CONTENTS

1. What
2. How
3. Results
4. Conclusion
WHAT

- Object retrieval system
- Unsupervised
HOW

- Encoding is key
- Projection
- Panoramic image

*Figure 14.8.* Cylindrical mapping around the y-axis. (a) The texture coordinate parameterization and image wrapping. (b) An example of cylindrical mapping.

*Figure 14.9.* Spherical mapping. (a) The texture-coordinate parameterization and image wrapping. (b) An evening-sky texture mapped to a dome using the spherical texture-coordinate generation function.

**POINT PLACEMENT**

![Diagram of a 3D model and point placement](image)

**Fig. 1** The cylinder used for acquiring the projection of a 3D model

\[
R = \frac{3}{2}d_{\text{mean}}
\]

\[
d_{\text{mean}} = \text{Average distances of the model's surface from the coordinate axes}
\]

**Fig. 2** The discretization of the lateral surface of the projection cylinder (points in orange) to the set of points \( s(\varphi_u, y_v) \)

The set of points \( s(\varphi_u, y_v) \)

\[
\varphi_u = u \times 2\pi / 2B \quad u = 0, 1, \ldots, 2B - 1
\]

\[
y_v = v \times H / B \quad v = 0, 1, \ldots, B - 1
\]

\[B = \text{Sampling rate} = 64\]
PROJECTIONS

- Two projections: $s_1(\varphi_u, y_v), s_2(\varphi_u, y_v)$
- $S1$ is for position
- $S2$ is for orientation

$s_1(\varphi_u, y_v) = \text{pos}(\varphi_u, y_v)$, which is the length from the center to the models surface

That means that the value of a point $s_1(\varphi_u, y_v)$ lies in $[0, R]$

$s_2(\varphi_u, y_v) = |\cos(\text{ang}(\varphi_u, y_v))|^n$

Fig. 3 The top-most cross section of the cylinder along with the corresponding rays emanating from the center of the cross section $c_{B-1}$
Fig. 4  (a) Pose normalized 3D model, (b) the unfolded cylindrical projection of the 3D model capturing the position of the surface (c) the unfolded cylindrical projection of the 3D model capturing the orientation of the surface
MULTIPLE PROJECTIONS

Fig. 5 (a) A 3D model of a cup and (b)–(d) the corresponding cylindrical projections $s_{1,t}(\varphi_t, y_t)$ using three cylinders each one aligned with the $z$, $y$ and $x$ coordinate axis, respectively.

Fig. 8 (a) A 3D model of a car and (b)–(d) the corresponding cylindrical projections $s_{1,t}(\varphi_t, y_t)$ and $s_{2,t}(\varphi_t, y_t)$ for $t = z$, $t = y$ and $t = x$, respectively.
PROJECTION MODIFICATIONS

- 2D Discrete Fourier Transform (DFT)
- 2D Discrete Wavelet Transform (DWT)
  - Mean
  - Standard deviation
  - Skewness
- Weighing factor

For each cylindrical projection \( s_{k,t}(\varphi_u, y_v) \) \( (k \in [1, 2] \) and \( t \in \{x, y, z\} \)), we compute the corresponding 2D Discrete Fourier Transform (DFT), which is given by:

\[
F_{k,t}(m, n) = \sum_{u=0}^{2B-1} \sum_{v=0}^{B-1} s_{k,t}(\varphi_u, y_v) \cdot e^{-2\pi j \frac{mu}{2B + \frac{nu}{B}}} \quad (3)
\]

where \( m \in [0, 2B - 1] \) and \( n \in [0, B - 1] \).

For each cylindrical projection \( s_{k,t}(\varphi_u, y_v) \) \( (k \in [1, 2] \) and \( t \in \{x, y, z\} \)), we compute the corresponding 2D Discrete Wavelet Transform (DWT), which is given by:

\[
W^{\psi}_{k,t}(j_0, m, n) = \frac{1}{\sqrt{2B \cdot B}} \cdot \sum_{u=0}^{2B-1} \sum_{v=0}^{B-1} s_{k,t}(\varphi_u, y_v) \cdot \psi_{j_0,m,n}(u, v) \quad (5)
\]

\[
W^{\psi}_{k,t}(j, m, n) = \frac{1}{\sqrt{2B \cdot B}} \cdot \sum_{u=0}^{2B-1} \sum_{v=0}^{B-1} s_{k,t}(\varphi_u, y_v) \cdot \psi_{j,m,n}(u, v) \quad (6)
\]

