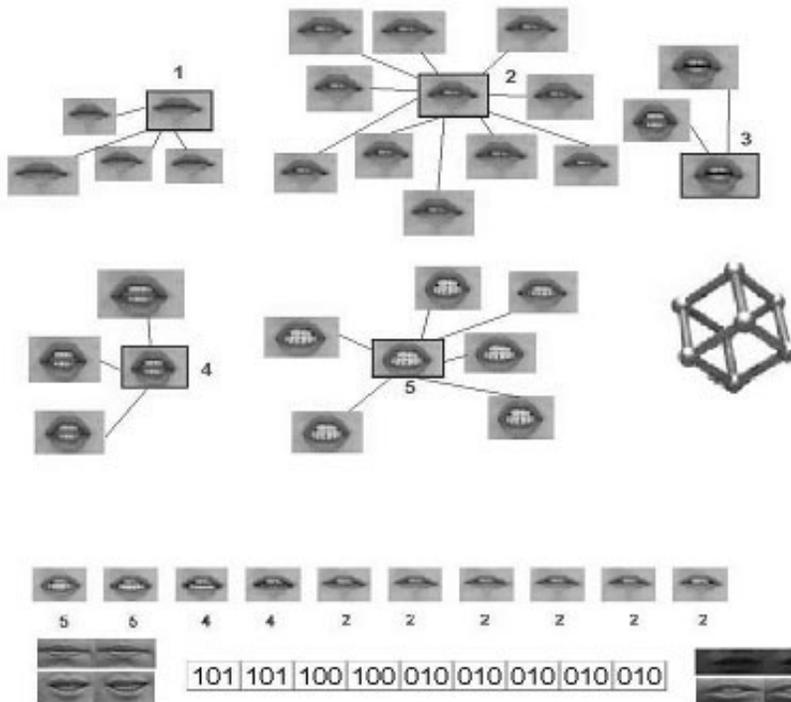


CODIFICACIÓN DE TELEFONÍA VISUAL

Empleando un conjunto de imágenes de imágenes



se utilizan métodos de cuantificación vectorial basados en el algoritmo de Lloyd-Max con el criterio de optimización min-max. Se presenta también un método de interpolación de imágenes intermedias empleando búsqueda de Viterbi para obtener secuencias óptimas.

PALABRAS CLAVES

Telefonía visual, Voz visual, Algoritmo de Lloyd-Max, Cuantificación Vectorial, Interpolación de Imágenes, cuantificador con criterio min-max.

ABSTRACT

This paper presents an encoding method for sequences of images taken from visual telephony scenes, which uses as codebook a set of natural images. The codebook is designed using vector quantization methods based on the Lloyd-Max algorithm with a min-max optimization criteria. A method for interpolation of images using the Viterbi's search algorithm for optimal performance is also presented.

Keywords

Visual Telephony, Image Interpolation, Vector Quantization, Lloyd-Max Algorithm, min-max quantization.

Pedro Vizcaya Ph.D.
 Ingeniero Omar Ayala
 Profesor Edson Bárcenas M.
 Profesor Rafael Carrillo
 Ingeniero Mauricio Díaz
 Ingeniero Ricardo Solano
 Profesora Carolina Soto M.Sc.
 Profesor Luis Valderrama
 Profesor Javier Villegas M.Sc.

Los autores trabajan con el Grupo de Bioingeniería, Análisis de Señales y Procesamiento de Imágenes de la Pontificia Universidad Javeriana de Bogotá (Colombia), Dirección: Cr 7 # 40-62. Correo electrónico para correspondencia: pvizcaya@javeriana.edu.co

RESUMEN

En este artículo se presenta un método de codificación de secuencias de imágenes de escenas de telefonía visual que emplea como libro de códigos un conjunto de imágenes naturales. En el diseño del libro de códigos

1. INTRODUCCIÓN

Los temas que se presentan en este artículo están enmarcados en un proyecto para el desarrollo de un Sistema de Transmisión de Telefonía Visual por un Canal de Muy Baja Capacidad, actualmente en realización por el grupo de Bioingeniería, Análisis de Señales y Procesamiento de Imágenes de la Facultad de Ingeniería, Pontificia Universidad Javeriana. En otros dos artículos se discuten temas relacionados con la transmisión y con la adquisición de las escenas de telefonía visual 7, 7. En este artículo se tratan específicamente los temas relacionados con la codificación y decodificación de secuencias de voz visual.

La codificación de imágenes normalmente se realiza en dos pasos: parametrización de las imágenes para reducir la dimensionalidad del espacio y selección de un subconjunto de imágenes, o libro de códigos, con algún criterio de optimización y en el dominio de los parámetros. Este libro de códigos se utiliza para cuantificar y codificar las imágenes nuevas. En nuestro caso la imágenes se parametrizan utilizando la transformada coseno discreta bidimensional de la región de interés, específicamente, la región de la boca. Las imágenes que formarán el libro de códigos se escogen dentro de un conjunto de imágenes de entrenamiento (corpus) tomadas de un video representativo de la escena de voz visual de alrededor de medio minuto (1000 imágenes a 30 cuadros por segundo).

