



INSTITUTO TECNOLÓGICO SUPERIOR DE MISANTLA

MAESTRÍA EN SISTEMAS COMPUTACIONALES

"DISEÑO DE UNA ARQUITECTURA
MULTIAGENTE DE SEGURIDAD
ACÚSTICA APLICADA AL ÁREA
PERIMETRAL DE UN INMUEBLE"

TESIS

QUE PARA OBTENER EL TÍTULO DE
MAESTRO EN SISTEMAS
COMPUTACIONALES

PRESENTA
JOSÉ ANTONIO HIRAM VÁZQUEZ
LÓPEZ

ASESOR
DR. LUIS ALBERTO MORALES
ROSALES

MISANTLA, VERACRUZ

SEPTIEMBRE, 2016



INSTITUTO TECNOLÓGICO SUPERIOR DE MISANTLA
DIVISIÓN DE ESTUDIOS PROFESIONALES
AUTORIZACIÓN DE IMPRESIÓN DE TRABAJO DE TITULACIÓN MAESTRÍA

FECHA: 07 de Septiembre de 2016.

ASUNTO: AUTORIZACIÓN DE IMPRESIÓN
DE TESIS.

A QUIEN CORRESPONDA:

Por medio de la presente se hace constar que el (la) C:

JOSÉ ANTONIO HIRAM VÁZQUEZ LÓPEZ

estudiante de la maestría en SISTEMAS COMPUTACIONALES con No. de Control 122T0563 ha cumplido satisfactoriamente con lo estipulado por el Lineamiento de Posgrado para la obtención del grado de Maestría mediante Tesis.

Por tal motivo se Autoriza la impresión del Tema titulado:

**DISEÑO DE UNA ARQUITECTURA MULTIAGENTE DE SEGURIDAD
ACÚSTICA APLICADA AL ÁREA PERIMETRAL DE UN INMUEBLE**

Dándose un plazo no mayor de un mes de la expedición de la presente a la solicitud del examen para la obtención del grado de maestría.

ATENTAMENTE


Dr. Luis Alberto Morales Rosales.
Presidente


M.I.A. Roberto Ángel Meléndez Armenta
Secretario


Dr. Ignacio Algreto Badillo
Vocal

Archivo.

VER. 01/03/09

F-SA--39

Resumen

En la actualidad existe un sin número de tipos de sistemas de vigilancia que aplican diversas estrategias encaminadas a maximizar la efectividad de la detección de situaciones potenciales que ponen en riesgo la seguridad de los inmuebles, y cada uno de ellos tienen deficiencias que reducen su margen de efectividad.

Una característica importante que debería estar incluida en los sistemas de vigilancia es la capacidad de censar los sonidos generados en el área cercana del perímetro de un inmueble, para realizar un análisis que permita detectar ruidos que son potencialmente sospechosos y que pueden representar una amenaza para el edificio que se resguarda.

En este documento, se describen las actividades realizadas para diseñar una arquitectura multiagente de seguridad acústica que atiende la necesidad actual de la sociedad de salvaguardar la seguridad de las personas que se encuentran dentro de un inmueble, o de los bienes materiales almacenados dentro de un edificio, ante el constante incremento del índice de inseguridad y violencia que se relaciona con el acceso no autorizado al área perimetral de los inmuebles.

Dentro de las tareas llevadas a cabo, primero se diseñó la etapa que realiza el pre procesamiento de los sonidos obtenidos para generar las características que describen el comportamiento de la señal acústica censada.

Se diseñó la estructura de un agente reactivo y de un agente deliberativo que componen a la arquitectura multiagente de seguridad acústica, y que tienen la finalidad de interpretar las características de los coeficientes matemáticos de la señal pre procesada, para distinguir entre eventos acústicos normales y eventos acústicos potencialmente sospechosos.

En ambos casos, la finalidad de cada uno de los agentes, es identificar ruidos que impliquen una amenaza probable de intrusión al área perimetral de un inmueble, para activar una alarma que permitirá corroborar con el uso de cámaras de video vigilancia la naturaleza potencial de la amenaza analizada.

Posterior al diseño, se llevó a cabo una etapa de adiestramiento para el agente reactivo y el agente deliberativo, utilizando un conjunto de sonidos seleccionados por sus características para el proceso de entrenamiento.

En consecuencia a la etapa de adiestramiento, se efectuó una serie de experimentos que pusieron a prueba las capacidades especificadas de cada uno de los componentes del sistema diseñado, con un grupo de sonidos elegidos para realizar pruebas y medir el grado de eficiencia de la arquitectura multiagente de seguridad acústica.

Finalmente se realizaron pruebas de cálculo y validación, para estimar estadísticamente la eficiencia de la arquitectura multiagente de seguridad acústica. En el caso del agente reactivo se obtuvo como resultado un rango de efectividad que oscila entre un 50% y 58%. Lo que significa que este agente reconoce al menos cinco de cada diez sonidos analizados. Para el agente deliberativo, se obtuvo como resultado un rango de efectividad que fluctúa entre un 66% y 74% que implica que este agente es capaz de reconocer al menos seis de cada diez sonidos.

Índice

Capítulo I.....	1
1.1. Introducción	2
1.2. Descripción del problema.....	4
1.3. Justificación	10
1.4. Objetivos	11
1.4.1. General	11
1.4.2. Específicos.....	11
1.5. Hipótesis	11
1.6. Metodología	11
1.7. Propuesta de solución	12
Capítulo II.....	14
2. Estado del arte.....	15
2.1. Sistemas de vigilancia basados en audio	15
2.2. Sistemas de vigilancia basados en video.....	18
2.3. Sistemas de vigilancia basados en audio y video	20
2.4. Descripción de características del estado del arte.....	21
Capítulo III.....	23
3. Descripción de la arquitectura de seguridad acústica aplicada al área perimetral de un inmueble.....	24
3.1. Pre procesamiento de la señal.....	25
3.2. Sistema multiagente	26
3.3. Agente Reactivo.....	26
3.4. Agente Deliberativo.....	27
Capítulo IV.....	29
4. Proceso de diseño	30
4.1. Formación del corpus de entrenamiento / validación	30
4.2. Pre procesamiento	31
4.3. Diseño del Agente Reactivo	32
4.4. Diseño del Agente Deliberativo	32
Capítulo V.....	35

5.	Análisis de los resultados	36
5.1.	Resultados del agente reactivo	36
5.2.	Resultados del agente deliberativo	38
5.3.	Contrastación de los resultados con la hipótesis	45
5.4.	Resultado.....	45
5.5.	Conclusiones	47
5.6.	Trabajo futuro.....	48
	Referencias	49

Índice de figuras

Figura 1. Arquitectura multiagente de seguridad acústica.....	24
Figura 2. Proceso de formación del <i>corpus</i> de trabajo.....	30
Figura 3. Diseño y desarrollo de agente reactivo.....	32
Figura 4. Diseño y desarrollo de agente deliberativo.....	33
Figura 5. Estadístico t-student mejor caso.....	40
Figura 6. Estadístico t-student caso promedio.....	42
Figura 7. Estadístico t-student peor caso.....	44
Figura 8. Mejora de la arquitectura multiagente de seguridad acústica.....	48

Índice de gráficas

Gráfica 1. Encuesta ENVIPE, año 2011.....	4
Gráfica 2. Encuesta ENVIPE, año 2012.....	4
Gráfica 3. Encuesta ENVIPE, año 2013.....	5
Gráfica 4. Encuesta ENVIPE, año 2014.....	5
Gráfica 5. Encuesta ENVIPE, año 2015.....	6
Gráfica 6. Encuesta ENVIPE, año 2015, delitos no denunciados.....	6
Gráfica 7. Encuesta ENVE, año 2011.....	7
Gráfica 8. Encuesta ENVE, año 2012.....	7
Gráfica 9. Encuesta ENVE, año 2015, delitos no denunciados.....	8
Gráfica 10. Resultados de pre procesamiento.....	31
Gráfica 11. Dispersión de los datos.....	32

Índice de tablas

Tabla 1. Principales características de trabajos relacionados.....	22
Tabla 2. Características de la arquitectura diseñada.	22
Tabla 3. Agente reactivo, mejor caso.	36
Tabla 4. Agente reactivo, caso promedio.	37
Tabla 5. Agente reactivo, peor caso.....	37
Tabla 6. Agente deliberativo, mejor caso.	39
Tabla 7. Agente deliberativo, caso promedio.	41
Tabla 8. Agente deliberativo, peor caso.	43
Tabla 9. Resultados generales.....	45

Capítulo I

1.1. Introducción

El índice nacional que mide la capacidad delictiva en el país se incrementa constantemente, así lo demuestra el INEGI, en dos instrumentos que aplica a nivel nacional, para estimar la cantidad de delitos perpetrados en diferentes modalidades. Una categoría de ilícito que aumenta incesantemente es el de robo a casa habitación, o a edificios de empresas y fábricas de la industria.

De acuerdo a la Encuesta Nacional de Victimización y Percepción sobre seguridad pública (ENVIPE) que se emplea anualmente desde el año 2011 para medir los delitos más representativos del fuero común, el valor del porcentaje de ilícitos perpetrados en casa habitación se acrecienta sin detenerse. En otro sentido la Encuesta Nacional de Victimización de Empresas (ENVE), usada desde el año 2011, con una periodicidad de dos años, demuestra el incremento del valor porcentual de los delitos cometidos a empresas o sus bienes, como robo o asalto de mercancía, dinero, insumos o bienes, y daños a las instalaciones, maquinaria y equipo.

Para evitar que un inmueble se vea afectado por crímenes que ponen en peligro no sólo los bienes y valores resguardados en ellos, sino también a las personas que los ocupan, se pueden implementar varias técnicas de seguridad y vigilancia. Algunas opciones disponibles que se pueden utilizar, son por ejemplo, el empleo de rejas electrificadas, el uso de sensores de movimiento, o sistemas de circuito cerrado de video vigilancia, entre otras.

Cada opción que se puede implementar, presenta un conjunto de limitantes que reducen el nivel de su eficiencia. Por ejemplo, en el caso de la solución que considera el video como un elemento que aporta mayor confianza en su desempeño, existe la posibilidad de que la imagen del video sea alterada, que la ubicación de las cámaras de video no cubran la totalidad de áreas importantes que deben ser vigiladas constantemente, que el personal encargado de vigilar el video obtenido por las cámaras desatienda su actividad por descuido, aburrimiento o desidia, y además carecen de la capacidad de alertar cuando existe un sonido inesperado que altera el entorno del área limítrofe del perímetro del inmueble que supervisan.

Con el fin de ofrecer una alternativa que disminuya los riesgos de la vigilancia asistida, en este trabajo se propone el diseño de una arquitectura de seguridad acústica multiagente autónoma, que combine el uso de sensores acústicos para conocer el estado acústico del medio ambiente del área perimetral de un inmueble. El uso de un agente reactivo y de un agente deliberativo para juzgar cuando el entorno sufre una perturbación que represente un ruido que simboliza la probabilidad de una intrusión a un área perimetral, que pone en peligro la integridad de las personas que ocupan un edificio, o de los bienes almacenados en el inmueble.

Este documento está dividido en cinco capítulos. En el capítulo uno, se explica la importancia del diseño de la arquitectura de seguridad acústica, partiendo de la necesidad de reducir el impacto de la delincuencia en la seguridad de edificios e inmuebles. En el capítulo dos, se analizan un conjunto de trabajos relacionados, que definen el marco teórico que acotan la relación de los alcances logrados anteriormente. En el capítulo tres, se describe el diseño de la arquitectura de seguridad acústica desde el punto de vista de los conceptos que permiten su funcionamiento. En el capítulo cuatro, se describe el diseño de la arquitectura de seguridad acústica desde el punto de vista del

funcionamiento de sus elementos. Y en el capítulo cinco, se analizan los resultados obtenidos en la etapa de experimentación y pruebas, además de presentar las conclusiones y el futuro del trabajo realizado.

1.2. Descripción del problema

En México el índice de delitos cometidos, tan sólo en el rubro de casas habitación ha incrementado considerablemente, generando un ambiente de inestabilidad social e inseguridad.

El Instituto Nacional de Estadística y Geografía (INEGI) realiza desde el año de 2011 la Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE) que mide los delitos más representativos del fuero común.

El sondeo reveló que en el año 2011 los robos cometidos a inmuebles de casa habitación se situaron en un 6.9% con respecto a todos los delitos considerados por la encuesta ENVIPE (Gráfica 1).



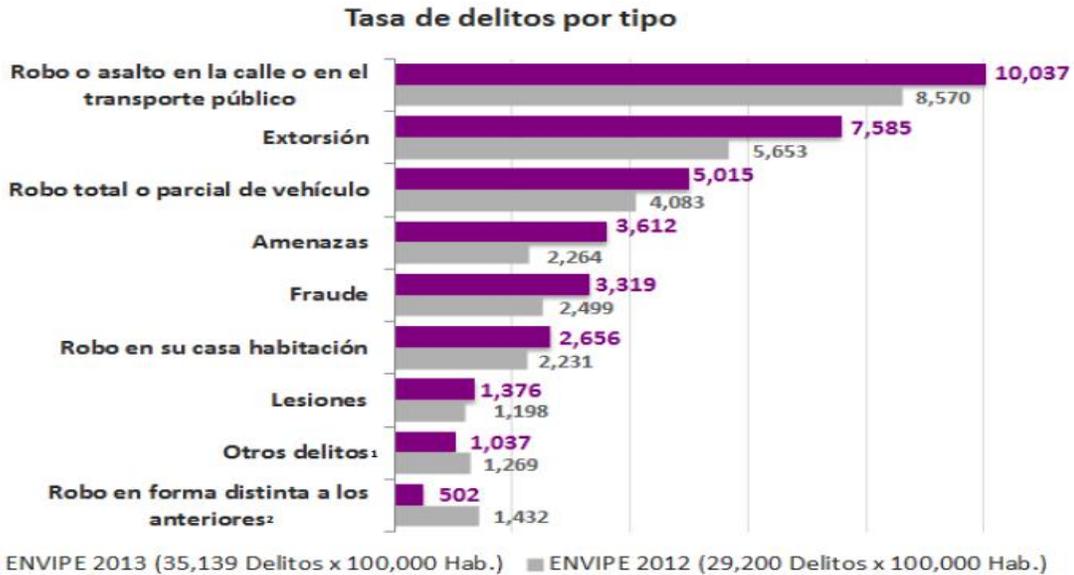
Gráfica 1. Encuesta ENVIPE, año 2011.

La representatividad del delito de robo, cometido en hogares en el año 2012, incrementó su porcentaje en 0.7% con respecto a la medición del año 2011, situándose en un 7.6% general en relación con todos los delitos considerados por la encuesta (Gráfica 2).



