Capítulo I: Introducción.

1.1 Actualidad y definición del problema.

La palabra domótica proviene de dos palabras, domus (casa) y de Informática. Este concepto hace referencia a la incorporación de tecnologías informáticas dentro de la casa; también se refiere a las comunicaciones que permiten gestionar y automatizar desde un mismo sistema cada una de las instalaciones de uso cotidiano en la casa. Cada una de estas tecnologías informáticas comunicadas proporciona una mejor calidad de vida, mejor conservación y cuidado del edificio a cada uno de los usuarios de la casa. Los tres objetivos principales de los sistemas domóticos son: mejorar la comodidad, seguridad y ahorrar energía [1].

Los humanos por naturaleza prefieren realizar sus actividades de la manera más fácil posible, es por eso que se día a día la tecnología evoluciona para automatizar y mejorar los procesos cotidianos. La domótica surgió en respuesta a muchas personas que requieren de una mayor automatización sobre la casa, es decir, la domótica nos permite tener un control fácil y automático sobre nuestra casa.

Actualmente existen sistemas domóticos totalmente funcionales como el X-10 Home Systems que es un sistema que ha sido instalado en varios lugares a lo largo de todo el planeta y ha sido probado por varios usuarios obteniendo resultados satisfactorios.

El área de la domótica es muy importante en nuestro país ya que aquí no se desarrolla mucha tecnología y es conveniente que nuestro país cuente con tecnología de este tipo para que empiece a competir con los países desarrollados.

Otro de los conceptos que debemos entender es el reconocimiento de voz, Esta tecnología nos facilita el control de dispositivos. Reconocimiento de voz es la conversión de palabras escritas, en el que si aplicamos el procesamiento de

lenguaje natural se obtiene el entendimiento de voz. Es además, una interfaz natural, rápida, flexible y accesible [2].

Uno de los principales objetivos que persigue el reconocimiento de voz es la interacción apropiada entre usuario-maquina por medio del habla. Este modelo de interacción es una manera muy natural de comunicación entre humanos pero a veces resulta difícil poder integrar un reconocedor de voz para la computadora.

Con la evolución de la tecnología se han creados reconocedores de voz muy completos y eficaces, por esa razón se exploran nuevas áreas en las que el reconocimiento de voz puede ser útil.

En la actualidad el mercado de dispositivos va más allá de un simple control de funciones del hogar de manera remota, ahora se trabaja en la posibilidad de manejar una casa por simples mensajes de voz, pero sin la necesidad de que el usuario lleve un micrófono [3].

Esta afirmación nos hace pensar que una de las áreas en las que realmente el reconocimiento de voz es útil, es la domótica. Al crear una tecnología con la cual se pueda controlar los dispositivos de una manera más natural, hace que la domótica tenga alcance mucho mayor en relación a la interacción usuario-maquina.

Al tener un sistema domótico para una casa demasiado amplia, la manipulación de los dispositivos se vuelve más compleja ya que se manejan un número elevado de componentes de la casa (detectores humo, luz, gas, etc.). Aunque el problema se reduce un poco al usar la voz para para dar algún comando, por ejemplo apagar o encender cada objeto (foco, televisión, etc.) en cada lugar que se encuentre.

No solo debemos de apoyarnos de la voz para dar los comandos. Hoy en día nuestra sociedad exige una manipulación más sencilla e inteligente sobre cada

uno de los dispositivos de la casa sin la necesidad de memorizar la lista completas de instrucciones por medio de voz.

Actualmente existen sistemas domóticos que responden mediante comandos de voz. Esta área se puede mejorar y ampliar con el fin de hacer más inteligentes a los sistemas domóticos por voz.

Una de las áreas en las que los sistemas domóticos por voz se pueden mejorar es el control de dispositivos por medio de la vista. La vista provee una de las formas más naturales para seguir objetos por esta razón es necesario implementar tecnologías que nos permitan analizar el movimiento de los ojos con el fin de controlar dispositivos de una manera sencilla.

En este proyecto de tesis se ampliaran los sistemas domóticos por voz implementando el análisis del movimiento de los ojos con el fin de controlar dispositivos en una casa inteligente, es decir, el usuario tendrá la capacidad controlar los dispositivos de la casa por medio de la voz y de la vista con el fin de reducir los comandos de voz para controlar un mismo dispositivo que se encuentre en varias partes de la casa.

1.2 Objetivo general.

Crear un sistema de control de dispositivos de una casa inteligente basado reconocimiento de voz aplicando el seguimiento de ojos hacia este dispositivo con el propósito de personalizar el acceso a este dispositivo.

