

Modelos Tridimensionales Utilizados para la Codificación de Análisis–Síntesis Basada en Objetos

Geovanni Martínez

Escuela de Ingeniería Eléctrica, Universidad de Costa Rica

2060, San José, Costa Rica

email: gmartin@pacuare.eie.ucr.ac.cr

RESUMEN

En este artículo se comparan los diferentes modelos tridimensionales propuestos en la literatura profesional para la codificación de análisis–síntesis basada en objetos. La codificación basada en objetos se utiliza para la compresión de señales de vídeo a bajos "bit–rates". La compresión de señales de vídeo es indispensable para ofrecer los servicios de videotelefonía y videoconferencia en la Red Digital de Servicios Integrados ISDN ("Integrated Services Digital Network"), telefonía celular e internet. Dichos modelos describen los objetos reales a través de modelos tridimensionales definidos por parámetros de forma, de movimiento y de color. Para señales de vídeo reales (CIF, 10 Hz) y una calidad de imagen de PSNR=38 dB, el modelo de objetos articulados resultó ser un 20% más eficiente que el modelo de objetos rígidos y un 10% más eficiente que el modelo de objetos flexibles.

1. INTRODUCCION

Para la transmisión de vídeo a bajos "bit–rates" se investiga actualmente el codificador de análisis–síntesis basado en objetos[5]. Un codificador basado en objetos divide cada imagen de una secuencia en objetos en movimiento y describe cada objeto mediante parámetros de forma, de movimiento y de color (ver Fig. 1). Los parámetros de color representan los valores de "luminance" y "chrominance" de la textura sobre la superficie del objeto. El significado exacto de los parámetros de forma y de movimiento depende del modelo utilizado. Los parámetros son estimados automáticamente por un analizador de imágenes. Para ello se evalúa la imagen actual y la última imagen transmitida. Por cada objeto se transmiten únicamente los parámetros de forma y de movimiento. Una reconstrucción de la imagen original se realiza haciendo uso de los parámetros actuales de forma y de movimiento, así como de los parámetros de color previamente transmitidos. Los parámetros que describen el color y la forma

bidimensional de aquellas regiones de la imagen actual, que no pudieron ser reconstruidas con suficiente calidad, deben ser transmitidos también al receptor. Dichas regiones reciben el nombre de objetos MF ("Model Failure") o regiones de modelado no satisfactorio. Debido a que la transmisión de parámetros de color requiere de un alto "bit–rate", el tamaño de los objetos MF debe ser mantenido lo más pequeño posible.

Hasta la fecha han sido propuestos tres modelos tridimensionales para la codificación de análisis–síntesis basada en objetos: el modelo de objetos rígidos tridimensionales R3D ("Rigid Three–Dimensional Objects")[6], el modelo de objetos flexibles tridimensionales F3D ("Flexible Three–Dimensional Objects")[7] y el modelo de objetos articulados tridimensionales A3D ("Articulated Three–Dimensional Objects")[2][3]. En este artículo se hará una comparación de esos modelos. Para ello se medirá y comparará la calidad de imagen resultante después de la codificación de algunas señales de vídeo reales. Para medir la calidad de imagen se utilizará el criterio objetivo denominado comúnmente en la literatura profesional como PSNR ("Peak Signal to Noise Ratio").

Este artículo está organizado como sigue: en la sección 2 se describen los modelos tridimensionales citados anteriormente. En la sección 3 se describe como se codifican y transmiten sus parámetros. En la sección 4 se describen los resultados experimentales. Un resumen del artículo se presenta finalmente en la sección 5.

2. MODELOS

En el modelo R3D la forma tridimensional de un objeto se describe mediante una red de triángulos (ver Fig. 2.a.i). Los vértices de la red de triángulos representan los parámetros de forma. La red de triángulos se calcula a partir de la silueta del objeto mediante una transformación. El movimiento de un objeto rígido se describe mediante seis parámetros: tres ángulos de rotación y un vector tridimensional de desplazamiento (ver Fig. 2.a.ii). Los parámetros

