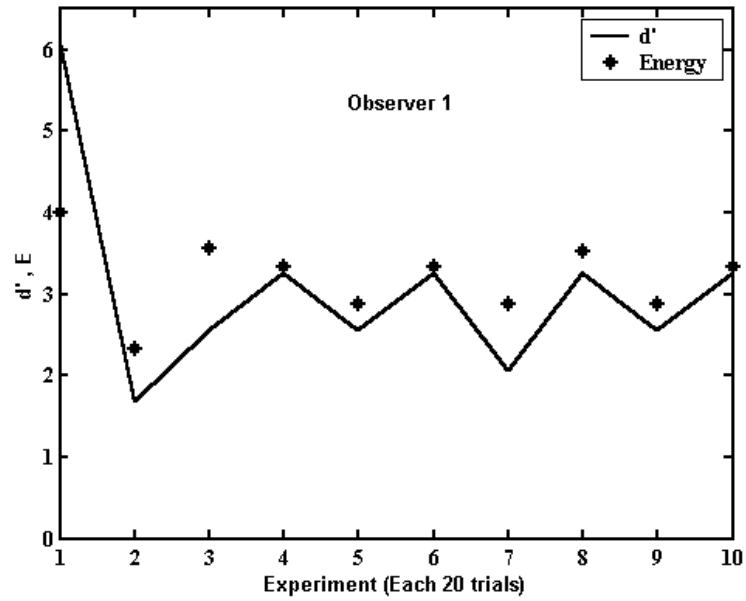
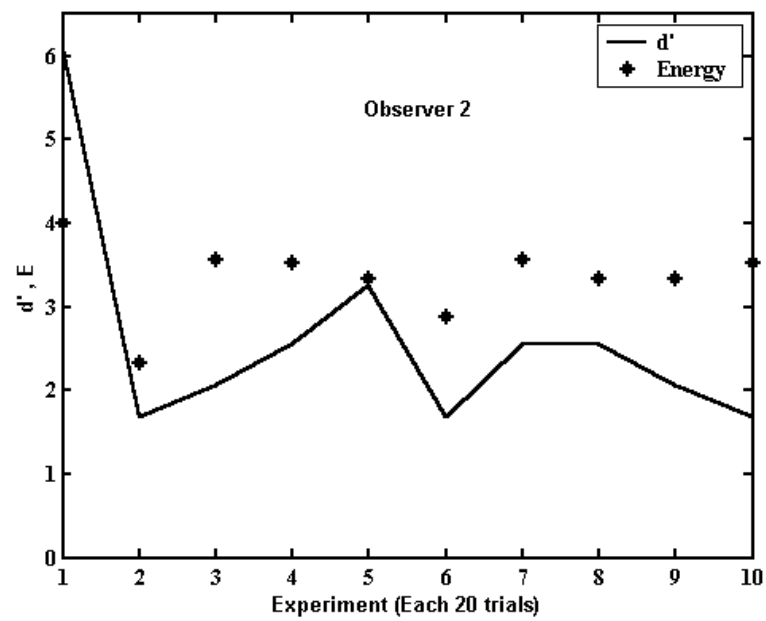


Figura 4.6ba. – Resultados de experimentos para un observador no experimentado.



b)

Figura 4.6b. – Resultados de experimentos para un observador experimentado.



Para poder determinar la naturaleza de los resultados obtenidos con el SED, en la Figura 4.7 se compara el comportamiento de uno de los observadores participantes a lo largo de un experimento tradicional de detección de señales (con una energía constante), con su comportamiento en un experimento con energía controlada por el sistema difuso.

