

# Simplificación robusta de nubes de puntos usando análisis de componentes principales y algoritmos genéticos

## Streamlining robust clouds point using principal component analysis and genetic algorithms

Nallig E. Leal. I.S.<sup>1</sup>, Esmeide A. Leal. I.S.<sup>1</sup> y John W. Branch.PhD<sup>2</sup>.

1. Facultad de Ingeniería de Sistemas, Universidad Simón Bolívar, Barranquilla, Colombia  
2. Universidad Nacional, Escuela de Sistemas, Medellín, Colombia  
Nallig\_eduardol@yahoo.co.mx; esmeide0226@yahoo.com; jwbranch@unal.edu.co

Recibido para revisión 22 de Septiembre de 2009, aceptado 23 de Octubre de 2009, versión final 14 de Diciembre de 2008

**Resumen**—Este artículo presenta un nuevo método de simplificación de nubes de puntos. El método propuesto, a diferencia de otros, no requiere la construcción previa de mallas poligonales y es robusto al ruido y a valores atípicos presentes en los datos. El método propuesto se compone principalmente de tres etapas. En la primera etapa, se segmenta la nube de puntos en regiones homogéneas, usando el algoritmo k-means. En la segunda etapa, se ajusta un plano de regresión de componentes principales robusto al ruido en cada cluster para determinar la tendencia local de los puntos. Finalmente, en la tercera etapa, usando un algoritmo genético se seleccionan los puntos de cada cluster cuyo plano de regresión de análisis de componentes principales minimice el ángulo con el plano de regresión del cluster. Resultados experimentales muestran que la distribución local y global de la nube de puntos original se mantiene.

**Palabras Clave**—Point cloud simplification, PCA, Genetic Algorithms.

**Abstract**—This paper presents a new method for robust point cloud simplification without previous reconstruction of polygonal meshes. The proposed method is robust to the noise and outliers. The proposed is roughly made up of three stages: The first one, segments the point cloud in homogeneous regions by using k-means clustering. The second one, fits a robust principal component analysis regression plane in each cluster, and the third one, by using a genetic algorithm, selects the points of each cluster whose principal component analysis regression plane minimize the angle with the regression plane of the cluster. It is shown that the local and global distribution of the original point cloud is kept.

**Keywords**—Simplificación de nubes de puntos, PCA, Algoritmos Genéticos.

### I. INTRODUCCIÓN

La reconstrucción 3D es un área de investigación activa por su amplia e importante variedad de aplicaciones en visión robótica, restauración de obras de arte, arqueología, supervisión y guiado de vehículos, medicina, ingeniería inversa y computación gráfica, entre otras. Reconstruir un objeto tridimensional, se refiere a obtener su representación en la memoria de un computador, manteniendo sus características de volumen y forma, partiendo de un conjunto discreto de datos muestreados de la superficie del objeto. La reconstrucción 3D es una tarea no trivial que, en general, involucra cinco etapas conocidas como: adquisición, registro, integración, segmentación y ajuste.

La etapa de adquisición de datos del proceso de reconstrucción 3D permite obtener un conjunto de puntos (nube de puntos) que describe la geometría de la superficie de un objeto real. A partir de dicho conjunto de puntos, se obtiene el modelo matemático del objeto. Gracias a los avances en la tecnología de adquisición de datos, particularmente en los scanner de rango láser, se pueden realizar muestreos de alta densidad sobre la superficie de objetos 3D, muestreos que pueden entregar miles e incluso millones de puntos para representar la superficie del objeto. La reducción de la cantidad de las muestras tomadas es la clave de las técnicas de procesamiento para el diseño de algoritmos escalables de visualización y modelamiento [7].

