Valle-Chavez, A.
Comparasion of New Techniques of Representation for 3D Shape Data Base Retrieval
International Journal of Combinatorial Optimization Problems and Informatics
Morelos, México

Available in: http://www.redalyc.org/articulo.oa?id=265224452008
Comparaison of New Techniques of Representation for 3D Shape Data Base Retrieval

A. Valle-Chavez
Computing Research Center
Instituto Politécnico Nacional
vchaveza10@sagitario.cic.ipn.mx

Abstract. The recent development of retrieval techniques and advances in 3D shape representation has induced a large amount of research. This paper presents a methodology with three options to retrieve shapes in 3D data bases, based on mapping shapes in 1D using representative sequences, which are matched using Fast Dynamic Time Warping distance. The three options are compared through a performance measure showing being better than the ones in literature.

Keywords: 3D Shapes, 3D Shape Matching, Similarity, PCA, Representation of 3D Shapes.

1 Introduction

The recent development of capturing techniques and 3D shapes storage has caused that the amount of these has grown significantly, which adds to the large number of 3D figures available on the Internet. On the other hand, indexing identification or words are options, just as search engines do with images, however they do not consider content extraction. Alternatively there are techniques which analyze more specific properties of the shape.

Some reported techniques the use of features extraction organized on the following groups: a) Histograms, b) 2D view, c) Graphs and d) Domain Changes. Group (a) consists on accumulators of global or local characteristics with the problem of not discriminating properties of objects in an efficient way [4, 5, 6]. Group (b), obtains data collection in a single shape, these techniques are highly discriminative [10, 11]. Techniques based on graphs have the property of codifying the shape topology, but these are difficult to obtain and require graphs matching [12, 13]. Group (d) uses the transformation of a representative space to other, having as advantage that are compact representations and efficient process offering the possibility of exactitude while searching similarity [7, 8, 9].

This paper presents a methodology with three options for the search of visual similarity among 3D figures, based on mapping figures to a 1D using representative sequences, which are matched through distance of Fast Dynamic Time Warping (FDTW).

A measure of results development of different methods of similarity search among 3D figures is also proposed. This performance measure is applied in the three methods exposed in this paper, obtaining a comparative of performance for the results obtained from 3D shape queries.

2 Definition of 3D Shape

A 3D shape representation is provided under the definition of polygon mesh, which is very popular for tridimensional models due to its simplicity. So, a tridimensional model is defined by a pair of ordered lists:

\[ M = \{ P, V \} \]

Where \( V = \{ v_1, v_2, \ldots, v_n \} \) is the list of vertex, and \( v_i = ( x_i, y_i, z_i ) \). \( P = \{ p_1, p_2, \ldots, p_r \} \) is the list of planar polygons, and \( p_r = \{ v_{n,1}, v_{n,2}, \ldots, v_{n,k_r} \} \) where \( k_r \) is the number of vertex in a polygon \( p_r \). If \( k = 3 \) for all \( p_r \), so the mesh is named triangle mesh.

Received Oct 30, 2011 / Accepted Jan 11, 2012
2.1 Data Set

Princeton Shape Benchmark (PSB) has 1814 shapes dividing it into training and test subsets of 907 shapes each [14]. The subset used in this paper is the test subset, see figure 1. PSB is a difficult data set because generally are non-regular and rustic models containing degenerations like non-connected triangles, also have position, size and different rotation degrees.

![Figure 1. Example of subset of figures from PSB.](image)

3 Invariant control of a 3D Shape

Generally, 3D models of a data base are in units, position and orientation in an arbitrary way. Due to the previous, techniques of invariants control shall be applied, in order to obtain the right properties for every condition.

3.1 Rotation Invariant

Rotating condition of a 3D model is a complicated problem compared to the scale and translation invariants. This is when the general technique of Principal Component Analysis (PCA) is employed, which estimates the main model axis, so that the model is rotated and its main axis is associated with a canonical coordinate system.

PCA is based on a statistical representation of random variables in the following way:

Let:

\[ P = P_1, P_2, \ldots, P_n \]  \hspace{1cm} (2)

Where

\[ P_i = (x_i, y_i, z_i) \in \mathbb{R}^3 \]  \hspace{1cm} (3)

This way (3) is the set of vertex on the surface of a 3D model. The objective of PCA is finding a mapping \( \tau = \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) so a 3D model in any stage of rotation be set in a normal state. The procedure is following specified.