El paradigma clásico para la selección del libro de códigos es hacer máxima la relación señal a distorsión (PSNR por sus siglas en inglés)7. El libro de códigos se escoge entre el conjunto de imágenes del corpus de manera que haga máxima su PSNR. Cada imagen del corpus es asignada a la imagen más cercana del libro de códigos, definiendo así clases. En este proceso se emplea tradicionalmente el algoritmo de Lloyd-Max7, el cual itera sobre dos condiciones necesarias para un cuantificador óptimo: (1) cada imagen del libro de códigos es el centro de momentos (centroide) del subconjunto de imágenes asignados a esa clase, y (2) cada imagen del corpus es asignada a la imagen del libro de códigos más cercana en el sentido de distancia euclidiana. Se ha demostrado que este algoritmo es localmente óptimo y, con algunas variaciones, converge

estadísticamente hacia el mínimo global 7. El algoritmo de Lloyd-Max asigna consistentemente más centroides a las regiones del espacio que tienen más muestras, pues esto significa que estas regiones tienen mayor probabilidad de ocurrencia, excluyendo imágenes atípicas (de baja probabilidad). En este trabajo se plantea un paradigma distinto para el diseño del cuantificador y es generar secuencias de imágenes que sean creíbles. Este concepto se explica en el numeral 2.

En el desarrollo de un sistema de transmisión de telefonía visual se puede utilizar un sistema de interpolación de imágenes para disminuir la capacidad de canal necesaria, enviando sólo una parte de la información e interpolando el resto, ó en la reconstrucción de escenas de telefonía visual en la que algunas imágenes no llegan al receptor debido a fallas en el canal o a limitaciones de la tasa de adquisición. Un sistema de interpolación de imágenes que genere trayectorias creíbles entre un par de imágenes a determinada velocidad también es de mucha utilidad en sistemas de síntesis de voz visual 7. Por ejemplo, se puede utilizar en un sistema de conversión de texto a voz visual 7, ó en un conversor de voz a voz visual. En el trabajo aquí presentado se propone un método en el cual se seleccionan las imágenes que mejor reconstruyen una secuencia dentro de una base de datos de imágenes limitada. Este método se presenta en el numeral 3.

2. SELECCIÓN DEL LIBRO DE CÓDIGOS

En este trabajo se plantea un paradigma distinto para la selección o diseño del libro de códigos (cuantificador): se desea que este incluya imágenes suficientes para generar secuencias creíbles, para lo cual es necesario que incluya también imágenes atípicas y no es deseable que incluya imágenes muy parecidas, como ocurre con el algoritmo de Lloyd-Max tradicional. En el contexto del presente trabajo en codificación para telefonía visual, imágenes creíbles son secuencias de imágenes con transiciones suaves entre ellas. En trabajos anteriores se ha establecido que la región de la boca es la más importante para establecer esta característica 7, 7. Para retener esta propiedad en el diseño del libro de códigos se plantea redefinir el criterio de optimización del cuantificador de la siguiente manera: hacer mínimo el error máximo entre las muestras

asignadas a cada representante y el representante correspondiente. El algoritmo de Lloyd-Max puede aprovecharse para resolver este problema redefiniendo el representante de cada subconjunto como el centro de la esfera de mínimo radio que incluye todas las muestras asignadas a ese representante (o circuncentro en 2-D) (condición de optimización (1)), en lugar del centroide. Este cambio permite que el algoritmo de Lloyd-Max converja hacia el conjunto de representantes que hagan mínimo la máxima distancia entre las muestras y su representante correspondiente. Sin embargo, el cálculo del circuncentro en un espacio n-dimensional no es trivial. Por esta dificultad aquí se plantea utilizar la distancia L_∞ en lugar de la L2 o Euclidiana, con la cual se puede calcular trivialmente la mínima caja rectangular que incluya todas las muestras y su centro.

Formalmente, la formulación de las dos condiciones anteriores es la siguiente:

Sea $X = \{ \mathbf{x}_n \}_{n=1, \dots, N}$ el conjunto de muestras de entrenamiento (corpus), en donde estas muestras son vectores de un espacio L-dimensional

$$\mathbf{x}_n = (x_n(1), \dots, x_n(l), \dots, x_n(L)), \mathbf{x}_n \in \mathfrak{R}^L,$$

$Y = \{ \mathbf{y}_k \}_{k=1, 2, \dots, K}$ el conjunto de códigos del libro, $Y_k = \{ \mathbf{x}_{k,i} \}_{i=1, 2, \dots, N_k}$ el subconjunto de N_k muestras del corpus asignadas al representante

\mathbf{y}_k , en donde $\mathbf{x}_{k,i} \in X$, $X = \bigcup_{k=1}^K Y_k$ y $Y_k \cap Y_l = \emptyset$

para todo $k \neq l$, es decir, el conjunto $\{ Y_k \}_{k=1, \dots, K}$ forman una partición de X .