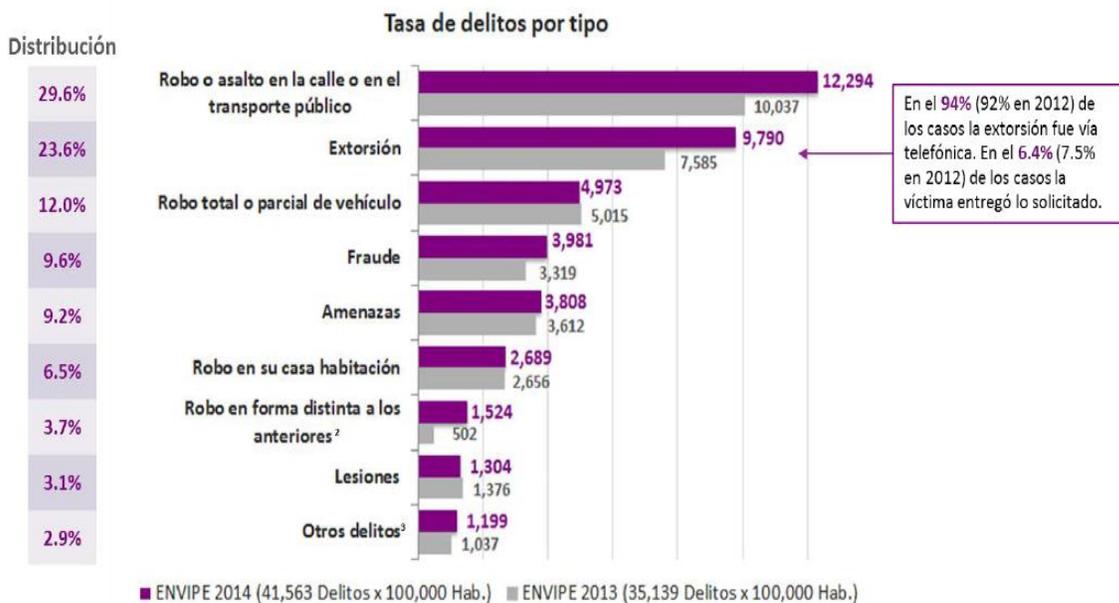
Gráfica 2. Encuesta ENVIPE, año 2012.

Para el año 2013, el delito de robo cometido a casa habitación creció en una proporción de 0.9% en cuanto al sondeo del año 2012, contabilizándose el registro de 2,656 robos en casa habitación, ubicándose en un valor general del 8.5% con respecto a todos los delitos considerados por la encuesta ENVIPE (Gráfica 3).



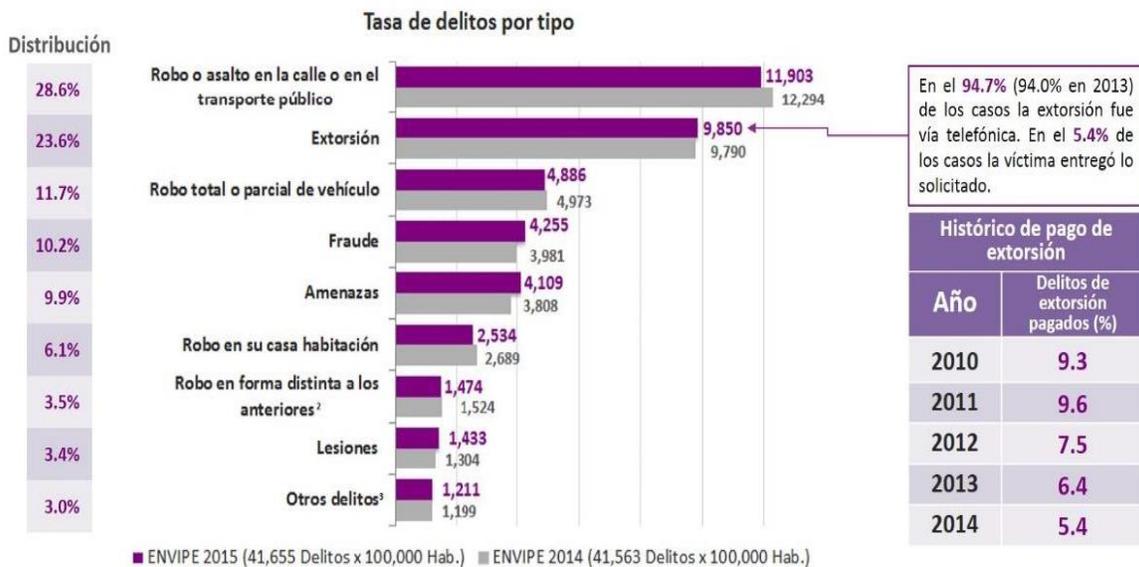
Gráfica 3. Encuesta ENVIPE, año 2013.

En el año 2014, el delito de robo cometido a casa habitación presentó un valor de 2,689 delitos cometidos, representado el 6.5% con respecto a todos los delitos considerados por la encuesta ENVIPE (Gráfica 4).



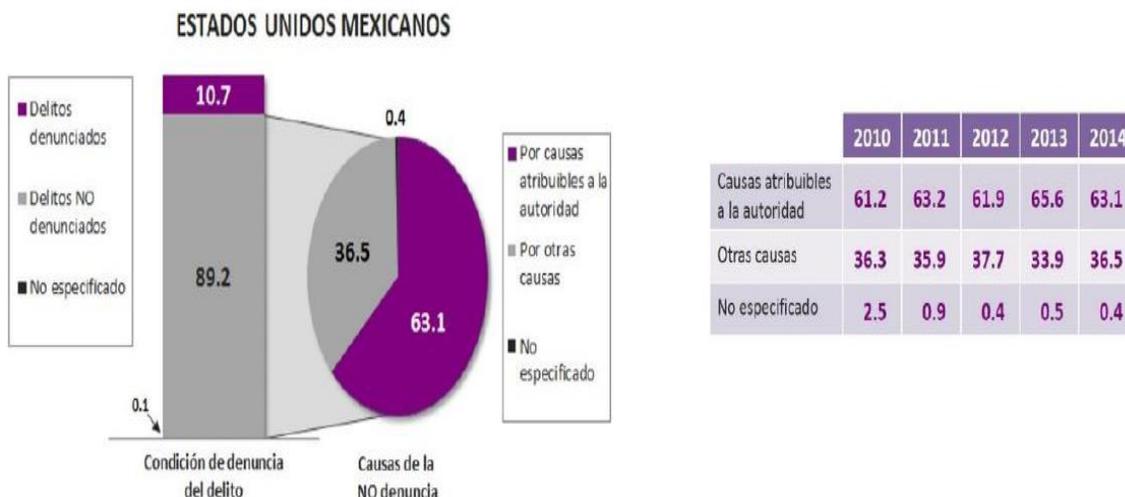
Gráfica 4. Encuesta ENVIPE, año 2014.

El sondeo reveló que en el año 2015 los actos ilícitos de robo denunciados a inmuebles de casa habitación se situaron en un 6.1%, presentándose un valor de 2,534 robos con respecto a todos los delitos considerados por la encuesta ENVIPE (Gráfica 5).



Gráfica 5. Encuesta ENVIPE, año 2015.

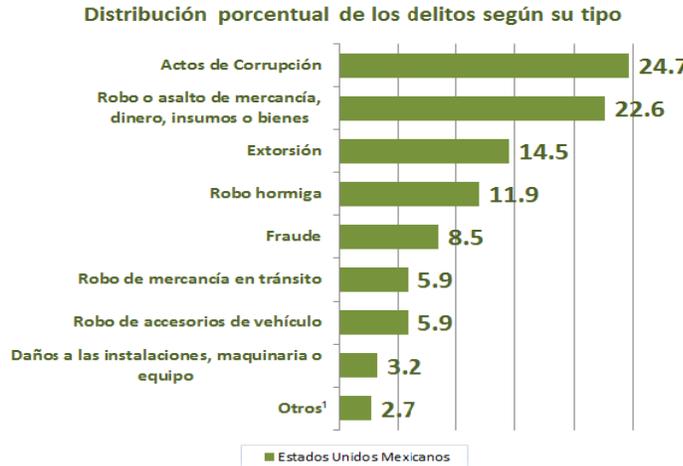
También existe la evidencia de que las cifras presentadas en los resultados anteriores pueden ser mayores, esto se concluye porque la encuesta ENVIPE también mide la estimación del número de delitos que no son denunciados, sólo en el año 2015 se considera que al menos el 80% de los ilícitos perpetrados no fueron presentados ante la autoridad competente (Gráfica 6).



Gráfica 6. Encuesta ENVIPE, año 2015, delitos no denunciados.

A partir del año 2011 con una periodicidad de cada dos años, el INEGI realiza la Encuesta Nacional de Victimización de Empresas (ENVE), describiendo que en el año 2011 se cometieron 4,324,529 delitos, de los cuales el delito de "Robo o asalto de mercancía, dinero, insumos o bienes" se contabilizó con un valor porcentual de 22.6%, y el delito de "Daños a las instalaciones, maquinaria o equipo" se reflejó con un porcentaje de 3.2% en comparación al total de delitos considerados para realizar la muestra (Gráfica 7).

La ENVE estima que durante 2011 se generaron 4,324,529 delitos asociados a 1,389,000 unidades económicas víctima (3.1 delitos por víctima).



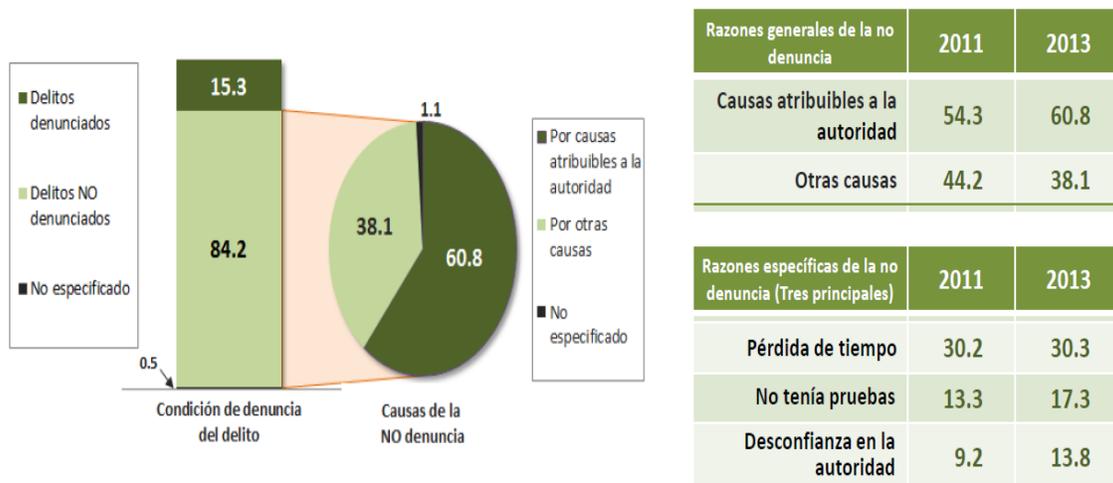
Gráfica 7. Encuesta ENVE, año 2011.

En la presentación de los resultados de la Encuesta Nacional de Victimización de Empresas (ENVE), en el año 2013 se cometieron 2.5 millones de delitos, de los cuales el delito de "Robo o asalto de mercancía, dinero, insumos o bienes", aumentó contabilizándose un valor porcentual total de 23.1%, y el delito de "Daños a las instalaciones, maquinaria o equipo" se reflejó con una disminución al porcentaje anterior con un valor de 1.7% en comparación al total de delitos considerados para realizar la muestra (Gráfica 8).



Gráfica 8. Encuesta ENVE, año 2012.

Como sucede con la Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE), la encuesta ENVE, también demuestra que existe una cantidad de ilícitos sin denunciar, estimando que sólo en el año 2015 no se presentaron a la autoridad pertinente al menos el 80% de los delitos cometidos (Gráfica 9).



Gráfica 9. Encuesta ENVE, año 2015, delitos no denunciados.

Los datos estadísticos mostrados anteriormente, describen la presencia ascendente del índice delictivo en la modalidad de robo o daño que incide tanto a inmuebles de casa habitación como a edificios de organizaciones o empresas.

Los resultados obtenidos de las encuestas aplicadas por el INEGI en los rubros de los delitos cometidos del fuero común y en empresas, revelan que a pesar de intentar proteger el área perimetral de los inmuebles para evitar intrusiones, los esfuerzos de prevención empleando estrategias de vigilancia son insatisfactorios, no lográndose contener el incremento de los índices de inseguridad.

Los medios y recursos que se tienen para implementar sistemas de vigilancia son deficientes e inadecuados ante el constante incremento en la sofisticación de las capacidades de la delincuencia organizada.

La inversión en complejos sistemas de seguridad con costos altos y privativos, condicionan su uso en sectores de la población con capacidad adquisitiva mayor, excluyendo a estratos sociales con menor solvencia económica situándolos en posición de mayor vulnerabilidad.

En la prevención de delitos que se pueden cometer en edificios e inmuebles de casa habitación, empresas, fábricas e instituciones de diversa índole, se pueden usar técnicas de seguridad que pueden ser consideradas desde elementales o básicas, hasta acciones bastante sofisticadas.

Algunas de las técnicas empleadas son el uso de perros guardianes, vigilancia por recurso humano, rejas electrificadas, y el empleo de sensores en circuitos cerrados de cámaras de video o micrófonos ambientales, entre otros.

También es posible implementar estrategias para coadyuvar en la eficiencia de los sistemas de seguridad. Estas estrategias pueden implicar buscar mejores puntos de ubicación para emplazar la posición de sus elementos y sensores, o el incremento del número de recursos que se emplean originalmente en la técnica.

Aunque existe un amplio abanico de posibilidades para implementar sistemas de seguridad y vigilancia, cada tipo, variante o combinación que se puede emplear para la protección de edificios e inmuebles enfrentan diversos tipos de deficiencias y retos que pueden ser clasificados bajo los siguientes criterios:

Recurso humano

- Es propenso a sufrir aburrimiento, distracción, o fatiga involuntaria.
- Presentar desatención o descuido de su área de responsabilidad.
- Verse influenciado por condiciones climatológicas o ambientales.
- Dar a conocer las rutinas y horarios que emplea para realizar los recorridos de vigilancia.
- Emplear criterios propios como intuición o experiencia para juzgar eventos extraordinarios.

