



Recomendaciones: Microfilmación de Conservación de Periódicos para Digitalización

Introducción

- La accesibilidad de los medios electrónicos y la larga perdurabilidad del microfilme son una buena combinación. La microfilmación de conservación debería durar 500 años, lo que proporciona a las frágiles colecciones transferidas a este soporte un futuro de seguridad archivística. Al mismo tiempo, la digitalización facilita el acceso a estas colecciones. Resulta posible combinar estos procesos de una forma más económica si se consideran los aspectos de la digitalización ya en el momento del proceso de microfilmación.
- Estas recomendaciones constituyen un apéndice a las Directrices de la IFLA para la Microfilmación de Conservación de Periódicos, y por ello harán referencia a estas Directrices y a las Normas Internacionales de Microfilmación de Conservación de Periódicos. Las recomendaciones no repetirán las Directrices o Normas, puesto que éstas constituyen la plataforma para una correcta microfilmación de conservación y su futura digitalización.
- El objetivo de estas recomendaciones ha sido facilitar el uso del microfilme como una plataforma para una futura digitalización y búsqueda a texto completo de las colecciones.

Dentro de poco, los periódicos en microfilme se convertirán en una parte de los programas de digitalización de nuestras bibliotecas en todas partes. Durante décadas, la microfilmación de conservación ha sido el medio de conservación y preservación de las colecciones de prensa en todo el mundo. La microfilmación constituye aún la elección natural para la preservación de periódicos en formato papel debido a la escasa calidad del papel y a su gran tamaño.

Digitalización de microfilme

Los microfilmes de conservación de nuestras colecciones presentan ventajas y desventajas relacionadas con su posible digitalización. Los microfilmes de alto contraste tienen una gama dinámica limitada, lo que en muchos casos es bueno. Es una buena característica especialmente si comparamos los resultados del escaneado de los originales y los microfilmes de texto gris en papel quebradizo o pardo. Se eliminan también sombras pequeñas, manchas y arrugas del papel. El texto es claramente separable de la base. Ello favorece las posibilidades de búsqueda a texto completo de los periódicos, que utiliza programas de Reconocimiento Óptico de Caracteres (OCR) en imágenes bitonales. La calidad de resolución que precisa el film es suficiente incluyendo índices de reducción de 16x y favorece un Índice de Calidad de 8.

Por otra parte, el material fotográfico nunca será tan bueno como la digitalización en escala de grises desde el original. Pero hay también una clara diferencia entre los diferentes tipos de copias, siendo los duplicados directos negativos los que dieron los peores resultados para las fotografías en los periódicos.

En el Proyecto Nórdico de Digitalización de Periódicos (TIDEN) se detectaron cuatro elementos que influyeron el resultado de la lectura OCR:

- el texto del original, la calidad del texto original/texto fuente, diseños complejos
- el índice de reducción de la cámara
- la fuente utilizada (Romana o Gótica, fuentes múltiples), el tamaño de la fuente
- la lengua (múltiples lenguas)

Sin embargo, la influencia de la marca comercial y resolución de la cámara, la polaridad y la generación de la película fueron marginales al comparar los resultados del OCR.

En cualquier caso, debe indicarse que no fue posible hacer una lectura OCR de las películas antiguas de los años 1950 y 1960, mientras que los micros de los años 80 basados en normas internacionales sí pueden leerse con OCR siempre que los periódicos fueran desencuadrados antes de su microfilmación.

Como base para las recomendaciones se han utilizado la experiencia y conocimientos del proyecto TIDEN, la Biblioteca Nórdica de Prensa Digitalizada de 1640-1900 y el proyecto AURORA para la Biblioteca Finlandesa de Prensa Digitalizada de 1771-1860.

Principios

Existen hoy muchas posibilidades de combinar microfilmación y digitalización. Es posible microfilmar primero y digitalizar a partir del microfilm. Otra posibilidad es digitalizar y preservar el contenido en película mediante COM (Computer output on microfilm, salida de ordenador en microfilm). También pueden utilizarse cámaras híbridas de microfilmación y digitalización para hacer ambos procesos de forma simultánea. En el caso de los periódicos, la primera opción es con mucho la más utilizada.

En estas recomendaciones nos centramos en el primer aspecto que destacan las Directrices de la IFLA para la Microfilmación de Periódicos. Se persigue una película de alta calidad que satisfaga la calidad de imagen y las necesidades de búsqueda avanzada de los usuarios de bibliotecas de prensa digitalizada.