Entonces el cuantificador debe cumplir las dos condiciones siguientes:

- (1) Cálculo del centro del cubo L-dimensional:

$$\mathbf{y}_k = (y_k(1), \dots, y_k(l), \dots, y_k(L))$$

$$\text{con } y_k(l) = \frac{1}{2} \left(\max_{i \in 1, \dots, N_k} \{ x_{k,i}(l) \} - \min_{i \in 1, \dots, N_k} \{ x_{k,i}(l) \} \right)$$

, y

- (2) Asignación de las muestras:

$$Y_k = \left\{ \mathbf{x}_n \in X \left| d_{L_\infty}(\mathbf{x}_n, \mathbf{y}_k) = \min_{i \in 1, \dots, K} \{ d_{L_\infty}(\mathbf{x}_n, \mathbf{y}_i) \} \right. \right\}$$

$$\text{con } d_{L_\infty}(\mathbf{x}, \mathbf{y}) = \max_{i=1, \dots, L} \{ |x(i) - y(i)| \}.$$

La coherencia entre la selección del

representante (centro del cubo L-dimensional) y la medida de distancia (L_∞) es esencial para asegurar que el algoritmo de Lloyd-Max efectivamente converja al límite esperado, es decir, hacer mínimo el máximo error (o min-max). La coherencia consiste en que la superficie del cubo L-dimensional es efectivamente el lugar de los puntos que equidistan del centro del mismo bajo la distancia L_∞ . Esta condición también las cumplen las parejas distancia-representante formadas por L2-centro de la esfera L-dimensional y L1(o Maniatan)-centro del rombo L-dimensional, con la diferencia de que el cálculo de los representantes en los últimos dos casos es más complejo. En el numeral 4 se muestran resultados de simulaciones y la aplicación de este paradigma a imágenes de voz visual.

3. INTERPOLACIÓN Y SELECCIÓN DE IMÁGENES

Una aproximación típica al problema de generar imágenes intermedias es la realización de una interpolación directa de cada uno de los píxeles, con el inconveniente de que se generan imágenes imposibles en las cuales, por ejemplo, se presentan dos bocas 7. En la Figura 1 se muestra un ejemplo de interpolación lineal entre dos imágenes, generando 4 imágenes intermedias.



Figura 1. Interpolación lineal de imágenes.

En algunos de los trabajos anteriores 7, 7 se utilizó la idea de hacer una interpolación sigmoideal para evitar la producción de este tipo de imágenes, con lo que la secuencia se observa más creíble (suave), debido a que las imágenes superpuestas tienen una duración más corta. Un ejemplo de interpolación sigmoideal es la tangente hiperbólica, la cual se ilustra en la Figura 2.

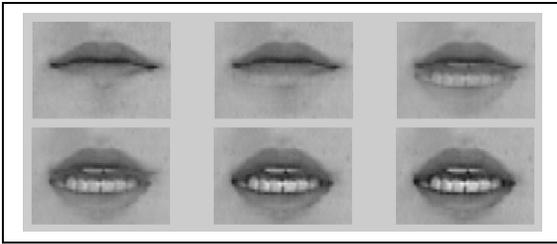


Figura 2. Interpolación sigmoïdal.

En otro trabajo anterior [7] se utilizaron técnicas de conformación (*morphing* en inglés) para realizar esta síntesis de imágenes intermedias, pero su costo computacional es más alto porque las imágenes deben procesarse a nivel de cada elemento (píxel).

En el trabajo aquí presentado, se propone un método en el cual se seleccionan las imágenes que mejor reconstruyen una secuencia entre las imágenes de una base de datos de imágenes [7]. La diferencia fundamental es, entonces, que se buscan las imágenes intermedias en un conjunto de imágenes y no se generan nuevas imágenes inexistentes para la generación del video. Por otro lado, para encontrar la combinación óptima de imágenes que reconstruya de forma creíble la secuencia se utiliza el algoritmo de Viterbi con la distancia modificada [7].

La utilización del algoritmo de Viterbi está relacionado con la optimización de la búsqueda dentro de una base de datos de la secuencia con mínima distancia acumulada. La modificación de la medida de distancia es una forma de asegurar la mayor credibilidad de las secuencias. A continuación se tratará este tema, antes de volver al tema de la selección de imágenes en la base de datos.

