THE CONSUMER DIGITAL VIDEO LIBRARY

Margaret H. Pinson and Stephen Wolf

National Telecommunications and Information Administration (NTIA)

Neha Tripathi

Nanyang Technological University, Singapore

Chin Koh

Nevion USA

ABSTRACT

The shortage of high quality video sequences has hampered video quality research for 20 years. Every Video Quality Experts Group (VQEG) test has encountered difficulty in obtaining quality source content. For some validation tests, this caused a multiple-year delay. To address this critical need, the newly launched Consumer Digital Video Library (CDVL) website (www.cdvl.org) accepts and shares contributions of high-quality video sequences. The site is designed to assist video researchers in the development and testing of new algorithms for video processing, coding, and both subjective and objective quality measurement.

1. INTRODUCTION

The Consumer Digital Video Library (CDVL) website (www.cdvl.org) provides pristine video sequences without visible artifacts that may be used freely for research and development purposes. These high quality uncompressed video clips are relevant for different types of video processing and quality measurement applications such as subjective video quality assessment, and objective video quality model development. Users may enrich the available CDVL database by contributing their videos.

The CDVL fills a critical industry need that was recognized in prior work [1], [2]. Video researchers and developers have found their progress stifled by the unavailability of high quality royalty-free test material. Content that is available for purchase may be of insufficient quality for the intended application, and the associated usage agreements may prevent researchers from publishing and/or sharing their research results. In addition, this purchased content may be expensive and difficult to locate, or may become unavailable for purchase when researchers have developed newer processing algorithms and would like to publish comparison results. Content owners want to protect their property from being used for commercial applications without their permission, and this has prevented otherwise willing companies from sharing their content for research and development purposes (i.e., non-commercial applications). The mechanism of content redistribution has also been problematic, since high quality uncompressed video scenes require large amounts of storage.

The CDVL provides solutions to all of these problems by supplying a standard legal agreement that protects contributors' rights for commercial applications while facilitating royalty-free distribution for research and development purposes. The CDVL's goal is to support and maintain a repository of content and knowledge that will facilitate and foster collaborative research and development in the area of consumer video processing and quality measurement. The CDVL website serves as a centralized repository and an automated redistribution mechanism. The CDVL website provides user forums, where researchers can discuss problems and solutions, and expects to provide access to subjectively rated videos in the near future. The

CDVL website is hosted by the National Telecommunications and Information Administration (NTIA), which is a part of the U.S. Department of Commerce. NTIA has a 20-year history of supporting video quality research, which will provide long term stability and support.

The CDVL website accepts and shares contributions of video content that are most relevant for determining the effectiveness of consumer video processing applications (e.g. corrective processing, reconstructive processing, enhancement, and re-formatting) and quality measurement algorithms. Video clips can be downloaded for the purpose of assessing visual quality using subjective or objective methods and the research results can then be shared with other CDVL users and published in the open literature. Video clips can be accessed through a browse and search interface that incorporates clip descriptors (e.g., resolution, scanning format, frame rate, chroma sampling structure, content type, coding complexity) as well as recommended usage guidance (e.g., users may find candidate 1080i video clips that are suitable for de-interlacing applications).

The CDVL fosters a collaborative approach to maintaining high-quality original source content as well as sharing research results in the areas of consumer video processing and quality evaluation. Typical industry uses for these high-quality video sequences include:

- Selecting video equipment and services
- Optimizing video system performance
- Improving video coding algorithms
- Developing objective video quality models
- Quantifying improvements to video processing algorithms
- Conducting subjective video quality tests
- Establishing minimum quality specifications for specific applications (e.g., emergency telemedicine, surveillance, firefighting)
- Other video research

The CDVL is overseen by a technical review committee that strives to maintain the most relevant video content, usage guidance, and website interface. Where appropriate, this technical review committee will also pursue standardization of video content and its usage guidelines in bodies such as the International Telecommunications Union (ITU). The committee also reviews all submitted video to ensure appropriate content before it is released to the CDVL user community.