1. The figure’s barycenter is the following:

\[ x_c = \frac{1}{N} \sum_{i=1}^{N} x_i, \quad y_c = \frac{1}{N} \sum_{i=1}^{N} y_i, \quad z_c = \frac{1}{N} \sum_{i=1}^{N} z_i \]  \hspace{1cm} (4)

2. A new point corresponding to the model’s center is determined:

\[ P_c = (x_c, y_c, z_c) \]  \hspace{1cm} (5)

3. For every point: \( P_i = (x_i, y_i, z_i) \in P \) is applied the transformation:

\[ P_{i'} = (x_i - x_c, y_i - y_c, z_i - z_c) \]  \hspace{1cm} (6)

4. A new set of points is determined:
5. From the new set $\mathbf{P}^\prime(\mathbf{P}_1^\prime, \mathbf{P}_2^\prime, ..., \mathbf{P}_n^\prime)$ the covariance matrix is calculated:

$$
\mathbf{C} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{P}_i^\prime \mathbf{P}_i^\prime^	op
$$

(8)

6. Matrix $\mathbf{C}$ has the property of being symmetric of real numbers, so its eigen values are not real negative. Then, to $\mathbf{C}$ eigen values are calculated, are ordered in a decreasing way, finding its correspondent eigen vectors. Eigen vectors are scaled to unit and we form a $\mathbf{R}$ rotating matrix. All points in $\mathbf{P}^\prime$ are rotated forming a new set of points:

$$
\mathbf{P}^\ast = \{ \mathbf{P}_i^\ast | \mathbf{P}_i^\ast = \mathbf{P}_i^\prime \mathbf{R}, \mathbf{P}_i^\prime \in \mathbf{P}^\prime, i = 1, ..., n \}
$$

(9)

At the end, the model is rotated so the $x$ axis maps the eigen vector with the highest eigen value, the $y$ axis is mapped with the eigen vector with the second highest eigen value and $z$ axis is mapped to eigen vector with the lowest eigen value.

### 3.2 Scale Invariant

The model's size is determined modifying its dimensions, where any point lies out of the open interval $(0,1)$ of real values. Having as reference the geometric center of the figure, any vertex lies out of the limits of a sphere with radius $1/2$. Defining $r_{\text{max}}$ as the maximum distance of a figure from its geometrical center, $r_{\text{max}}$, and it is calculated as follows:

$$
r_{\text{max}} = \sqrt{\text{max}(x_i^\ast)^2 + \text{max}(y_i^\ast)^2 + \text{max}(z_i^\ast)^2}
$$

(10)

Finally obtaining the scale factor:

$$
\mathbf{s} = \frac{1}{2r_{\text{max}}}
$$

(11)

Hereinafter for every point in $\mathbf{P}^\ast$, transformation is performed and the set of scaled point is obtained:

$$
\mathbf{P}^\ast = \{(sx_i^\ast, sy_i^\ast, sz_i^\ast) | (sx_i^\ast, sy_i^\ast, sz_i^\ast) = \mathbf{P}_i^\ast \in \mathbf{P}^\ast, i = 1, ..., n \}
$$

(12)

### 3.3 Translation Invariant

In $\mathbf{P}^\ast$ the same topological structure is maintained like the original model. Finally all figure's data are moved to a reference frame with values within the open interval $(0,1)$, with its fixed center on the point $m_c = (1/2, 1/2, 1/2)$.

### 4 Construction of Representative Sequences

Representative sequences have as objective to perform a mapping of figures 3D to 1D in a usefull way for the comparison criteria. To obtain sequences, intersections of triangles of the normalized figure with rays directed to the geometrical center of the 3D model [15] are calculated. In every trajectory the model's representation is given by a series of distances taken from an origin $0(x_o, y_o, z_o)$ with direction $b(x_o, y_o, z_o)$, searching the intersection with the closest mesh triangle, if there is no intersection, distance is considered as zero. The origin of the ray is controlled in 3 different ways describing trajectories. Each of them was adjusted with a control value ($t$) to obtain 512 values, that means, every pseudocode tests 512 origins searching one intersection, see figure 3.