La simplificación de puntos, o simplificación de superficies, provee un medio para generar aproximaciones requeridas de

una superficie dada, que use pocos puntos muestrados a partir de la superficie real. Estas aproximaciones deberían seguir la superficie o el conjunto de puntos original tan cerca como sea posible [7]. Diferentes enfoques para realizar la simplificación de una nube de puntos han sido presentados [3], [4], [6], [7], [9], [10] los cuales, en términos generales, presentan inconvenientes como la alteración de la distribución local y global del conjunto de puntos originales, dependencia de parámetros de usuario y necesidad de realización de pasos intermedios como la construcción de mallas poligonales.

En este artículo un nuevo método para simplificación de nubes de puntos es presentado. A diferencia de otros métodos, nuestro método no requiere pasos intermedios de construcción de mallas poligonales, es robusto al ruido y a valores atípicos del conjunto de datos, y mantiene tanto la distribución local como global de los datos. Nuestro método primero determina regiones homogéneas (clusters) en las cuales la simplificación de puntos no elimine detalles finos o datos representativos de la nube de puntos; luego, ajusta a cada cluster un plano de regresión de componentes principales robusto al ruido y a los valores atípicos, para determinar las tendencias locales del conjunto de datos. Finalmente, utilizando un algoritmo genético, se seleccionan a partir de cada cluster conjuntos reducidos de puntos cuyo plano de regresión de análisis de componentes principales minimice el ángulo con el plano de regresión del cluster.

Lo que queda de este artículo está organizado como sigue. En la sección 2, se presentan trabajos relacionados que tratan el tema de la simplificación de nubes de puntos. En la sección 3, se presentan de manera detallada cada una de las etapas de nuestro método. En la sección 4, se presentan los resultados experimentales obtenidos. En la sección 5 se presentan las conclusiones y el trabajo futuro.

## II. TRABAJOS RELACIONADOS

En 2001 Linsen [6] presentó un método de reducción de puntos basado en entropía. Linsen asocia una medida del contenido de la información de cada punto de entrada y subsecuentemente remueve puntos caracterizados por baja entropía. El algoritmo produce resultados visualmente hablando, pero no da garantía en la densidad de los puntos de salida. Puntos distribuidos no uniformemente resultarán en puntos reducidos con distribución diferente a la del conjunto original. Puntos distribuidos uniformemente, pueden resultar en puntos reducidos con distribución no uniforme.

Dey et al. en 2001 [3] presentaron un algoritmo con garantía de densidad controlada por el usuario. El algoritmo, con la asistencia de "Cocone" y el concepto del tamaño característico local [1], detecta redundancia en los puntos de entrada. Este algoritmo es sensible a cambios locales de curvatura y resulta

costoso por el cálculo de diagramas de Voronoi que debe realizar para efectuar la reducción.

En el mismo año, Boissonnat y Cazals [2] presentaron un algoritmo que toma un subconjunto inicial de puntos, a partir de los puntos originales, y usa su triangulación 3D de Delaunay para definir una función de distancia signada sobre el conjunto. La función de distancia es usada para ampliar el conjunto inicial hasta que un número significativo de puntos caiga dentro de una tolerancia de error de aproximación definida por el usuario. Posteriormente, el conjunto ampliado es triangulado nuevamente para construir una superficie de malla; si dicha superficie no alcanza la condición de error, puntos adicionales son insertados de forma iterativa, ordenados por su distancia a la superficie más cercana. El método emplea reconstrucción y simplificación de mallas simultáneamente, resultando costoso por la construcción y el mantenimiento de las mismas.

En 2002 Pauly et. al. [9] presentaron varios algoritmos de reducción de puntos. El método de reducción iterativa, basado en el error cuadrático, produce conjuntos reducidos de puntos con bajo error de aproximación, pero en tiempo de ejecución es muy sensible al tamaño del conjunto de puntos original. La aproximación basada en la técnica de simulación de partículas también resulta en conjuntos de puntos de bajo promedio de error, pero en general, su cálculo es relativamente ineficiente. El método de agrupamiento incremental uniforme es computacionalmente eficiente pero no mantiene los detalles finos del conjunto original. La aproximación basada en agrupamiento jerárquico es eficiente computacionalmente en términos de velocidad y memoria, pero presenta errores de aproximación relativamente altos.