de color de un objeto se definen mediante la proyección de una imagen real sobre la superficie de su red de triángulos (ver Fig. 2.a.iii). Mientras el modelo R3D permite únicamente la descripción de objetos rígidos, el modelo F3D permite también una deformación local tangencial a la superficie de los objetos. Esa deformación se describe mediante vectores de deformación (ver Fig. 2.b.i). El resto de los parámetros del modelo F3D se definen como en el modelo R3D. Según el modelo de objetos A3D los objetos reales pueden ser articulados y sus componentes moverse en forma diferente. Cada componente se describe con sus propios parámetros de forma, de movimiento y de color (ver Fig. 2.c). La forma de un componente se asume rígida y se representa mediante una red de triángulos (ver Fig. 2.c.i). Los vértices representan los parámetros de forma de un componente. La forma de los componentes se obtiene dividiendo una primera forma rígida en componentes[3][4]. Esa primera forma rígida se calcula, al igual que en el modelo R3D, de la silueta del objeto mediante una transformación. Los componentes de un objeto articulado se suponen acoplados unos con otros a través de articulaciones esféricas. Las articulaciones se consideran parte de la forma de un objeto articulado (ver Fig. 2.c.i) y sus posiciones se representan mediante vectores de posición tridimensional. El movimiento tridimensional de un componente se describe también mediante seis parámetros (ver Fig. 2.c.ii). Los parámetros de color de un componente se definen también mediante la proyección de una imagen real sobre la superficie de su red de triángulos (ver Fig. 2.c.iii).

3. CODIFICACION

Los parámetros estimados de cada objeto son codificados y finalmente transmitidos al receptor. La silueta del objeto real se codifica mediante "polygon/spline approximation". Para codificar la subdivisión en componentes, a cada vértice de la red de triángulos se le asigna un número que indica a cual componente pertenece. Los números son posteriormente codificados mediante "run length coding". Por cada articulación se codifican las tres coordenadas de su vector de posición. Cada coordenada es cuantizada primero en forma uniforme y posteriormente transmitida en PCM. En los modelos R3D y F3D los parámetros de rotación y translación de cada objeto son cuantizados en forma uniforme y posteriormente transmitidos en PCM. En el caso del modelo A3D los parámetros de rotación y translación del componente más grande y únicamente los parámetros de rotación del resto de los componentes son cuantizados en forma uniforme y luego transmitidos en PCM. Para transmitir los vectores de deformación del modelo F3D a cada vértice de la red de triángulos se le asigna un número que indica si contiene o no un vector de deformación. Los números son posteriormente codificados mediante "run length coding". Los componentes de los vectores de deformación se cuantizan en forma uniforme y

luego se les elimina la redundancia mediante "entropy coding". La forma bidimensional de los objetos MF se codifica también mediante "polygon/spline approximation". Los parámetros de color se codifican mediante la transformada discreta del coseno DCT ("Discrete Cosinus Transformation") para regiones de forma arbitraria.

4. RESULTADOS EXPERIMENTALES

Los codificadores basados en objetos así como el codificador basado en bloques según el standard H.263[1] fueron aplicados a las secuencias de prueba "Claire" y "Miss America" con formato CIF y un índice reducido de repetición de imágenes de 10Hz. Cada secuencia muestra un objeto articulado (persona) constituido por los componentes: cabeza y hombros. Los componentes se mueven en forma diferente. El tamaño de las regiones de modelado no satisfactorio para las cuales se deben transmitir los parámetros de color, fue en promedio de un 4% del área total de la imagen para el modelo R3D y de un 3% para los modelos F3D y A3D. Mediante los modelos A3D y F3D se logra una mejor descripción de la imagen actual que haciendo uso del modelo de objetos R3D. Esto genera regiones más pequeñas de modelado no satisfactorio, para las cuales se deben transmitir parámetros de color. En la Fig. 3 se muestran los valores PSNR obtenidos para diferentes "bit-rates" haciendo uso del codificador basado en bloques según el standard H.263 y los codificadores basados en objetos: R3D, F3D y A3D. Según este criterio el codificador según el standard H.263 es de 1 a 1.5 dB mejor que los codificadores basados en objetos. Esto se debe a que los codificadores basados en objetos no transmiten aquellos errores de reconstrucción que fueron ocasionados por pequeños errores de forma y de posición. Esto no reduce la calidad de imagen pero si reduce el valor PSNR. Para una calidad de imagen de PSNR=38 dB, el codificador basado en el modelo A3D alcanza un "bit-rate" aproximadamente de 50 kbit/s. Para la misma calidad de imagen el codificador basado en el modelo F3D y el codificador basado en el modelo R3D alcanza un "bit-rate" de 56 kbit/s y 64 kbit/s, respectivamente. El modelo A3D es más eficiente que el modelo R3D debido a que con él se logra una mejor descripción de la forma y del movimiento de los objetos reales, particularmente para cuando los componentes se mueven en forma diferente. El modelo A3D es más eficiente que el modelo F3D debido principalmente a que los vectores de deformación del modelo F3D requieren de un alto "bit-rate" para ser transmitidos.