Uno de los objetivos de la interpolación de imágenes es generar una trayectoria creíble entre la pareja de imágenes. De la observación de secuencias naturales adquiridas para diferentes personas y de los videos generados que parecían poco creíbles, se pudo concluir que la credibilidad de una secuencia de video se puede asociar con la suavidad de la misma. Es decir, una secuencia de imágenes se considera más creíble si no se presentan saltos bruscos entre una imagen y la siguiente. La modificación que se propone a la definición de medida de distancia, y que se explica a continuación, tiene como objetivo que la secuencia final no presente saltos bruscos.

El cálculo de la medida de distancia se explicará mediante un ejemplo sencillo. En este caso la distancia de una secuencia se define como la suma de las distancias de los pasos. El ejemplo utiliza la Figura 3.

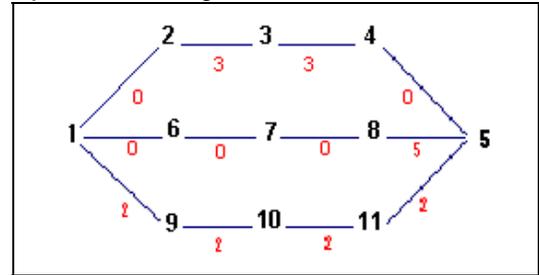


Figura 3. Ilustración de la medida de distancia.

La Figura 3 muestra tres caminos posibles con tres imágenes intermedias, entre la imagen 1 y la imagen 5, y se quiere escoger el camino de mayor credibilidad, es decir, la trayectoria en el que las transiciones sean más suaves. Los nodos representan las imágenes y los números bajo las ramas son la distancia original entre las imágenes respectivas. Si se escoge el camino de mínima distancia entre la imagen 1 y la 5 con la medida original, se escogería el camino medio que tiene un costo de 5 en comparación con las distancias del camino superior (6) y del inferior (8). Sin embargo, este es el camino que tiene el salto más brusco (5) entre la imagen 8 y la 5.

Una forma de castigar los caminos que tienen saltos bruscos es elevar las distancias a alguna potencia mayor que 1 [7]. Siguiendo con el ejemplo, si cada distancia se eleva a la potencia 1.5, la distancia acumulada por el camino superior (10.4) es menor a la del camino medio (11.2) y a la del inferior (11.3). Si, en cambio, se eleva la distancia a la potencia 2, el camino inferior resulta con la menor distancia (16) con relación al camino superior (18) y al medio (25). En consecuencia, ajustando la potencia se pueden seleccionar caminos más suaves entre la imagen inicial y la final para un número de imágenes intermedias dado.

En general, la modificación que se hizo a la definición de distancia entre una secuencia de imágenes es como sigue:

$$d_{1,N} = \sum_{k=1}^{N-1} (d_{k,k+1})^p$$

en donde $d_{i,j}$ es la distancia entre las imágenes i e j , N es el número de pasos y p es la potencia. Las imágenes se escogen entre un

conjunto de $M > N$ imágenes de la base de datos.

Las imágenes inicial y final y el número de pasos son la entrada para este proceso. El factor p se ajusta heurísticamente para obtener secuencias creíbles. Es importante notar que si se requieren muchas imágenes intermedias entre una imagen y otra, la potencia debe aumentar lo suficiente para garantizar que la secuencia de menor distancia sea una secuencia creíble (suave). Sobre este punto se ha observado que p debe ser proporcional a N .

El otro problema importante es seleccionar la secuencia de imágenes intermedias de mínima distancia modificada acumulada en la base de datos de M imágenes. En general existen del orden de $MN-1$ secuencias posibles para N pasos. Para determinar la secuencia de mínima distancia, se necesitaría la evaluación directa de cada distancia, lo cual sería un proceso altamente ineficiente. Para solucionar este inconveniente se utilizó el algoritmo de Viterbi⁷, el cual es la técnica más eficiente y óptima conocida de programación dinámica para encontrar el camino de costo mínimo desde un nodo inicial conocido hasta un nodo final también conocido. El algoritmo consiste en evaluar todos los caminos posibles para llegar a cada uno de los nodos pero conservar únicamente el camino de menor costo para llegar a cada uno de ellos en cada etapa. De esta forma no es necesario almacenar todas las secuencias sino solo las M secuencias sobrevivientes.