The representative sequence is created through a trajectory given by the order in which distances are taken to the object from an origin, which are provided by the following pseudocodes:
1) Trajectories on a cube.

\begin{verbatim}
Initialize \( y_0 \leftarrow 1/2 \)
Initialize \( z_0 \leftarrow 1/2 \)
Repeat for \( y_0 = 0, \ldots, (t-1) \) in intervals of \( 1/t \)
    Initialize \( z_0 \leftarrow 1, x_{pos} \leftarrow 0 \)
    Repeat for \( i = 0, \ldots, 2(t-1) \) in intervals of \( 1 \)
        Assign \( X_0 = x_{pos} \)
        Assign \( X_0 = x_{pos} \)
        If \( i \geq t-1 \)
            Assign \( x_{pos} = x_{pos} - 1 \)
            Assign \( z_0 = 0 \)
        Else
            Assign \( x_{pos} = x_{pos} + 1 \)
        Return \((X_0, y_0, z_0), (X_0, y_0, z_0)\)
\end{verbatim}

2) Trajectories on a cylinder.

\begin{verbatim}
Initialize \( y_0 \leftarrow 1/2 \)
Initialize \( z_0 \leftarrow 1/2 \)
Initialize \( r_c \leftarrow 1/2 \)
Repeat for \( \theta = 0, \ldots, 2(t-1) \) in intervals of \( 1/t \)
    Repeat for \( h = 0, \ldots, (t-1) \) in intervals of \( 1/t \)
        Assign \( X_0 \leftarrow h \)
        Assign \( Y_0 \leftarrow r_c \cos(\theta \pi) + y_c \)
        Assign \( Z_0 \leftarrow r_c \sin(\theta \pi) + z_c \)
        Assign \( X_0 \leftarrow h \)
    Return \((X_0, y_0, z_0), (X_0, y_0, z_0)\)
\end{verbatim}

3) Trajectories on a sphere.

\begin{verbatim}
Initialize \( X_0 \leftarrow 1/2 \)
Initialize \( Y_0 \leftarrow 1/2 \)
Initialize \( Z_0 \leftarrow 1/2 \)
Initialize \( r_c \leftarrow 1/2 \)
Repeat for \( \theta = 0, \ldots, (t-1) \) in intervals of \( 1/t \)
    Repeat for \( \phi = 0, \ldots, 2(t-1) \) in intervals of \( 1/t \)
        Assign \( X_0 = r_c \cos(\theta \pi) \sin(\phi \pi) + X_c \)
        Assign \( Y_0 = r_c \cos(\theta \pi) + y_c \)
        Assign \( Z_0 = r_c \sin(\theta \pi) \sin(\phi \pi) + z_c \)
    Return \((X_0, y_0, z_0), (X_0, y_0, z_0)\)
\end{verbatim}

In figure 2 it is shown the results of previous procedure with examples of the origin point of every way in which trajectories are performed. The implementation may perform 38471 intersections about a model of 14400 triangle polygons in a period close to 1 second [16].
5 Dynamic Time Warping (DTW)

DTW is a technique that has been widely employed in voice recognition. The characteristic of the technique is that it allows a non-linear mapping of a sequence in terms of another minimizing base distance between them that usually corresponds to Euclidean distance. DTW is proper to computing similarity among sequences that are out of sync in terms of the other, or when one of them presents or omits segments.

Having two sequences \( Q \) and \( C \) specified in (13), DTW is defined as (14).

\[
Q = \{q_1, q_2, \ldots, q_p\}, C = \{c_1, c_2, \ldots, c_p\}
\]

Where it is met:

\[
\begin{align*}
\text{DTW}(\emptyset, \emptyset) &= 0 \\
\text{DTW}(Q, \emptyset) &= \text{DTW}(\emptyset, C) = \infty \\
\text{DTW}(Q, C) &= D_{\text{base}}(h(Q), h(C)) + \begin{cases} \\
\text{DTW}(Q, r(C)) \\
\text{DTW}(r(Q), C) \\
\text{DTW}(r(Q), r(C)) \\
\end{cases}
\end{align*}
\]

Where:

\[
D_{\text{base}}(q, c) = (q - c)^2
\]
The previous is the traditional way of defining DTW, a temporary complexity technique provided by an expression $O(3^n)$, being untreatable for large size data.