1.3 Objetivos específicos.

- Análisis bibliográfico.
- Selección de prototipo de sistema domótico con reconocimiento de voz.
- Selección de método de procesamiento de imágenes para seguir los ojos de un usuario.
- Instalación y pruebas de software comercial necesario.
- Diseño del sistema
- Implementación del modulo de reconocimiento de voz

- Implementación del modulo de reconocimiento de imágenes para seguir los ojos de un usuario.
- Integración del sistema.
- Pruebas y correcciones.
- Evaluación y conclusiones.

1.4 Alcances y limitaciones.

1.4.1 Alcances.

En esta tesis de licenciatura se obtendrá un paquete completo de software para la manipulación de dispositivos para una casa inteligente por medio de la voz analizando el nivel de atención del usuario al dispositivo. Los módulos del paquete que se obtendrán serán:

- Modulo de reconocimiento de comando de voz para una casa inteligente.
- Modulo de reconocimiento de imágenes para seguir los ojos de un usuario.

Finalmente en esta tesis se obtendrá el manual de usuario y documento de tesis.

1.4.2 Limitaciones.

Por el tiempo que se tiene para realizar este proyecto de tesis solo se podrá simular algunos dispositivos domóticos, el número de comandos de voz será limitado, el sistema no será basado en Web.

Otra de las limitantes que se tiene por el hardware que se utilice, es la calidad de detección de ojos que dependerá totalmente del equipo con el que se cuente.

Se partirá de un modulo de control de dispositivos electrónicos en una casa inteligente a través de voz.

1.5 Hardware y Software a utilizar.

Debido a que el software que se realizará es un prototipo y no una implementación completa. El hardware y software que se menciona adelante será suficiente para este proyecto de tesis.

1.5.1 Hardware.

- Computadora Laptop: MacBook con procesador Intel core 2 Duo a 2.16 Ghz, memoria ram de 1024 M y con disco duro de 120 GB.
- Cámara de video: WebCam para puerto usb para obtener frames de imágenes reales.
- Bocinas integradas en la Laptop.
- Micrófono integrado en la Laptop.
- Dispositivos X10

1.5.2 Software.

- Java SE Development Kit (JDK): Se usará para hacer la codificación del programa en el ambiente de programación Java.
- Microsoft Windows XP Profesional version 2002.
- Modulo de control de dispositivos electrónicos en una casa inteligente a través de voz realizado en una tesis de licenciatura en la UDLAP.
- Matlab: Para el análisis de técnicas de procesamiento de imágenes.
- Internet: Para la búsqueda de información y búsqueda de software complementario para esta tesis.

1.6 Estado del problema.

Dentro del área de domótica se ha realizado varios esfuerzos para que se llegue a un sistema en el cual se puede obtener total comodidad para usar

cada uno de los dispositivos de la casa. Ahorrar es otra de las metas de la domótica que se debe tomar en cuenta al implementar el sistema ya que la gente gasta miles de peso en su implementación. Por último, la gente busca que un sistema domótico sea seguro, aunque a veces es difícil decidir el nivel de seguridad que se le implementara al sistema, es necesario que se implemente funciones básicas como avisos al usuario sobre el estado de los dispositivos, cambios en cada uno de los dispositivos de la casa, etc.

En el mercado existen varios protocolos que nos ayudan a implementar los sistemas domóticos. X-10 es uno de los protocolos que se usan comúnmente en las aplicaciones domóticas. Fue diseñado en Escocia entre los años 1976 y 1978 con el objetivo reducir costos de instalación de sistemas domóticos. Al usar las líneas de eléctricas de la vivienda, no es necesario tender nuevos cables para conectar dispositivos [4].

Existen tres tipos de dispositivos X-10, los que pueden transmitir ordenes, lo que solo reciben órdenes y las que pueden enviar/recibir órdenes. La ventaja de este protocolo es que pueden usar de manera libre ya que X10 no es un protocolo registrado por lo tanto no es necesario pagar derechos de autor [5].

Jini es un API desarrollada por Sun Microsystem permite conectar dispositivos y comunicarlos entre sí, compartiendo todos sus servicios. Jini es un sistema distribuido que se evita la configuración de la red para que los dispositivos se conecten [6].