Para permitir la búsqueda a texto completo de las colecciones, los originales de los periódicos deben ser de gran calidad. El objetivo es conseguir imágenes de alto nivel y microfilmar nuestras colecciones una sola vez, evitando así volver a microfilmar cada vez que la tecnología exija mejores imágenes. Hay que perseguir la mejor calidad de imagen digitalizada y la mejor posibilidad de búsqueda a texto completo y los menores gastos de conversión.

Estas recomendaciones intentan proporcionar a las bibliotecas algunas sugerencias para sus propios planes de producción y directrices para futuras digitalizaciones de microfilmes. Con el microfilme como plataforma es posible convertir nuestras colecciones a cualquier otro soporte tecnológico de los siglos futuros cuando sea necesario.

Las cuestiones relativas a criterios de selección y de propiedad intelectual no se tratarán aquí, puesto que son temas de mantenimiento de las colecciones. Tampoco se considerarán los aspectos de requisitos bibliográficos,

indización de ediciones de periódicos, suplementos, control de calidad y elección del centro microfilmador, puesto que ya han sido tratados en las Directrices de la IFLA.

Proceso de Preparación

Concienciación

Es posible conseguir una búsqueda a texto completo del microfilm escaneado tan buena como del escaneado de los originales. La resolución de un índice de reducción de 35 mm de microfilm de grano fino equivale a 12.5 millones de píxels, lo que supone una buena plataforma para el usuario en años futuros. Sin embargo, no siempre se utilizan las posibilidades de la película debido a la falta de una buena práctica del laboratorio durante la preparación y los procesos de microfilmación. Para aumentar la calidad se precisa un programa de calidad que puede desarrollarse durante la ejecución de los procesos. Es importante elevar la concienciación del personal implicado.

Calidad del original

La calidad de los originales es un aspecto crucial para su uso futuro. A la hora de preparar los periódicos es esencial separar las páginas de mala calidad y sustituirlas por otras copias si fuera posible. En el caso de material roto y quebradizo deberán conservarse las páginas, especialmente en el área de la información. Deben eliminarse las arrugas.

Los procedimientos de preparación deberán tener como objetivo la lectura OCR de los archivos de imágenes digitalizadas del microfilme. Esto incluye que las páginas sean alisadas y que las columnas de texto sean rectas, incluso los márgenes internos. En los procedimientos de preparación esto supone:

- desencuadernar los periódicos
- microfilmar los periódicos encuadernados en un soporte para libros bajo un cristal (en este caso la distancia en el margen interno deberá ser 1,5 cm)

A la hora de digitalizar los microfilmes es importante la identificación de los números de los periódicos, de las páginas, ediciones y suplementos. También sería importante advertir de la necesidad de volver a microfilmar algunas páginas. Este proceso podría prepararse en esta fase y desarrollarse durante la microfilmación. Si la biblioteca decide indizar el contenido de la película se facilitará también el uso futuro de las imágenes escaneadas.

Proceso de Microfilmación

Mesa soporte de originales

Existen niveles más y menos sofisticados para acometer las necesidades del proceso de escaneado. Normalmente los escáneres de microfilmación tienen problemas a la hora de identificar los distintos tamaños de las páginas en la película. Ello se debe a que los escáneres suelen reconocer un nuevo fotograma en un sitio indicado, en su esquina o margen izquierdo. Los problemas de identificación aparecen cuando cambia el

formato del periódico o de su suplemento. Los suplementos también pueden ser filmados en una dirección distinta que la del periódico original. A la hora de filmar periódicos con doble página, el cambio en el número del periódico puede causar problemas si se filma separada la primera página de un nuevo número. En este caso, el escáner no puede reconocer el fotograma en el que la parte izquierda está vacía y por consiguiente está en negro.

Para poder identificar cada imagen en la película, la mesa soporte puede marcar la parte izquierda del área del fotograma con una tira de 2-3 cm de ancho de cinta adhesiva blanca, que servirá para indicarle al escáner el comienzo de un fotograma. Un sistema más sofisticado consiste en colocar una barra blanca en el lado izquierdo, que puede incluir toda o parte de la siguiente información:

- información relativa al tamaño de la unidad (en cm o pulgadas)
- códigos de barras para símbolos fijos relativos a: cambios en la numeración del periódico, (página), nueva edición (si se filman ediciones cambiadas), suplemento nuevo, repetición de página, original defectuoso, etc. Esta información podría también indizarse hasta cierto punto mediante la función blip de algunas cámaras (inclusión de un "blip" o mancha que sirve de referencia para localizar un fotograma determinado o para una posible digitalización del microfilme de forma automatizada).
- un indicador de escala de grises.

Proceso de microfilmación

La tira o barra blancas arriba mencionadas pueden también servir de ayuda al operador de la cámara para mantener rectas las páginas del periódico durante el proceso. Si se ha movido el periódico, lo que es difícilmente detectable por el ojo humano, es complicado para el escáner encontrar el umbral correcto, que se encuentra a 1-3 mm en dirección izquierda-derecha.

