An Authoring Solution for a Façade-Based AR Platform: Infrastructure, Annotation and Visualization

Guillaume GALES
Eric McCLEAN
John MCDONALD
Motivation

Input Image

Augmented Image
Overall System

Mobile client side

Server side

Database

Authoring client

Façade extraction

Façade Matching

Widget Retrieval

Rendering
Façade Extraction

Authoring solution
  - Façade database
  - 3D annotation

Visualization

Conclusion, perspectives and discussion
Façade Extraction
Façade Extraction

Key of the authoring solution and visualization
- Façades are the frame of references of 3D content
- Augment a view with 3D content in urban environment from a single image

Input: image of a façade
Output: homography between the façade and its image

Advantages:
- Robust matching (invariant to rotation and perspective changes)
- Façade normalization (used to build a representation of the environment)
- Camera pose estimation (used by the visualization system)
Façade Extraction
Authoring Solution
For each street, take pictures of façades

Automatic façade extraction and matching

- Geometrical constraint makes matching robust
- Invariance to rotation and perspective changes
Façade Database

Input

Database
Candidate selection

- GPS coordinates
- Bag-of-word description for selecting candidates

Similarity constraint
3D Annotation

- Extension of façade to 3D: frame of reference
3D Annotation

- Desktop application
  - Easy and intuitive interface to use for non expert users
  - Predefined list of 3D models
3D Annotation

- Mobile Application (ongoing work)

Long Tap to add widgets  Manual Accurate Geolocalisation
Visualization
The infrastructure is flexible in that we can add different users submit an image from their mobile, thus providing information about that location. Finally, we use images of the facades to build a map of the environment, as well as a frame of reference to link 3D coordinates and bag-of-words make our solution fast and suitable for a large amount of data. Furthermore, we do not need the user to define the frame of reference for that particular street. This frame is then extended to 3D, and the system provides a representation suitable for augmented reality applications.

In this paper we propose an infrastructure that makes authoring usable augmented reality applications, an efficient creation within that frame of reference, we can easily add 3D content for use in urban environments. It provides frames of reference for the environment as well as a mechanism to match new content on top of the facades and hence opens many possibilities for augmented reality applications in urban environments.

We start by creating an image database where the images of the facades of the same street are stitched together to provide a 3D view of the building. The user submits an image of the building facade with their mobile phone, then the system can place the user image relative to the scene. A video showing these steps is available at: http://www.cs.nuim.ie/research/vision/data/ismar2017.

The infrastructure of the image database described in this paper has many advantages:

- Flexible and expandable
- Scalable content model
- Easy and intuitive authoring
- Easy, intuitive and flexible

For example, we can position a cup of coffee in front of the doors of a building to indicate there is a coffee shop inside this building. Furthermore, once the system has learnt the position of the facade of a building, it is possible to retrieve the coordinates of the augmented content associated with this facade.

Since there is an homography between the image of the in-situ image and the put facade and the image stored in the database, we can easily add 3D content. Our system is based on the fact that, in an urban environment, many of the facades are approximated by planes. Furthermore the system provides a representation suitable for use in urban environments. It provides frames of reference for the environment as well as a mechanism to match new content on top of the facades and hence opens many possibilities for augmented reality applications in urban environments.

Our goal is to create a platform that makes it easy for both manual and automatic content authoring.

Keywords: Augmented Reality, Infrastructure, Authoring, Data Retrieval

1. Introduction

2. Adding content

3. Retrieving content

4. Conclusion

Figure 1: Overall system. The infrastructure of the image database is based on the fact that, in an urban environment, many of the facades can be approximated by planes. This space is used as a frame of reference to store content. For example, we can position a 3D coffee cup in front of the doors of a building to indicate there is a coffee shop inside this building.

The process of adding new streets is as follows:

- We use images of these facades to build a map of the environment.
- We then use this map to create a 3D model of the environment.
- We then use this 3D model to position the augmented content.

Data retrieval is then extended to 3D by using the normal of the plane as the third dimension. Next, given any particular location, we can easily add 3D content. Our system is based on the fact that, in an urban environment, many of the facades can be approximated by planes.
**Camera Pose Estimation**

- intrinsics are known
- extrinsics are given by the façade extraction algorithm (homography between plane and its image decomposed into rotation and translation)

\[
x = K \begin{bmatrix} R & t \end{bmatrix} \begin{bmatrix} 0 & 0 & sh & sh \\ 0 & h & h & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}
\]

\[
= H \begin{bmatrix} 0 & 0 & sh & sh \\ 0 & h & h & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}
\]
Results
Conclusion,
Perspectives and
Discussion
Conclusion, Perspectives & Discussion

- Easy and intuitive authoring solution for AR applications based on façade extraction
- Extension to real time computation on mobile
  - Use of GPU and accelerometer
- Extension to automatic dynamic authoring
  - Link with OpenStreetMap through GPS coordinates
- Collaborative authoring client (and database updates)
  - HTML5
  - Mobile version
Acknowledgment

- Research presented in this paper was funded by a Strategic Research Cluster grant (07/SRC/I1169) by Science Foundation Ireland under the National Development Plan. The authors gratefully acknowledge this support.
- Thank you for your attention