Recurso Tecnológico

- Ininterrupción del suministro eléctrico.
- Costos privativos de la tecnología a usar.
- Complejidad en la configuración del hardware.
- Rangos de tolerancia en situaciones climatológicas o ambientales.
- Ubicación en un lugar y posición adecuada de sensores.
- Complejidad en la interconexión inherente del diseño de la arquitectura de las redes de sensores.

Los sistemas de vigilancia actuales no consideran escuchar los sonidos en el entorno, con la intención de analizarlos, cuantificar los valores que genera el espectro de su señal con respecto a ciertos umbrales y clasificarlos, para valorar la probabilidad de que un ruido represente una amenaza potencial, y poder llevar a cabo la comprobación visual de la naturaleza del ruido.

Otro elemento que no contemplan los sistemas de seguridad, es la capacidad de recordar cuando un ruido se ha presentado y representó una amenaza potencial, primero para descartar los sonidos del ambiente que se puede considerar comunes para disminuir las falsas alarmas derivadas de una mala interpretación; y segundo para centrar su atención en todos los sonidos desconocidos que sí pueden representar una nueva amenaza de intrusión.

1.3. Justificación

El siglo XXI, es una época que se distingue por cambios profundos que marcan la evolución de la sociedad.

Estas transformaciones incrementan la complejidad que existe entre los factores que describen la relación de la interacción humana en el ámbito social, económico y cultural.

La diversidad de los perfiles culturales y económicos que pueden estar presentes en el medio ambiente de una ciudad se magnifica por un ecosistema social cambiante y agresivo.

En la heterogeneidad de los entornos urbanos existe un signo oscuro, representado por la inseguridad prevaleciente en variadas modalidades, que contribuyen a formar un estado de caos e incertidumbre.

En este sentido, los edificios de inmuebles destinados a actividades diferentes como la industria, empresarial o casa habitación no están exentos de sufrir diferentes tipos de ilícitos, y aunque es posible aplicar técnicas y estrategias de seguridad para darles protección, presentan carencias que los hacen vulnerables.

Un aspecto sensible de los sistemas de seguridad que se pueden implantar para restringir el acceso al espacio de un edificio, es la falta de inteligencia para poder juzgar cuando un ruido generado en el exterior representa la posibilidad de un ataque o intrusión.

Es por eso, que se propone el diseño de una arquitectura de seguridad acústica que sense los sonidos generados en el medio ambiente cercano al área perimetral de un inmueble, para clasificar la naturaleza de los ruidos.

Cuando la arquitectura considere que existe la potencialidad de que se presente un evento de delincuencia, activar una alarma para que se realice una inspección visual a través de cámaras de video vigilancia, corroborar la existencia de la intención voluntaria de penetrar al interior de un inmueble sin autorización, y en consecuencia dar aviso a las autoridades correspondientes.

1.4. Objetivos

1.4.1. General

Diseñar una arquitectura de seguridad acústica multiagente basada en la detección y clasificación de eventos acústicos sospechosos, para identificar intrusiones al área perimetral de un inmueble.

1.4.2. Específicos

- Diseñar el procedimiento que se empleará en el pre procesamiento de los ruidos del ambiente, para obtener los coeficientes que representan las características acústicas sobresalientes de la señal analizada.
- Diseñar el comportamiento de un agente reactivo de la arquitectura de seguridad acústica, para clasificar los ruidos generados en el ambiente y determinar si se trata de una amenaza potencial de intrusión.
- Diseñar el comportamiento de un agente deliberativo de la arquitectura de seguridad, para para clasificar los ruidos generados en el ambiente y determinar si se trata de una amenaza potencial de intrusión.

1.5. Hipótesis

Es posible reconocer la mayoría de los ruidos generados en el ambiente, dentro de los cuales es probable que exista una amenaza de intrusión al área perimetral de un inmueble usando una arquitectura multiagente, clasificando los sonidos con un agente reactivo y con un agente deliberativo.

1.6. Metodología

Los esquemas de seguridad de circuitos cerrados de vigilancia que funcionan empleando cámaras de video vigilancia no están diseñados para reaccionar al presentarse ruidos sospechosos que se generan alrededor de su entorno.

Para atender este problema se propone el diseño de una arquitectura de seguridad acústica que analice los ruidos del área perimetral de un inmueble para disminuir los espacios de tiempo en que algunas zonas quedan desprotegidas y son blancos potenciales de que se cometa un ilícito de ingreso o penetración no autorizada al inmueble.

- a) Se plantea que la estructura de la arquitectura de seguridad acústica esté conformada por un grupo de sensores acústicos. Cada sensor está asociado con una etapa de pre procesamiento que proporciona a un agente reactivo la información necesaria para realizar la detección de amenazas.

- b) El pre procesamiento se realiza aplicando una transformada wavelet, que calcula los coeficientes que describen el espectro que representa los momentos en los que la señal obtenida en el ambiente sufre un cambio significativo. El patrón que se genera se aplica en la siguiente etapa de la arquitectura, que consiste en un agente reactivo que toma la señal, y la compara con los datos que se emplearon durante su entrenamiento para minimizar la posibilidad de que la señal sea ubicada de forma incorrecta y se genere una falsa alarma.
- c) Cuando el agente reactivo determina que la señal sí representa una amenaza, envía información al componente de la arquitectura que se encarga de activar una alarma para que se coordine la ubicación de una cámara en un punto específico del perímetro del inmueble. En el caso de que el agente reactivo no pueda determinar si la señal representa o no una amenaza, entonces activa la siguiente etapa de procesamiento de la arquitectura.
- d) En la siguiente etapa de la arquitectura se activa un agente deliberativo que se encarga de examinar los resultados que fueron generados por la etapa de pre procesamiento. Cuando, el agente deliberativo tiene agrupados los resultados que describen el ambiente, aplica el reconocimiento de patrones, comparándolos con los datos que se utilizaron en su etapa de entrenamiento, para reducir la posibilidad de interpretar de forma equivocada un evento acústico que genere una falsa alarma.
- e) En el caso de que persista la posibilidad de que la señal sensada represente una amenaza potencial, recurre a su modelo simbólico para determinar las acciones que se deben de seguir y enviar la información necesaria al componente de la arquitectura que se encarga de coordinar la ubicación de una cámara en un punto estratégico del perímetro del inmueble.

1.7. Propuesta de solución

Se propone el diseño de una arquitectura de seguridad acústica multiagente para analizar y clasificar los ruidos sensados en el ambiente del entorno de un inmueble, y reconocer la probabilidad de que exista una amenaza de intrusión al área perimetral de un inmueble. Para realizar esta labor, la arquitectura contará con tres elementos básicos:

- Una etapa de pre procesamiento que extraerá los coeficientes matemáticos que describen las características de los ruidos que fueron sensados alrededor del inmueble.
- Un conjunto de agentes reactivos que estarán asociados en una relación "1:1" con los sensores acústicos instalados en el área perimetral del inmueble, que implementarán la técnica estadística de un clasificador ingenuo de Bayes para estimar si los coeficientes obtenidos por la etapa de pre procesamiento representan la existencia de una amenaza al área perimetral de un inmueble.
- Un agente deliberativo que agrupará la información proveniente de los agentes reactivos, que implementará una red neuronal multicapa de tipo *feed-forward*

backpropagation, para que en caso de que el agente reactivo no sea capaz de establecer la probabilidad de una amenaza, analice los coeficientes del espectro del sonido sentido y determine si realmente existe o no una amenaza potencial para el inmueble.

Capítulo II

2. Estado del arte

Los sistemas de vigilancia, son un componente tecnológico de gran importancia para la sociedad en general, por esa razón, cada día se exploran nuevas formas, técnicas y estrategias para su diseño, construcción, implementación y operación.

A continuación se analizan trabajos que se han realizado en la misma área de desarrollo tecnológico que la arquitectura multiagente de seguridad acústica desarrollada en este trabajo.

2.1. Sistemas de vigilancia basados en audio

Tan Teng Teng *et al.* en [1], desarrollaron un sistema de vigilancia basado en la detección de sonidos anormales a través del uso de un sistema embebido en un microcontrolador para mejorar un sistema de circuito cerrado de video vigilancia que incorpora un solo micrófono para el sensado de los sonidos. El proceso completo se realiza en tres pasos, el primero consiste en la detección de sonidos utilizando coeficientes cepstrales en las Frecuencias de Mel; la segunda etapa consiste en el procesamiento del audio, en el que se aplicó la transformada de Fourier para descomponer la señal de audio en función del tiempo y frecuencia. Para analizar el componente periódico de la señal emplearon la transformada discreta de Fourier, y para realizar el análisis en el dominio de la frecuencia se empleó la transformada rápida de Fourier. La clasificación se realizó utilizando un árbol de decisiones. Para la experimentación se utilizaron sonidos de gritos humanos y cristales rotos, con una tasa de efectividad promedio de 88%.

Debmalya Chakrabarty y Mounya Elhilali en [2], proponen el desarrollo de un sistema aplicado en vigilancia, para detectar eventos de sonido anormales. Para detectar las características de un conjunto de datos sin etiquetar y generar un modelo acústico se usó una máquina restringida de Boltzmann (RBM), la cual se entrenó usando 400 unidades de sonido, obteniendo la representación de los atributos y los pesos correspondientes para cada uno de ellos en el plano espectro temporal. Posteriormente se utiliza un modelo de mezcla dinámica creado a partir de agregar una variable de la mezcla que tiene q posibles estados, para descomponer el espacio temporal global de un evento anormal en múltiples trayectorias, cada una perteneciente a un sub evento específico. En la detección de eventos de sonido anormales, se procesan los pesos obtenidos en el análisis RBM de una señal de audio, y se utiliza una puntuación de probabilidad logarítmica para realizar un modelado dinámico, con lo cual el sistema debe aprender a decidir si el evento se clasifica como un evento normal o anormal. Posteriormente se formó un conjunto de datos heterogéneo para evitar sesgos, que se dividió aleatoriamente en dos sub conjuntos, repartiendo el 80% de los datos para entrenamiento, y el 20% restante de los datos para la etapa de prueba. La eficiencia del sistema de vigilancia oscila en valores entre el 65.77% y 93.11%.

Roneel V Sharan y Tom J Moir en [3], desarrollaron un trabajo de vigilancia, que emplea la técnica de análisis basada en el uso de una matriz de co-ocurrencia de nivel de gris que captura la distribución de valores concurrentes a una distancia determinada, para formar un espectrograma único en escala de grises que muestra las frecuencias dominantes de energía y ruido. Los valores obtenidos que marcan la intensidad de los datos de la imagen del espectrograma en escala de grises, son normalizados entre valores de [0,1]. La etapa

de experimentación se realizó con el procesamiento de señales con una ventana de Hamming de 512 puntos y 50% de solapamiento. La base de datos que se empleó consta de 1143 sonidos distribuidos en 10 clases, los cuales fueron divididos en dos grupos usando la técnica de validación cruzada, que generó dos grupos de sonidos para entrenamiento que contiene dos terceras partes del total de sonidos disponibles y para pruebas el tercio restante de los sonidos totales. Para la clasificación se empleó un support vector machine. Los resultados obtenidos en la etapa de clasificación oscilaron entre el 70% y el 86%.

Bochra Bouchhima *et al.* en [4], desarrollaron un sistema de clasificación de sonidos de corta duración, aplicado al contexto de la vigilancia. La base de datos empleada incluye sonidos agrupados en dos clases perfectamente separadas para maximizar las diferencias entre ellas, la primera contiene sonidos de seres humanos como risas, expresiones de asombro o miedo. La segunda clase está representada por sonidos de animales, por ejemplo de gatos maullando. La extracción de las características de los sonidos se realiza, primero, pre procesando la señal con un filtro Weiner, que elimina los ruidos ambientales y normalizando el resultado para evitar sesgos significativos. La siguiente tarea genera un vector que contiene las características seleccionadas del audio, incrementando la varianza entre clases y reduciendo las variaciones intra clases. El vector de características obtenido se divide por validación cruzada en dos grupos, uno que representa el 70% de los datos obtenidos y que se emplea para entrenamiento, el resto de los datos se emplea para pruebas y representa el 30% de los datos obtenidos. La clasificación de los datos se realiza con SVM que incluye el algoritmo de *sequential minimal optimization*. Los resultados de clasificación alcanzan al menos el 90% de efectividad sobre las dos clases de sonidos estudiadas.

Tom J. Moir *et al.* en [5], proponen un trabajo en el que diseñaron un sistema de vigilancia de audio, que emplea un método que extrae características de los sonidos representadas en imágenes en el plano tiempo-frecuencia. La investigación se basa en analizar dos diferentes tipos de representaciones de imágenes en escalas de grises; la primera trabaja sobre una imagen representada en escala de grises lineal, a la cual se le aplica una transformada discreta de Fourier para obtener el valor absoluto de cada trama que compone a la imagen y su normalización para obtener los valores de intensidad que describen a la imagen; la segunda trabaja con una imagen en escala de grises log de manera similar al procesamiento anterior, sólo que representan la energía del sonido analizado normalizando los valores de los datos antes del procesamiento. Para la etapa de experimentación emplearon 10 clases bien definidas y estandarizadas de sonidos. La clasificación se realiza comparando el trabajo de un SVM y un algoritmo KNN. En el caso de la clasificación con SVM obtuvieron un 98.16% de efectividad, y en el caso del algoritmo KNN la efectividad de clasificación fue de 96.59%.

Jens Schröder *et al.* en [6], proponen un sistema de detección de eventos acústicos, en donde el proceso se orienta a mejorar la reducción del ruido de una señal acústica a partir de las características del propio ruido, con ese fin la entrada de la señal acústica es pre procesada usando estadísticas mínimas para la estimación del ruido y atenuación logarítmica espectral de amplitud para la reducción del ruido. Posteriormente, se extraen las características del evento acústico aplicando un banco de filtros Gabor de tipo 2D para representar los patrones de modulación en el plano espectro-temporal. Finalmente se aplica un conjunto de herramientas de las cadenas ocultas de Markov para construir una red de reconocimiento de voz dividida en tres categorías: a nivel de palabra, a nivel de modelo y a nivel de cadenas ocultas de Markov.