2. INITIAL AND ANTICIPATED CONTENT

Before launching the CDVL, NTIA loaded a substantial amount of video content onto the website. These scenes are categorized into "databases" of similar content, to aid in user searches. The content available at launch includes:

- ANSI T1.801.01 Standard Test Sequences (525-line) [3]
- ITU-R Rec. BT-802 Standard Test Sequences (525-line) [4]
- ITU-R Rec. BT-802 Standard Test Sequences (625-line)[1] [4]
- Broadcast Television
- NTIA Source Sequences
- Public Safety Communications Research (PSCR) Crowd Scenes
- PSCR Event Identification
- PSCR Target Identification

The ANSI, ITU-R 525-line and ITU-R 625-line databases contain the highest quality digitization available of these standard test sequences, in their original format (525-line/NTSC or 625-line/PAL).

The Broadcast Television database contains video content originally intended for broadcast television. At launch, this database contained advertisements made available by Big Green Rabbit[©]. In the near future, we expect to be able to include the Sveriges Television (SVT) HDTV content previously made available to VQEG. Companies involved in broadcast television are encouraged to make their content available for this database, so that research results better reflect the quality of source scenes used for broadcast television.

The NTIA Source Sequences database contains the source sequences made available by NTIA to VQEG prior to 2007. As time permits, NTIA will upload video sequences made available by NTIA to VQEG during and after 2007. The NTIA dataset contains a wide variety of content and resolutions, including sports, video teleconferencing, home video, animation, simulated news, music video, movie footage, and general television (both standard and high definition). We also expect to make new content available where content types appear to be missing (e.g., simulated news footage).

---

[1] CDVL contains only a sub-set of these public domain scenes. If you have a copy of the missing ITU-R Rec. BT-802 625-line source, please upload these to CDVL.

The Public Safety Communications Research (PSCR) program provides research, development, testing and evaluation to foster nationwide first-responder communications interoperability (see www.pscr.gov). Three PSCR datasets were made available to facilitate research and development in areas of interest to the public safety community (e.g., fire fighters, police, and emergency telemedicine). The first dataset, PSCR Crowd Scenes, contains HDTV crowd scenes filmed from a stationary tripod, which simulate crowd surveillance footage through a stationary camera. These crowd scenes include shots of a crowd walking down a hallway, and a crowded stadium at different focal lengths. The intention of the latter footage is that the user can determine both the quality and the zoom needed to perform recognition tasks. The second dataset, PSCR Event Identification, contains simulated crimes and police events, such as shoplifting, bank robbery, and simulated police stops. The third dataset, PSCR Target Identification, contains a series of repeatable events that differ by small details (e.g., the same person walking left to right, carrying a different object each time they cross the screen). These scenes were designed for target recognition subjective testing, as described in ITU-T Recommendation P.912 [5].

NTIA expects to make the following content available as time permits:

- VQEG Full-Reference Television (FR-TV) phase I subjective test (DMOS available) [6]
- NTIA source content made available to VQEG from 2007 to 2009
- Simulated news footage
- Portions of the VQEG HDTV subjective tests (when released to the public)
- PSCR smoke and fire footage from controlled burns
- Joint Verizon/NTIA HDTV subjective test that compares MPEG-2 and H.264 both with and without packet loss (video sequences & MOS) [7]

3. SUBMITTING VIDEOS

The CDVL website allows registered users to share high quality source video sequences with other researchers and developers. By uploading video sequences to the CDVL, users can encourage video research and development with a minimum of effort.

One purpose of this site is the redistribution of very high-quality (e.g., uncompressed) video sequences that contain minimal visual artifacts. This site can also be used to redistribute subjectively-rated video experiments and audio-video experiments. The preferred format for uploaded video is either uncompressed 8-bit UYVY AVI or uncompressed 8-bit RGB AVI. The CDVL technical committee is currently investigating suitable uncompressed 10-bit file formats.

Before submitting videos, the contributor must agree to a video posting license agreement. To read this agreement, select the "Submit Videos" link on the left side of the CDVL site, and then click the hyperlink in the sentence "Click here to start the submission process." After logging into the CDVL website and digitally signing the video posting agreement, the contributor will be able to describe and then upload the video sequence that the contributor wishes to submit.

