INTRODUCCIÓN

ESTADO DEL ARTE

La implantación de las tecnologías digitales ha posibilitado la transmisión y almacenamiento de grandes cantidades de información multimedia de forma eficiente, sin pérdida de calidad y a muy bajo coste. De todos es conocida la expansión que ha experimentado el formato CD (Compact Disc) de audio en los últimos años, así como la reciente aparición del DVD (Digital Video/Versatil Disc) para consumo de vídeo en formato digital, que cuenta con grandes perspectivas de crecimiento de mercado. Por otro lado, la televisión digital, tanto vía satélite como por cable, es ya algo cotidiano. Esta invasión de nuevos medios digitales ha sido posible gracias a la creciente escala de integración en los circuitos digitales y en soportes de almacenamiento físico, que ha conllevado un considerable abaratamiento de los equipos electrónicos para procesado y almacenamiento de información multimedia.

A todo esto hay que añadir el crecimiento sin precedentes de la red Internet que ha provocado que cantidades ingentes de información estén al alcance de un gran número de usuarios. Tanto es así, que el tráfico de información que fluye diariamente por Internet ha crecido de forma exponencial en los últimos años.

También el tipo de información transferida ha ido evolucionando rápidamente. De simples ficheros de datos se ha pasado a disponer de imágenes digitales y secuencias de audio y video que hasta pueden ser difundidas en tiempo real.

Esta nueva situación ha abierto las puertas a interesantes perspectivas para la utilización de la red Internet como plataforma para el desarrollo de actividades de comercio electrónico. De hecho es éste un sector que está experimentando un gran auge en la actualidad, y para el que se espera un gran crecimiento en años sucesivos. La red Internet se ha convertido no solamente en un medio atractivo para la realización de contratos de compra y transacciones comerciales, sino en una plataforma eficaz para la distribución comercial de información multimedia que ofrece grandes ventajas a los usuarios en cuanto a comodidad y precios, y que por tanto, pueden a su vez suponer un impulso sustancial al consumo de material audiovisual.

Todas estas grandes perspectivas que ofrecen las tecnologías digitales no están exentas, sin embargo, de nuevos peligros. Concretamente, los creadores de obras audiovisuales ven amenazados sus derechos de propiedad intelectual ya que es

posible realizar de forma masiva copias exactas de un trabajo almacenado en formato digital sin apenas esfuerzo y a un coste cada vez más bajo. Un claro ejemplo lo tenemos en las grabaciones de música en formato CD, cuyo contenido puede ser fácilmente copiado a un CD virgen por un coste mucho menor que el de un original. Problemas similares plantea el video digital en formato DVD. Cuando las creaciones audiovisuales sólo podían llegar al consumidor en formato analógico, el autor podía estar relativamente tranquilo porque el coste requerido para realizar una copia y la pérdida de calidad inherente al proceso hacían que en la mayoría de los casos resultase más conveniente comprar un original. A esto hay que añadir que la red Internet, además de ofrecer interesantes perspectivas para la distribución comercial de obras multimedia, al mismo tiempo constituye un atractivo canal para la transferencia de copias ilegales de obras en formato digital.

Esta seria amenaza contra la salvaguarda de los derechos de propiedad intelectual ha supuesto un gran freno a la implantación definitiva del comercio electrónico con obras en formato digital a través de redes abiertas como Internet. Los creadores no están interesados en la posibilidad de vender sus obras electrónicamente porque una vez entregada una copia, ésta puede ser replicada y redistribuida indiscriminadamente sin posibilidad alguna de control.

Se hace necesario, pues, proporcionar herramientas tecnológicas que garanticen al autor cierto grado de control sobre el uso y abuso de sus obras vendidas a través de una red. En este sentido, las tecnologías digitales tienen la ventaja de que permiten la introducción de mecanismos de control de forma más simple y eficiente que en el caso de los sistemas analógicos.

Las técnicas tradicionalmente utilizadas para restringir el acceso a la información, basadas en mecanismos de encriptación, no solucionan el problema planteado, pues, una vez extraída la versión en claro de la obra para su consumo, nadie impide que ésta pueda ser copiada ilegalmente. Es decir, los algoritmos de encriptación sólo resuelven el problema de la copia ilegal por acceso no permitido durante la distribución, pero una vez finalizada ésta, nos encontramos en la misma situación que al principio.