5.1 Fast Dynamic Time Warping (FDTW) in its optimized way of DTW

FDTW is a non-recursive approach to DTW. The matching solution of two sequences with FDTW is the minimum total distance and the adjustment route (warpath), see figure 4. The adjustment route indicates the relation among the elements of both sequences provided by the found alignment. This distance is within a range of 0 to infinite; where distance $d$ near to 0 indicates similarity and to infinite dissimilarity. It is important to note that FDTW has some restriction because the space of solutions that can be found is reduced [2, 3].

The algorithm of solution to sequences alignment with FDTW is described:

- Initialize $r_{i,j} \leftarrow 0$ for $i = 2, \ldots, n$ and $j = 2, \ldots, m$
- Assign $r_{1,1} \leftarrow d_{base}(q_i, c_j)$
- Repeat for $i = 2, \ldots, n$
  - Calculate $r_{i,j} \leftarrow d_{base}(q_i, c_j) + r_{i-1,j}$
- Repeat for $j = 2, \ldots, m$
  - Calculate $r_{i,j} \leftarrow d_{base}(q_i, c_j) + r_{i,j-1}$
- Repeat for $i = 2, \ldots, n$
  - Repeat for $j = 2, \ldots, m$
    - Calculate $u \leftarrow d_{base}(q_i, c_j) + r_{i-1,j-1}$
    - Calculate $v \leftarrow d_{base}(q_i, c_j) + r_{i,j-1}$
    - Calculate $w \leftarrow d_{base}(q_i, c_j) + r_{i-1,j}$
    - Assign $r_{i,j} \leftarrow \min\{u, v, w\}$
- Assign $d \leftarrow r_{n,m}$

It is clear that complexity of such algorithm is provided by a quadratic polynomial term, that in terms of implementation a comparison has approximately 0.3 ms in a PC Pentium 4 (CPU with 2.4 GHz., 3 GB of RAM).

![Figure 4. Process of sequences alignment using FDTW: (a) Original sequences. (b) Alignment process. (c) Alignment found.](image)
6 3D Shape Matching Methodology based on Representative Sequences

Methodology is comprised of one stage of pre-process (see figure 5), that is about generating the three representative sequences for every 3D model (section 4). Sequences are stored in a common data base associating the model identifier and one of the 3 techniques of representation.

The following is to perform a matching using FDTW, this is carried on taking every sequence of the data base and determining its similarity value according to the rest of them, note that this matching is made among elements with origin in the same representation technique.

At the end of all matches we have a distance matrix per every representation technique. Matrices were stored in a data base, identifying the two compared objects associating its correspondent distance.

Once stored in the data base, it is possible to perform a consult of any model by identifier and to obtain an ordered list from lowest to highest according to the distance value.

Through a series of consults it is possible to observe particular characteristics of every technique within matching.

![Figure 5. Matching Methodology using FDTW between two 3D shapes. Sequences obtained from: (a) Trajectories on a cube. (b) Trajectories on a cylinder. (c) Trajectories on a sphere. This result are distance values on (1), (2) y (3).](image)

7 Techniques Comparison of 3D Shapes’ Representation

In most of recovering application of 3D shapes, the results are evaluated based on the coincidence of predefined groups by designers of data sets. Since data bases are used in different systems, classifications on these data also vary, creating the need of using methods to compare different techniques of 3D representation of shapes.

Princeton Shape Benchmark (PSB) provides a data base which intends a defined base classification, including a tool to measure performance.
If a comparison method of figures calculates a distance between two figures, being positive and minor when figures are similar and that the distance is large when they are not similar, so a numeric value may use the common performance to compare the method with others, under the principle that they work with the same data base of 3D figures.

For a method of comparison and a data base of 3D models having \((m_1, m_2, K, m_h)\), a matrix consisting in distances between pairs of models can be calculated. For any figure \(q \in M\), the list of most similar models \(h\) may be recovered from the distances matrix.