4.RESULTADOS

4.1 Selección del libro de Códigos.

La utilización del paradigma min-max para la selección del libro de códigos se ilustra primero con una simulación y después con imágenes. La simulación se realizó generando muestras bidimensionales de dos variables aleatorias gaussianas bivariadas con matriz de covarianza diagonal descritas como sigue:

$$f_i(\mathbf{x}|\boldsymbol{\mu}) = \mathcal{CN}(\mathbf{x} | \boldsymbol{\mu}_i, \mathbf{C}_i)$$

con

$$\boldsymbol{\mu}_1 = (-5, 2), \quad \mathbf{C}_1 = 1.6 * \mathbf{I}$$

$$\boldsymbol{\mu}_2 = (1, -3), \quad \mathbf{C}_2 = 5 * \mathbf{I}$$

De la distribución f_1 se tomaron 200 muestras y de la distribución f_2 se tomaron 1000 muestras. Se empleó el algoritmo de Lloyd-Max min-max

con la distancia L_∞ y con el centro del cuadrado como representante de cada clase, para encontrar 4 clases. También se entrenó con el algoritmo de Lloyd-Max original, con distancia L_2 y el centro de momentos o centroide como representante de cada clase, para su comparación. Los resultados se ilustran en la Figura 4 para el Lloyd-Max original y en Figura 5 para el Lloyd-Max modificado.

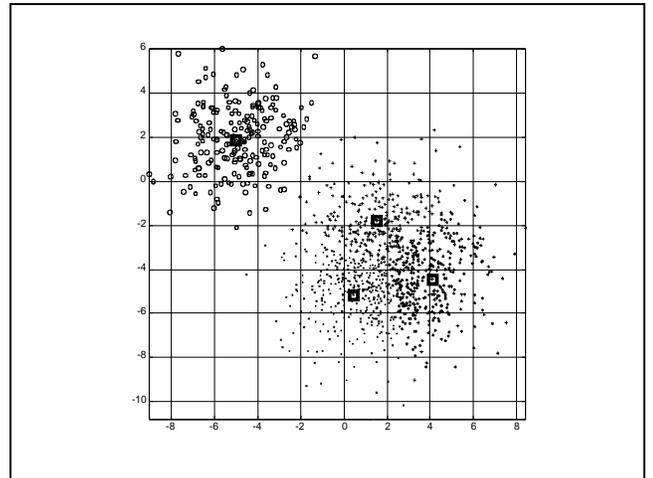


Figura 4. Algoritmo de Lloyd-Max original.

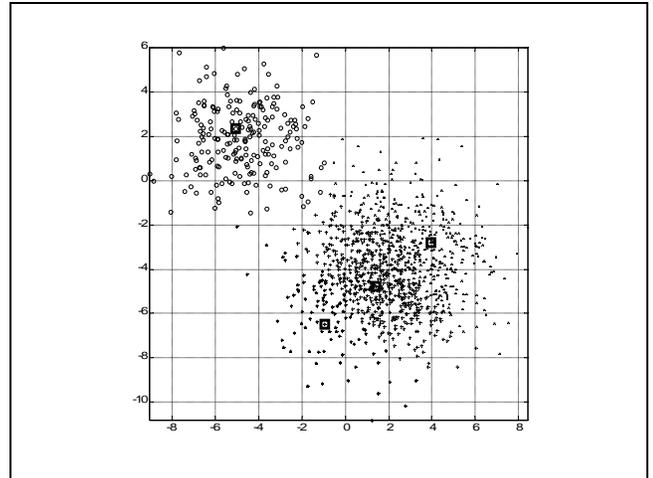


Figura 5. Algoritmo de Lloyd-Max min-max.

En estas figuras se aprecia como con el algoritmo modificado los representantes se escogen más separados entre sí, con mayor cubrimiento de las muestras con relación al algoritmo original. Los resultados numéricos se muestran en la Tabla 1 para la distancia L_∞ y en la Tabla 2 para la distancia L_2 .

La Tabla 1 ilustra claramente que la distancia máxima L_∞ es mayor para el algoritmo Lloyd-

Max original (5.6) y menor para el propuesto min-max (5.0). Consistentemente, la Tabla 2 muestra que la distancia euclidiana media es mayor para el algoritmo minimax con relación al algoritmo Lloyd-Max original.

Tabla 1. Distancia L^∞ .

	Lloyd-Max original	Lloyd-Max min-max
Clase 1	3.8809	4.6154
Clase 2	4.6037	4.3967
Clase 3	5.6450	5.0041
Clase 4	4.1722	4.0573

Tabla 2. Distancia L_2 (euclidiana)

	Lloyd-Max original	Lloyd-Max min-max
Clase 1	3.4098	4.0203
Clase 2	3.5945	4.6741
Clase 3	3.4853	5.2723
Clase 4	4.7095	5.3703

En la Figura 6 se comparan los dos métodos de obtención del libro de códigos aplicados a la selección de imágenes de la región de la boca, en la forma como se ha aplicado por el grupo para codificación de voz visual. En este caso el algoritmo se aplica sobre la parametrización generada por la transformada coseno discreta bidimensional calculada sobre 1000 imágenes de 64x128 originales, filtradas y decimadas por 4 en las dos dimensiones. Del total de 16x32 coeficientes, se seleccionó una región de 8x16, equivalente a un total de 128 parámetros. En este experimento se seleccionan 8 imágenes para el libro de códigos.