Por tanto, se hace patente que la única forma de proveer un mecanismo que permita, una vez finalizado el proceso de distribución, cierto grado de control sobre la utilización de una obra, es hacer que ésta vaya acompañada por información que sólo el autor pueda verificar y que no pueda ser eliminada fácilmente por personas no autorizadas para ello. Esto debe hacerse, además, teniendo en cuenta que la

información añadida no debe alterar de manera sustancial el contenido de la obra, es decir, la copia entregada al consumidor debe ser una fiel réplica del original. El hecho de que esta información de control no pueda ser eliminada fácilmente nos lleva a pensar que debe ir inmersa dentro de la propia obra de forma imperceptible, para que el destinatario pueda utilizar ésta como si fuera la original. Este concepto enlaza directamente con las técnicas de protección de los derechos de autor estudiadas en el presente documento y que reciben el nombre de técnicas de sellado digital con marcas de agua aplicables a documentos de texto, imágenes estáticas, señales de audio y/o vídeo.

No obstante, y aunque su estudio queda fuera de nuestro trabajo, las imágenes (también las señales de vídeo) pueden contener marcas de agua visibles, utilizadas, fundamentalmente, para su etiquetado y/o clasificación.

OBJETIVOS

Como resultado del planteamiento propuesto, surge la necesidad de estudiar y desarrollar técnicas que permitan proteger los derechos de autor y evitar posibles abusos por parte de usuarios no autorizados.

Una interesante solución propuesta recientemente para la protección de la propiedad intelectual en obras en formato digital la constituye el uso de las técnicas de sellado invisible, conocidas también como técnicas de inserción de marcas de agua (del término en inglés “watermarking”). Este conjunto de mecanismos podría enmarcarse, en una primera aproximación, dentro de la disciplina conocida como esteganografía, que estudia cómo introducir información oculta dentro de un medio o mensaje, aparentemente inofensivo, sin levantar sospechas.

Puesto que las técnicas de sellado digital están siendo objeto de una creciente atención, (dada su capacidad de proteger datos en entornos abiertos y poco controlados, donde la criptografía no puede ser aplicada con éxito), un estudio detenido de las características más importantes de las mismas será uno de los puntos fundamentales del presente proyecto.

Otros aspectos de interés a desarrollar serán:

Implementación de diferentes algoritmos de sellado con marcas de agua, aplicados al tratamiento de imágenes estáticas. También se desarrollarán los correspondientes algoritmos de detección y extracción de las señales de marcado con el fin de poder comparar tanto las técnicas de inserción como de detección.

Estudio de las características de robustez de las diferentes técnicas de sellado implementadas, mediante la aplicación de una variedad de ataques y/o técnicas de procesado de imágenes comunes a las versiones marcadas, que incluyen el empleo de bancos de prueba disponibles para el estudio de técnicas de watermarking.

Comparación de resultados y conclusiones.

Planteamiento de problemas abiertos y posibles líneas de investigación futuras.

Todos los puntos enumerados se irán desarrollando en los diferentes apartados de la memoria del proyecto. Indicar aquí que entre las técnicas de detección de marcas de agua que se estudiarán se encuentra una basada en el uso del procesado de análisis de componentes independientes, ICA, técnica a la que se le dedicará un capítulo completo del proyecto, dados su creciente interés y utilidad en muchas aplicaciones actuales de procesado de señal.

ESTRUCTURA DE LA MEMORIA

El presente proyecto ha sido estructurado en cinco bloques, cada uno de ellos relacionado con aspectos fundamentales tratados en el documento. La primera parte se dedica al planteamiento y descripción del problema de watermarking o sellado digital de imágenes con marcas de agua, en general. Está dividida en cuatro capítulos, en el primero de los cuales se aporta una introducción general que nos sitúa en los aspectos más importantes de las técnicas de watermarking, dedicándose los tres restantes a un estudio más detallado de las diferentes etapas de un proceso completo de sellado digital con marcas de agua, como son la etapa de inserción de la marca de agua en el documento original (en nuestro caso una imagen), la de realización de ataques a la imagen sellada y la de recuperación de la marca de agua. El último capítulo de este primer bloque guarda cierta relación con la segunda parte del documento, dedicada al estudio de las características fundamentales del Análisis de Componentes Independientes usado, en nuestro caso, para la recuperación de la marca de agua. En el tercer bloque del proyecto se describen los diferentes algoritmos de watermarking implementados, con mención de los resultados más significativos obtenidos en cuanto a robustez, grado de imperceptibilidad de la marca de agua, eficiencia computacional, etcétera…así como una comparativa de los mismos según criterios. Se enumeran conclusiones y problemas abiertos para futuros proyectos en la cuarta parte del documento, para finalmente dedicar el último de los bloques a documentación adicional y anexos.

