



Web-based Real-time Video Indexing

White Paper

Abstract

This paper presents a new approach to video indexing. Traditionally used as a means for video asset reuse, indexing can also be used to create TV shows and build stories faster. This is a major advantage for creators of shows that have a short time-to-air, such as reality TV and news shows. The paper explains how to use indexing in this way, by redrawing the production workflow and moving the indexing process upstream to the actual shoot, in what is called real-time indexing. Next, it looks at Web-based search and story-building tools, compressed video and natural language descriptions as ways of supporting real-time indexing.



What is indexing?

Indexing is the act of describing video content. Depending on what the indexing is meant to capture, it can describe the action, the people interacting, the dialog, the image structure, the camera settings, etc. It can describe an aspect of the production as a whole, for example it can indicate the producer, director, year of production, etc., or it can provide information on particular segments and even specific images within a segment.

The purpose of indexing is to make it possible to find a specific clip that corresponds to certain criteria, such as those mentioned above: director, action, etc. Just as a library index pinpoints the exact location of a book by author, publisher, year of publication, etc, a video index pinpoints the exact location of a video segment according to specific criteria.

Underlying the indexing and tools used to build a video index is, of course, the search engine. The video search engine pinpoints video content that corresponds to desired criteria.

A video index thus contains descriptions of video content and may have different levels of detail. At the most basic level, it contains the title of a production and the cassette ID of the original material. At a more detailed level, it contains many hierarchical levels of descriptions: descriptions of shows, of each episode within a show and of each segment or scene within the show expressed as start and end time codes, thumbnail images, and, of course, the video itself. An index is self-contained in the sense that all information required to differentiate one shot from another is contained within it.



Common terminology

The term “**indexing**” is often used interchangeably with the term “logging.” However, the term “**logging**” is often used to refer to a textual description with associated time codes. The log is therefore not self-contained, since users have to go back to the original material to see the shot. Logging can be performed using such tools as Avid’s Media Logger, FileMaker or even a simple spreadsheet.

The term “**indexing**” is also used loosely to refer to digital asset management (DAM). Although indexing is at the heart of DAM, the latter encompasses many other functions, such as rights management, collaborative work (tracking assets that are checked in and checked out), transcoding from one format to another, tracking asset usage, etc. Indexing is the key to DAM, but indexing refers strictly to video description. DAM usually requires large infrastructures, while indexing can be performed on a simple workstation or notebook. Moreover, DAM usually implies committing to long-term, company-wide guidelines for asset storage, use and workflow. Indexing does not dictate workflow as a whole; it is a step in asset usage, and therefore workflow can easily be adapted to reflect a change within an organization.



The uses of indexing

We have defined indexing as the art of organizing video information so as to be able to find specific content quickly. **So what concrete use does indexing have?**

Traditional indexing for reuse

Indexing traditionally applies to the reuse of video assets. Reuse can mean:

- reselling a shot to a third party, as footage libraries do
- redistributing content for a different platform
 - Today, there is increased interest in redistribution because of the diversity of platforms and associated client profiles. For example, a television broadcast may be reformatted or repackaged for Web distribution, mobile TV, cell phones, etc. Because client profiles are different for each platform, interest is generated among new advertisers, creating potential for increased revenues.
- repackaging a shot for use in a new production
 - To be cost-efficient, producers can reuse a shot of a helicopter view of Montreal, for example, rather than shooting it again.

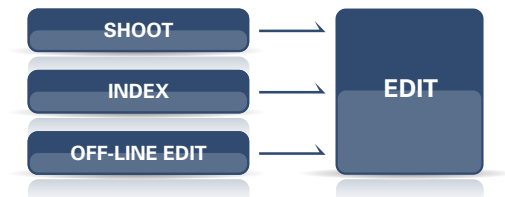
Rethinking workflow: indexing for creation

As we have seen, indexing for reuse refers mainly to taking advantage of existing assets, which are for the most part, finished material. In this article, however, we look at indexing from a different angle: as a production tool for creating an original show. In this context, indexing not only applies to raw material, it is also intertwined with the production process. This is why we call this type of indexing “real-time indexing.”

Traditional Linear Workflow



Parallel Workflow Using Real-Time Indexing





Real-time indexing

Real-time indexing is important in unscripted productions—shows in which action is largely improvised and unpredictable and for which no scripts exist.