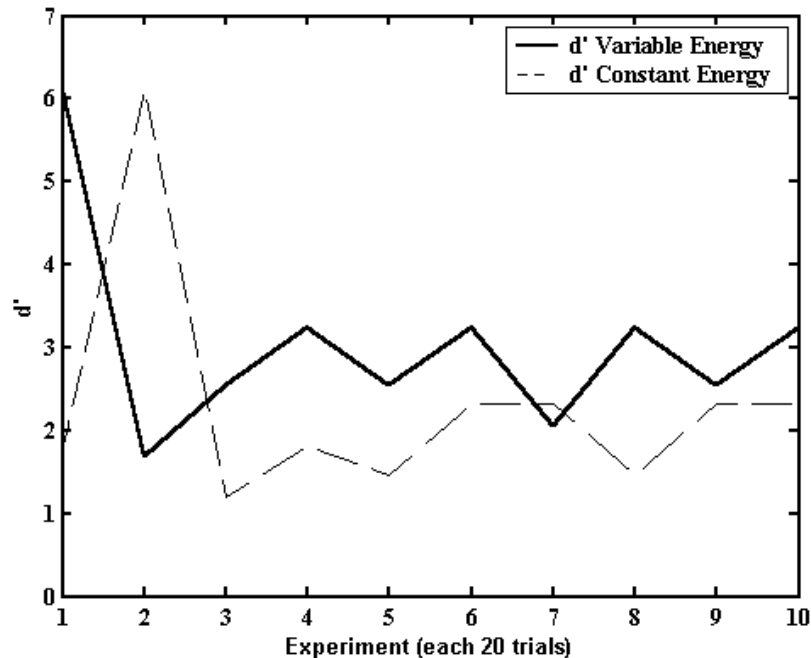


Figura 4.7. – Comparación de d' en experimentos controlados por nuestro sistema difuso (energía variable) y no controlado (energía constante)

Analizando los datos del experimento tradicional cada 20 presentaciones y, como se puede ver en la Figura 4.7, el comportamiento del observador en ambos experimentos tiene características similares. La oscilaciones del desempeño de los observadores, alrededor de un índice determinado, ocurren en ambos casos. Se esperaba que las variaciones de d' en experimentos con energía de señal fija fueran más pequeñas que aquellas ocurridas en los experimentos adaptivos, pero no fue el caso.

En la Figura 4.8 se muestran las variaciones de d' en un experimento tradicional para un observador experimentado (600 presentaciones en una sola sesión). Los efectos de la fatiga y pérdida de atención son notorios a la mitad del experimento. Es verdad que, al menos en éste caso, variaciones significativas en ambas direcciones ocurrieron tal que sus efectos sobre el índice final se cancelaron. Sin embargo, no existe la necesidad de gastar tiempo en éste tipo



de experimentos cuando estos pueden ser desarrollados en una forma más eficiente, de la forma en que el presente trabajo lo propone.

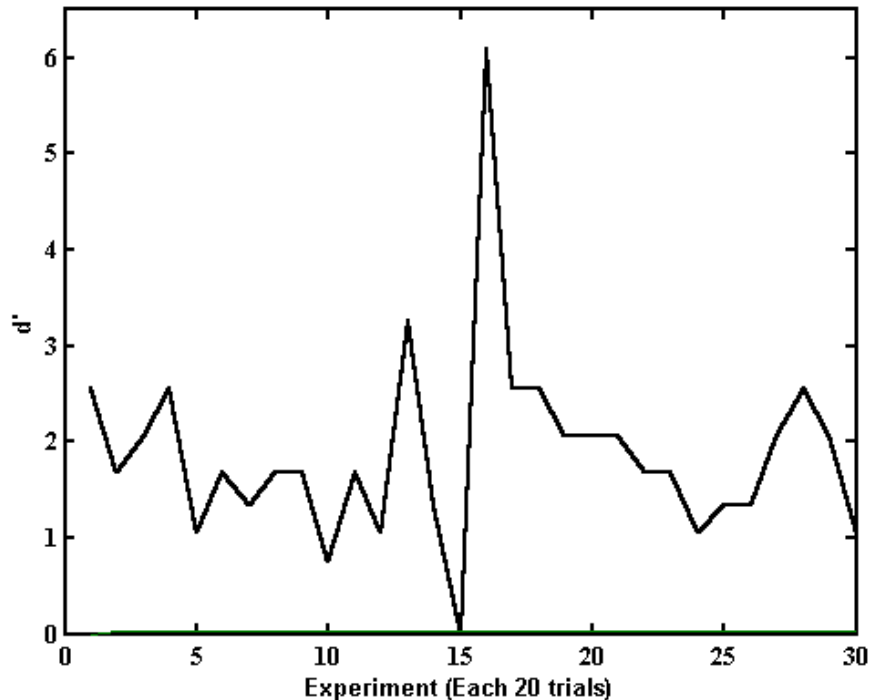


Figura 4.8. – Variaciones de d' cada 20 presentaciones para un observador experimentado en un experimento tradicional de 600 presentaciones.

Es posible tomar como índice de desempeño “final”, el promedio de los índices de desempeño en cada “experimento corto”. Este punto es tratado mas adelante en este capítulo.