Stavros Ntalampiras e Ilyas Potamitis en [7], proponen el desarrollo de un sistema de audio vigilancia aplicado a la conservación de ambientes de reservas naturales empleando la señal acústica captada por un solo micrófono con la finalidad de detectar actividades humanas ilícitas que pongan en peligro el equilibrio de un ecosistema protegido. En la etapa de experimentación se emplearon sonidos de origen humano, como voces, sonidos de disparos, o ruidos de auto o motocicleta, y también sonidos de origen distinto como ruido Gaussiano, canto de aves, sonidos de lluvia y sonidos de viento. También se generaron mezclas de los sonidos disponibles con el objetivo de reforzar la calidad durante la etapa de experimentación. La etapa de reconocimiento se realizó suponiendo que cada sonido presenta un patrón acústico único que se obtiene a partir de la distribución de su energía en el contexto de su frecuencia. El patrón que representa ese modelo es identificado con el uso de algoritmos de reconocimiento de patrones estadísticos. En este trabajo se utilizó el enfoque de los modelos ocultos de Markov, en donde cada estado se representa usando un modelo de mezcla Gaussiana con una matriz diagonal de covarianza. La precisión general de reconocimiento es igual un 81.5%, de manera individualizada. La tasa de eficiencia para las cuatro clases empleadas en experimentación es como sigue: sonidos de auto igual a 75.8%, sonidos de motocicleta igual a 70.1%, sonidos de conversaciones igual a 100% y sonidos de disparos igual a 80.2%.

Andrzej Glowacz y Grzegorz Altman en [8], desarrollaron un sistema de vigilancia que orienta sus actividades al análisis, detección y clasificación de sonidos que representan una amenaza. Las señales de audio son captadas por micrófonos ubicados en un área de estudio, después de captar los datos, son enviados por una red de transporte hasta un nodo central que se encarga de realizar la tarea de procesamiento. La aplicación detecta y clasifica amenazas especificando un nivel de peligro y manda la información a un centro de gestión de crisis. La extracción de las características sirve para formar un vector 8 de coeficientes cepstrales de frecuencias de Mel. Se usa un algoritmo que genera un grupo de entrenamiento con la extracción de características de coeficientes de un marco simple, que entre otras bondades, tiene la cualidad de resolver el problema de necesitar el uso de un SVM multi clase para señales con parámetros de diferente longitud. En el proceso de formación del vector de características se divide la señal de sonido en varias subclases. En el proceso de clasificación, la SVM encuentra un hiperplano que separa linealmente las sub clases en un espacio dimensional superior. En la etapa de experimentación se realizaron 24 pruebas, obteniendo una precisión de clasificación por parte del SVM que tiene una fluctuación de valores, cuando se usan 82 señales de referencia la tasa de éxito es de 75%, y cuando se emplean 139 señales de referencia, la tasa de éxito es de 79.19%.

J. Moragues *et al.* en [9], analizan el problema de la detección y reconocimiento acústico para las aplicaciones de vigilancia en el contexto de ambientes ruidosos con un origen de diversa naturaleza. Para descubrir qué está ocurriendo un evento acústico, emplean un detector de energía múltiple (MED) por su capacidad que tiene de adaptarse la longitud variante de una novedad acústica y en consecuencia ajustar el tamaño de la ventana de análisis. En la etapa de clasificación, primero se toma en cuenta la duración de la longitud del evento identificado por el MED para determinar si el origen analizado se considera como impulsivo o no impulsivo. Con el conjunto de la información recopilada se genera un vector de características representativas. Para el modelado de los eventos acústicos se trabaja con un enfoque supervisado, tomando como base el vector de características obtenido en la etapa anterior, en la etapa de selección se aplican reglas de decisión de un

clasificador bayesiano con la intención de asignar a las clases de sonido impulsivo o no impulsivo pre establecidas una nueva fuente de sonido detectado.

Burak Uz Kent y Buket D. Barkana en [10], proponen un sistema de vigilancia que busca detectar situaciones anómalas analizando el cambio de tono dentro de un rango de valores calculados a través del uso de una función de auto correlación, de una señal de entrada de audio. El proceso de preparación de los datos para su análisis, requiere primero la extracción y después la selección de las características usando la técnica de coeficientes de predicción lineal, para generar un vector que agrupa un conjunto de características basadas en relaciones públicas que describen diferentes eventos acústicos. El proceso de clasificación se realiza utilizando un SVM para analizar tres clases de sonidos, con las que previamente fue entrenado. La tasa de efectividad de reconocimiento obtenida se estima en un intervalo de valores entre un 79% y un 92%.

Bo Hang y Ruimin Hu en [11], proponen el diseño de un modelo de vigilancia de audio con características estéreo que analiza las señales de audio espacial y su factor de energía. El objetivo es detectar eventos causados por un origen de sonido con movimiento rápido, para la localización espacial de la fuente que los origina. Un factor importante a tomar en cuenta en el análisis, es la variación de las condiciones del fondo del entorno que predispone las características de la señal de audio que se obtiene de los sensores. La señal de audio extraída se convierte del dominio del tiempo al dominio de la frecuencia, la señal resultante se divide en 24 sub-bandas, y para cada una de ellas se calcula el parámetro de audio espacial. El seguimiento de la fuente de sonido requiere analizar el cambio de los valores los parámetros espaciales del audio para calcular el nuevo lugar de ubicación. El modelo mostró que es posible detectar y seguir eventos de audio en entornos complejos. Logrando detectar fuentes de sonidos con movimiento rápido y eventos que tienen valores altos de energía como gritos y disparos.

Cholé Clavel *et al.* en [12], proponen un sistema multimedia de vigilancia aplicada en ambientes ruidosos que analiza las firmas acústicas de una cantidad específica de armas de fuego y desarrollaron una investigación basada en el supuesto de que el análisis de ambientes ruidosos puede aportar información valiosa para detectar cuando se presenta una situación anormal que pone en riesgo a la comunidad durante el desarrollo de un hecho. En particular, analizan el potencial que implica complementar los sistemas de vigilancia convencionales que basan su funcionamiento en el análisis de imágenes obtenidas a través de video, con la incorporación del análisis de información de audio. Para el proceso de entrenamiento, cada clase que se construye sigue un modelo de mezclas gaussianas. La extracción de las características de audio consiste en el análisis de ventanas de audio con una longitud de marcos o ventanas de 20 ms y 50% de solapamiento, de este procesamiento se obtienen las características de audio más populares y más probables. La detección de ruidos anormales se hace empleando reglas de decisión de máximo a *posteriori*.

2.2. Sistemas de vigilancia basados en video

Meng-da Hsieh *et al.* en [13], proponen el desarrollo de un sistema de vigilancia basado en imágenes de video representadas en 3D. El proceso requiere primero generar imágenes 2D, la corrección de la distorsión, y posterior mapeo a 3D. A continuación las imágenes 3D se asocian a un conjunto de cámaras de video para realizar la detección y rastreo de objetos. Para la detección de objetos, se realiza el análisis de cada fotograma

de la secuencia de video captada en primer plano y el fondo. Para detectar objetos sobresalientes se realiza un análisis, filtrando la parte correspondiente del plano de cada fotograma utilizando un algoritmo de diferencia temporal para obtener un modelo del fondo inicial. El modelo generado se usa para detectar objetos sobresalientes en primer plano, a través de la regla de actualización de fondo. El proceso de seguimiento requiere la observación de cada fotograma, y el etiquetado aspectos importantes de cada bloque, como el área analizada, cantidad de objetos que la componen, color RGB. El seguimiento de un sujeto se reconoce en consecuencia de los siguientes cinco estados de transición: "entrada, salida, agrupación, oclusión y fracción". Cada objeto sólo puede estar en un estado a la vez y se modifica con su movimiento en la escena de video. Con la información recabada es posible realizar el reconocimiento de personas en movimiento.

Supavadee Aramvith *et al.* en [14], proponen diseño de una aplicación de vigilancia que procesa y analiza video sin compresión. La primera actividad es el proceso de detección identificando porciones del cuerpo humano mediante segmentación para obtener una máscara que representa características y patrones del sujeto sospechoso, estas características se mantendrá como datos en bruto para ser empleadas en etapas posteriores y en formato XML para su almacenamiento e intercambio con otros sistemas. Para el análisis del sujeto detectado se utilizó el método de extracción de fondo adecuado por su baja complejidad y eficiencia en sistemas que trabajan en tiempo real, además de realizar la eliminación del ruido de la imagen usando operadores morfológicos como la dilatación y la erosión. Cada sujeto detectado, será registrado y etiquetado para su seguimiento usando el método de búsqueda de ventana, logrando manejar hasta 256 personas en una escena. Para analizar y predecir el movimiento del sujeto de estudio, se utiliza el método de regresión lineal, en el supuesto de que el movimiento que realiza es una línea recta. Toda la información obtenida genera un vector de características que describe elementos como la existencia de gafas para sol, mochila o el tiempo que invierte en su recorrido, que se emplea para evaluar la situación y ubicar el sistema en alguna de las siguientes cuatro posibilidades: sin alarma, alarma baja, alarma intermedia y alarma alta. El trabajo propuesto se encuentra en la etapa de refinamiento y servirá además de localizar sujetos sospechosos en escenas de video, para la construcción de una base de datos que servirá en el seguimiento de la investigación de algún crimen.

Nahum Kiryati *et al.* en [15], proponen un sistema de video vigilancia en tiempo real para la detección de movimiento anormal. El desarrollo se enfrenta con tres retos: el primero, es la fiabilidad que busca reducir a una tasa baja la incidencia de falsas alarmas al detectar eventos irregulares, el segundo es la caracterización de un movimiento normal, y el tercero es la detección de un movimiento anormal empleando la potencia limitada de cálculo. El procedimiento de detección usa un algoritmo que obtiene información del vector que contiene cortes de las secuencias de imágenes o macrobloques de movimiento que se generan en el proceso normal de compresión de video estándar. Durante la tarea de detección, se calcula el vector de características de cada trama de video, cuando la probabilidad de la ocurrencia de k con respecto las k tramas consecutivas de video están por debajo de un umbral T , la trama k -enésima se considera un movimiento anormal.

Wu Chao *et al.* en [16], describen el diseño de una arquitectura multigigante para un sistema distribuido de video vigilancia, compuesto por una red de equipos, en donde cada terminal está diseñada como un agente responsable de procesar información que proviene del conjunto de sensores de video que administra, y sólo transmite por una red IP los resultados obtenidos en tiempo real. Otro equipo es empleado como el centro de intercambio de información y control de unidades, para asegurar que la información de

control sea transmitida automáticamente en el momento que se genere un evento que requiera una rápida atención. El procesamiento de la señal de video captada, es tratado por la unidad de análisis de video inteligente, que realiza el reconocimiento de los patrones de objetos y su comportamiento. Los hechos detectados se transfieren al motor de razonamiento que busca elementos con conductas o intenciones dudosas, con lo que se disparará una alarma o se activará el control de otros equipos de la red ip distribuida de video vigilancia.

Yufeng Chen *et al.* en [17], proponen el desarrollo de un sistema de seguridad multiagente para al análisis de imágenes de captadas en secuencias de video de un grupo de personas y realizar el seguimiento de un sujeto identificado por el contorno de la región de la cabeza. Un agente de análisis se encarga del inicio del proceso que se necesita para el estudio de las señales generadas por tomas de video, que fueron analizadas usando la transformada wavelete Haar. De los datos resultantes de la etapa anterior se obtiene información referente al sujeto en estudio, como velocidad de desplazamiento y dirección. Un agente de seguimiento utiliza un SVM que fue entrenado con muestras pre seleccionadas, y que toma como datos de entrada, los valores almacenados en un vector de características generadas por la etapa anterior, que se utiliza para reconocer las acciones o intenciones del sujeto. En la etapa de experimentación se concluyó que es posible realizar el seguimiento de un sujeto dentro de un grupo de personas en tiempo real.

Hugo Dias *et al.* en [18], proponen el desarrollo de un sistema de vigilancia distribuido que detecta y rastrea elementos dentro de un espacio determinado. El sistema se compone por varios agentes de vigilancia y un agente central que mantiene comunicación con el resto del equipo y los datos del proceso. Cada agente de vigilancia se integra de una computadora con comunicación inalámbrica, una cámara de video y un lenguaje llamado VigiLANG desarrollado para la comunicación entre agentes, basado en el intercambio de mensajes. El sistema funciona con dos etapas. La primera realiza el procesamiento de las imágenes captadas, identificando los colores de las tomas de video, relacionándolas a una clase para formar una tabla de consulta asociada a valores RGB de píxeles y la segmentación de la imagen considerando dicha tabla. En la segunda, el sistema utiliza las imágenes en escala de grises para llevar a cabo la detección de movimiento. Para eso se calcula la media de los valores de los píxeles de los componentes RGB, y el resultado se normaliza entre los valores de cero y uno. Con esa información se comprueba si el resultado obtenido es mayor a un umbral establecido, entonces se ha detectado un movimiento. Con la información generada se realiza el seguimiento de un objeto detectado. El aspecto más complejo de todo el trabajo es el procesamiento de las imágenes, lo cual, además condiciona la velocidad de captura de fotogramas de video. Al inicio del desarrollo el tiempo era alto y la tasa de captura era baja, al final del desarrollo la velocidad de procesamiento se estableció en 1.1 ms y la tasa de captura es de 60 fotogramas por segundo.

2.3. Sistemas de vigilancia basados en audio y video

Patiana Intani y Teerapong Orachon en [19], analizan el diseño de un sistema de alertas de vigilancia que combina el análisis de video y de audio para detectar eventos que clasifican como controvertidos. Ambas señales se captan por el sistema de circuito cerrado de video cámaras y micrófonos. La detección de intrusos puede hacerse analizando el componente video o el componente sonido. En el caso del sonido, el

sistema procesa los cambios de los niveles de energía de la señal de audio para detectar una posible situación de delincuencia y enviar una notificación al observador del sistema de circuito cerrado de video cámaras. En el caso del componente video se analiza el cambio en el número de píxeles de la imagen de fondo y el cuadro de la imagen de referencia para identificar un evento sospechoso y enviar una notificación al observador del sistema de circuito cerrado de video cámaras. El rango de alcance del sistema de vigilancia se encuentra delimitado por el área que abarca la vista y que oscila entre cinco y diez metros.

G. Valenzise *et al.* en [20], describen un sistema de audio y video vigilancia que analiza el audio generado en una plaza para detectar eventos anómalos y ubicar el origen del sonido, con la intención de orientar una cámara de video. El proceso de extracción de características utiliza una combinación de las técnicas filtro híbrido y métodos de envoltura con la finalidad de formar un vector de características que representa la distribución de la energía a lo largo de intervalos situados en el tiempo. La clasificación de los sonidos de interés en el trabajo se realiza a través de dos clasificadores de modelo de mezclas Gaussianas que funcionan en paralelo. Cada clasificador es entrenado por separado con muestras de cada una de las clases utilizadas en el desarrollo del trabajo. Para ubicar el origen de un evento acústico se utiliza un algoritmo que calcula la correlación lineal de mínimos cuadrados y la diferencia de tiempo que existe entre el tiempo en que es percibido y la ubicación del micrófono que lo escucha. Los experimentos realizados dieron como resultado una precisión de detección de eventos acústicos del 93%, con una tasa de rechazo del 5%.