En 2003 Moenning y Dodgson [7] presentaron una técnica de simplificación que no requiere de la construcción previa de una superficie. Usando el muestreo rápido de puntos marchantes más lejanos para superficies implícitas y nubes de puntos, idearon un algoritmo de simplificación sensible a detalles finos con garantía de densidad controlada por el usuario, el cual es computacionalmente eficiente en términos de uso de memoria y de fácil implementación. La debilidad de la propuesta de Moenning y Dodgson es la dependencia de parámetros externos proporcionados por el usuario, además de ser sensible al ruido y a los valores atípicos presentes en los datos.

En 2004 Ferrari et. al. [4] propusieron un procedimiento de aprendizaje que puede ser visto como método de optimización iterativa para encontrar un conjunto reducido de puntos (llamado Vector de Referencia), el cual minimiza una medida adecuada de error para representar una nube de puntos. Mediante el particionamiento del dominio de los datos en regiones disjuntas llamadas hiperbóxeles, se logró un tiempo de procesamiento lineal. El procedimiento propuesto es adecuado para implementación en hardware paralelo. Este

procedimiento es dependiente de parámetros, y se deben analizar los datos para obtener los valores de parámetros adicionales requeridos por el procedimiento.

En 2005 Rovira et. al. [10] presentaron un nuevo método de muestreo para reducción de puntos basado en líneas globales de distribución. El método utiliza la técnica denominada discrepancia estrella, como medida de la calidad del muestreo. El conjunto reducido de puntos obtenido con este método mantiene la distribución del conjunto de puntos original.

En términos generales, los métodos de reducción de puntos antes mencionados, presentan inconvenientes como la alteración de la distribución local y global del conjunto de puntos originales, dependencia de parámetros de usuario, sensibilidad al ruido y a valores atípicos, presentes con frecuencia en los datos, además de pasos previos de preprocessamiento como la construcción de mallas poligonales. Podrían considerarse enfoques heurísticos, dada su aplicabilidad en la solución de problemas no determinísticos, que sean independientes de la intervención del usuario, robustos al ruido y a los valores atípicos, al tiempo que mantienen la distribución de los datos.

### III. MÉTODO PROUESTO

Nuestro método de simplificación de nubes de puntos está compuesto de tres etapas con metas intermedias específicas, para realizar la simplificación manteniendo las características de los datos originales, las cuales se presentan a continuación.

#### A. Agrupamiento o Clusterización

El agrupamiento tiene como objetivo determinar regiones de puntos homogéneas; regiones en las cuales, se puede realizar el proceso de reducción sin pérdida de información significativa, gracias a las condiciones de homogeneidad del grupo. El agrupamiento es realizado por medio del algoritmo k-means; aunque, esta etapa es abierta a otras técnicas de agrupamiento. La aplicación del k-means encuentra grupos o regiones con características similares, como lo ilustra la Figura 1.

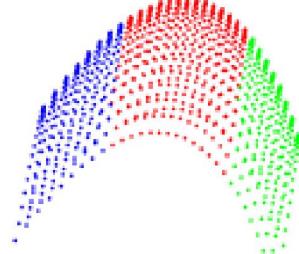


Figura 1. Regiones homogéneas encontradas por el k-means

Los puntos de la Figura 1 fueron muestreados a partir de la función  $Z = X^2 + Y^2$ . Se puede apreciar la homogeneidad de las regiones encontradas por el k-means, de ahí que la

reducción de puntos realizada en cada uno de ellos no de lugar a pérdida de información significativa de las mismas.

