DESIGN OF A MULTIMEDIA DATA MANAGEMENT SYSTEM THAT USES HORIZONTAL FRAGMENTATION TO OPTIMIZE CONTENT-BASED QUERIES

Marcos Joaquín Rodríguez-Arauz, Lisbeth Rodríguez-Mazahua, Mario Leoncio Arrioja-Rodríguez, María Antonieta Abud-Figueroa, Silvestre Gustavo Peláez-Camarena, Luz del Carmen Martínez-Méndez

Division of Research and Postgraduate Studies
Tecnológico Nacional de México/Instituto Tecnológico de Orizaba
Content

- Introduction
- Design of the multimedia data management system
- Results
- Conclusion
- References
Introduction

- Multimedia database features become transcendent when the number of stored multimedia objects increases and such big challenges begin to appear [Silberschatz et al. 2011].

- The importance of fragmentation lies in the fact that it allows minimizing the number of access to irrelevant data, thus reducing the response time and the execution cost of the queries [Özsu y Valduriez, 2020].

- Efficient Content-based Image Retrieval (CBIR) techniques are a must for the optimal use of multimedia databases [Prinka and V. Wasson, 2017].
Design of the multimedia data management system

The system architecture was designed based on the Model View Controller architectural model (MVC) which distributes the system components in a way that facilitates its maintenance and is represented abstractly in Figure 1.

Figure 1. Application architecture.
Design of the multimedia data management system

- Horizontal fragmentation

The images collection is fragmented considering the attribute type according to the [Fasolin et al. 2013] method, which gives as result the horizontal fragments named buildings, events, people, and equipment.

Figure 1. Application architecture.
Design of the multimedia data management system

- Figure 2 shows the use case diagram of the Web application and announces the functional requirements. There are five actors.
Design of the multimedia data management system

- The conceptual model in this work is represented by the physical diagram of the database on Figure 3. The images collection is fragmented considering the attribute type according to the [Fasolin et al. 2013] method.
Design of the multimedia data management system

- A CBIR System
  
  Once a valid file is uploaded, the system will analyze the image and retrieve the most similar content allocated in the specific fragment type of the image itself.

Figure 4. Content-Based Image Retrieval diagram of the process model.
Results

- It is compared the execution cost of four content-based queries in the multimedia database of the proposed system without fragmentation (NF) vs. horizontally fragmented (HF) using the cost model described by [Rodríguez-Mazahua et al, 2014].

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Number of Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events</td>
<td>756</td>
</tr>
<tr>
<td>Equipment</td>
<td>63</td>
</tr>
<tr>
<td>People</td>
<td>886</td>
</tr>
<tr>
<td>Buildings</td>
<td>158</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table I. Predicates Used by Queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>q1</td>
</tr>
<tr>
<td>q2</td>
</tr>
<tr>
<td>q3</td>
</tr>
<tr>
<td>q4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table II. Fragments of the MMDB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragments</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>fr1</td>
</tr>
<tr>
<td>fr2</td>
</tr>
<tr>
<td>fr3</td>
</tr>
<tr>
<td>fr4</td>
</tr>
</tbody>
</table>

\[
\text{ITAC}(q_j) = \begin{cases} 
  f_j(\text{card}_k - \sum_{\pi_i \in P_j sel_{\tau}}) & \text{if} \quad n_{\pi_j} \geq 1 \\
  \text{otherwise} & 
\end{cases} 
\]

\[
P_j = \{p_k | PUM(q_j, p_k) = 1 \land FPUM(f_{\pi_k}, p_k) = 1 \} 
\]
Results

Figure 5. Shows the comparison of the execution cost of the queries.
Conclusion

- This work presented the design of a system for multimedia data management.
- The system uses horizontal fragmentation to improve the execution cost of content-based queries.
- The cost of each query using horizontal fragmentation is significantly lower than when a fragmentation method is not used.
- Future work includes the implementation of the system using the selected technologies and the incorporation of a vertical fragmentation method for even better optimization of the content-based queries.
References

THANK YOU!

email: marcos.joroar@gmail.com