The following information documents the submitted video and helps users find the most appropriate video sequence for their application. The first seven fields are shown in Figure 1, as follows:

- **Title:** A concise yet descriptive name for the video clip that will be displayed to the CDVL user (i.e., not too cryptic).
- **Description:** A brief description of the video clip's content and unique characteristics of interest to CDVL users. For example, the description may identify an intended application (e.g., testing de-interlacing algorithms, investigating quality needed for emergency telemedicine), or list unusual scene characteristics (e.g., fine detail, random movement of small objects).
- **Video Specifications:** Description of the video production equipment and process, such as the camera and lens used to shoot the video, camera settings like shutter speed, recording medium, capture system (if non-transparent), and editing suite (e.g., to add scene cuts, transitions, create overlays).
- **Audio Specifications:** Description of the audio production equipment and process, such as the microphone used, recording medium, and editing suite.
- **Processing History:** Identifies further encodings of the video after initial production, prior to current version (e.g., secondary compression, format conversion, corrective or enhancement processing).
- **Other Details:** Any other details of importance the submitter wants to include, such as credit to the contributing organization, or reference to related video sequences.

Figure 1.   Top portion of the CDVL video submission page.

Figure 2.   Bottom portion of the CDVL video submission page.

- **Copyright Restrictions:** Choose "Public Domain" or "Research and Development".[2] Video that does not fit either category should not be submitted.

The bottom portion of the "Submit Videos" documentation screen contains several drop-down lists, shown in Figure 2. In all cases, if the appropriate answer is not listed, select "other." The drop-down lists describe the video sequence as follows:

- **Resolution & Scanning Format:** Specifies the name for each resolution that is in common use (e.g., CIF, QCIF, VGA) along with the image size (pixels horizontally, then rows vertically), and whether the video is interlaced or progressive, for example "HDTV 1080i (1920 x 1080 interlaced)."
- **Frame or Field Rate:** Specifies the frame rate if progressive or the field rate if interlaced, both abbreviated as "fps."
- **Chroma Sampling:** Specifies the chroma sampling format used to store this video sequence (e.g., 4:4:4 for RGB, 4:2:2 for YUV with color information sub-sampled by two horizontally).
- **Content Type:** Specifies a broad category of content type (e.g., teleconferencing, animation, sports).
- **Coding Complexity:** Indicates the scene complexity and thus whether a video coder is likely to have an easy, moderate, or difficult time producing a high quality coded sequence at a given bit-rate.

---

[2] The acceptable types of copyright restrictions for CDVL are defined on the CDVL website.

- **Origin and Quality:** Specifies the quality of the video (excellent, good, or impaired) and whether the video sequence is original (e.g., captured directly from the camera) or has been processed further (e.g., format conversion, secondary compression).
- **Special Attributes:** specifies content characteristics that may be of interest. The available attributes are:
  - Flesh tones
  - High contrast, for example strong black/white edges
  - High motion, such as an object moving across the screen in 1-second
  - Night scenes and dimly lit content
  - Random camera motion, such as from a shoulder held camcorder
  - Rich colors and scenes containing a wide variety of colors
  - Water or fire, which contain random motions that are difficult for coders
- **Audio Type:** specifies whether video has embedded audio, and if so the approximate type of audio (speech, music, or multiple sources such as background noise or speech with music).

Once the video sequence is documented, the contributor will press the "Submit" button at the bottom of the screen and be given the option to install a web-browser plug-in application that uploads the video sequence. The video will be reviewed by the technical review committee and site administrator to ensure that only appropriate content is posted. The review process ensures that the CDVL contains only appropriate content, and may take several days.

4.   FIND & DOWNLOAD VIDEO

The CDVL website allows registered users to find and download high-quality video sequences for research and

development purposes. Some of these video sequences have audio. Subjectively rated video sequences will be made available for download soon.

Before searching the CDVL database and downloading videos, the user must agree to a database content user agreement that specifies appropriate uses for the video. To read this agreement, select the "Find Videos" link on the left side of the CDVL site, and then click the hyperlink in the sentence "Click here to start searching." After logging into the CDVL website and digitally signing the database content user agreement, users will be directed to the CDVL database search page.

Search criteria, shown in Figure 3, include the option to enter key words and select items from the pull-down lists. Currently, the title and description text boxes shown in Figure 1 will be searched for the key words. Most of the pull-down lists are identical to those entered when submitting a video (e.g., Resolution & Scanning Format, Audio Type). There is also one additional pull-down list:

- **Dataset Name:** The technical review committee organizes the CDVL video sequences into "datasets", to assist searching. The dataset may be the source of the video content (if one organization has contributed a lot of content), the name of a subjectively rated dataset of video sequences, or a group of similar content.