En la parte (a) de la Figura 6 se muestra claramente que varias de las imágenes seleccionadas con Lloyd-Max original son de la boca cerrada (imágenes 1 a 5), puesto que esta posición es la más frecuente, mientras que la selección con Lloyd-Max minimax incluye solo una de esta clase y una muestra variada de bocas abiertas con y sin dientes.

4.2 Interpolación y selección de imágenes.

Los resultados de estos procesos se evaluaron empleando una base de datos de 856 imágenes, las cuales fueron parametrizadas con la transformada coseno discreta bidimensional y, finalmente, se escogieron los 100 primeros coeficientes.

En un primer experimento se generaron secuencias de imágenes entre una imagen

inicial y otra final, empleando la técnica de búsqueda de Viterbi y potencia p igual a 2 en la definición de la distancia. La Figura 7 muestra el resultado de la interpolación entre el mismo par de imágenes utilizando el nuevo método.

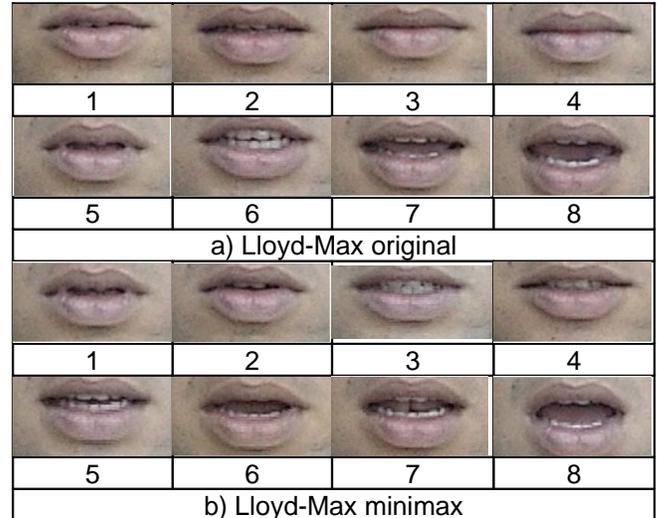


Figura 6. Libros de códigos de imágenes.



Figura 7. Interpolación de imágenes usando una base de datos de imágenes.

Estas imágenes deben compararse con las presentadas en la Figura 1 y en la Figura 2 en donde se presenta el problema de las imágenes superpuestas. Para evaluar el funcionamiento del sistema de interpolación se realizaron algunos experimentos consistentes en tomar videos naturales y simular la pérdida periódica de imágenes intermedias. Es decir, realizar un proceso de decimación consistente en:

$$y[n] = x[kn]$$

donde k es el factor de decimación, $x[n]$ es la imagen n -ésima de la secuencia de entrada y $y[n]$ es la imagen n -ésima de la secuencia de salida. Posteriormente se realizó un proceso de interpolación de las imágenes perdidas y finalmente se calculó la distancia entre la secuencia original y la secuencia resintetizada mediante interpolación.

Para la evaluación del error en la

reconstrucción se definió el índice de calidad de reconstrucción como:

$$iqor = 10 * \log\left(\frac{1}{N} \sum_{i=1}^N \frac{\text{var}(x_i)}{\text{var}(x_i - y_i)}\right) \text{ (dB)}$$

en donde, N es el número de imágenes de la secuencia de video, x_i es la i-ésima imagen de la secuencia original, y_i es la i-ésima imagen de la secuencia resintetizada. Aunque este indicador no está relacionado con credibilidad de la secuencia interpolada, sí permite relacionar los resultados del proceso con los métodos tradicionales de medida basados en distorsión y determinar la sensibilidad con relación a las variables libres.

Los resultados del experimento de decimación se muestran en la Figura 8.

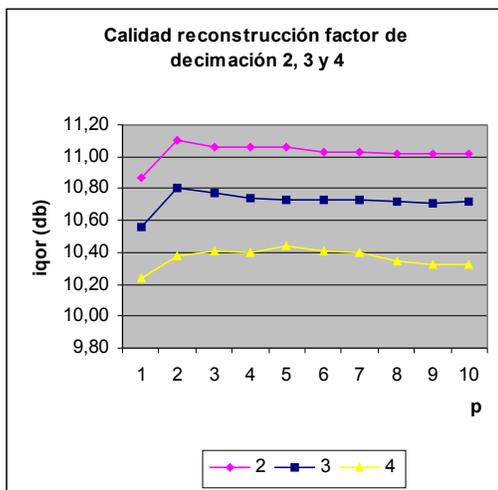


Figura 8. Experimento de decimación.