Jini se tiene un conjunto de componentes que proporcionan la infraestructura necesaria para tener los servicios en un sistema distribuido, un modelo de programación que tiene un soporte e incita a una producción de servicios distribuidos y los servicios tiene la misma funcionalidad para cada miembro del sistema [7]. En la figura 1.1 nos muestra la información de cómo Jini hace realiza la búsqueda de dispositivos, unión de dispositivos a la red, buscar los clientes, registrar a los clientes, recepción de mensajes para uso del dispositivo y finalmente para que cada cliente pueda usar cada uno de los dispositivos disponibles

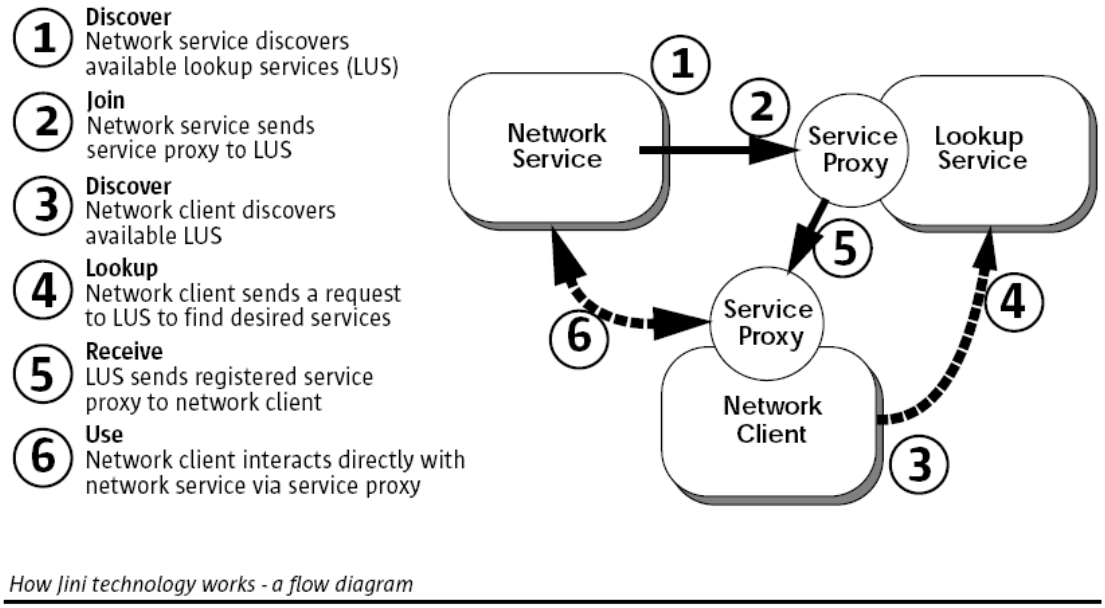

Figura 1.1 Forma en la que trabaja Jini [8].

Otra de las áreas que debemos de analizar es el reconocimiento de voz, en esta parte se han realizado varios avances que nos han ayudado a llegar a reconocedores de voz de alta calidad.

El “Realize voice” es un buen ejemplo de estos avances en reconocimiento de voz, este software nos ayuda a ejecutar una gran cantidad de comandos dentro del entorno Windows. Este software es funcional y eficaz reconociendo varios comandos básicos en el manejo de una computadora (cerrar ventana, maximizar, minimizar, etc.) el único problema es que solo funciona para un entorno totalmente en el idioma inglés [9].

El “Via Voice” es un software creado por IBM el cual fue diseñado para dispositivos móviles, este software permite un reconocimiento total de discursos acelerando el proceso de creación de documentos. Este software viene con un vocabulario de más de 20, 000 palabras en varios idiomas. Este es un software muy completo si se requiere hacer tareas como dictarle un documento a la computadora [10].

Otro del software de dictado es el “SpeechMagic” de Philips, este software también está totalmente completo, este software posee un soporte de 15,000

usuarios trabajando al mismo tiempo y un tiempo de dictado máximo de 1,500 horas. Este software posee más de 20 idiomas disponibles con terminología general, sobre el área de medicina y leyes. Todo esto hace al software de philips un gran candidato para gente que trabaja en hospitales [11].

El CSLU Toolkit es una librería que contiene un conjunto de herramientas que nos permiten explorar, aprender e investigar dentro de la interacción usuario-maquina mediante la voz [12].

El análisis hecho por NavigAvile en el cual nos muestran dos componentes de software. El primero es “Dragon Dictate” y el segundo es “Vt simple Speaking gold”. Este análisis se llevo a cabo con personas que tenían alteraciones en el habla.

Más allá de los resultados sobre cuál es el mejor software en reconocer la voz de una persona con dificultades en el habla, este análisis nos lleva a obtener ver los resultados de la evolución de los sistemas de reconocimiento de voz. Estos sistemas han evolucionado tanto que pese a las dificultades que muestre una persona al pronunciar algo nos demuestran que obtienen un resultado satisfactorio.