After clicking the "Search" button, a list of videos matching the search criteria are presented, as shown in Figure 4. The search results are presented in five columns, as follows:

- **Frame Preview:** Thumbnail of one sample frame from the video sequence.
- **Video Title and Description:** The video title, followed by the first portion of the video description.
- **Resolution and Scanning Format:** as provided by contributor (see Section 3).
- **Frame or Field Rate:** as provided by contributor (see Section 3).
- **Size (Mb):** size of video sequence file, in Megabytes.

Figure 3. The CDVL video search form.

Figure 4. Example CDVL search results.

This screen has two clip reviewing capabilities, which enable users to select only the most relevant videos for download. This facilitates efficient use of time and transmission bandwidth, since the video sequences are large, uncompressed AVI files. First, click on a thumbnail image displayed in the "Frame Preview" column, like those shown in Figure 4, to display one full resolution video frame. This enables users to obtain a general idea of the spatial attributes of the video scene. The description of the video sequence is presented below this full resolution sample frame.

Second, the video title hyperlink, like those shown in Figure 4, opens a web page containing all the available information for the video sequence, as described above under "Submitting Videos." Additionally, the user is given the option of playing a sub-sampled, compressed version of the video sequence, using a Windows Media Player™

plug-in[3]. This enables users to obtain a general idea of the temporal attributes of the video scene. To download a sequence, the user selects the "Download" button at the bottom of the detailed information page.

5. FORUMS

The CDVL website provides various forums. The "General Discussions" forum is intended for general discussion of the CDVL usage, questions, suggestions, and feedback. Questions not answered in this paper should be directed to that forum.

The CDVL website also provides forums designed to promote discussions and maximize the effectiveness of the available video sequences. The "Subjective Video Quality Testing" forum is for discussions involving subjective testing procedures and subjectively rated datasets that will be available in the CDVL. The "Objective Video Quality Models" forum is for discussions involving objective video quality model development. The "Video Correction, Enhancement, and Conversion" forum is for discussions involving corrective & reconstructive processing, enhancements, and reformatting & conversion (e.g., which video sequences are particularly suited to a particular endeavor). The "Video Coding" forum is for discussions involving the use of the CDVL scenes for codec development, tweaking, and testing. The "Public Safety and Surveillance" forum is for discussions involving surveillance, public safety scenes, and their applications.

6. CONCLUSION

The effectiveness of the CDVL depends in large part upon contributions and participation of users. Users are encouraged to make source video content available through the CDVL and to initiate forum discussions that are relevant to the industry. Please provide comments on the website and feedback on the available videos through the user forums.

7. ACKNOWLEDGEMENTS

The authors would like to thank Dr. Jorge Caviedes at Intel Corporation and Dr. Sanjit Mitra at the University of

8. REFERENCES

[1] C.C. Koh, J.E. Caviedes, and S.K. Mitra, "Content management in a consumer digital video library design," in Proceedings of VPQM07, Scottsdale, AZ, January 2007.

[2] C.C. Koh, J. Caviedes, and S.K. Mitra, "Test video selection for video processing research and development," in Proceedings of ICCE 2008, Las Vegas, NV, January 2008.

[3] ANSI T1.801.01-1995, "American National Standard for Telecommunications – Digital Transport of Video Teleconferencing/Video Telephony Signals - Video Test Scenes for Subjective and Objective Performance Assessment,"
https://www.atis.org/docstore/product.aspx?id=24816

[4] ITU-R Recommendation BT.802-1, "Test Pictures and Sequences for Subjective Assessments of Digital Codecs Conveying Signals Produced According to Recommendation ITU-R BT.601-4,"
http://www.itu.int/rec/R-REC-BT.802-1-199407-I/en

[5] ITU-T Recognition P.912, "Subjective Video Quality Assessment Methods for Recognition Tasks,"
http://www.itu.int/rec/T-REC-P.912-200808-I/en

[6] ITU-T COM 9-80-E, "Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment,"
http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseI/
June, 2000.

[7] M. H. Pinson, S. Wolf, and G. Cermak, "HDTV Subjective Quality of H.264 vs. MPEG-2 With and Without Packet Loss," to be published in *IEEE Transactions on Broadcasting*, March, 2010.

---

[3] Certain commercial hardware and software are identified in this report to specify adequately the technical aspects of the website. In no case does such identification imply recommendation or endorsement by the National Telecommunications and Information Administration, nor does it imply that the material or equipment identified is necessarily the best available for this purpose.