#### B. Estimación de la tendencia de los puntos en cada grupo

Una vez establecidos los grupos, se determina su tendencia (tendencias locales de los puntos sobre la superficie) empleando una nueva variante ponderada del análisis de componentes principales, la cual asigna factores de peso a los datos mediante repartición de pesos inversamente proporcional a la distancia a la media del centro del grupo. Esta nueva variante ponderada no requiere parámetros externos asignados por el usuario (como es el caso de los factores de peso exponenciales que requieren de la especificación del factor de suavidad  $\sigma$  por parte del usuario), ya que los pesos son calculados a partir de los datos mismos. La propuesta de cálculo de pesos consiste en realizar una repartición de la suma total de las distancias de los puntos a la media, inversamente proporcional a la distancia a que los puntos se encuentran de ésta; es decir, los puntos más lejanos tendrán menos influencia sobre el plano de regresión que los puntos más cercanos. Los factores de peso se calculan según ecuación 1.

$$W_i = \frac{1}{d_i \cdot \sum_{j=1}^n \frac{1}{d_j}} \quad (1)$$

donde  $d_i$  es la distancia del  $i$ -ésimo punto a la media y  $n$  es el número de puntos que participan en el cálculo del plano de regresión. El empleo del análisis de componentes principales ponderado (WPCA), hace a nuestro método robusto al ruido y a los valores atípicos. Se puede apreciar la corrección del bias (Figura 2b) generado por los outliers o valores atípicos (puntos encerrados en círculos) presentes en los datos (Figura 2a). El plano de ajuste dado por el análisis de componentes principales será la referencia de la tendencia local de los datos que se tendrá para el proceso de selección de puntos. Para más detalles del WPCA propuesto diríjase a [5].

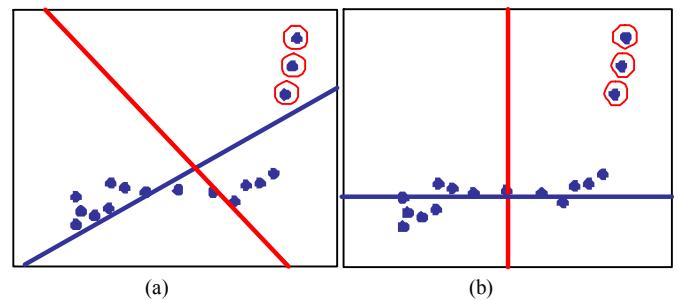
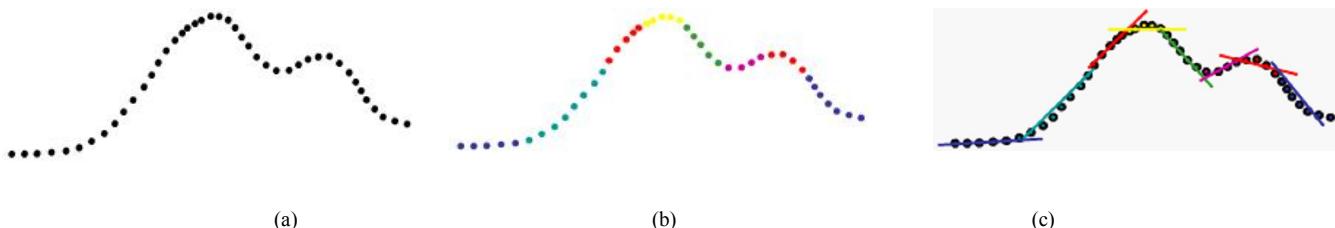


Figura 2. Análisis de componentes principales. a) Componentes principales de los datos influenciados por los valores atípicos. b) Corrección de la tendencia generada por los valores atípicos, luego de aplicar a los puntos factores de pesos calculados con la ecuación 1.

La Figura 3c muestra de tendencia local de los datos dada por los planos WPCA calculados con la variante propuesta. Se aprecia la homogeneidad de los grupos de datos. Gracias a la

homogeneidad de los puntos en cada grupo, la reducción realizada en cada uno de ellos no ocasiona o reduce la pérdida de información representativa.