Para hacer una comparación mas exhaustiva de los resultados obtenidos con el sistema difuso con respecto a los obtenidos en experimentos tradicionales, se aplicó a los datos la prueba no-parámetrica Wilcoxon [25]. La hipótesis nula fue $H_0 : \mu = \mu_0$, donde μ fue el promedio de los índices de desempeño d'_i ($i = 1, 2, \dots, 9$) a lo largo del experimento adaptivo; y μ_0 fue el índice de desempeño obtenido para el mismo observador en un experimento tradicional con 600 presentaciones y con una energía de la señal fija, cercana al umbral de detección. La hipótesis alternativa fue $H_a : \mu \neq \mu_0$. De acuerdo con la prueba de Wilcoxon, para todos los observadores (ver Tabla 4.1), el índice promedio de desempeño obtenido en experimentos cortos adaptivos es muy similar a los índices obtenidos con experimentos largos tradicionales. Es importante recalcar nuevamente que la duración de los experimentos



adaptivos fue aproximadamente un cuarto de la duración de los experimentos tradicionales con 600 presentaciones.

Observer	Data	S_+	$S_\alpha (\alpha=0.049)$	$H_0 : \mu=\mu_0$
1	$\mu = 3.049, \mu_0 = 2.802$	32	47	Accepted
2	$\mu = 2.616, \mu_0 = 1.914$	43	47	Accepted
3	$\mu = 2.390, \mu_0 = 2.136$	22	47	Accepted
4	$\mu = 3.390, \mu_0 = 3.202$	28	47	Accepted

S_+ = estadístico de prueba obtenido de los datos, S_α = estadístico de signo para $\alpha = 0.49$.

Tabla 4.1 .- Prueba Wilcoxon para d' obtenida en forma tradicional y utilizando nuestro sistema difuso.

4.4.- DISCUSIÓN.

En la psicofísica, como en otras aplicaciones, la duración de los experimentos es un parámetro importante. Por otro lado, en los experimentos de percepción visual es importante reducir la variación de los resultados y éste problema se ha atacado tradicionalmente diseñando éste tipo de experimentos que incluyan un número de presentaciones tan grande como sea posible [3,4,5,6]. No obstante, los investigadores en esta área saben que, en general, la fatiga y las pérdidas de atención afectan los resultados de experimentos largos, por esta razón, la duración de los experimentos limita la aplicabilidad de las técnicas y teorías relacionadas. El tiempo típico de realización de un experimento largo es de aproximadamente una hora. Con sistema difuso presentado el tiempo promedio de realización de los experimentos fue de 15 minutos aproximadamente y las pruebas de Wilcoxon muestran que los resultados de ambos tipos de experimentos son muy similares.

Tradicionalmente, las oscilaciones de los índices de desempeño que ocurren a lo largo de cada experimento (con energía de señal fija) no son tomadas en cuenta. No obstante, tales oscilaciones pueden contener información importante acerca de las capacidades de percepción de los observadores. Con el sistema desarrollado es más fácil acceder a esta información de la forma en como lo hicimos con el índice propuesto (el promedio de los índices de desempeño calculados cada 20 presentaciones a lo largo de un experimento). Esto es, con el sistema propuesto, es posible obtener mejores resultados.



El sistema aquí propuesto fue desarrollado para ser utilizado con observadores experimentados y no experimentados. No obstante, los observadores no experimentados, en general, requieren niveles de energía de la señal mas grandes que aquellos requeridos por observadores no experimentados. Cuando se tuvo que definir el posible rango de energía producido por el sistema se decidió utilizar una energía mínima relativamente alta . Esta es la razón por la que, cuando un observador experimentado fue probado (Figura 4.6) el sistema no pudo inferir una energía de la señal que llevaría al observador a obtener el 80% de respuestas correctas. No obstante, se pudo evaluar a éste observador con un experimento adaptivo corto.

Los sistemas de control difusos son normalmente aplicados a situaciones que involucran variables que son afectadas por un número limitado de factores. En el caso de observadores humanos en una tarea de detección visual, muchos factores pueden tomar un papel importante en su desempeño (fatiga, sed, hambre, estado de salud, emociones, etc.). Como una buena alternativa se propone un sistema difuso para controlar el desarrollo de experimentos de percepción de tal manera que, con experimentos adaptivos cortos, se puedan obtener mejores evaluaciones que aquellas obtenidas con los experimentos tradicionales largos.



BIBLIOGRAFÍA.