Examples of unscripted shows are sports events, news feeds, and reality shows. These types of productions have a high ratio of raw to finished material. Just imagine the amount of tape required to get the best five-second clip of Kobe Bryant's slam dunk. For one reality show that used Lanterna Magica's system, 39 cameras captured action for over 12 hours to create a single one-hour show.

Unscripted shows are also often characterized by short time-to-air; they have to be put together in a matter of hours. TV news shows are a good example of this. As new items come in from different sources, they have to be viewed, filtered, have portions of them selected and then edited as quickly as possible. Reality TV has the same short time-to-air: to give viewers a real-time flavor, action is shot and raw footage has to be viewed, selected and edited in a matter of hours.

Unlike in a traditional workflow, in which indexing is done after the fact, real-time indexing is done while shooting is underway. Therefore at any given moment, the index is up to date. The only way to select the best material from huge amounts of footage in the least amount of time possible is through real-time indexing.



Advantages of real-time indexing

A traditional workflow has an extra indexing process at the end of the shooting period. Whether the indexing is done from tape or from a digital file, taking the extra time means wasting valuable creation time and potentially diminishing the quality of the show. With real-time indexing, not only do you save this time, but you can also start editing as soon as new material is shot.

Locate shots and build a story

Even a great index is useless without a good search tool. The search tool's main function is to pinpoint the exact video segment that corresponds to the search criteria. However, its use is not limited to locating the right segment, tape ID or time code. In fact, the search tool can be an important value-added tool if used during the editing process.

Let's see how.

When shots are selected within the search tool, they can be assembled into stories, roughly cut and rearranged, and an edit decision list (EDL) can be produced. When coupled with a story-building tool, the search tool can therefore do full-fledged off-line editing. More than that, it allows off-line editing to be done as soon as material is shot. The linear time boundaries between shot selection, rough cutting and editing suddenly vanish. Real-time indexing not only supports the production process, it is part of the creation process, up to the fine-tuning of transitions in a non-linear editing (NLE) environment.



Advantages of real-time indexing

Advantages of moving creation up in the workflow

The advantages of real-time indexing are that:

- it allows up-to-the-minute shots to be searched through and played instantly, dispensing with the need to sift through tapes
- it lets creators start putting shows together right away, without waiting until the end of a shooting session
- directors can change a story as better shots come along or as they choose a new angle in response to recent action
- it makes it possible to rough cut a story and therefore dispenses with the need for an off-line editing station
- it saves time on capture, as only required material from the EDL needs to be captured—nothing else

All in all, real-time indexing moves story building earlier in the workflow and therefore has a major impact on time-to-air.



Web-based search

In the context described above, indexing is performed live by specialized operators using dedicated workstations and software. But the end-user, the person who actually uses the index, is, of course, the person looking for shots and creating the story, namely the director, and by extension, the post staff. Why is it important for these users to have a search engine/story-builder that is Web-based? Because it fits perfectly into their creative vision. Here's how.

Ideally, directors want complete control over content, including assessing raw material, selecting the perfect shot, playing with the shot sequence, getting a feel for the sequence, modifying it and feeling free to play with different scenarios. They also want to bring their vision to life without the bottlenecks of staff or machine shortages. In a perfect world, they and their post staff want to work remotely, from home or from another city, as if they were at the shooting location. Plus, they want to be able to dispatch search jobs to many people at once and see what results they come up with. How does the Web help?

A Web-based search engine/story-builder:

- is accessible through a simple Web browser from your laptop, with no special hardware or software
- is accessible remotely from home, from another city, or from wherever
- is accessible to an unlimited number of simultaneous users
- lets you send your shot assembly to colleagues by email

That's what a Web-based search engine provides! It completely supports the creative vision of the people making the show.



Compressed vs. uncompressed video

Even though the cost of storage drops every six months, we believe compressed video is the most appropriate choice for indexing. In fact, compressed video offers many advantages over uncompressed video. Here are just a few:

- Compressed video can be manipulated (browsed, played, etc.) without special hardware.
- Using compressed formats is the only way to exchange and distribute video over standard networks and protocols.
- Compressed video allows the entire production, from day one, to be stored on and accessible from a standard disk.