where \( m \in [0, 2B - 1] \), \( n \in [0, B - 1] \), \( j \geq j_0 \) denotes the scale of the multi-level DWT, \( j_0 \) is the starting scale and \( \varphi_{j_0,m,n}(u, v) \), \( \psi_{j,m,n}(u, v) \) denotes the scaling and wavelet function, respectively. The \( W^{\psi}_{k,t}(j_0, m, n) \) approximation-scaling coefficients correspond to the low-pass subband of the transform at the starting scale \( j_0 \). The \( W^{\psi}_{k,t}(j, m, n) \) detail-wavelet coefficients correspond to the vertical, horizontal and diagonal subbands. We take the absolute values of the coefficients and normalize to their \( L_1 \) norm, which are now denoted as \( \tilde{W}^{\psi}_{k,t}(j_0, m, n) \) and \( \tilde{W}^{\psi}_{k,t}(j, m, n) \).
**Fig. 6** (a) A typical 2D Fourier transform of a cylindrical projection; (b) The Fourier coefficients that lie inside the area of the ellipsoid are discarded to reduce dimensionality.

\[
\mathbf{s}_F = (\tilde{F}_{1,x}, \tilde{F}_{2,x}, \tilde{F}_{1,y}, \tilde{F}_{2,y}, \tilde{F}_{1,z}, \tilde{F}_{2,z})
\]

**Fig. 7** 2-level wavelet transformation of a cylindrical projection of an airplane (the image is shown in negative colors).

\[
\mathbf{s}_W = (\tilde{V}_{1,x}, \tilde{V}_{2,x}, \tilde{V}_{1,y}, \tilde{V}_{2,y}, \tilde{V}_{1,z}, \tilde{V}_{2,z})
\]
ALIGNMENTS

- Two alignment schemes: CPCA and NPCA
- Different objects are better aligned with either

\[ p_l = (s_{F,l}, s_{W,l}), l \in \{c pca, n pca\} \]
\[ P = (p_{c pca}, p_{n pca}) \]

Fig. 3. (a), (d) columns: alignment using NPCA; (b), (c) columns: alignment using CPCA. The objects of columns (a), (b) are better aligned using NPCA while those of columns (c), (d) are better aligned using CPCA.
DETERMINING SIMILARITY

- Mahattan / Canberra
- Local Relevance Feedback (LRF)

\[
d_l(p_l, \hat{p}_l) = L_1(s_{F,l}, \hat{s}_{F,l}) + D_{can}(s_{W,l}, \hat{s}_{W,l})
\]

\[
D(P, \hat{P}) = \min_l d_l(p_l, \hat{p}_l)
\]
RESULTS

- Noise
- CPCA vs NPCA

Table 2: Effect of noise in the determination of principal axes for the CPCA and NPCA rotation normalization methods within the PSB dataset

<table>
<thead>
<tr>
<th>Alignment method</th>
<th>$\sigma = 0.01$</th>
<th>$\sigma = 0.03$</th>
<th>$\sigma = 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPCA</td>
<td>6°</td>
<td>6.6°</td>
<td>7.4°</td>
</tr>
<tr>
<td>NPCA</td>
<td>10.4°</td>
<td>14°</td>
<td>16.8°</td>
</tr>
</tbody>
</table>

Table 1: Characteristics of the evaluation datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#models</th>
<th>#classes</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCCC</td>
<td>472</td>
<td>55</td>
<td>generic</td>
</tr>
<tr>
<td>NIST</td>
<td>800</td>
<td>40</td>
<td>generic</td>
</tr>
<tr>
<td>WM-SHREC</td>
<td>400</td>
<td>20</td>
<td>generic</td>
</tr>
<tr>
<td>MPEG-7</td>
<td>1300</td>
<td>135</td>
<td>generic</td>
</tr>
<tr>
<td>PSB</td>
<td>907</td>
<td>92</td>
<td>generic</td>
</tr>
<tr>
<td>ESB</td>
<td>866</td>
<td>48</td>
<td>CAD</td>
</tr>
</tbody>
</table>

Fig. 9: (a) Example 3D objects and (b)–(d) the effect of different degrees of additive Gaussian noise on their surface

Fig. 10: Demonstration of the effect of different amounts of additive Gaussian noise on a depth buffer
RESULTS

Fig. 12 Performance evaluation of the PANORAMA descriptor when using: (a) $s_1(\phi_x, \gamma_y)$ (PANORAMA $\rightarrow$ Pos), $s_2(\phi_x, \gamma_y)$ (PANORAMA $\rightarrow$ Ori) or both (PANORAMA); (b) a single cylinder aligned with the $z$ coordinate axis (PANORAMA $\rightarrow$ $z$-projection) and a set of three perpendicular cylinders aligned with the object's principal axes (PANORAMA).
RESULTS

Fig. 13 Precision-recall plots comparing the proposed PANORAMA descriptor against the 2D/3D Hybrid, DESIRE, LF and SH-GEDT descriptor in various 3D model datasets. The comparison includes the combination of the PANORAMA descriptor with local relevance feedback (LRF).
### RESULTS

Fig. 14 Examples of queries within the PSB dataset and the corresponding top 5 retrieved models using the PANORAMA descriptor. The retrieved objects are ranked from left to right in decreasing order of similarity.
# RESULTS

Table 3: Quantitative measures scores for the proposed PANORAMA, 2D/3D Hybrid, DESIRE, LF and SH-GEDT methods for the CCCC, NIST, WM-SHREC, MPEG-7, PSB and ESB datasets.

