Shot Boundary Detection (SBD)

# Purpose of the Exercise

The purpose of this exercise is acquiring the practice in opportunities for video content analysis. An example of video content analysis is automatic Shot Boundary Detection (SBD). SBD is commonly used in case of creating video summarizations.

# Needed Knowledge

Before starting exercise, one should possess knowledge in the following topics:

1. SBD basics (why it is used)
2. SBD methods (general information)
3. Applications for SBD used during the exercise

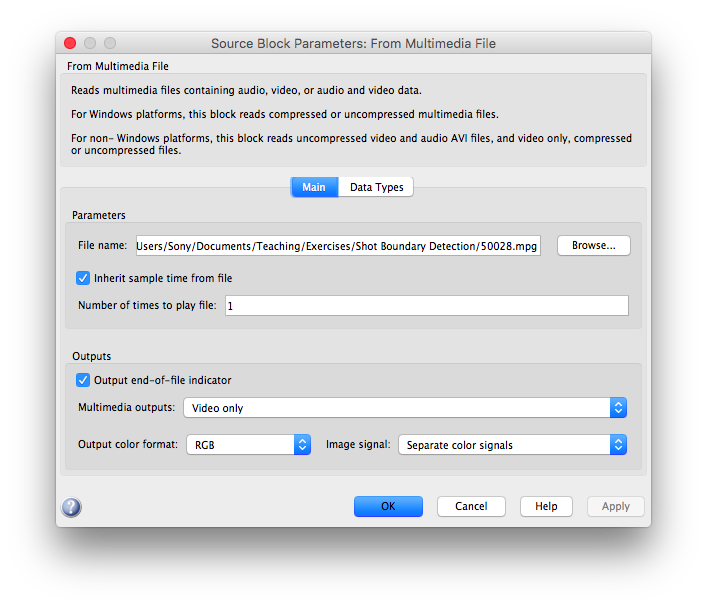
# Work Environment

We will use two answers to the question of SBD. The solutions are **sbd\_r\*.mdl/slx** (MATLAB/Simulink models available at the server) and **Virtual Dub** [2] (running in the Windows environment).

# Execution of the Exercise

The first task is the preparation of a video, which will be used for testing SBD systems. You may use files with videos; alternatively, it is possible to find video files (having numerous, easy to distinguish, shots) on the Internet. The proposes files are preferred as they are accompanied by manually created reference shot positions (**ref\_\*.mat** files, variable: manualPastFNum). Please be aware that not all video formats and codecs are handled by the programs used for the exercise. The videos should not be too long as the exercise duration is limited. In the case of long videos, it is acceptable to analyze only a part of a video. The audio track is not used; thus, also it is not needed.

You may want to **turn off the audio** in the „From Multimedia File” block:



If the reference shot positions are not available, you should oversee the video to obtain real shot boundaries. You may use the **Virtual Dub** application for this purpose. The **Virtual Dub** application displays a frame number of each currently watched video frame as well as allows for precise adjustment of playing position within a video. Please note that different programs may use numbering of frame numbers starting from 0 or 1, in case of incompatibility, you should apply the appropriate translation.

The next steps are a try of automatic SBD in the video, and then, a determination of the accuracy of SBD.

## Virtual Dub

Using the function of automatic SBD in the **Virtual Dub** application (Figure 1), please try to obtain shot boundaries.



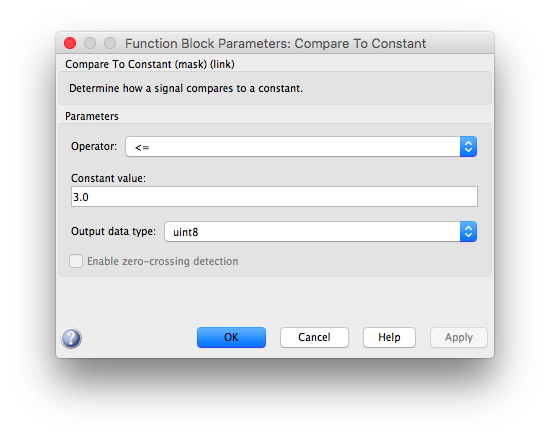
**Figure 1. Automatic SBD buttons in Virtual Dub.**

## The SBD models

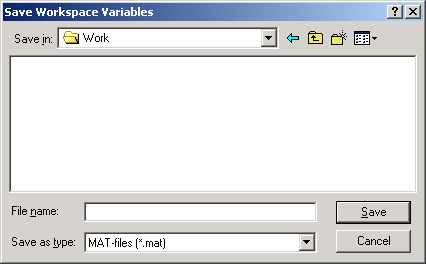
The SBD model allows for automatic SBD in video files. The model analyses luminance changes and then detects shot boundaries everywhere, where it finds significant changes. Using the model, please try to obtain shot boundaries. To see shot boundaries, having the model window focused, please click on the start button to start a simulation.