Capítulo 1. Introducción a las técnicas de sellado digital con marcas de agua

1.1 Introducción

Desde principios de los noventa las técnicas de sellado digital de imágenes mediante marcas de agua (del término en inglés “watermark”) han gozado de un interés creciente en aplicaciones de inserción de información escondida en datos multimedia. La primera conferencia académica de este tema se celebró en 1996 [Barni01]. Las marcas de agua digitales tienen, principalmente, tres campos de aplicación: monitorización de datos, protección del copyright y autentificación. Los primeros métodos de sellado mediante marcas de agua fueron propuestos por Caronni y aplicados a imágenes digitales en 1993, aunque ya documentos anteriores incluían la idea de insertar información escondida en las imágenes para conseguir la protección de los derechos de autor [Kutter99]. Desde entonces, la idea de sellado digital mediante marcas de agua ha sido extendida a otros tipos de datos tales como audio y video.

En este capítulo se van a introducir conceptos fundamentales relacionados con el problema de watermarking. Se presenta además, una breve introducción histórica.

1.2 Introducción histórica

Puesto que el concepto de sellado con marcas de agua se encuentra muy relacionado con la llamada esteganografía (como se verá en el apartado siguiente donde se definirán ambos términos formalmente), una introducción histórica del término watermarking se reduce a la mención de las técnicas más curiosas de esteganografía llevadas a cabo a lo largo de la historia, [Barni01], [Johnson], [Johnson98].

El concepto de esteganografía era ya usado en la Antigüedad, así Herodoto relata cómo los griegos recibieron mensajes escondidos bajo la cera de tablas de madera escritas con textos que nada hacían sospechar de la existencia de mensajes ocultos (para ver esta información sólo era necesario fundir la cera envolvente) que alertaban sobre las intenciones de Xerxes, y describe, además, una forma de burlar la seguridad mediante la incrustación de

mensajes en textos usando claves secretas, debido a Aeneas el Tacticiano. Otro ingenioso método consistía en afeitar la cabeza del mensajero y tatuar o grabar el mensaje en ella, luego se dejaba que el pelo creciera, de manera que el mensaje no fuera detectado hasta que se le volvía a afeitar la cabeza. Kanh relata una tradicional táctica china que consistía en insertar un código de ideograma en un lugar acordado previamente de un documento; la misma idea fue utilizada en la Europa medieval con sistemas en los que un papel o madera templada era colocada sobre el mensaje, con un texto aparentemente inofensivo encubriendo la información secreta.

Tales sistemas únicamente cobran sentido cuando existe un oponente. Este adversario puede ser pasivo, que simplemente observa el tráfico, o activo, que intenta modificarlo. Un caso famoso ocurrió en 1586 cuando la reina Mary de Escocia intentó conspirar para asesinar a la reina Elizabeth de Inglaterra, con la intención de destruir el trono. Sin embargo el cifrado que ella usó fue interceptado, la policía secreta de Inglaterra obtuvo los nombres de los señores involucrados en la conspiración, procediendo a su arresto y ejecución.

El estudio de este concepto dentro de la literatura científica se debe a Simmons, quien el 1983 formuló el llamado “Problema de los prisioneros”, [Cachin98]. En este escenario, Alice y Bob son dos prisioneros que desean diseñar un plan para escapar, pasando todos sus mensajes por el guardián de la prisión, Willie. Si éste detecta cualquier mensaje encriptado o codificado destruirá el plan, enviándolos a celdas incomunicadas. De manera que ellos deben encontrar la forma de esconder el texto cifrado en el interior de mensajes aparentemente inocuos. Al igual que en el campo de la criptografía, suponemos que el mecanismo usado para cifrar puede ser conocido por el guardián, que únicamente desconoce la clave secreta que Alice y Bob comparten.