2.4. Descripción de características del estado del arte

El proceso de revisión de trabajos relacionados con el estado del arte de la arquitectura multiagente de seguridad acústica desarrollada en esta tesis reveló los siguientes resultados:

- La técnica que más se utilizó en el procesamiento de los datos es *Mel frequency cepstral coefficients* (Tabla 1), en cambio el componente de la arquitectura multiagente de seguridad acústica encargado del procesamiento utiliza la herramienta matemática Transformada Wavelet (Tabla 2).
- La técnica de clasificación empleada mayoritariamente es *support vector machine* (Tabla 1); el componente multiagente de la arquitectura de seguridad acústica que se usó para diferenciar los sonidos analizados para el agente reactivo es la técnica estadística Naive Bayes (Tabla 2), para el agente deliberativo desarrollado se utilizó la técnica de clasificación de red neuronal *feed-forward backpropagation* (Tabla 2).
- Los resultados obtenidos en los trabajos revisados oscilan en el rango de 65% a 100% de efectividad para reconocimiento de los datos procesados, en cambio los resultados obtenidos por la arquitectura multiagente de seguridad acústica se ubican en un rango de efectividad entre 66.5% y 74.5%.

- Para los trabajos revisados se emplearon *corpus* de entrenamiento y prueba con una tendencia homogeneizada que incluyó un número constante de clases, con características mutuamente excluyentes entre sí.
- El *corpus* de entrenamiento y prueba utilizado en este trabajo tiene un número constante de clases, pero la naturaleza de los sonidos analizados es heterogénea, existiendo entre ellos una alta dispersión de los coeficientes matemáticos procesados (Capítulo IV, Gráfica 11).

Número	Herramienta de procesamiento de los datos	Técnica de clasificación	Porcentaje de efectividad
1	<ul style="list-style-type: none"> • Mel frequency cepstral coefficients • Transformada de Fourier 	<ul style="list-style-type: none"> • Árbol de decisiones 	88%
2	<ul style="list-style-type: none"> • Máquina restringida de Boltzman • Modelo de mezcla dinámica 	<ul style="list-style-type: none"> • Puntuación de probabilidad logarítmica 	65.77% 93.11%
3	<ul style="list-style-type: none"> • Matriz de co-ocurrencia 	<ul style="list-style-type: none"> • Support vector machine 	70% - 86%
4	<ul style="list-style-type: none"> • Filtro Weiner 	<ul style="list-style-type: none"> • Support vector machine 	90%
5	<ul style="list-style-type: none"> • Espectrograma en escala de grises 	<ul style="list-style-type: none"> • Support vector machine • K-nearest neighbors 	98.16% 96.59%
6	<ul style="list-style-type: none"> • Filtro Gabor 2D 	<ul style="list-style-type: none"> • Hidden Markov model 	----
7	<ul style="list-style-type: none"> • Mel frequency cepstral coefficients • Integración perceptual de paquetes Wavelet 	<ul style="list-style-type: none"> • Hidden Markov model • Gaussian mixture model 	70.1% - 80%
8	<ul style="list-style-type: none"> • Mel frequency cepstral coefficients 	<ul style="list-style-type: none"> • Support vector machine 	75% - 79.19%
9	<ul style="list-style-type: none"> • Multiple energy detector 	<ul style="list-style-type: none"> • Reglas de decisión de Bayes 	96%
10	<ul style="list-style-type: none"> • Coeficientes de predicción lineal 	<ul style="list-style-type: none"> • Support vector machine 	79% - 92%
11	<ul style="list-style-type: none"> • Interaural level difference 	----	70% - 100%
12	<ul style="list-style-type: none"> • Mel frequency cepstral coefficients 	<ul style="list-style-type: none"> • Gaussian mixture model • Reglas de decisión a posteriori 	Tasa de rechazo = 10%

Tabla 1. Principales características de trabajos relacionados.

Número	Herramienta de procesamiento de los datos	Técnica de clasificación	Porcentaje de efectividad
1	<ul style="list-style-type: none"> • Transformada Wavelet 	<ul style="list-style-type: none"> • Naive Bayes • Neural network 	66.5% - 74.5%

Tabla 2. Características de la arquitectura diseñada.

Capítulo III

3. Descripción de la arquitectura de seguridad acústica aplicada al área perimetral de un inmueble

El incremento del índice delictivo que existe en el país que afecta a la seguridad de los espacios físicos destinados a inmuebles, la creciente sofisticación de las técnicas criminales que se emplean para irrumpir y penetrar el área perimetral de edificios para cometer ilícitos, y que los sistemas de seguridad que se emplean para la protección de estos inmuebles regularmente no utilizan sensores acústicos para conocer los ruidos generados en su ambiente cercano, la capacidad que tienen para juzgar cuando un evento producido en el exterior representa una amenaza potencial de peligro es limitada. Por eso se propone el diseño de una arquitectura de seguridad acústica que está compuesta por dos elementos (Figura 1).

- Una etapa de pre procesamiento que se utiliza para realizar un análisis matemático de los sonidos que se obtienen del sensado del ambiente, para describir las características que representan a la señal obtenida.
- Un sistema multiagente que tienen la finalidad de considerar cuando un conjunto de sonidos obtenidos desde el medio ambiente, contienen información que puede representar una condición de peligro. El sistema multiagente está integrado por dos tipos diferentes de componentes:
 - Agente reactivo.
 - Agente deliberativo.

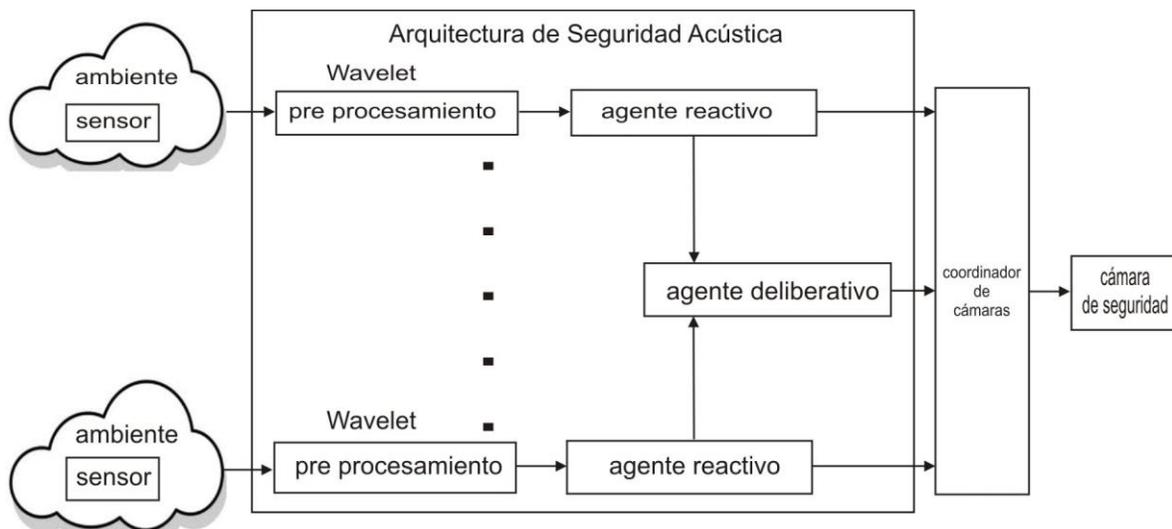


Figura 1. Arquitectura multiagente de seguridad acústica.

3.1. Pre procesamiento de la señal

Cuando un sensor acústico capta los sonidos generados en el medio ambiente, se obtiene una mezcla rica en ruidos con diferentes características, para que el sistema multiagente pueda juzgar los datos recopilados es necesario procesarlos para preparar la señal, eliminar impurezas presentes que no son significativas y conservar la información relevante.

La señal que se analiza presenta un comportamiento no estacionario, por lo que se necesita una herramienta de procesamiento matemático que pueda realizar un análisis que descomponga la señal en ventanas con regiones de tamaño variante.

Para realizar ese trabajo se utiliza una transformada wavelet que se emplea para representar y segmentar un conjunto de señales no estacionarias multiresolución que incluyen variaciones con picos abruptos, en el espacio "tiempo – frecuencia".

El análisis multiresolución se describe como una sucesión de subespacios cerrados V_j , $j \in \mathbb{Z}$, de $L^2(\mathbb{R})$.

La transformada wavelet madre que se emplea, es la función wavelet Haar que se describe de la siguiente manera:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 0.5, \\ -1 & 0.5 \leq t < 1. \\ 0 & \text{de otra forma} \end{cases}$$

Y la función escalar $\phi(t)$ se describe como:

$$\phi(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & \text{de otra forma} \end{cases}$$

El uso de la transformada wavelet genera una matriz compuesta por coeficientes matemáticos, que representan el espectro de la señal, a la que se ha filtrado el componente estático, eliminando el ruido de fondo sin que exista una degradación apreciable de la señal.

La matriz de los coeficientes matemáticos representan los puntos de quiebre de la señal. Estos puntos simbolizan el momento en el cual existe un cambio brusco en la intensidad de la señal obtenida del ambiente. Estas perturbaciones del entorno caracterizan la tendencia de los puntos de quiebre que definen el espectro del patrón de comportamiento del ruido censado.

Para estandarizar la distribución de los resultados obtenidos se normalizan los coeficientes matemáticos para reducir la dispersión que existe entre los valores. La técnica de normalización estadística aplicada es "min – max", descrita por la ecuación:

$$Y = \min + \frac{X - \text{mínimo_original}}{\text{máximo_original} - \text{mínimo_original}} (\max - \min)$$

3.2. Sistema multiagente

Un sistema multiagente se compone de un conjunto de agentes que cooperan entre sí para realizar una tarea coordinada. En esta tesis se presenta un sistema multiagente inteligente para la detección de intrusiones mediante un análisis acústico de señales.

Para ello, el sistema multiagente se alimenta de la señal captada por los sensores acústicos, la cual es procesada por la transformada wavelet Haar, siendo necesario que los coeficientes matemáticos resultantes sean juzgados por el sistema multiagente para que determine cuando un ruido puede representar a un evento potencialmente sospechoso.

3.3. Agente Reactivo

Un agente reactivo es un ente que se apoya en el uso de sensores para percibir los estímulos que se generan en el entorno en el que existe, y que responde en consecuencia, usando uno o varios actuadores para transmitir la decisión que toma.

El agente reactivo propuesto en este trabajo, que compone la arquitectura multiagente de seguridad acústica implementa la técnica estadística de clasificación supervisada Naive Bayes, para detectar un evento anómalo a partir de un modelo que predice la probabilidad de que este suceso ocurra. El proceso que sigue el clasificador Naive Bayes se describe en la siguiente ecuación:

$$valor = \underset{i = 1}{\operatorname{argmax}} P(c_i) \cdot \prod_{j = 1}^m P(a_j | c_j)$$

Para que el agente reactivo pueda catalogar correctamente cada elemento que analiza, es necesario llevar a cabo un proceso de entrenamiento fuera de línea del clasificador Naive Bayes. Para ese propósito se eligió un conjunto de sonidos que pueden o no representar una amenaza.

Después de llevar a cabo la actividad de entrenamiento, se realiza el proceso de predicción que se emplea para medir la capacidad del modelo para clasificar el conjunto de datos de interés.

En este trabajo el conjunto de agentes reactivos que componen la arquitectura se encargan de procesar el espectro de la señal captada en el ambiente, en primera instancia compara las características de la señal obtenida con los datos que reconoce.

El resultado del proceso de comparación puede determinar que la señal analizada pueda ser considerada o no como una amenaza potencial, evitando de esa manera la generación de una falsa alarma (falso positivo).

Cuando la comparación de los elementos particulares del espectro de la señal mediante el clasificador Naive Bayes determina que el evento acústico se considera una amenaza potencial se activa una alarma para que se corrobore de manera visual la naturaleza del evento acústico.

3.4. Agente Deliberativo

Un agente deliberativo, es una entidad que posee un conjunto de conocimientos, los cuales compara con el estado actual del entorno, para evaluar cual debe de ser la decisión que tomará y que determinará la forma en la que este actuará en consecuencia.

En el proceso del diseño de la arquitectura de seguridad acústica, el agente deliberativo se construyó a partir del modelo BDI, el cual define un conjunto de creencias, intenciones y deseos para que pueda llevar a cabo las tareas que se le encomienden.

En relación a la problemática estudiada en este trabajo, las creencias del agente parten del hecho de conocer a *priori*, durante un proceso de entrenamiento fuera de línea, los tipos de eventos acústicos que pueden representar un ruido común generado en el área perimetral del inmueble.

Los deseos del agente deliberativo, consisten en la necesidad de tener la capacidad de detectar cuándo el estado inicial de las creencias que tiene el agente deliberativo sobre el medio ambiente es modificado por alguna alteración.

Las intenciones del agente deliberativo, se representan por contar con la capacidad de identificar cuándo la modificación del ambiente pasó de ser un sonido normal a un ruido que pueda ser considerado potencialmente sospechoso.

El modelo deliberativo BDI, se implementó en una red neuronal con una arquitectura tipo *feed-forward backpropagation*, que usa el método de aprendizaje supervisado para entrenar a la red neuronal siguiendo los pasos que describen a continuación:

- Uso de un modelo de entrada, que se difunde de las capas de entrada, oculta y salida, hasta generar una salida determinada.
- Análisis del resultado obtenido, validando el grado de error cometido por cada neurona de la capa de salida.
- Ajuste de los pesos de cada neurona para corregir los errores cometidos en sentido contrario a la propagación inicial del entrenamiento, es decir, desde la capa de salida hasta la capa de entrada.

En inicio, para poder trabajar con la red neuronal es necesario crearla, proceso en el que se define, el número de neuronas por capa, y la función de transferencia que será aplicada por cada capa.

Para que el agente deliberativo pueda clasificar correctamente la señal que recibe, y que contiene la descripción de las características de los coeficientes matemáticos obtenidos en la etapa de pre procesamiento, debe de pasar por un proceso de entrenamiento supervisado fuera de línea, en el que se le muestra un patrón de inicio que contiene el modelo que describe el conjunto de conocimientos o creencias que regirán su comportamiento.

Posteriormente, cuando la red ya conoce los valores que acotan el caso de estudio, se realiza el proceso de simulación, para valorar la capacidad que tiene la red neuronal de reconocer datos similares que procesará posteriormente.

Cuando la red neuronal realiza la actividad de comparación del conocimiento que adquiere durante la etapa de entrenamiento, con respecto a los datos que procesa en la etapa de reconocimiento, se pueden presentar dos posibles situaciones, que la señal que recibe sea considerada como un ruido común del ambiente, o que la señal contenga sonidos que potencialmente representan una amenaza de intrusión al área perimetral del inmueble que se está custodiando.