On the one hand, then, the key benefit of compressed video indexing systems is that standard hardware and Web-based interfaces can be used. On the other hand, the key benefit of uncompressed video indexing systems is that they lead directly into editing, without the need to capture from your NLE system. What we advocate is that you do both and capture compressed and uncompressed formats in parallel! You simply use the compressed format for video manipulation and exchange and then use the resulting EDL to import uncompressed formats directly into your NLE. It's the best of both worlds!

If you would like to read more about this type of solution, please refer to the white paper "HDV: Shoot, Index, Edit" at www.lanternamagica.com.



Automatic vs. manual indexing

Examples of automatic indexing

Automatic indexing takes on many forms, all of which rely on signal processing for describing a specific aspect of video content. For example, face recognition indicates on which frame a template pattern appears. Face recognition provides good results in contexts where the video under analysis has relatively still images. Face recognition is difficult to achieve within real-time constraints, unless there is serious processing power at hand.

Voice recognition or speech-to-text systems provide good results in environments where there is a single speaker, background noise is weak and the system has been trained for that particular speaker. These systems can be useful for continuous news feeds or sportscasts where the system will automatically pick specific keywords such as "goal" or "penalty."

Scene detection is another aspect of automatic indexing. By way of frame luminosity variance analysis, scene detection automatically inserts cue marks at scene cuts. In non-real-time indexing, the resulting marks help the user navigate through the material and quickly find the segment to index. However, this approach is irrelevant with real-time indexing.



Automatic vs. manual indexing

Limits of automatic indexing

Automatic indexing is therefore useful only in very specific contexts and is limited in the aspects of video content it can describe. The decision to invest in automatic indexing has to be based on what you are trying to describe and whether the characteristics of the video material make that possible. We have found that in the great majority of cases, video content cannot be described in any way other than manually, because actions, emotions, intent, expressions, etc. simply cannot be recognized automatically. The knowledge base required to distinguish or correlate these descriptions is simply beyond the realm of current technology. Imagine automatically indexing a dinner party where everyone is speaking at once! Clearly, manual indexing remains, and will remain for years to come, the only viable solution.

Another approach: keywords and metadata standards

What an indexing system can provide, however, is structure. The indexing approach we have taken involves free text descriptions augmented by a user-defined keyword list. This provides flexibility as well as structure and consistency since the system automatically suggests keywords as you type. Depending on the type of material to be indexed and what you choose to describe, predefined keywords will relate to the subject as, for example, “goaltender,” “shot” and “penalty” relate to a sports event. Keywords not only allow for precise searches, they also ensure homogeneous indexing no matter who indexes or when—“goaler” will not be used instead of “goaltender.”

Descriptive metadata standards, such as the Society of Motion Picture and Television Engineers (SMPTE) metadata dictionary (see also MXF), are gaining wide acceptance. These standards use descriptor value pairs such as “Original Title/Survivor” and are useful for describing universal characteristics of a video. We have chosen to support SMPTE metadata for productions as a whole, but still emphasize free text descriptions, which are more natural. Until the widespread acceptance of a metadata standards and their support by NLE systems, natural language remains the most user-friendly approach.



Conclusion

We have seen how indexing can become part of the production workflow as a live, real-time process rather than being performed after the fact. This type of indexing, real-time indexing, allows for faster time-to-air, which is key for news shows and reality TV. The traditional benefits of indexing, which are mainly to do with reusing video assets, have been expanded to include the creation of a production itself. In effect, with a good search and story-building tool, the index can be used to select shots, assemble them and create rough cuts, all as soon as the first frame has been shot.

We have also seen how a Web-based video search/story-building tool increases the accessibility of video material, since a simple laptop from anywhere will do. Moreover, we have taken an approach that favors manual indexing using natural language descriptions and keywords, because in most cases automating indexing delivers irrelevant results and because metadata standards confine descriptions, whereas natural language enriches them.



Lanternamagica develops video logging systems and offers indexing services to help its customers produce shows more efficiently and maximize the value of their video assets.

For more details please contact us at:

4000 St. Ambroise St., Loft 283, Montreal, Quebec H4C 2C7

T (514) 807-5296 F (514) 807-5705 info@lanternamagica.com

lanternamagica.com