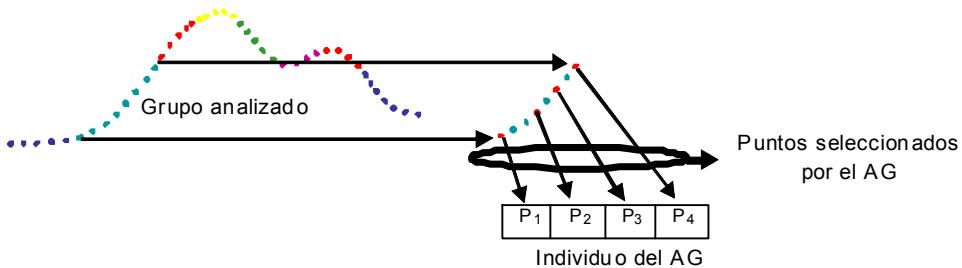
**Figura 3.** Análisis de tendencia de los datos. a) Nube de puntos. b) Puntos agrupados. c) Planos de ajuste que describen la tendencia local de los puntos

### C. Selección de puntos con la misma tendencia del grupo

La selección de puntos con la misma tendencia del grupo se lleva a cabo por medio de un algoritmo genético. El algoritmo genético selecciona puntos cuyo WPCA forme un ángulo menor que  $\epsilon$  (donde  $\epsilon$  es un nivel de tolerancia previamente especificado) con el plano WPCA de ese grupo.

Si el ángulo entre el plano WPCA del grupo original y el de los puntos seleccionados es menor que  $\epsilon$ , la tendencia de los puntos seleccionados debe ser muy similar a la tendencia de los puntos del grupo original. El Algoritmo Genético (AG) que realiza la selección se configura de la siguiente manera:

**Individuo:** puntos seleccionados del grupo original. La Figura 4 ilustra el individuo de la estrategia

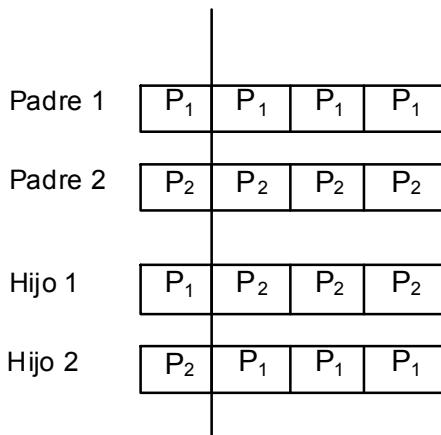


**Figura 4.** Configuración de los individuos del AG

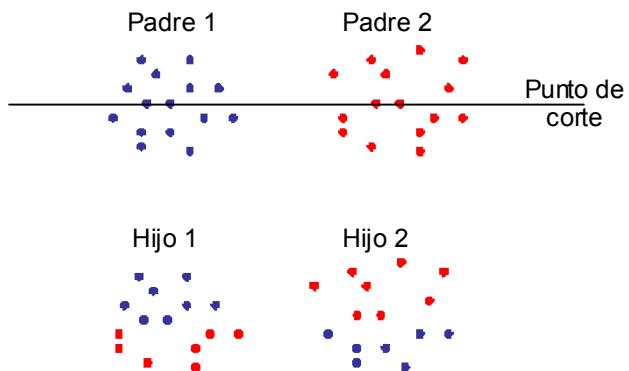
**Operador de recombinación:** Los individuos del AG representan un conjunto de puntos seleccionados a partir de los grupos encontrados, como lo ilustra la Figura 4. Dado que la ubicación de los puntos en el individuo es aleatoria, la recombinación se puede llevar a cabo de dos maneras: 1) Seleccionar aleatoriamente puntos de cruce y emplear 1-point crossover, como lo ilustra la Figura 5. 2) Determinar un punto

de corte en el plano de ajuste y transferir genes de un individuo a otro según si estos se encuentran por encima o por debajo del punto del corte sobre el plano de ajuste WPCA, como lo muestra la Figura 6.