En la Universidad de las Américas, Puebla se desarrolló un proyecto de tesis en el cual se implementó un sistema domótica con reconocimiento de voz usando la tecnología de X10 para conectar los dispositivos de la casa, para el reconocimiento de voz se utilizo la librería CSLU Toolkit. Se partirá de este software para agregar la parte del análisis de los ojos del usuario con el fin de hacer una mejora a los sistemas domóticos existentes [14]

El movimiento de los ojos es muy diferente a las entradas convencionales de la computadora por eso importante analizar cada una de las técnicas de procesamiento del canal de ojos para poder llegar al objetivo requerido. Para analizar los ojos se ha investigado y creado técnicas que analizan la retina con el fin de saber hacia qué punto está viendo los ojos [2].

Uno de los métodos más sencillo para analizar el movimiento de los ojos es la grabación electrónica, esta utiliza electrodos colocados alrededor de los ojos para medir la diferencia de los cambios potenciales que existen en la orientación entre la cornea y la retina sin embargo este método solo nos permite analizar si el ojo se movió y no la posición exacta a donde el ojo se movió [10].

Uno es los métodos más prácticos para analizar el moviendo de los ojos es el procesamiento de imágenes. Ya que se puede analizar cada uno de las imágenes del usuario estudiando la luminosidad del iris o la reflexión de la parte delantera de la cornea. Para esta técnica, se puede usar solo el movimiento de los ojos o se puede complementar usando el movimiento de la cabeza con respecto a la de los ojos [12].

El sistema Eyegaze se utilizo en algunos experimentos para generar un conjunto de puntos de movimiento de la mirada sobre la localización de una cámara. En este experimento se minimizo el movimiento de la cabeza usando una pinza en la barbilla [3].

En la figura 1.2 se muestra el modelo de este sistema Eyegaze en el cual se analiza el movimiento de los ojos y se almacena cada imagen en una base de datos para después utilizarse con el fin de tener un panorama más amplio de otros tipos de movimiento similares.

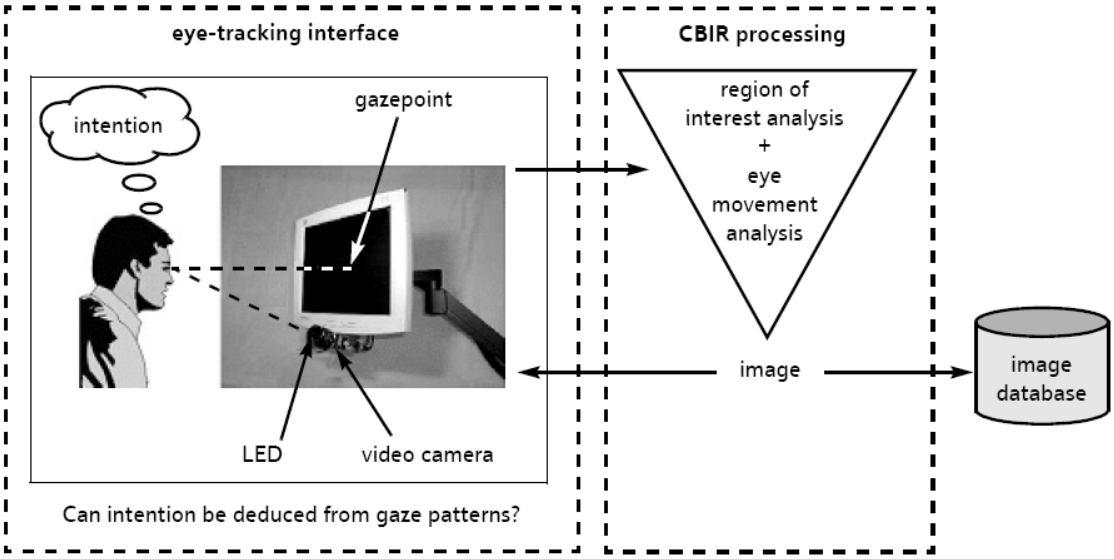

Figura 1.2 Modelo de Eyegaze [3].

1.7 Conclusiones.

En este capítulo se reviso cada uno de los avances hechos en las tres áreas diferentes: Domótica, procesamiento de voz y análisis de seguimiento de ojos. Cada una de estas áreas se complementan entre si y se ayudan con la finalidad de dar un producto que resuelve las necesidades del usuario final.

En el caso de domótica por voz se vio que existen sistemas realizados por otras personas y que no es necesario crear estos sistemas, sin embargo se pueden ampliar con la finalidad de hacer más fácil el control de dispositivos en una casa inteligente. En esta tesis se ampliaran los sistemas domóticos existentes con el fin darle al usuario una mayor comodidad y control de los dispositivos de la casa.