La formulación del Problema de los Prisioneros de Simmons fue sólo un ejemplo de información escondida en textos inofensivos. Fue un ruso el que consiguió que la comunidad académica prestara atención en el desarrollo de una aplicación crítica y a la vez clasificada: pruebas de control para la verificación de existencia de armas nucleares. Los EEUU y la URSS querían colocar sensores de manera que proporcionaran cierta información (como el número de misiles), pero que no revelaran otros tipos de información (como la localización de los misiles, por ejemplo). Esto forzó un cuidadoso estudio acerca de los modos en que el equipo de un determinado país podría pasar bajo cuerda datos olvidados

para las facilidades de monitoreo de otros países, de manera que se pudiera disponer de esa información que ciertos países no deseaban proporcionar.

Por tanto, aunque la aplicación es diferente, los conceptos y técnicas empleadas en watermarking tienen su antecedente en procedimientos conocidos y utilizados ya en tiempos remotos.

1.3 Conceptos fundamentales

En esta sección se definen, de una manera formal, los conceptos más importantes relacionados con el tema objeto de estudio.

1.3.1 Sellado con marcas de agua. Justificación y alcance

El sellado con marcas de agua de un documento cualquiera, ya sea un texto, una imagen o una señal de audio o video, es una técnica que trata de encajar una señal llamada *marca de agua* (w) dentro de la misma información a proteger, considerando, para ello, diferentes aspectos tales como invisibilidad y robustez de la señal añadida. Así, la marca de agua insertada deberá no ser perceptible, es decir, no debe degradar la calidad visual de la imagen, pero además, deberá tener la capacidad suficiente como para resistir a diferentes procesados y/o ataques que pudiera sufrir el documento original y por tanto la propia señal de marca de agua.

Para su realización se determina, en primer lugar, la información que se va a incluir en la imagen, como por ejemplo la identificación del propietario, caso de que la técnica de sellado digital sea utilizada para la protección de los derechos de autor y/o de copia. Notar a este respecto que en el presente proyecto se estudian propiedades de robustez resultado de aplicar una señal de marca de agua a una imagen, quedando fuera de los límites del mismo la relación que pueda existir entre la señal de marca de agua y el posible mensaje asociado. Es decir, nuestra señal de marca de agua sería como la salida del codificador de fuente en un sistema de comunicaciones.

A pesar de que las técnicas de sellado con marcas de agua pueden ser aplicadas a cualquier tipo de información multimedia, sea audio, video o imágenes fijas, nuestro trabajo de investigación se ha centrado en el estudio del problema asociado al marcado de imágenes estáticas, la razón principal radica en la existencia, en este caso, de bancos de prueba de libre difusión para comprobar la calidad y robustez de la técnica. Además, la protección de

imágenes despierta un gran interés comercial de cara a la distribución de las mismas en formato digital a través de redes abiertas como Internet.

1.3.2 Esteganografía y encriptación

Pero las técnicas de sellado con marcas de agua pueden verse dentro de un concepto aún más amplio como es el de *esteganografía,* término derivado del griego que significa literalmente *"mensaje encubierto"* y se refiere al arte de insertar información (de cualquier tipo) en diferentes documentos, ya sean textos, imágenes, señales de audio o video, de tal manera que resulte difícil detectar mencionada información secreta.

El proceso de inserción del mensaje secreto se realiza, normalmente, mediante el uso de una *clave*; sin ella resultaría difícil, para una tercera persona, detectar o borrar el material insertado. Una vez el objeto original tiene el mensaje secreto insertado, este será referido como *estego objeto*. En el caso que nos ocupa, puesto que el objeto que porta el sello será una imagen estática, una vez marcada será conocida con el nombre de *estego imagen*.

Se han desarrollado una gran cantidad de diferentes técnicas para esconder la información en una imagen digital. Algunos de los métodos más comunes serán brevemente descritos en el capítulo siguiente. Cada una de estas técnicas puede ser aplicada con diferentes grados de variación a diferentes ficheros de imagen.

A pesar de hacer uso de la misma idea, la de esconder cierta información en un documento, existe una diferencia notable entre la *esteganografía* y las técnicas de sellado digital y viene determinada por el objetivo que persigue cada uno de los mecanismos mencionados. Así, mientras en la *esteganografía* el mensaje importante es el que se encuentra escondido en la imagen, texto o señal de audio usada para evitar crear sospechas sobre la existencia de un mensaje oculto, en las técnicas de watermarking la finalidad principal es proteger el documento sellado, en nuestro caso la imagen digital, de manera que lo verdaderamente importante es el documento en sí, no la información secreta añadida.