**Operador de mutación:** la mutación de un individuo se refiere al cambio de uno de sus puntos, por otro no incluido en el, que pertenezca al grupo. Como lo ilustra la Figura 7.



**Figura 5.** Recombinación caso 1



**Figura 6.** Recombinación caso 2

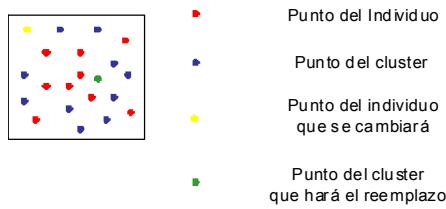


Figura 7. Operador de mutación

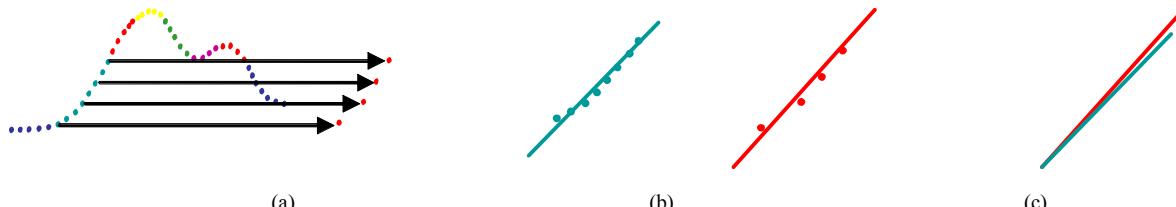


Figura 8. Función de aptitud. a) Puntos seleccionados por el algoritmo genético a partir del cluster analizado. b) Planos de ajuste WPCA a los puntos del grupo y de los puntos seleccionados por el AG respectivamente. c) Ángulo entre los planos de ajuste que minimizará el AG

#### IV. RESULTADOS EXPERIMENTALES

En esta sección se presentan los resultados experimentales obtenidos con el método propuesto. Las pruebas se realizaron en un PC con procesador Intel Pentium M de 1.6GHZ y 512MB de memoria RAM. Los aplicativos de prueba se desarrollaron en Microsoft Visual C++ 6.0 y el despliegue de gráficos se realizó en MATLAB 6.5. Las imágenes de rango empleadas se obtuvieron del repositorio de la Universidad de Ohio y de la página Web de Alexander Belyaev del instituto Max Plank ([www.mpi-sb.mpg.de/~belyaev](http://www.mpi-sb.mpg.de/~belyaev)).

##### A. Diseño experimental

En la simplificación de puntos interesa probar que la distribución local y global del conjunto de puntos reducido sea igual que la del conjunto original. Puesto que en la prueba no intervienen factores o bloques, se realiza un diseño completamente aleatorizado que involucra las dos siguientes pruebas de hipótesis.

##### Prueba 1

$$H_0 : d_{i,o} \neq d_{i,r}$$

$$H_a : d_{i,o} = d_{i,r}$$

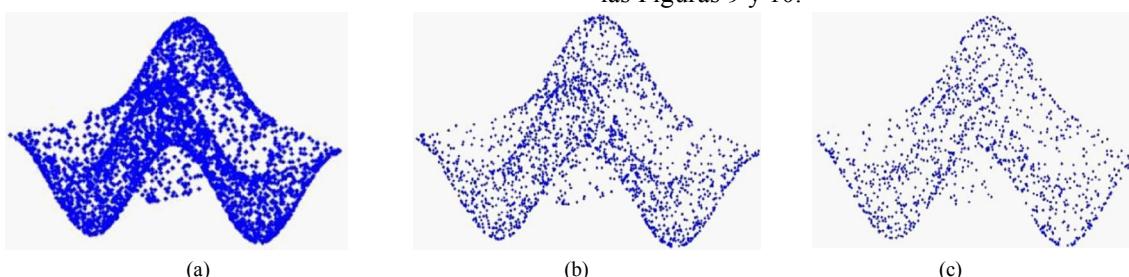


Figura 9. Simplificación de una nube de puntos sintética, función seno(x)seno(y). a) Nube original, 3600 puntos. b) Nube reducida al 64% - 2300 puntos. c) Nube reducida al 40% - 1437 puntos