The most common qualitative measure that indicates the development of a technique of 3D shapes representations is the precision-recall curve. In order to obtain the precision-recall curve is used as input a distance matrix and a classification file provided by PSB.

Representations’ techniques used to compare our results are shown in table 1, it is indicated the type of procedure used for invariants (RTS: rotation, translation, scale), the type of information extraction, numeric transformation performed and representation type (FV: Feature Vector) [18]. Reference techniques correspond to the ones with better results considering precision-recall curve.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Prep.</th>
<th>Extraction</th>
<th>Numeric transf.</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth Buffer</td>
<td>RTS</td>
<td>Image</td>
<td>2D DFT</td>
<td>FV</td>
</tr>
<tr>
<td>Voxel</td>
<td>RTS</td>
<td>Volumetric</td>
<td>No</td>
<td>Histogram</td>
</tr>
<tr>
<td>Silhouette</td>
<td>RTS</td>
<td>Image</td>
<td>Sampling + DFT</td>
<td>FV</td>
</tr>
<tr>
<td>Rays-SH</td>
<td>RTS</td>
<td>Image</td>
<td>Sampling + Harm. Esf.</td>
<td>FV</td>
</tr>
<tr>
<td>Complex-SH</td>
<td>RTS</td>
<td>Image</td>
<td>Sampling + Harm. Esf.</td>
<td>FV</td>
</tr>
<tr>
<td>3DDFT</td>
<td>RTS</td>
<td>Volumetric</td>
<td>3D DFT</td>
<td>FV</td>
</tr>
<tr>
<td>Ray Moments</td>
<td>RTS</td>
<td>Surface</td>
<td>Sampling</td>
<td>FV</td>
</tr>
<tr>
<td>Rotational Invariant</td>
<td>RTS</td>
<td>Volumetric</td>
<td>Sampling</td>
<td>Histogram</td>
</tr>
</tbody>
</table>

### 8 Results

The comparative analysis of proposed representation techniques, precision-recall are used as a performance measure.

For a query it is summarized that \(q\), which is a member of class \(C\) with a size \(|C|\), precision (vertical axis) is the reason of relevant elements \(k_q\) (elements belonging to the same class of the query) according to the number of recovered figures \(k_{ret}\). Recall (horizontal axis) is the ratio for relevant elements \(k_q\) according to the query class size \(|C|\). Ideally, this curve must be a horizontal line with unitary value precision.

High precision means less non relevant elements, while low precision indicates more non relevant elements. Frequently precision is inversely proportional to recall, it is common to decrease recall indicator to increase precision. Recall measures how complete or sensible a technique of representation is. High recall means less non relevant, while low recall indicates more non relevant elements. Improving recall may decrease precision, since being exact is complicated while sampling space increases.

Figure 6 shows the results of proposed techniques comparison according to techniques in table 1.
It is observed in figure 6 that the 3 proposed representations show an advantage in terms of precision of 0.7 for a recall of 0.1 for trajectories on a cube, 0.715 for trajectories on a sphere and 0.72 for trajectories on a cylinder. The technique with a better general performing for a recall of 0.05, 0.1, 0.15, 0.2 and 0.25 was for trajectories on a cylinder, always presenting a better precision indicator.

9 Conclusions

It was shown through the implementation of mapping a 3D shape to a representative 1D sequence and the use of FDTW as a similarity distance, it is an approach to matching problem and the search of similarity among 3D shapes.

There are proposed techniques for representative sequences and a powerful distance function is used, the latter is determined under the observation of problem’s nature, a fact the classic 3D matching methods consider the less important.

Considering that a 3D model may be in different rotation conditions, scale and translation, direct comparison of entities turns complicated. At this moment an invariant control method is employed in order to employ a proper condition of the figure.

Results are indicatives of methodology’s potential before a set of heterogeneous shapes, since better results in terms of precision compared with 8 techniques of literature are shown. The latter provides the possibility of creating aggregates to methodology, unifying the three techniques of trajectories in order to improve results and finally creating a complete 3D shape retrieval system.

Thanks to

PhD. J. Figueroa Nazuno for his unconditional support in this paper.
References