En los resultados se pudo observar que el índice de calidad en la reconstrucción es más alto para factores de decimación bajos. Sin embargo, con un factor de decimación de 4 ó 5 la calidad de la reconstrucción sigue siendo subjetivamente bastante aceptable.

La Figura 8 también muestra el efecto de la potencia para diferentes factores de decimación. Si las imágenes entre quienes se está interpolando no son muy lejanas, potencias del orden de 2 funcionan bien. Sin embargo, en el caso de factor de decimación igual a 5, la potencia que da mayor calidad en la reconstrucción es 5.

5.CONCLUSIONES

En este artículo se ha mostrado que la

modificación propuesta en el uso del algoritmo de Lloyd-Max es efectivamente útil para generar libros de códigos con el criterio de hacer mínimo el máximo error, en lugar de hacer mínimo el error cuadrado medio. Se ha mostrado también que el uso de la distancia L_∞ conjuntamente con el centro del cubo L-dimensional son una alternativa válida para emplear con el algoritmo y permite además un cálculo eficiente del representante y de la distancia entre muestras y representantes. Se mostró igualmente que el libro de códigos escogido con este método tiene menos imágenes similares y no se sesga por la frecuencia de las imágenes.

En el artículo también se ha mostrado un método efectivo para interpolación de imágenes empleando una base de datos de imágenes, en lugar de procesamiento, y aplicando el algoritmo de Viterbi sobre una distancia modificada para identificar secuencias creíbles. Se demostró que la modificación de la distancia entre imágenes permite encontrar secuencias de imágenes creíbles en el sentido de que la transición entre ellas es suave. Es necesario observar que si la potencia a la cual se eleva la distancia entre las imágenes p se deja ir hasta infinito, el método propuesto converge a la secuencia más suave (con pasos más homogéneos) posible. Para aplicaciones en tiempo real, por ejemplo, esta secuencia puede entonces precalcularse entre todas las imágenes del libro de códigos para distinto número imágenes intermedias.

6.RECONOCIMIENTOS

Este trabajo ha sido financiado en parte por el Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología "Francisco José de Caldas" –COLCIENCIAS– bajo el contrato RC. No. 35-2003 entre COLCIENCIAS y la Pontificia Universidad Javeriana.

7.REFERENCIAS BIBLIOGRÁFICAS

- [1] BAPTISTE C., SOTOMAYOR M. Segmentación y Parametrización Automática de Rostros en videos. Trabajo Final de Grado, Pontificia Universidad Javeriana, 2002
- [2] BÁRCENAS E., GALÁN J., SOTO C., URBINA J., VÁSQUEZ S., VIZCAYA P., (2001), Visual speech synthesis in Spanish using an optical flow algorithm. In: Proceedings of IASTED International

- Conference on Visualization, Imaging and Image Processing, 577-583, 2001.
- [3] BESKOW J., Talking heads – communication, articulation and animation, TMH-QPRS 2, Swedish Phonetics Conference, Nasslingen, 1996.
- [4] CHEN T y RAO R, Audio-visual integration in multimodal communication, Proceedings of IEEE, Special Issue on Multimedia Signal Processing, 837-852, 1998.
- [5] EZZAT E., GEIGER G., POGGIO T. Trainable Videorealistic Speech Animation, ACM Transactions on Graphics, 21(3): p. 388-398, 2002.
- [6] LINDE Y., BUZO A., GRAY R. An Algorithm for Vector Quantizer Design, IEEE T-COM, Vol .28, No. 1, pp. 84-95, January, 1980.
- [7] MACHADO J., VIZCAYA P., SANTA D. Visual Speech Synthesis Using A Real Time Parametric Approach, Memorias del VII Simposio de Tratamiento de Señales, Imágenes y Visión Artificial, pp. 104-109, Bucaramanga, Colombia, noviembre, 2002.
- [8] MUÑOZ M., SOTO C. y VIZCAYA P. Avances en síntesis de voz visual y sus aplicaciones. En: VIII Simposio de tratamiento de señales, imágenes y visión artificial, Medellín, Colombia, 2003.
- [9] NASRABADI N., KING R. Image Coding Using Vector Quantization: a Review, IEEE T-COM, vol 36, No. 8, pp 957-971, august, 1988.
- [10] PROAKIS J. "Digital Communications", McGraw-Hill, 4th edition, 2001.
- [11] SOTO, C. Generación de Corpus para Síntesis de Voz Visual, Trabajo de Investigación de Maestría, P. Universidad Javeriana, 2004.
- [12] VIZCAYA P., AYALA O., BÁRCENAS E., SOTO C., CARRILLO R., VALDERRAMA L., VILLEGAS J., SOLANO R. Segmentación de Características Faciales en Tiempo Real para Secuencias de Telefonía Visual. Memorias del IX Simposio de Tratamiento de Señales, Imágenes y Visión Artificial. Manizales (Colombia), Septiembre 2004.
- [13] VIZCAYA P., AYALA O., BÁRCENAS E., SOTO C., CARRILLO R., VALDERRAMA L., VILLEGAS J., SOLANO R. Transmisión de Secuencias Codificadas de Telefonía Visual Usando RTP. Memorias del IX Simposio de Tratamiento de Señales, Imágenes y Visión Artificial. Manizales (Colombia), Septiembre 2004.