<table>
<thead>
<tr>
<th>Method</th>
<th>CCCC NN (%)</th>
<th>FT (%)</th>
<th>ST (%)</th>
<th>WM-SHREC PANORAMA + LRF</th>
<th>ST (%)</th>
<th>PSB PANORAMA + LRF</th>
<th>ST (%)</th>
<th>MPEG-7 PANORAMA + LRF</th>
<th>ST (%)</th>
<th>ESB PANORAMA + LRF</th>
<th>ST (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PANORAMA + LRF</td>
<td>87.9</td>
<td>66.3</td>
<td>81.2</td>
<td>95.7</td>
<td>74.3</td>
<td>83.9</td>
<td></td>
<td>95.7</td>
<td>74.3</td>
<td>83.9</td>
<td></td>
</tr>
<tr>
<td>PANORAMA</td>
<td>87.4</td>
<td>60.2</td>
<td>75.8</td>
<td>95.7</td>
<td>67.3</td>
<td>78.4</td>
<td></td>
<td>95.5</td>
<td>64.2</td>
<td>77.3</td>
<td></td>
</tr>
<tr>
<td>2D/3D Hybrid</td>
<td>82.8</td>
<td>55.6</td>
<td>70.0</td>
<td>91.7</td>
<td>53.5</td>
<td>67.3</td>
<td></td>
<td>92.3</td>
<td>52.6</td>
<td>66.2</td>
<td></td>
</tr>
<tr>
<td>DESIRE</td>
<td>79.8</td>
<td>50.2</td>
<td>63.1</td>
<td>87.0</td>
<td>44.7</td>
<td>58.5</td>
<td></td>
<td>87.0</td>
<td>44.7</td>
<td>58.5</td>
<td></td>
</tr>
<tr>
<td>LF</td>
<td>75.7</td>
<td>45.9</td>
<td>59.9</td>
<td>55.3</td>
<td>31.0</td>
<td>41.4</td>
<td></td>
<td>55.3</td>
<td>31.0</td>
<td>41.4</td>
<td></td>
</tr>
<tr>
<td>SH-GEDT</td>
<td>90.4</td>
<td>71.5</td>
<td>84.1</td>
<td>87.2</td>
<td>65.5</td>
<td>75.9</td>
<td></td>
<td>86.1</td>
<td>59.6</td>
<td>70.7</td>
<td></td>
</tr>
<tr>
<td>PANORAMA</td>
<td>90.6</td>
<td>63.4</td>
<td>77.5</td>
<td>87.2</td>
<td>61.8</td>
<td>73.1</td>
<td></td>
<td>86.1</td>
<td>59.6</td>
<td>70.7</td>
<td></td>
</tr>
<tr>
<td>2D/3D Hybrid</td>
<td>88.1</td>
<td>55.6</td>
<td>72.1</td>
<td>86.1</td>
<td>59.6</td>
<td>70.7</td>
<td></td>
<td>86.4</td>
<td>57.7</td>
<td>67.7</td>
<td></td>
</tr>
<tr>
<td>DESIRE</td>
<td>83.7</td>
<td>50.9</td>
<td>64.9</td>
<td>80.2</td>
<td>51.7</td>
<td>61.9</td>
<td></td>
<td>80.2</td>
<td>51.7</td>
<td>61.9</td>
<td></td>
</tr>
<tr>
<td>LF</td>
<td>84.1</td>
<td>43.9</td>
<td>56.0</td>
<td>83.7</td>
<td>50.3</td>
<td>59.4</td>
<td></td>
<td>80.3</td>
<td>40.1</td>
<td>53.6</td>
<td></td>
</tr>
<tr>
<td>SH-GEDT</td>
<td>76.5</td>
<td>40.5</td>
<td>53.7</td>
<td>80.3</td>
<td>40.1</td>
<td>53.6</td>
<td></td>
<td>80.3</td>
<td>40.1</td>
<td>53.6</td>
<td></td>
</tr>
</tbody>
</table>
CONCLUSION

- Novel 3D shape descriptor
- Superior performance
- Increased efficiency
- Can be beneficial in other object retrieval systems