El concepto de *esteganografía* no debe ser confundido con el de *criptografía*, donde se transforma el mensaje de manera que la persona que lo intercepta no sea capaz de descifrarlo, pero sí de detectarlo. Tal protección es, a menudo, insuficiente. La detección de un tráfico de mensajes cifrados tiene

implicaciones obvias, infunde sospechas, mientras que si el mensaje estuviera escondido no se detectaría nada anormal.

1.3.3 Analogías con sistemas de comunicaciones

Las técnicas de sellado invisible de imágenes se pueden modelar como un sistema de comunicaciones tradicional [Barni01], donde el transmisor sería la parte del sistema encargada de generar la marca de agua a insertar en el documento original y el receptor lo constituiría el detector de la marca de agua, siendo el canal de comunicaciones representado por la propia imagen y los posibles ataques. Sin embargo, a diferencia de los canales tradicionales de comunicaciones, las técnicas de sellado invisible con marcas de agua deben prestar una especial atención para mantener la fidelidad del trabajo original, es por ello que se consideran muchas restricciones en el momento de la inserción de la marca de agua. Esta restricción de fidelidad junto con los requisitos de que la marca sea robusta (es decir, que resista a una serie de ataques comunes como pueden ser las técnicas de procesado de señal aplicadas al trabajo), y segura (que no pueda ser borrada por ataques que persigan mencionado objetivo), lleva a muchos desarrolladores a considerar el marcado de agua como una transmisión digital, y a usar técnicas de modulación como la de espectro expandido, (los sistemas de comunicaciones de espectro ensanchado transmiten muy poca potencia en cualquier frecuencia individual y presentan buenas propiedades frente a la interferencia lo que los hace adecuados para aplicaciones de watermarking). Véase capítulo 4 para más información.

1.4 Conclusiones

A la luz de lo visto hasta el momento se puede concluir que para disponer de un estudio riguroso del problema será necesario realizar un completo análisis de los siguientes elementos:

- **De la imagen**. Es conveniente disponer de un modelo de la imagen, donde se indicará el formato más adecuado de la misma de cara a su protección frente a los ataque principales que pueda sufrir.
- **Modelo del sistema visual humano**, que nos permita disponer de conocimientos suficientes de manera que se inserte la marca de agua en la localización más adecuada de cara a obtener una buena eficiencia del método de protección en cuanto a robustez e invisibilidad de la marca.

- **Modelo de la técnica de watermarking a emplear**. No todas las técnicas de watermarking son igualmente robustas a los diferentes ataques posibles. Las dificultades asociadas con este problema pueden ser ilustradas mediante la consideración de los compromisos relacionados con la robustez de la marca de agua, su imperceptibilidad y la capacidad de información.

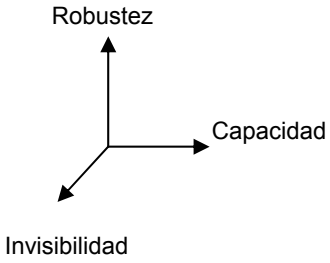

FIGURA 1.1. Esquema simplificado que ilustra el espacio de todo sistema de watermarking. Una técnica de watermarking no puede simultáneamente maximizar estas tres cantidades, que tienden a trabajar de forma opuesta [Heileman98].

- **Modelos de ataques y resultados de aplicarlos a la imagen protegida**. El estudio de los resultados se realizará adoptándose algunas medidas de calidad de la imagen decodificada y estableciéndose un umbral que delimite si la imagen es apta o no dependiendo del grado de alteración que haya sufrido con respecto a la original.
- **Modelo de las técnicas de detección y extracción** de la marca de agua.

Aspectos a los que se les irá dando forma en los diferentes capítulos de la memoria.

A partir del parecido encontrado entre los sistemas de comunicaciones tradicionales y cualquier proceso de watermarking, se observa que su estudio puede dividirse en tres etapas fundamentales:

- **Generación** de la marca de agua (transmisor de información).
- Posibles **ataques** (transmisión a través del canal).
- **Recuperación** de la marca (decodificación de la información en el receptor).