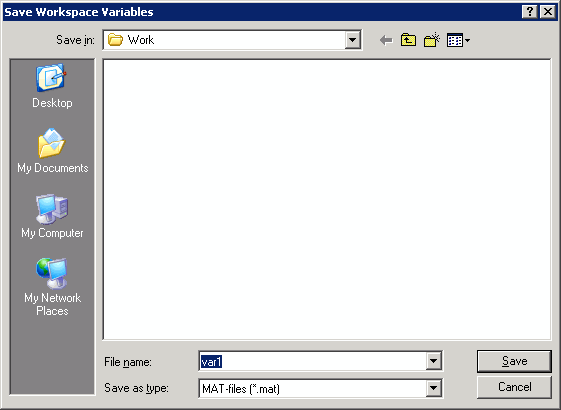
A video sequence is played out in two video viewers (current and recent frames), accompanied by a viewer showing their mutual luminance differences as well as a running scope if a shot is detected; it is shown at the area as a peak. The default sensitivity (a comparison block) works quite well, but it may be needed to adjust it so to find all shots:



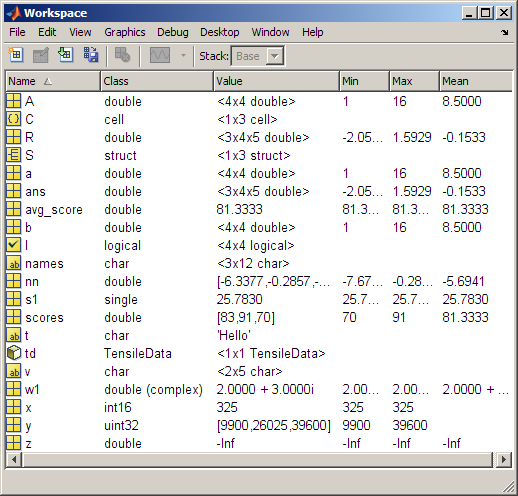


Please experiment with the sensitivity parameter, and please try to reconnect the model connections to use AE or PSNR as a comparison method. In the variable autoPostFNum, it is possible to find the shot beginnings[[1]](#footnote-1).





To open the Workspace browser, select **Desktop** > **Workspace** in the MATLAB desktop or type [workspace](http://www.mathworks.com/access/helpdesk/help/techdoc/ref/workspace.html) at the Command Window prompt:



# Next Steps

Please use methods presented in [1], [3] and [5] for determining the accuracy of SBD (in particular, please try to assess the accuracy of SBD using Precision and Recall metrics [5]). If time permits, after testing SBD for video content with easily detectable shot boundaries, please try downloading the video (or videos) where shot boundaries are not so visible. Please do the tests for these videos as well.

# Psycho-Physical Test

We kindly ask you to do the test located at <http://amis.kt.agh.edu.pl>. You can provide your student ID as the username, to leave a trace of your performance.

The purpose of this test is to find out your individual opinion on the quality of the presented video summarization. Please assess how well we prepared the resume. Please consider only summarization quality. Each outline of the video has about a minute. The test length is nine video summaries. At the end of the video summary, you are asked to rate the entire video summary. We prepared the final quality assessment on a discrete scale: bad, poor, average, good, excellent. Please select the appropriate answer and press the button to go to the next sequence. Please treat each view of the summary of the video as an independent. Video summaries cannot be undone or paused. It is not possible to display the video summary again. The total time spent on the experiment is 10 minutes, of which most of the time it will take to watch the video summaries. Please remember that this is your subjective opinion. There are no better or worse answers.

We are interested in your advice regarding the quality of summarizations!

# Report

In a report (if required – please check) please consider methods presented in [1], [3] and [5] for determining the accuracy of SBD.

# Additional Exercise

You are encouraged to write your very own Shot Boundary Detector (SBD). The task is to write it without using any dedicated libraries for this purpose (e.g., [PySceneDetect](https://pyscenedetect.readthedocs.io/en/latest/)). However, usage of supplementary libraries (like NumPy or Pandas for Python) is allowed.