Cuando la red neuronal que representa al agente deliberativo, determina que la señal no contiene sonidos que simbolizan una amenaza, desestima la información que recibe, para reducir la generación de falsas alarmas.

Cuando la red neuronal que representa al agente deliberativo, determina que la señal contiene sonidos que describen la existencia de una amenaza de acceso no autorizado al área perimetral de un inmueble, entonces activa una alarma para que se verifique visualmente a través de una cámara de video vigilancia la naturaleza de la amenaza.

Capítulo IV

4. Proceso de diseño

En este apartado se explica el proceso que se siguió para diseñar la arquitectura multiagente de seguridad acústica, describiendo las características de cada uno de sus elementos desde el punto de vista de su conceptualización, su estructura interna y componentes, su funcionamiento interno, la forma de asociarse con otros componentes de la arquitectura, y los experimentos realizados para comprobar los resultados esperados de cada elemento.

4.1. Formación del corpus de entrenamiento / validación

Para la integración del conjunto de sonidos que se emplearon para realizar los procesos de entrenamiento y validación del sistema multiagente se analizó un universo de archivos disponibles, que están organizados en las siguientes categorías de efectos de sonido:

- Ambiente de lugares de reunión común.
- Armas de fuego.
- Aviones, automóviles y ferrocarril.
- Domésticos.
- Eléctricos y beeps.
- Hogar.
- Mecánica y máquinas

Después de escuchar los sonidos, fueron seleccionados 40 archivos que fueron divididos en dos categorías que representan sonidos comunes, y ruidos potencialmente sospechosos.

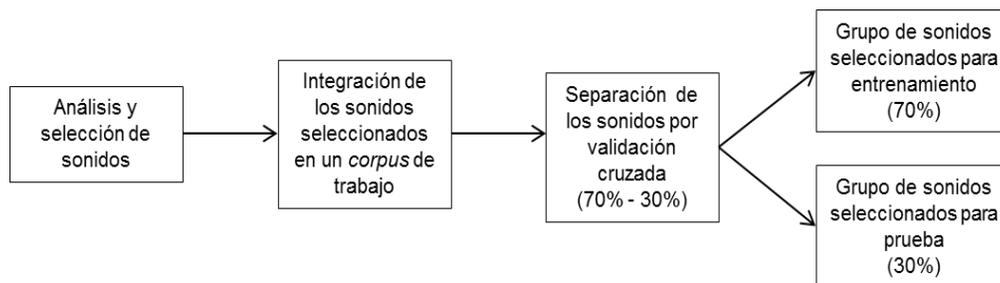


Figura 2. Proceso de formación del *corpus* de trabajo.

El *corpus* de sonidos que se conformó con ambas categorías se dividió en dos grupos para formar el conjunto de datos de entrenamiento y el de prueba, usando la técnica de validación cruzada, definiendo para el grupo de entrenamiento 28 archivos y para el grupo de prueba 12 archivos (figura 2).

Con el conjunto de archivos de entrenamiento se formó una matriz de 28 filas, una fila por cada archivo, y de 40 columnas que representan las características que se extrajeron del archivo original que se analiza de cada uno de los sonidos.

Con el conjunto de archivos de prueba se formó una matriz de 12 filas, una fila por cada archivo, y de 40 columnas que representan las características que se extrajeron del archivo original que se analiza de cada uno de los sonidos.

4.2. Pre procesamiento

Las matrices obtenidas para entrenamiento y prueba que se formaron, contienen información que representa el estado acústico del medio ambiente del perímetro de un inmueble.

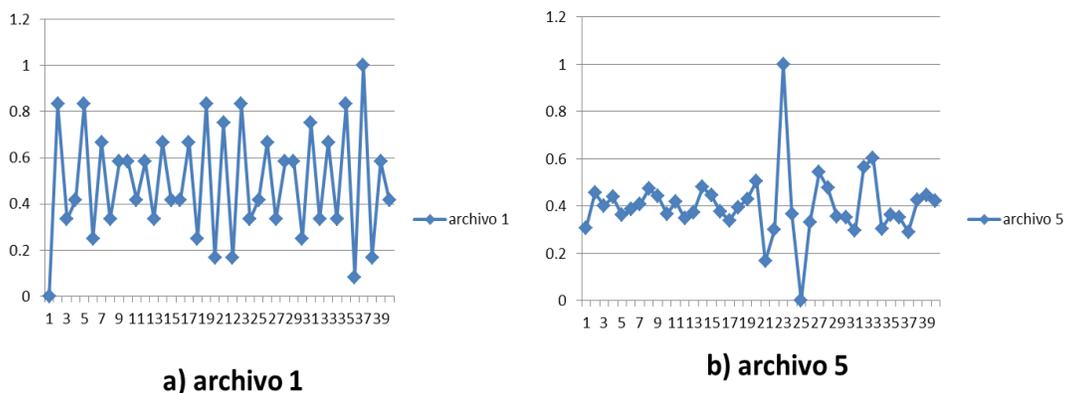
Para poder determinar el momento en el que se alteran las condiciones acústicas del entorno, primero es necesario realizar un análisis matemático que permita conocer el instante en el que se presenta una perturbación del ambiente.

El análisis matemático se realiza aplicando la transformada wavelet Haar, a cada una de las filas de las matrices, para detectar el momento en el que existe un cambio abrupto en los valores de los coeficientes matemáticos, conocido como valor "pico".

Como se planteó en la sección 1.4.2 (Gráfica 10), el resultado del análisis usando la transformada wavelet Haar genera un conjunto de coeficientes matemáticos que contienen valores que describen el comportamiento de la señal acústica, los cuales son almacenados en una matriz con las mismas dimensiones de los arreglos bidimensionales originales que fueron analizados.

Es precisamente el conjunto de los coeficientes matemáticos representados por los valores pico, los datos que se analizan para determinar si los trastornos acústicos generados en el medio ambiente representan un peligro potencial de intrusión.

El último paso de la etapa de pre procesamiento, consiste en normalizar los valores almacenados en las matrices resultantes para generar una distribución uniforme y evitar una alta dispersión de los coeficientes matemáticos.



Gráfica 10. Resultados de pre procesamiento.

4.3. Diseño del Agente Reactivo

El agente reactivo se diseñó a partir de un clasificador ingenuo Bayesiano (Figura 3), que emplea para la etapa de entrenamiento fuera de línea, una matriz de 28 filas y 40 columnas que recibe de la etapa de pre procesamiento y un vector columna de 28 celdas que se emplea para etiquetar el significado que representa cada fila de la matriz de entrenamiento, generando como resultado un objeto que contiene información del proceso de adiestramiento.

Posteriormente se comprueba la capacidad que tiene el objeto generado para clasificar nuevos elementos que se proporcionan en una matriz que se emplea para la etapa de prueba compuesta por 12 filas y 40 columnas.

De la etapa de prueba se obtiene un objeto que se emplea para generar una matriz de confusión y determinar cuántos elementos fueron clasificados correctamente como "verdadero positivo", "verdadero negativo", y cuántos elementos fueron ubicados erróneamente como "falso positivo" y "falso negativo".

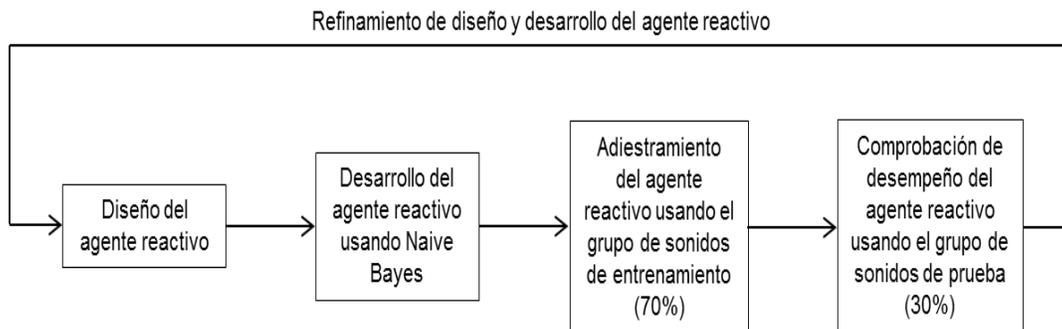
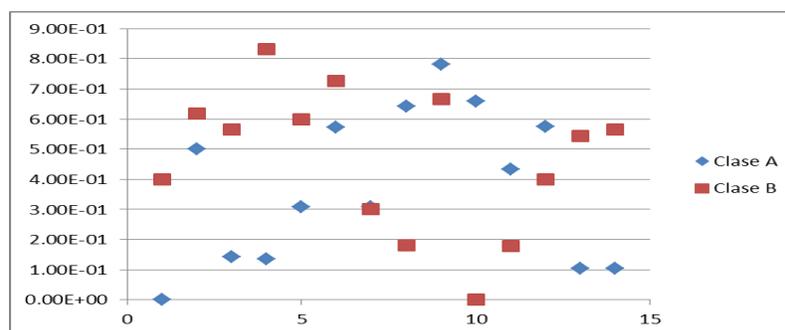


Figura 3. Diseño y desarrollo de agente reactivo.

4.4. Diseño del Agente Deliberativo

El agente deliberativo se programó en dos versiones.

Primero se diseñó bajo el esquema de una máquina de vectores soporte (SVM), sin embargo en la etapa de prueba se demostró que la naturaleza del problema que se analiza tiene una característica de no ser linealmente separable, por lo que el uso del SVM quedó descartado (Gráfica 11).



Gráfica 11. Dispersión de los datos.

En segundo lugar, se utilizó una red neuronal de tipo *feed-forward backpropagation* que siguió un proceso de tres etapas para su diseño (Figura 4).

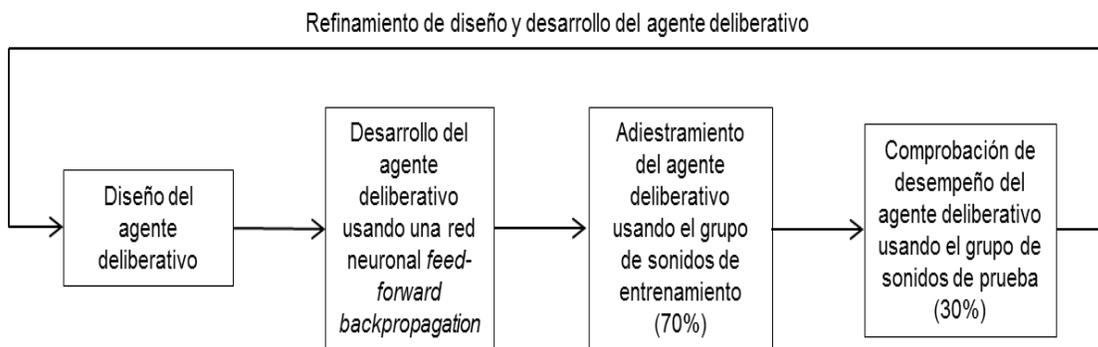


Figura 4. Diseño y desarrollo de agente deliberativo.

- Creación de la red neuronal: La etapa de creación de la red neuronal, genera como resultado un objeto, que para ser formado se emplean cuatro argumentos que preparan a la red para la etapa de entrenamiento:
 - Una matriz de 28 filas y 2 columnas, que contienen el valor mínimo y máximo de cada fila de la matriz original de entrenamiento compuesta por 28 filas y 40 columnas.
 - El valor del número de capas que tendrá la red neuronal, tomando en cuenta sí únicamente contará con una capa de entrada y una capa de salida, o sí incluirá algún número de capas ocultas.
 - La función de transferencia con la que cada capa de la red neuronal trabajará.
 - La función de entrenamiento.
 - En el caso específico de este trabajo se eligió una arquitectura de red neuronal de tipo *feed-forward backpropagation*, con tres capas ocultas por las siguientes razones:
 - Se eligió una red tipo *feed-forward backpropagation*, por la capacidad de tomar los datos de entrenamiento y compartirlos hacia las capas superiores. Cuando se tiene el resultado de la capa de salida se analiza el error de clasificación generado, y se transmite hacia atrás con la intensidad de reajustar los valores de los pesos de cada neurona.
 - La elección del número correcto de capas ocultas, depende inicialmente del tipo de problema que se atiende, aunque “el número adecuado de unidades ocultas no está bien comprendido” [21], pero se recomienda que “cuando se analizan datos que tienen el comportamiento de una función continua se recomienda usar una

sola capa". En caso contrario, "cuando se analizan datos con un comportamiento que se describe con una función discontinua se recomiendan al menos dos capas" [21].

- Entrenamiento: En la etapa de entrenamiento, que se realiza fuera de línea, se lleva a cabo el adiestramiento supervisado de la red, para que sea capaz de reconocer los patrones que están definidos en el modelo descrito por la matriz de entrenamiento. Para el entrenamiento se utilizan tres argumentos:
 - El objeto generado en la etapa de creación de la red neuronal.
 - El conjunto de adiestramiento que consiste en una matriz de 28 filas y 40 columnas.
 - El conjunto de prueba, que consiste en una matriz de 12 filas y 40 columnas.
- Simulación: La simulación es la tercera etapa del diseño, en la que se corrobora el comportamiento que tiene la red, para evaluar su capacidad adquirida para reconocer en el conjunto de prueba, los rasgos con los que fue instruida la red neuronal. En este proceso se utilizan dos argumentos:
 - El objeto que se obtuvo de la etapa de adiestramiento.
 - El conjunto de prueba, formado por una matriz de 12 filas y 40 columnas.

Capítulo V

5. Análisis de los resultados

Con la intención de medir el nivel de efectividad que la arquitectura multiagente de seguridad acústica tiene cuando lleva a cabo el proceso de clasificación de sonidos, se realizaron experimentos que permitieron medir en un ambiente controlado la capacidad que tiene el trabajo propuesto en esta tesis de reconocer entre eventos acústicos normales y potencialmente sospechosos.

Los experimentos se realizaron en los siguientes componentes de la arquitectura multiagente de seguridad acústica que se encargan de realizar el proceso de clasificación acústica:

- Agente reactivo.
- Agente deliberativo.

5.1. Resultados del agente reactivo

El agente reactivo se implementó utilizando la técnica estadística de clasificación Naive Bayes, y se evaluó en tres condiciones distintas de exigencia.

Mejor caso.

En este experimento se utilizó una matriz de prueba de 12 filas y 40 columnas, cada fila representa un archivo de sonido que contiene únicamente una señal acústica sin mezcla o edición digital.

El desempeño del agente reactivo se muestra en la matriz de confusión representada en la tabla 3.