5. CONCLUSIONES

Una comparación de los modelos tridimensionales encontrados en la literatura profesional muestra, que para la misma calidad de imagen el modelo de objetos articulados A3D es un 20% más eficiente que el modelo de objetos rígidos R3D y un 10% más eficiente que el modelo de objetos flexibles F3D.

AGRADECIMIENTO

El autor agradece al Prof. Musmann de la Universidad de Hannover, Alemania, por su invaluable consejo.

6. REFERENCIAS

- [1] ITU-T Recommendation H.263, "Video Coding for Narrow Telecommunication Channels at <64 kbit/s", 1995.
- [2] G. Martínez, "Analysis-Synthesis Coding Based on the Model of Articulated Three-Dimensional

- Objects", Picture Coding Symposium 1999 (PCS'99), Portland, Oregon, pp. 213-216, 1999.
- [3] G. Martínez, "Analyse-Synthese-Codierung basierend auf dem Modell bewegter, dreidimensionaler gegliederter Objekte", Ph.D Thesis, University of Hannover, Germany, 1998.
- [4] G. Martínez, "Shape Estimation of Articulated 3D Objects for Object-Based Analysis-Synthesis Coding (OBASC)", Signal Processing: Image Communications, Vol. 9, No. 3, pp. 175-199, 1997.
- [5] H. G. Musmann, M. Hötter, J. Ostermann, "Object-Oriented Analysis-Synthesis Coding of Moving Images", Signal Processing: Image Communication, Vol. 1, No. 2, pp. 117-138, 1989.
- [6] J. Ostermann, "Object-Based Analysis-Synthesis Coding Based on the Source Model of Moving Rigid 3D-Objects", Signal Processing: Image Communication, No. 6, pp. 143-161, 1994.
- [7] J. Ostermann, "Object-Based Analysis-Synthesis Coding (OBASC) Based on the Source Model of Moving Flexible 3D-Objects", IEEE Trans. on Image Processing, Vol. 3, No. 5, 1994.