1. H. H. Barrett, Objective assessment of image quality: effects of quantum noise and object variability, *JOSA*, 7, (1990), 1266-1277.
2. J. Bezdek, Fuzzy Models -what are they and why?, *IEEE trans. On Fuzzy Systems*, vol. 1, no.1, (1993), 1-5.
3. A. Burgess and B. Colborne, Visual Signal Detection. IV. Observer Inconsistency, *J. Optics Society American A*, 5, (1988) 617-627.
4. A. Burgess, Comparison of receiver operating characteristic and forced choice observer performance measurement methods, *Med. Phys.*, 22, (1995), 643-655.
5. A. Burgess, Image quality, the ideal observer, and human performance of radiologic decision tasks, *Acad. Radiol.*, 2, (1995), 522-526.
6. A. Burgess, On sampling statistics in Observer Performance Studies, *Science and Engineering of Medical Imaging*, SPIE vol. 1137, (1989), 190-197.
7. J. P. Egan, A.I. Schulman, and G.Z. Greenberg, Operating characteristics by binary decisions and by ratings, *J. Acoust. Soc. Am.*, 31,(1959), 768-773.
8. J. P. Egan, *Signal Detection Theory and ROC Analysis*, Academic Press, New York, 1975.
9. S. Galichet and L. Fulloy. Fuzzy Controllers: Synthesis and equivalences. *IEEE Trans. Fuzzy Syst.*, vol. 3, no 2, (1995) 140-148.
10. D. M. Green and J.A. Swets, *Signal Detection Theory and Psychophysics*, John Wiley and Sons, New York, 1966.
11. J.L. Hall, Hybrid adaptive procedure for estimation of psychometric functions. *JOSA*, 69, (1981), 1763- 1769.
12. N. Infelise, A clear vision of fuzzy logic, *Control Engineering*, vol. 38, no. 9 (1993), 28-30.
13. A. Kandel, *Fuzzy Expert Systems*, CRC Press, Tampa Florida, 1991.
14. B. Kosko, *Fuzzy Engineering*, Prentice Hall, U.S.A, 1997.
15. B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice Hall, U.S.A., 1992.
16. Madigan and D. Williams, Maximum likelihood psychometric procedures in two-alternative forced choice: Evaluation and recommendations. *Perception and Psychophysics*, 42, (1987), 240-249.
17. J. Markowitz and J.A. Swets, Factors affecting the slope of empirical ROC curves:



- comparison of binary and rating responses, *Percept. Psychoph.*, 2, (1967), 91-100.
18. C.E. Metz, ROC methodology in radiological imaging, *Invest. Radiol.*, 21, (1986), 720 - 733.
 19. S. Russell and P. Norving, *Inteligencia Artificial: un enfoque moderno*, Prentice Hall, México, D.F., 1996.
 20. J.P. Sánchez and Y. Beltran, *Los Sistemas Expertos: Una Metodología de Programación*, Macrobit Corporation, México, 1990.
 21. F.J. Sanchez-Marin, D.L. Wilson, Y. Srinivas, and K.N. Jabri, Quantitative Image Quality Analysis of a Non-Linear Spatio-Temporal Filter. *IEEE Trans. Image Processing*. vol. 10, no. 2 (2001) 288-295.
 22. W.A. Simpson, The step method: A new adaptive psychophysical procedure. *Perception and Psychophysics*, 45 (1989), 572-576.
 23. M. Sugeno, A Fuzzy Logic Based Approach to Qualitive Modeling, *IEEE trans. On Fuzzy Systems*, vol 1, no. 1, (1993), 7-31.
 24. J. A. Swets, *Signal Detection and Recognition by Human Observers*, John Wiley and Sons, New York, 1964.
 25. J. A. Swets, *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics*, Lawrence Erlbaum Associates, New Jersey, 1996.
 26. A.B. Watson, A.B. and Pelli, D.G., QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics*, 33, (1983), 113-120.
 27. G. Weibenberg, J. Schumaker, and D. Oltman, *Statistic: An Intuitive Approach*, Wadsworth, California, USA, 1981
 28. D. J. Weintraub and H. W. Hake, Visual Discrimination, an interpretation in Terms of Detectability Theory, *Journal of the Optical Society of America*, 52, (1962) 1179-1184.
 29. K.A. Wen, C.Y. Lu, and M.C. Tsai. The transform Image Codec based on Fuzzy control and human visual system. *IEEE Trans. Fuzzy Sys*, vol. 3, no 3, (1995) 253-259.
 30. C.D. Wickens, *Engineering psychology and human performance*, New Jersey, Prentice Hall, 2000.
 31. P. Xue, C.W. Thomas, G.C. Gilmore, and D.L. Wilson, An adaptive reference/test paradigm: Application to pulsed fluoroscopy perception. *Behavior Research Methods, Instruments and Computers*, 2 (1998), 332-348.



32. L. Zadeh, Fuzzy Sets, information and control, Academic Press, 3, vol. 8, (1965), 338-353.



Centro de Investigaciones en Óptica, A. C.
Universidad de Guanajuato

