CONSTRUCT A NARRATIVE ABOUT A BUILDING USING VIDEO

A Prototype For Computer Aided Building Documentation Video Production System

XIAOLEI CHEN, ATIBA PHILLIPS
Dept. of Architecture, University of California, Berkeley
Lilychen@uclink.berkeley.edu

Abstract. This paper proposes a prototype for a computer aided building documentation video production system, which is based on building characteristics and certain cinematic rules and best practices, to assist non-technical video producers without significant prior filmmaking knowledge in the process of producing video documentaries about buildings and potentially furthermore, the recorded media can be retrieved and re-sequenced for multiple purposes.

1. Motivation

Architects and building professionals have a strong motivation to document buildings, for their own future design reference and professional need, as well as for teaching to upcoming students in the building professional fields. Occasionally, for some famous structures, these documentaries may be produced for public consumption.

Currently the vast majority of building documentaries are produced by architects using regular still cameras. The architect physically visits building site and selects based on his own subjective reasoning the shots that he wishes to take for his pre-defined purpose.

There are problems with this approach, for instance, it is difficult for still pictures to show the target as well as its context. Only limited annotations (or metadata) are currently held per picture / presentation and hence a large workload is placed on the presenter to construct a narrative context for the series of still shots given to him.

On the other hand, as video technology becomes cheaper and more ubiquitous, there is a growing tendency to document buildings with video,
which proves many advantages as compared to the still photo documentation. For instance, it is much easier to show context along with target and convey continuity of landscape and building with video. It has significant ascendance in constructing a narrative about a building via its time-based sequence. However, the new challenge that an architect or building professional may face is insufficiency of film production expertise, which has a significant impact in the quality of documentation product.

The goal of this project is to create an interactive prototype of a software application that assists non-technical video producers without significant prior filmmaking knowledge in the process of producing video documentaries about buildings and potentially furthermore, the recorded media can be retrieved and re-sequenced for multiple purposes.

2. Goals

As understated but extremely important cultural artifacts, buildings, though vary much in appearance, do have many common descriptors for their various features. Concomitantly, the video production process also has many rules and best practices that have evolved and that should be applied to the video used to document buildings. Furthermore, after shooting is done, there is also a constrained set of typical sequences used to construct a narrative about a building. So, there is a growing need and opportunity to anchor these cinematic rules, as well as these common building descriptors, in a computational system that can work to assist non-traditional video producers (such as architects) to achieve high quality video products within a guided video documentation process.

We propose an expert multi-media information system, which will assist an architect or a building professional in the production of building video documentaries by:

- automating the pre-production process by using a series of pre-annotated templates and automating the application of cinematic rules
- using the system’s in-built cinematic knowledge to assist the user during production
- using automatically generated and user provided metadata, to automatically assemble a “cinematically correct” edit of the user video
- leveraging the annotations on each shot to allow complex searching, retrieval and re-purposing of generated media assets

Advantage of an Automated Building Video Documentation System:
• More efficient time/effort-wise (fewer shots but greater effect); little video production skill required
• Video can be shot as narrative thus reducing the need for a skilled presenter
• Captured video is easier to be annotated, both with text and voice, and can be distributed as a multimedia package
• Annotated video assets are easier to be re-used/ re-purposed beyond original intentions (e.g. to tell a different story to a new audience)

Benefits of Intended System:
• Pre-production (Prototype Focus)
  o Sequence Shot proportion Best Practices (for Narrative structure)
    • We propose several possible narrative structures for the video product, e.g. standard vs. whole Building Dominated vs. detail dominated with a difference in the percentage of time distributions
    • We provide annotated list of all possible shots using shot categories already established in the architecture world emanating (most notably) form the Art and Architecture Thesaurus.
  o Cinematic Help
    • Generates shot plan suggestion and attendant film length
    • Pre-calculates and suggests shot sequencing (narrative flow)
    • Allows for a movie preview (or shot sequence preview) before actual shooting
    • Allows re-calculation / re-generation based on user requirements / constraints
• Production
  o Cinematographic Best Practices
    • Suggested shot types and camera movement / timing per type
    • Suggested appropriate usage scenarios per shot type
    • Suggested time (shot duration) per shot type
    • Suggested shot framing through the use of visual icons and sample shots
  o Mise-en-Scene Best Practices
    • Best time of day to shoot given building orientation and desired shot effect (i.e. lighting)
• Best date to shoot given building location / orientation and the proposed shooting dates (seasonality and lighting concerns).

• **Post Production**
  - Automatic generation of final video based on pre-defined shot sequences. Drastically reduced user burden.

• **Other Possible system Benefits (mainly in retrieving data and re-constructing the narrative)**
  - The search and retrieval of specific video sequences based on high-level user semantic queries since rich metadata is linked with each shot
  - Re-combination / re-use of existing assets into new sequences for alternate purposes, e.g. generating a slide show about a building based on the key frame of selected shots
  - Intuitive and powerful editing facilities which alter the look and feel of held video assets to achieve alternate viewer affects in different audience contexts.