		Valores predichos	
		Positivo	Negativo
Valores reales	Positivo	1	5
	Negativo	0	6

Tabla 3. Agente reactivo, mejor caso.

Considerando los resultados mostrados en la tabla 4, el agente reactivo tuvo una eficiencia del 58.3% de clasificación general al reconocer 7 de 12 elementos de la matriz de prueba. La eficiencia particular de reconocimiento en cada caso es como sigue:

- 1 elemento positivo clasificado correctamente.
- 6 elementos negativos clasificados correctamente.
- 5 elementos que fueron clasificados como falso negativo, que tendrían que estar considerados como positivos.
- 0 elementos que fueron clasificados como falso positivo.

Caso promedio

En este experimento se utilizó una matriz de prueba de 12 filas y 40 columnas, cada fila representa la mezcla de dos archivos de sonido tomados del *corpus* de entrenamiento y un archivo de sonido que simboliza el entorno común de una calle de medio ambiente urbano.

El desempeño del agente reactivo se muestra en la matriz de confusión representada en la tabla 4.

		Valores predichos	
		Positivo	Negativo
Valores reales	Positivo	3	3
	Negativo	3	3

Tabla 4. Agente reactivo, caso promedio.

Considerando los resultados mostrados en la tabla 5, el agente reactivo tuvo una eficiencia del 50% de clasificación general al reconocer 6 de 12 elementos de la matriz de prueba. La eficiencia particular de reconocimiento en cada caso es como sigue:

- 3 elementos positivos clasificado correctamente.
- 3 elementos negativos clasificados correctamente.
- 3 elementos que fueron clasificados como falso negativo, que tendrían que estar considerados como positivos.
- 3 elementos que fueron clasificados como falso positivo, que tendrían que estar considerados como negativos.

Peor caso

En este experimento se utilizó una matriz de prueba de 12 filas y 40 columnas, cada fila representa la mezcla de doce archivos de sonido tomados del *corpus* de entrenamiento y un archivo de sonido que simboliza el entorno común de una calle de medio ambiente urbano.

El desempeño del agente reactivo se muestra en la matriz de confusión representada en la tabla 5.

		Valores predichos	
		Positivo	Negativo
Valores reales	Positivo	1	5
	Negativo	0	6

Tabla 5. Agente reactivo, peor caso.

Considerando los resultados mostrados en la tabla 6, el agente reactivo tuvo una eficiencia del 58.3% de clasificación general al reconocer 7 de 12 elementos de la matriz de prueba. La eficiencia particular de reconocimiento en cada caso es como sigue:

- 1 elemento positivo clasificado correctamente.
- 6 elementos negativos clasificados correctamente.
- 5 elementos que fueron clasificados como falso negativo, que tendrían que estar considerados como positivos.

- 0 elementos que fueron clasificados como falso positivo.

Considerando los resultados anteriores se puede concluir, que de forma general el agente reactivo es capaz de reconocer al menos el 50% de los elementos que analiza. Tiende a confundirse cuando clasifica elementos que son verdaderos positivos, ubicándolos en las categorías de falso positivo o negativo, con una efectividad máxima del 25%.

5.2. Resultados del agente deliberativo

El agente deliberativo se implementó utilizando la técnica de clasificación de inteligencia artificial red neuronal, con una arquitectura de tipo *feed-forward backpropagation*, misma que se evaluó en tres condiciones distintas de exigencia.

Mejor caso

En este experimento se utilizó una matriz de prueba de 12 filas y 40 columnas, cada fila representa un archivo de sonido que contiene únicamente una señal acústica sin mezcla o edición digital.

La comprobación de la efectividad del agente deliberativo se realizó a través de la prueba estadística t de student.

En la tabla 6, se muestra el rendimiento del agente deliberativo al analizar el comportamiento del elemento de prueba 1.

Obteniendo los siguientes valores:

- Número de pruebas: 30
- Media aritmética del porcentaje de eficiencia: 52.3
- Desviación estándar: 41.39
- t - calculada: -2.042
- Grados de libertad: 29
- Nivel de significancia: 0.05
- t - tabla: 2.045
- Nivel de eficiencia al clasificar el 100% de los elementos: 68%

Número de pruebas.	% de Eficiencia	x	X
1	100	48	2304
2	93	41	1681
3	18	-34	1156
4	90	38	1444
5	100	48	2304
6	93	41	1681
7	93	41	1681

8	55	3	9
9	60	8	64
10	100	48	2304
11	18	-34	1156
12	5	-47	2209
13	0	-52	2704
14	3	-49	2401
15	5	-47	2209
16	95	43	1849
17	15	-37	1369
18	38	-14	196
19	18	-34	1156
20	100	48	2304
21	83	31	961
22	100	48	2304
23	98	46	2116
24	0	-52	2704
25	15	-37	1369
26	3	-49	2401
27	15	-37	1369
28	95	43	1849
29	58	6	36
30	3	-49	2401

Tabla 6. Agente deliberativo, mejor caso.

Operaciones de cálculo:

Media aritmética:

$$\bar{x} = \frac{\sum x}{N} = \frac{1569}{30} = 52.3$$

Desviación estándar:

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}} = \sqrt{\frac{49688.3}{29}} = \sqrt{1713.38} = 41.39$$

t - student:

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{52.3 - 68}{\frac{41.39}{\sqrt{29}}} = \frac{-15.7}{\frac{41.39}{5.385}} = \frac{-15.7}{7.686} = -2.042$$

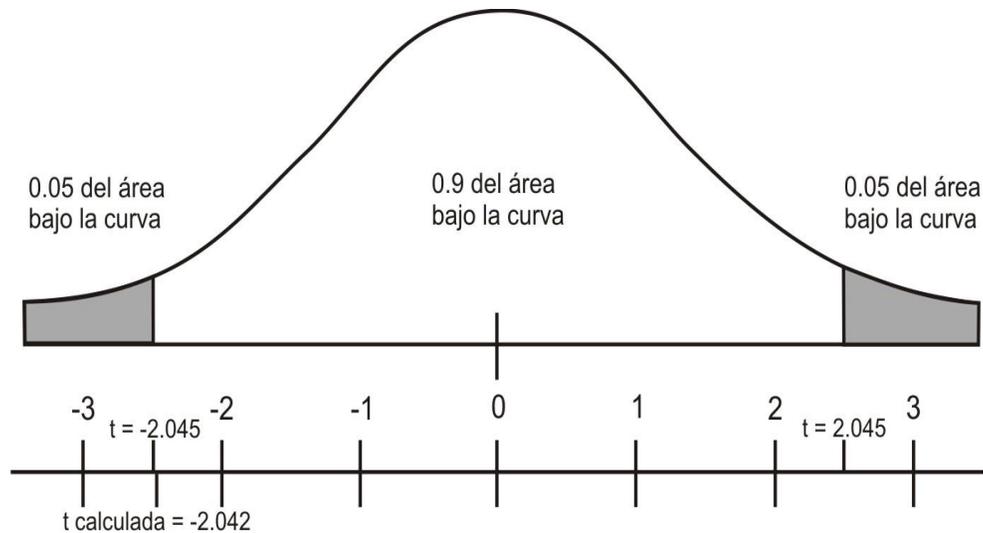


Figura 5. Estadístico t-student mejor caso.

Como el valor de la t - calculada es menor al valor de t que marca la tabla correspondiente, se concluye, que en el experimento del mejor caso, la efectividad del agente deliberativo es del 68%, es decir, que de cada 100 elementos analizados, 68 elementos son bien clasificados (Figura 5).

Caso promedio

En este experimento se utilizó una matriz de prueba de 12 filas y 40 columnas, cada fila representa la mezcla de dos archivos de sonido tomados del corpus de entrenamiento y un archivo de sonido que simboliza el entorno común de una calle de medio ambiente urbano.

La comprobación de la efectividad del agente deliberativo se realizó a través de la prueba estadística t de student.

En la tabla 7, se muestra el rendimiento del agente deliberativo al analizar el comportamiento del elemento de prueba 1.

Obteniendo los siguientes valores:

- Número de pruebas: 30
- Media aritmética del porcentaje de eficiencia: 58.6
- Desviación estándar: 42.25
- t - calculada: -2.026
- Grados de libertad: 29
- Nivel de significancia: 0.05
- t - tabla: 2.045
- Nivel de eficiencia al clasificar el 100% de los elementos: 74.5%

Número de pruebas.	% de Eficiencia	x	X
1	0	-58.6	3433.96
2	98	39.4	1552.36
3	80	21.4	457.96
4	100	41.4	1713.96
5	98	39.4	1552.36
6	100	41.4	1713.96
7	88	29.4	864.36
8	0	-58.6	3433.96
9	88	29.4	864.36
10	0	-58.6	3433.96
11	98	39.4	1552.36
12	0	-58.6	3433.96
13	0	-58.6	3433.96
14	95	36.4	1324.96
15	50	-8.6	73.96
16	100	41.4	1713.96
17	93	34.4	1183.36
18	35	-23.6	556.96
19	18	-40.6	1648.36
20	90	31.4	985.96
21	100	41.4	1713.96
22	93	34.4	1183.36
23	93	34.4	1183.36
24	55	-3.6	12.96
25	60	1.4	1.96
26	100	41.4	1713.96
27	18	-40.6	1648.36
28	5	-53.6	2872.96
29	0	-58.6	3433.96
30	3	-55.6	3091.36

Tabla 7. Agente deliberativo, caso promedio.

Operaciones de cálculo:

Media aritmética:

$$\bar{x} = \frac{\sum x}{N} = \frac{1758}{30} = 58.6$$

Desviación estándar:

$$S = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n - 1}} = \sqrt{\frac{51785.2}{29}} = \sqrt{1785.69} = 42.25$$

t - student:

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{58.6 - 74.5}{\frac{42.25}{\sqrt{29}}} = \frac{-15.9}{\frac{42.25}{5.385}} = \frac{-15.9}{7.845} = -2.026$$

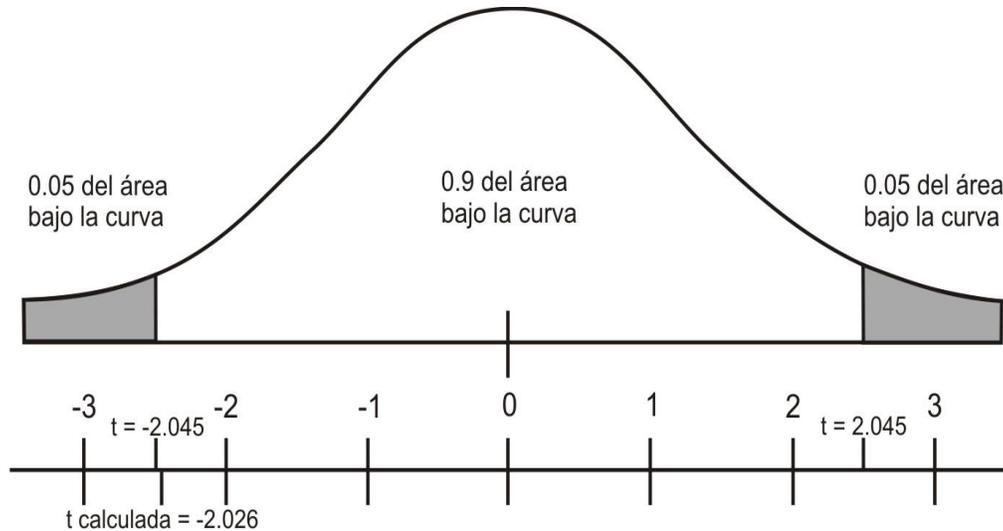


Figura 6. Estadístico t-student caso promedio.

Como el valor de la t - calculada es menor al valor de t que marca la tabla correspondiente, se concluye, que en el experimento del mejor caso, la efectividad del agente deliberativo es del 74.5%, es decir, que de cada 100 elementos analizados, 74.5 elementos son bien clasificados (Figura 6).

Peor caso

En este experimento se utilizó una matriz de prueba de 12 filas y 40 columnas, cada fila representa la mezcla de doce archivos de sonido tomados del *corpus* de entrenamiento y un archivo de sonido que simboliza el entorno común de una calle de medio ambiente urbano.

La comprobación de la efectividad del agente deliberativo se realizó a través de la prueba estadística t de student.

En la tabla 8, se muestra el rendimiento del agente deliberativo al analizar el comportamiento del elemento de prueba 2.

Obteniendo los siguientes valores:

- Número de pruebas: 30

- Media aritmética del porcentaje de eficiencia: 51.9
- Desviación estándar: 38.9
- t - calculada: -2.021
- Grados de libertad: 29
- Nivel de significancia: 0.05
- t - tabla: 2.045
- Nivel de eficiencia al clasificar el 100% de los elementos: 66.5%

Número de pruebas.	% de Eficiencia	x	X
1	100	48.1	2313.61
2	100	48.1	2313.61
3	98	46.1	2125.21
4	90	38.1	1451.61
5	38	-13.9	193.21
6	3	-48.9	2391.21
7	100	48.1	2313.61
8	3	-48.9	2391.21
9	28	-23.9	571.21
10	30	-21.9	479.61
11	0	-51.9	2693.61
12	88	36.1	1303.21
13	0	-51.9	2693.61
14	75	23.1	533.61
15	98	46.1	2125.21
16	35	-16.9	285.61
17	25	-26.9	723.61
18	13	-38.9	1513.21
19	83	31.1	967.21
20	28	-23.9	571.21
21	88	36.1	1303.21
22	0	-51.9	2693.61
23	90	38.1	1451.61
24	95	43.1	1857.61
25	88	36.1	1303.21
26	25	-26.9	723.61
27	5	-46.9	2199.61
28	78	26.1	681.21
29	43	-8.9	79.21
30	10	-41.9	1755.61

Tabla 8. Agente deliberativo, peor caso.