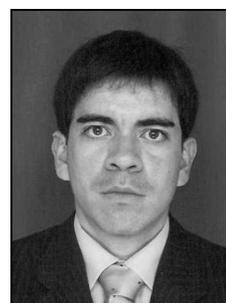
AUTORES



Pedro Raúl Vizcaya Guarín. obtuvo su grado como Ingeniero Electrónico en la Pontificia Universidad Javeriana, Bogotá, en 1980, el grado de MSc in EE del Instituto Politécnico Rensselaer, Troy, NY, EEUU, en 1986 y el grado de PhD en EE de la misma institución en 1998. Trabaja por Profesor Asociado en la Facultad de Ingeniería, Pontificia Universidad Javeriana, Bogotá, desde 1980. Actualmente es el Director de la Maestría en Ingeniería Electrónica de esta institución. Su dirección es Facultad de Ingeniería, Pontificia Universidad Javeriana, Carrera 7 No. 40-62, Bogotá, Colombia, (571) 3208320, ext. 5328.
pvizcaya@javeriana.edu.co



Omar Ricardo Ayala Peñuela. Ingeniero Electrónico (2004) de la Pontificia Universidad Javeriana, Bogotá (Col). Estudiante de la Maestría en Automatización y Robótica en la Universidad de Dortmund, Dortmund (Alemania).
oraporion@hotmail.com



Edson Nairo Bárcenas Ortiz. recibió su título de Ingeniero Electrónico de la Pontificia Universidad Javeriana en 2001 y su título como Master in Wireless Systems and Related Technologies del Politecnico di Torino (Italia) en 2002. Actualmente se encuentra trabajando como profesor de hora cátedra e investigador de tiempo parcial en el Área de Análisis de Señales y Procesamiento de Imágenes de la Pontificia Universidad Javeriana.
ebarcenas@javeriana.edu.co



Rafael Eduardo Carrillo Rangel. Recibió su grado como Ingeniero Electrónico en el 2003 y es candidato a Magíster en Ingeniería

Electrónica de la Pontificia Universidad Javeriana. Actualmente es auxiliar de investigación y profesor de cátedra en el Departamento de Ingeniería Electrónica de la misma universidad. Trabaja en investigación en las áreas de procesamiento de señales y reconocimiento de patrones.

rcarrillo@javeriana.edu.co



Mauricio Díaz Melo.

Ingeniero Electrónico (2004) de la Pontificia Universidad Javeriana, Bogotá, Colombia. Estudiante de la Maestría en Ingeniería Electrónica en la misma universidad. Actualmente se desempeña como

auxiliar de investigación en el proyecto "Telefonía Visual por canales de muy baja capacidad".

mauricio.diaz@javeriana.edu.co



Carolina Soto Quintero.

Recibió su grado como Ingeniera Electrónica y como Magistra en Ingeniería Electrónica de la Pontificia Universidad Javeriana en el 2001 y 2004 respectivamente. Actualmente es profesora asistente en el

Departamento de Ingeniería Electrónica de la misma universidad. Trabaja en investigación en las áreas de procesamiento de imágenes, señales y reconocimiento de patrones.

carolina.soto@javeriana.edu.co



Ricardo Solano Mulett.

Ingeniero Electrónico de la Universidad del Norte en el 2002 y estudiante de Maestría en Ingeniería Electrónica de la Pontificia Universidad Javeriana. Actualmente se desempeña como auxiliar de investigación del grupo de

Bioingeniería, Análisis de Señales y Procesamiento de Imágenes en la misma institución.

rsolano@javeriana.edu.co



Luis Rafael Valderrama.

Recibió su grado como Ingeniero Electrónico de la Pontificia Universidad Javeriana en el 2001. Actualmente es asistente de investigación en las áreas de procesamiento de imágenes, señales y reconocimiento de

patrones. Trabaja como profesor de cátedra de las asignaturas de Medios de transmisión de radio frecuencia II y circuitos lógicos en la misma universidad.



lvalderrama@javeriana.edu.co

Javier Villegas Plazas. Ingeniero electrónico de la Pontificia universidad Javeriana (2000). Magíster en ingeniería electrónica y de computadores de la universidad de los Andes (2002). Profesor de cátedra e investigador de la Pontificia Universidad Javeriana de Bogotá.

jvillegas@javeriana.edu.co