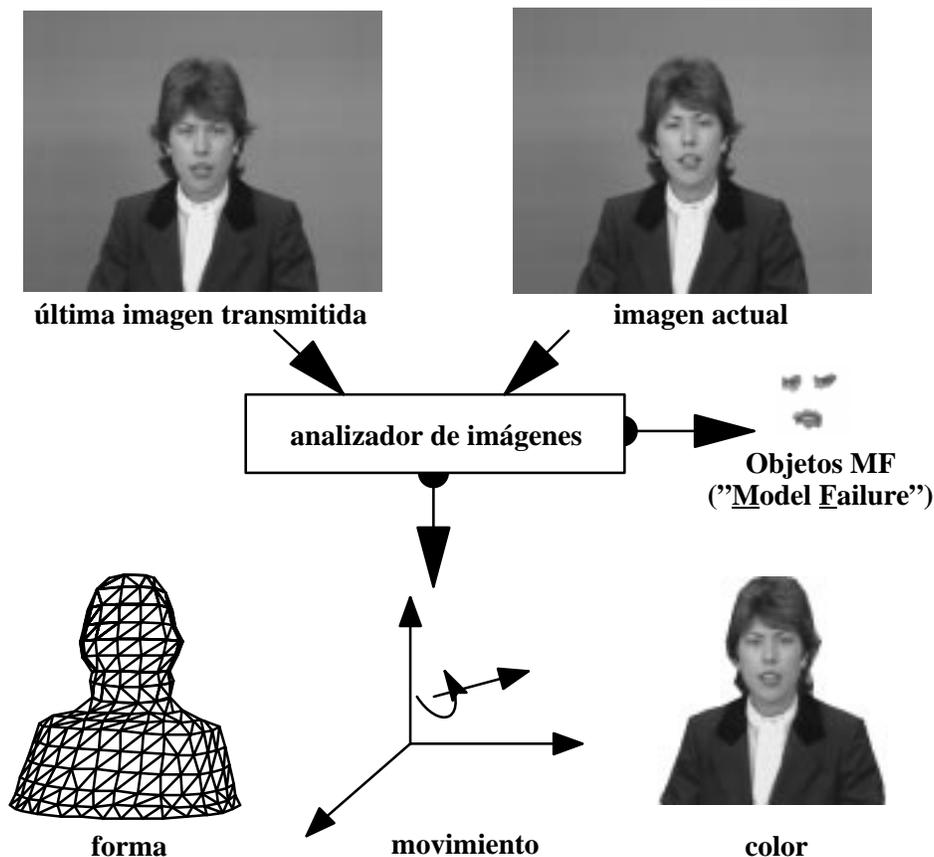


Fig. 1 Codificación de análisis-síntesis basada en objetos

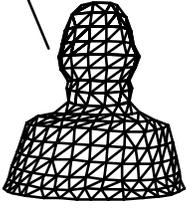
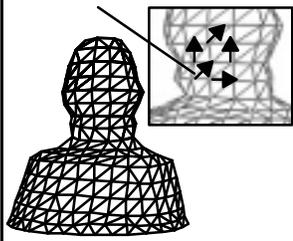
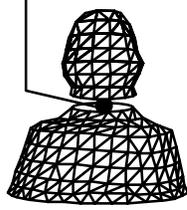
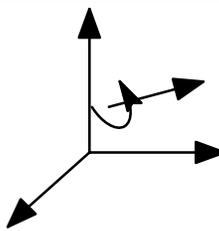
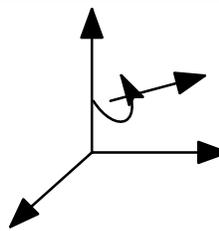
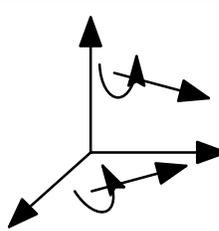
modelo parámetros	objetos rígidos tridimensionales (R3D)	objetos flexibles tridimensionales (F3D)	objetos articulados tridimensionales (A3D)
forma i	red de triángulos 	vectores de deformación 	articulación esférica 
movimiento ii			
color iii			
	a	b	c

Fig. 2 Modelos tridimensionales utilizados para la codificación de análisis-síntesis basada en objetos: a) modelo de objetos rígidos tridimensionales (R3D), b) modelo de objetos flexibles tridimensionales (F3D) y c) modelo de objetos articulados tridimensionales (A3D).

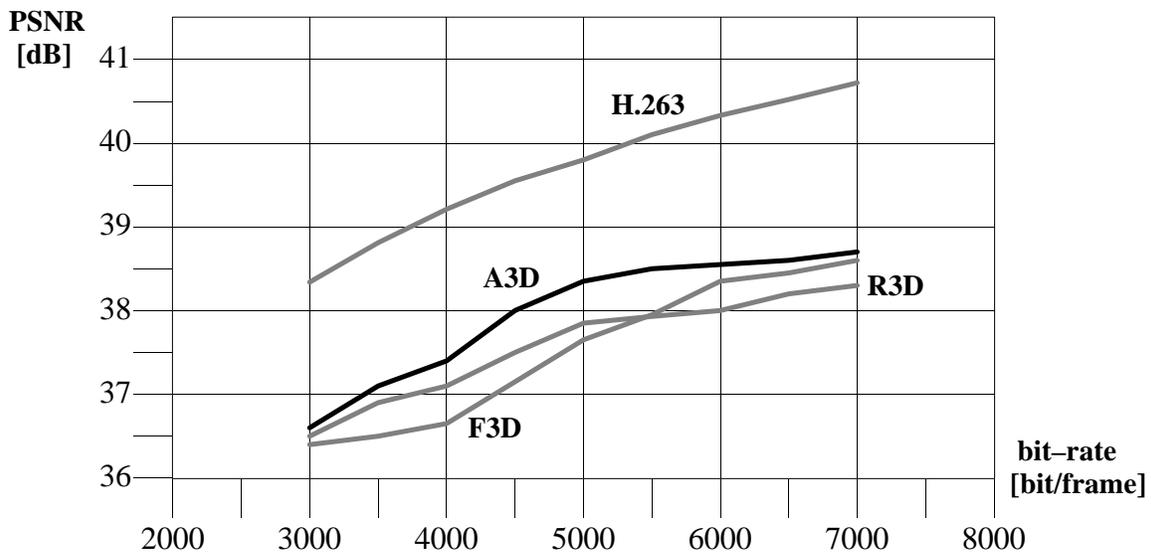


Fig. 3 PSNR obtenido para diferentes "bit-rates" con el codificador basado en bloques (H.263) y los codificadores basados en objetos rígidos tridimensionales (R3D), objetos flexibles tridimensionales (F3D) y objetos articulados tridimensionales (A3D).