Operaciones de cálculo:

Media aritmética:

$$\bar{x} = \frac{\sum x}{N} = \frac{1557}{30} = 51.9$$

Desviación estándar:

$$S = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n - 1}} = \sqrt{\frac{44002.7}{29}} = \sqrt{1517.33} = 38.9$$

t - student:

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{51.9 - 66.5}{\frac{38.9}{\sqrt{29}}} = \frac{-14.6}{\frac{38.9}{5.385}} = \frac{-14.6}{7.223} = -2.021$$

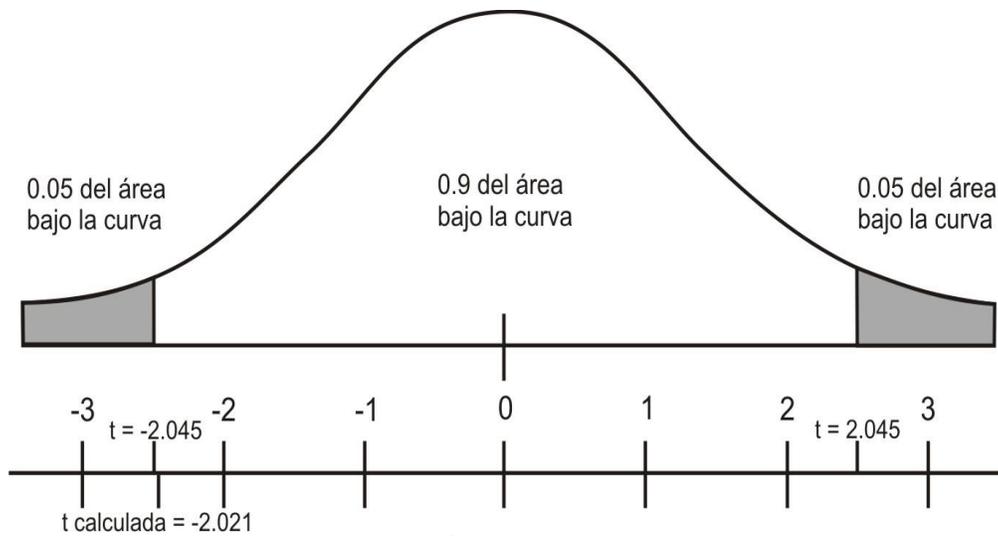


Figura 7. Estadístico t-student peor caso.

Como el valor de la t - calculada es menor al valor de t que marca la tabla correspondiente, se concluye, que en el experimento del mejor caso, la efectividad del agente deliberativo es del 66.5%, es decir, que de cada 100 elementos analizados, 66.5 elementos son bien clasificados (Figura 7).

5.3. Contrastación de los resultados con la hipótesis

Hipótesis general: "Es posible reconocer la mayoría de los ruidos generados en el ambiente, dentro de los cuales es probable que exista una amenaza de intrusión al área perimetral de un inmueble usando una arquitectura multiagente, clasificando los sonidos con un agente reactivo y con un agente deliberativo".

H_0 = Es posible reconocer la mayoría de los ruidos generados en el ambiente.

H_1 = No es posible reconocer la mayoría de los ruidos generados en el ambiente.

Como se puede apreciar en la tabla 9, de manera general, la arquitectura multiagente de seguridad acústica, tiene un rango de eficiencia que oscila entre el 50% y 74.5% de efectividad al momento de reconocer uno de los doce sonidos que integran el *corpus* de prueba que se empleó para llevar a cabo el proceso comprobación de su funcionamiento.

El agente deliberativo es el último elemento de la arquitectura multiagente de seguridad acústica que entra en funcionamiento al tratar de detectar un sonido sospechoso en el área perimetral de un inmueble, cuando determina que el ruido no representa una amenaza, es desechado para seguir sensado el ambiente.

Con base en ese funcionamiento, se toma como referencia el valor de 66.5% para medir la efectividad total de la arquitectura multiagente de seguridad acústica.

Concentrado de resultados					
Agente Reactivo (efectividad)			Agente Deliberativo (efectividad)		
Mejor caso	Caso promedio	Peor Caso	Mejor caso	Caso promedio	Peor Caso
58.3%	50%	58.3%	68%	74.5%	66.5%

Tabla 9. Resultados generales.

5.4. Resultado

En este trabajo se desarrolló una arquitectura de seguridad acústica que clasifica sonidos captados en el medio ambiente que existe en el área perimetral de un inmueble ubicado en una zona urbana. Para realizar las actividades se diseñó un proceso que está integrado por tres etapas, cada una de las cuales analiza un conjunto de sonidos desde diferentes perspectivas.

- La etapa de pre procesamiento, trabaja con base en la aplicación de la herramienta matemática transformada Wavelet, que recibe los sonidos como son originados en el ambiente, y extrae sus características o rasgos representativos.
- La etapa de clasificación se realiza inicialmente por un agente reactivo que fue desarrollado en base en un Naive Bayes, y que considera el conjunto de sonidos con el que fue entrenado para tomar la decisión de activar una alarma debido a que reconoció un ruido perteneciente al conjunto de datos generados en la etapa anterior, y que tiene probabilidad de ser sospechoso.

- La etapa de clasificación se realiza también por un agente deliberativo que fue desarrollado en base en una red neuronal, el cual se activa cuando el agente reactivo no puede determinar la potencialidad de una amenaza. El agente deliberativo analiza los datos que recibe de las etapas anteriores y los juzga tomando en cuenta la información con la que fue entrenado para precisar si se trata de una amenaza potencial, y en consecuencia activar una alarma de seguridad.

Con base en los experimentos realizados durante las pruebas, se midió que la arquitectura multiagente de seguridad acústica es capaz de reconocer al menos el 66.5% de los sonidos que analiza (Tabla 9), por lo tanto existe evidencia para aceptar la hipótesis H_0 y rechazar la hipótesis H_1 .

5.5. Conclusiones

El diseño de la arquitectura multiagente de seguridad acústica requirió de la realización de procesos que incluyen actividades de análisis, diseño lógico y físico, pruebas y experimentación de los componentes, comprobación y validación estadística de los resultados.

Una de las primeras tareas que se realizaron fueron la selección y formación de una colección de cuarenta sonidos, que se agruparon usando validación cruzada, en dos categorías que describen eventos acústicos de la vida cotidiana y ruidos que tienen potencialidad de representar situaciones sospechosas de intención de intrusión al área perimetral de un inmueble, que se emplearon en las etapas de adiestramiento supervisado, y en la etapa de prueba del agente reactivo y del agente deliberativo.

Para mejorar la funcionalidad de la colección de sonidos en los procesos de entrenamiento y prueba, se lleva a cabo la tarea de pre procesamiento, aplicando una transformada Wavelet, con la intención de obtener los coeficientes matemáticos que representan el comportamiento de los eventos acústicos, sin que exista pérdida significativa de información relevante en las siguientes etapas.

El agente reactivo que forma parte de la arquitectura multiagente de seguridad acústica, se diseñó empleando la técnica estadística de clasificación Naive Bayes. Se entrenó de manera supervisada usando el grupo de sonidos que se formó para su adiestramiento.

El agente deliberativo que forma parte de la arquitectura multiagente de seguridad acústica se diseñó usando una red neuronal *feed-forward backpropagation*, con tres capas ocultas, realizando un adiestramiento supervisado con el grupo de sonidos que se formó para entrenamiento.

Se aplicaron treinta experimentos a la arquitectura multiagente de seguridad acústica, usando el grupo de sonidos que se formó para prueba, realizando ensayos para las situaciones del mejor caso, el caso promedio y el peor caso.

El agente reactivo demostró a través de su matriz de confusión generada, que tiene la capacidad de reconocer al menos el 50% de los sonidos empleados en la etapa de prueba.

El agente deliberativo demostró a través de la prueba estadística t de student, que tiene la capacidad de reconocer los sonidos que se usaron en la etapa de prueba, en un rango de efectividad que oscila entre el 66% y el 74%.

Con los experimentos y las pruebas estadísticas realizadas, se comprobó, que la arquitectura multiagente de seguridad acústica tiene al menos una efectividad calibrada de reconocimiento de sonidos del 66%, que corresponde a los resultados en el peor de los casos del agente deliberativo.

En conclusión el rendimiento general de la arquitectura multiagente de seguridad acústica desarrollada en este trabajo, se encuentra dentro de los parámetros de los valores obtenidos en los trabajos revisados en el estado del arte.

5.6. Trabajo futuro

Aunque los procesos modernos de ingeniería están dedicados al diseño, construcción e implementación de sistemas con un alto nivel de rendimiento, los productos obtenidos tienen un valor real de eficiencia menor a la efectividad total estimada.

En consecuencia, existen procesos en los que se analiza y revisa el desempeño de los componentes de un sistema, con la intención de identificar las partes que son susceptibles de recibir un tratamiento de optimización, con el que se logre una mejora significativa de rendimiento del producto.

En el caso particular de la arquitectura multiagente de seguridad acústica desarrollada en este trabajo, no se encuentra exenta de esta máxima, por lo que se proponen las siguientes mejoras, para acrecentar sus capacidades y en consecuencia incrementar el grado de su efectividad (Figura 8).

- Implementar un banco de memoria de recuerdos permanente para almacenar nuevos sonidos.
- Incorporar a la memoria de almacenamiento permanente los sonidos que el agente reactivo o el agente deliberativo identifiquen de eventos acústicos significativos.
- Reforzar el aprendizaje del agente reactivo y deliberativo empleando los nuevos sonidos que se incorporen al medio de almacenamiento permanente.
- Usar los sonidos recopilados en la memoria de almacenamiento permanente para que el agente deliberativo los consulte cuando se presente un sonido similar procesado anteriormente.

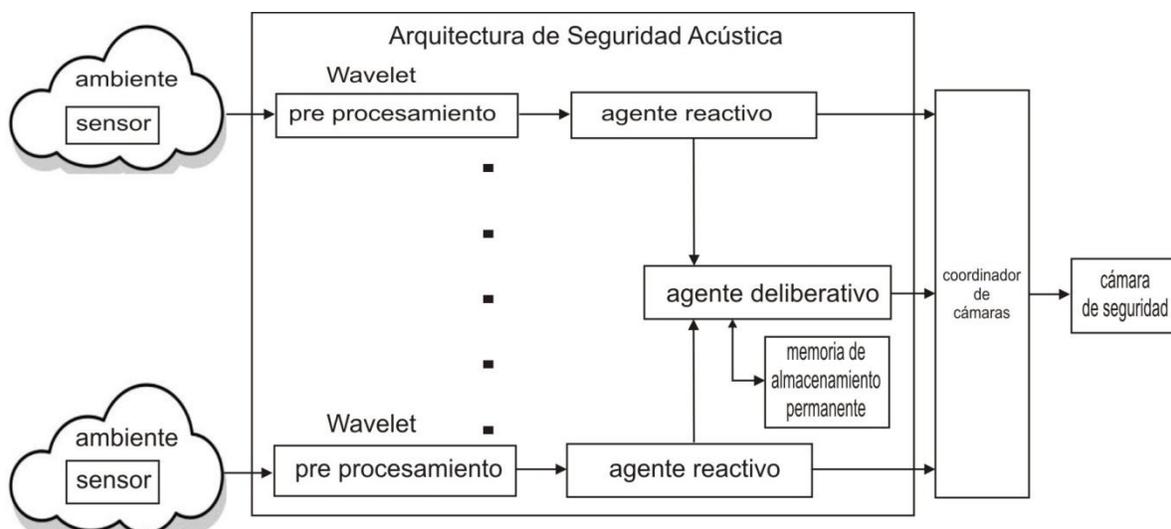


Figura 8. Mejora de la arquitectura multiagente de seguridad acústica.

Referencias

- [1]. T. T. Teng, L. T. Sze, O. L. Yeng, "Abnormal sound analytical surveillance system using microcontroller", International Colloquium on Signal Processing & its Applications, IEEE, 2016.
- [2]. D. Chakrabarty, M. Elhilali, "Abnormal sound event detection using temporal trajectories mixtures", International Conference on Acoustics, Speech and Signal Processing, IEEE, 2016.
- [3]. R. V. Sharan, T. J. Moir, "Robust audio surveillance using spectrogram image texture feature", International conference on acoustics, speech and signal processing (ICASSP), IEEE, 2015.
- [4]. B. Bouchhima, R. Amara, M. Turki, "Classification of short-duration sounds for environmental monitoring", 1st International Conference on Advanced Technologies for Signal and Image Processing, IEEE, 2014.
- [5]. T. J. Moir, R. V. Sharan, "Audio surveillance under noisy conditions using time-frequency image feature", 19 th International Conference on Digital Signal Processing, IEEE, 2014.
- [6]. J. Schröder, B. Cauchi, M. R. Schädler, N. Moritz, "Acoustic event detection using signal enhancement and spectro-temporal feature extraction", ASP Challenge on detection and classification of acoustic scenes and events, IEEE, 2013.
- [7]. S. Ntalampiras, I. Potamitis, "Detection of human activities in natural environments based on their acoustic emissions", 20th European signal processing conference, IEEE, 2012.
- [8]. A. Glowacz y G. Altman, "Automatic threat classification using multiclass svm from audio signals", 17th International conference on emerging technologies & factory automation, IEEE, 2012.
- [9]. J. Moragues, A. Serrano, L. Vergara, J. Gosálbez, "Acoustic detection and classification using temporal and frequency multiple energy detector features", International conference on acoustics, speech and signal processing (ICASSP), IEEE, 2011.
- [10]. B. Uzkent, B. D. Barkana, "Pitch-range based feature extraction for audio surveillance systems", Eighth international conference on information technology, IEEE, 2011.
- [11]. B. Hang, R. Hu, "Spatial audio cues based surveillance audio attention model", International conference on acoustics, speech and signal processing, IEEE, 2010.
- [12]. C. Clavel, T. Ehrette, G. Richard, "Events detection for an audio-based surveillance system", International conference on multimedia and expo, IEEE, 2005.
- [13]. M. Hsieh, S. Chen, Y. Cai, Y. Chen, J. Y. Chiang, "Immersive surveillance for total situational awareness", International computer symposium, IEEE, 2010.

- [14]. S. Aramvith, S. Pumrin, T. Chalidabhongse, S. Siddhichai, "Video processing and analysis for surveillance applications", International Symposium on Intelligent Signal Processing and Communication Systems, IEEE, 2009.
- [15]. N. Kiryati, T. Riklin, Y. Ivanchenko, S. Rochel, "Real-time abnormal motion detection in surveillance video", Pattern recognition, 19th international conference on, IEEE, 2008.
- [16]. W. Chao, X. M. Jun, "Multi-agent based distributed Video surveillance system over ip", International symposium on computer science and computational technology, IEEE, 2008.
- [17]. Y. Chen, Z. Zhong, K. Lee, Y. Xu, "Multi-agent based surveillance", International Conference on Intelligent Robots and Systems, IEEE, 2006.
- [18]. H. Dias, J. Rocha, P. Silva, C. Ledo, L. P. Reis, "Distributed surveillance system", portuguese conference on artificial intelligence, IEEE, 2005.
- [19] P. Intani, T. Orachon, "Crime warning system using image and sound processing", 13th international conference on control, automation and systems, IEEE, 2013.
- [20] G. Valenzise, L. Gerosa, M. Tagliasacchi, F. Antonacci, A. Sarti, "Scream and gunshot detection and localization for audio-surveillance systems", Conference on advanced video and signal based surveillance, IEEE, 2007.
- [21] S. Russell, P. Norving, "Inteligencia Artificial, Un enfoque moderno", Pearson Prentice Hall, 2007.