3. Process Description

See process flow diagram of building video documentation system in Figure 1.

4. A Walk Through of Screen Shots

4.1 BUILDING INFORMATION

This is the initial screen presented to the user. Other than general building information, there are two items of particular interest on this screen.

4.1.1. Upload CAD Drawing

This entry allows the user to make available to the system a 3D model of the building to be documented. The system will use 3D model to generate a first preview to the user based on the cinematic information to be provided.

4.1.2. Building Orientation

The user is given a compass diagram which she will use to indicate the orientation of the front face of the building to the cardinal points. (e.g. front face faces East). This metadata will later be used to provide suggestions to the user about when is the best time of day to take the shots that the system will propose, in order to achieve to the optimal lighting for narrative clarity.
Figure 1. Process Diagram
4.2. VIDEO INFORMATION

In this screen there are three main items:

4.2.1. Proposed Video Length

Here the user constrains the output video duration by entering a minute value which the system will use as target when designing its suggested shot plan.
4.2.2. Video Style Type

The user must specify to the system the video “style” or the shot emphasis which the system will use as a bias when generating the initial shot plan. There are three default narrative types:

- Standard (45%-55%-5%)
- Whole building Dominated (65% - 30% - 5%)
- Detail Dominated (30% - 65% - 5%)

(The percentage values indicate the approximate proportion of time which will be spent on whole building shots, detail shots, and ending shots respectively in the final video.)

4.2.3. Shooting Window

For this option the user is presented with two calendar images and asked to mark off a date range over which she will be able to actually shoot the video. The system will use this information, coupled with the orientation and location information to assist the user by suggesting optimal shooting dates / times per shot to obtain the best shot taking conditions.
4.3. SHOT PLAN

The shot plan is presented as listing of possible shots, and annotations such as shot description, shot duration and type. Here the user is allowed to customize the plan by adding or deleting shots or altering shot elements such as type or duration. A preview of a new customized movie is then made available by clicking the “Preview Customization” hyperlink.
4.4. STORY BOARD

The story board screen presents the shots to the user in the chronological sequence of the intended film. For each shot it will display the cinematic shot properties, camera information and suggested shooting time and duration. It will also allow users to customize these parameters to be more relevant to their particular building situation should this become necessary. Shot samples are also provided (by clicking the view samples button), so that users can get an intuitive sense of the time construction of movement shots (pans for e.g.), as well as a sense of optimal framing for both still and movement shots.
4.5. PDA SCREEN

After the user has downloaded the information to her PDA and is on site, the PDA screen (which can be sorted by sequence chronology or by suggested shooting time), guides the user along the shot capture process and also maintains the link between the system and the video production stage.

![Figure 9. Shot Plan—PDA Screen](image)

5. System Components

5.1. SHOT PLAN ALGORITHM

At the heart of the system is a shot plan algorithm, which translates the user supplied building information into a series of shot descriptions which, output to the user as the shot plan. The algorithm acts on both user supplied information (i.e. video length, and video proportion type) as well as system supplied data (such as sequence and proportion definitions, cinematic shot features and shot “appropriateness” filters) to automatically generate the suggested shot plan.
This algorithm dynamically re-calculates the video output parameters based on real time user input; i.e. when, during customization the user adds or deletes a shot, or amends shot properties (such as a move from a pan to a still), the algorithm recalculates the video shot sequence and total shot and film length based on the new and changing user input.

This algorithm exists both on the user PDA and on the main system. It is available in the main system as a tool for the user via the “Shot Plan” screen. The user may shoot the plan as suggested or may choose to modify the plan using the algorithm within the interactive PDA session where it is also available.

These processes take the burden off of the user, and serve to add significant value in terms of time taken in pre-production planning, and post production editing, but also in relieving the requirement of extensive technical video production / editing expertise.

5.2. 3D VIDEO PARSER

This module parses the various layers of the input CAD model and identifies the metadata for each layer. Through this process it is able to identify distinct building features and view and angles. Thus in conjunction with the shot plan algorithm which specifies which shots are to be taken as well as the length and type of each shot, it uses the CAD metadata file to render the shots and come up with a preview video which can be viewed and adjusted by the user.

5.3. VIDEO PARSER

This system component parses the downloaded digital footage to detect the start and end of shots. It identifies the satisfactory shots as indicated by the user and re-assembles the footage in the chronological narrative sequence identified by the shot plan algorithm and updated by the user. This system component almost totally negates the editing burden or necessity for the user.

References

Bloch, G.R. From Concepts to Film Sequences, Yale University Department of Computer Science, 1987
Isenhour, J.P. The Effects of Context and Order in Film Editing. AV Communications Review, 23 (1)
Rosenblum, R., Karen, Ph.D., R. When the Shooting Stops, the Cutting Begins: A Film Editor's Story. DaCapo Press, May 1988