En las películas antiguas, el área de película entre los fotogramas puede ser demasiado pequeña para que el escáner divida los fotogramas. A la hora de microfilmarse deberán utilizarse áreas iguales.

La reducción del periódico influye directamente sobre la calidad de la imagen y los resultados del OCR. Deberá utilizarse el menor índice de reducción posible. 10x es mejor que 14x, y 14x mejor que 16x. Cuando el índice de reducción es 20x o superior a 24x no puede ejecutarse OCR.

La microfilmación de números o suplementos en direcciones distintas puede ser causa de problemas a la hora de hacer disponibles las imágenes escaneadas.

Para los volúmenes encuadernados de prensa se utilizan cámaras con soporte de libros. En estos casos, los periódicos se microfilman bajo una plancha de cristal. Si la manipulación de volúmenes gruesos resulta problemática, los responsables de la biblioteca deberán buscar posibles soluciones.

Los volúmenes sin encuadernar se microfilman en mesas de microfilmación tradicionales.

Aspectos técnicos

A la hora de microfilmear para digitalizar es importante la iluminación homogénea en cada fotograma. Los cambios de densidad entre fotogramas exigirán manejos especiales del operador del escaneado o se reducirán los resultados de la lectura OCR. Densidades diferentes en las exposiciones precisan ajustes manuales de los parámetros del escaneado y obligan a interrumpir el proceso en cada imagen. Ello resulta en un coste mucho mayor en tiempo y dinero que el escaneado de la película sin interrupciones.

Las lecturas de densidad deben ajustarse a las normas, pero es preferible que sean inferiores a las que permiten las normas, es decir, .90-1.20. Estas lecturas pueden variar entre diferentes escáneres de microfilme. Las lecturas de densidad deben estar comprendidas entre 0.2 dentro de una imagen y dentro de la película.

Procedimiento Esquemático

Microfilmación para digitalización

- Ceñirse a las normas nacionales e internacionales relativas a microfilmación y a las directrices de la IFLA.

Microfilmación para digitalización y búsqueda a texto completo

- Ceñirse a las normas nacionales e internacionales relativas a microfilmación y a las directrices de la IFLA y a su suplemento.
- Por lo que respecta a OCR, los aspectos más importantes para la Biblioteca son la calidad del original, la filmación de volúmenes des encuadernados y la reducción.

Microfilmación para digitalización y búsqueda a texto completo para procesos automáticos de escaneado futuros

- Ceñirse a las normas nacionales e internacionales relativas a microfilmación y a las directrices de la IFLA y a su suplemento.
- Por lo que respecta a OCR, los aspectos más importantes para la Biblioteca son la calidad del original, la filmación de volúmenes des encuadernados y la reducción.
- Para los procesos de escaneado automático, las lecturas de densidad deben estar comprendidas entre 0.9 y 1.2 y no deben diferir más de 0.2 entre cada imagen. Debe colocarse una tira blanca en la mesa de la cámara para indicar el comienzo de la imagen. Debe utilizarse el blip, o indización legible por máquina mediante códigos de barras, para los números de los periódicos, (páginas), suplementos, etc.

Aspectos Económicos

En un principio, el desarrollo del programa de calidad precisará de financiación extra, que se utilizará para su planificación y ejecución en la biblioteca y en el centro de microfilmación. El resultado será un programa de microfilmación que puede ser la base de los futuros programas de digitalización. El impacto financiero debe ser razonablemente suficiente.

Aspectos que pueden influir sobre los costes:

- Una reducción de 16x puede variar el programa si se han utilizado reducciones superiores
- Desencuadrar los periódicos puede suponer de media a una hora por volumen
- Si se utilizan códigos de barras, el cambio de éstos exigirá mayor atención del operador de cámara y puede ralentizar el proceso de microfilmación.
- Lecturas nuevas de densidad pueden elevar las exigencias de calidad y obligar a tomar la decisión de volver a filmar.

Resultados:

- La OCR permitirá la búsqueda a texto completo de las colecciones de periódicos
- El proceso de escaneado puede ser automatizado
- Los números, páginas, suplementos pueden ser identificados automáticamente.

Suplementos (no incluidos)

1. Filmación de hojas sueltas
2. Filmación mediante cámara Zeutchel con soporte de libros
3. Cámara de microfilmación con tira blanca y código de barras
4. Texto leído mediante OCR: buen ejemplo y ejemplos problemáticos
5. Dibujo esquemático del reconocimiento de la exposición de la película
6. Búsqueda a texto libre