**Función de aptitud:** ángulo entre el plano WPCA del grupo y el plano WPCA de los puntos seleccionados de éste. El algoritmo genético minimiza el ángulo entre los dos planos, dando como resultado el subconjunto de puntos del grupo que mejor aproxime la distribución de la totalidad de puntos de este. La Figura 8 ilustra el proceso.

donde  $d_{i,o}$  es la distribución local del  $i$ -ésimo grupo de la nube de puntos original agrupado por el método propuesto y  $d_r$  la distribución local del grupo correspondiente en la nube reducida

##### Prueba 2

$$H_0 : D_o \neq D_r$$

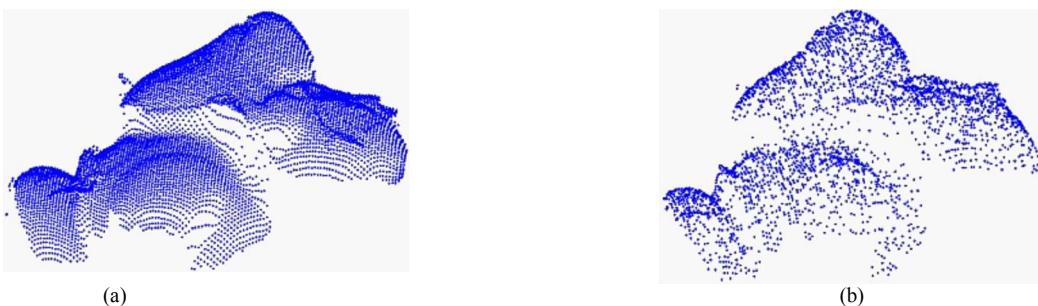
$$H_a : D_o = D_r$$

donde  $D_o$  es la distribución global de la nube de puntos original y  $D_r$  la distribución global de la nube reducida.

##### B. Resultados experimentales

La configuración del algoritmo genético planteado para la reducción de puntos es como sigue: Población de 20 individuos, los cuales intervienen en su totalidad en el proceso de recombinación y se hacen mutar con una probabilidad de

mutación de  $\frac{1}{n}$ , donde  $n$  es el tamaño del individuo. El número de generaciones por cada prueba realizada es de 100. El método de reducción de puntos propuesto en esta investigación se aplicó tanto en imágenes sintéticas como en imágenes de rango, obteniéndose los resultados mostrados en las Figuras 9 y 10.



**Figura 10.** Simplificación de una nube de puntos obtenida a partir de un objeto real. a) Nube original, 9114 puntos. b) Nube reducida al 40% - 3644 puntos

El ángulo medio entre el plano WPCA de los grupos de la nube de puntos original, y el plano WPCA de los grupos correspondientes de la nube reducida fue de 0.0001 grados. La realización de la prueba 1 planteada en la sección 4.1 arrojó un valor  $P = 0.00001$ , lo cual quiere decir que se puede rechazar la hipótesis nula; es decir, la distribución local de los puntos de la superficie de la nube de puntos original, es la misma que la de la nube de puntos reducida. El ángulo medio entre el plano WPCA de la nube original y el plano WPCA de la nube reducida fue de 0.001 grados. La prueba 2 del apartado 4.1 arrojó un valor  $P = 0.0001$ , por lo que se puede rechazar la hipótesis nula; es decir, la distribución global de los puntos de la nube original es la misma que la de los puntos de la nube reducida. El proceso de simplificación de puntos al aplicarlo sobre 30 imágenes tardó en promedio 190 segundos por cada 14K puntos procesados, divididos en 20 grupos reducidos al 40%.

Se aprecia que se mantiene la continuidad de los bordes de la superficie; es decir, el método de reducción, en términos generales, mantiene suficientes puntos sobre los bordes de la superficie, por lo que se mantiene su topología. La baja eliminación de puntos en los bordes de la nube, obedece a la influencia que éstos tienen en la orientación del plano de regresión PCA, de ahí que los individuos del algoritmo genético tiendan a incorporar dichos puntos, manteniéndose así la continuidad de los bordes de la nube.

## V. CONCLUSIONES Y TRABAJO FUTURO

Un nuevo método para simplificación de nubes de puntos fue presentado en este artículo. El método mantiene las características de distribución local y global del conjunto original. El método, al utilizar una variante ponderada del análisis de componentes principales, resulta robusto al ruido y a valores atípicos presentes en los datos. El cálculo de los factores de peso para estimar los planos WPCA no depende de parámetros externos proporcionados por el usuario; ya que, el cálculo se realiza a partir de los datos mismos. Adicionalmente, a diferencia de otros métodos relacionados en la literatura, el cálculo de los factores de peso realizado

por el método propuesto no requiere de funciones exponenciales y solo usa operaciones aritméticas simples, lo cual lo hace computacionalmente eficiente. El método propuesto preserva los detalles finos de la superficie original; esto se debe principalmente, a la reducción local realizada en los clusters y al empleo del WPCA.

Como trabajo futuro se podría realizar una investigación orientada a determinar en forma automática el número de grupos en que se debe dividir la nube de puntos. En la actualidad, este parámetro es fijo ingresado por el usuario, lo cual afecta el desempeño del método.

## REFERENCIAS

- [1] Amenta, N., Choi, S., Dey, T. and Leekha, N. A simple algorithm for homeomorphic surface reconstruction. Proceedings of the 16th Annual ACM Symposium on Computational Geometry. Hong Kong. 2000. pp. 213-222.
- [2] Boissonnat, J. and Cazals, F. Coarse-to-fine surface simplification with geometric guarantees. Proceedings of EUROGRAPHICS 2001. Manchester, UK. 2001. pp. 490-499
- [3] Dey, T., Giesen, J. and Hudson, J. Decimating Samples for Mesh Simplification. Proceedings of the 13th Canadian Conference on Computational Geometry. Waterloo, Canada, 2001, pp. 85-88.
- [4] Ferrari, S., Frosio, I., Piuri, V., and Borghese, N. Enhanced Vector Quantization for data reduction and filtering. Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT). 2004.
- [5] Leal, E., Leal, N. Point cloud denoising using robust principal component analysis. International Conference on Computer Graphics Theory and Applications. Setúbal, Portugal. Febrero 22 – 25 de 2006.
- [6] Linsen, L. Point cloud representation. CS Technical Report 2001-3, Universitat Karlsruhe, Germany, 2001.
- [7] Moenning, C. and Dodgson, N. Fast Marching farthest point sampling for implicit surfaces and point clouds. Technical Report No. 565. Computer Laboratory University of Cambridge, UK, 2003.
- [8] Pauly, M. Point Primitives for Interactive Modeling and Processing of 3D Geometry. Doctoral Thesis. Federal Institute of Technology (ETH) of Zurich. 2003.
- [9] Pauly, M., Gross, M. and Kobbelt, L. Efficient Simplification of Point-Sampled Surfaces. Proceedings of the 13th IEEE Visualization Conference. Boston, USA, 2002.
- [10] Rovira, J., Wonka, P., Castro, F., and Sbert, M. Point Sampling with Uniformly Distributed Lines. Eurographics Symposium on Point-Based Graphics. 2005.