The recommended method for the detection of shot boundaries is the Differential Mean Squared Error (dMSE) approach. It is not directly mentioned in the presentation accompanying this instruction, but it merely refers to calculating MSE between each two (2) consecutive video frames and then taking the derivative. Of course, you can use any other method if it guarantees accuracy as good as this one.

We recommend first to test your code before you submit it to us. The best way to do it would be to utilize the [PySceneDetect](https://pyscenedetect.readthedocs.io/en/latest/) library. You can also try to push it manually (by watching video samples and noting downtimes of shot boundaries), but the library mentioned is sufficiently accurate to do the same more effectively.

The task is worth **13** points, and a deadline is the end of the semester. We expect a source code, which we could test on our video samples. We prefer Python 3 but accept the language of your choice if Python is not your thing.

# References

1. Browne P., Smeaton A. F., Murphy N., O’Connor N., Marlow S., Berrut C.: *‘Evaluating and Combining Digital Video Shot Boundary Detection Algorithms’*, Proc. of the Fourth Irish Machine Vision and Information Processing Conference, <http://doras.dcu.ie/336/1/imvip_2000.pdf>, Belfast, North Ireland, 2000.
2. Lee A.: ‘Welcome to virtualdub.org! - virtualdub.org’, [*http://www.virtualdub.org/*](http://www.virtualdub.org/).
3. Leszczuk M., Papir Z.: ‘Accuracy vs Speed Trade-off in Detecting Shots in Video Content for Abstracting Digital Video Libraries’, *Proc. Joint International Workshop on Interactive Distributed Multimedia Systems / Protocols for Multimedia Systems IDMS-PROMS’2002*, <http://link.springer.com/chapter/10.1007/3-540-36166-9_16>, Coimbra, Portugal, 2002.
4. Leszczuk M.: ‘Sekwencyjna SBD’, <http://www.kt.agh.edu.pl/pl/node/554>, Krakow 2005.
5. ‘Information retrieval – Wikipedia, the free encyclopedia’, <http://en.wikipedia.org/wiki/Information_retrieval>
6. Au O.: ‘Visually Detectable Halftone Image Watermarking’, http://www.ee.ust.hk/~eeau/demo\_halftoneWatermark\_B.htm
7. Cole E.: ‘Hiding in Plain Sight: Steganography and the Art of Covert Communication’, Wiley Publishing, Inc., ISBN 0471444499, 2003.
8. Digimarc Corporation: ‘Digimarc Corporation’, http://www.digimarc.com/
9. ‘Directory of Books, Journals & Conferences on Digital Watermarking and Digital Watermarking Assessment Tools’, <http://knowledgebase.aegisdrm.com/knowledgebase_digital_watermarking_drm.htm>
10. Wayner P.: ‘Disappearing Cryptography, Second Edition – Information Hiding’, Morgan Kaufmann, ISBN 1558607692, 2002
11. Adjeroh D. A., Lee M. C.: „Scene-Adaptive Transform Domain Video Partitioning”, IEEE Transactions on Multimedia, 2004, vol. 6, no. 1, pp. 58-69.
12. Aghabari Z., Kaneko K., Makinouchi A.: „Content-Trajectory Approach for Searching Video Databases”, IEEE Transactions on Multimedia, 2003, vol. 5, no. 4, pp. 516-531.
13. Benini S., Xu L. Q., Leonardi R.: „Using Lateral Ranking for Motion-Based Video Shot Retrieval and Dynamic Content Characterization”, Proc. Fourth International Workshop on Content-Based Multimedia Indexing CBMI’2005, 2005, Riga, Latvia.
14. Boavida F., Monteiro E., Orvalho J.: Protocols and Systems for Interactive Distributed Multimedia, Springer-Verlag, Berlin 2002.
15. Browne P., Smeaton A. F., Murphy N., O’Connor N., Marlow S., Berrut C.: „Evaluating and Combining Digital Video Shot Boundary Detection Algorithms”, Proc. of the Fourth Irish Machine Vision and Information Processing Conference, Belfast, North Ireland, 2000.
16. Duplaga M., Juszkiewicz K., Leszczuk M., Marek M., Papir Z.: „Design of Medical Digital Video Library”, Proc. Fourth International Workshop on Content-Based Multimedia Indexing CBMI’2005, 2005, Riga, Latvia.
17. Fan J., Elmagarmid A. K., Zhu X., Aref W. G., Wu L.: „Class View: Hierarchical Video Shot Classification, Indexing, and Accessing”, IEEE Transactions on Multimedia, 2004, vol. 6, no. 1, s. 70-86.
18. Farag W. E., Abdel-Wahab H.: „A new paradigm for the analysis of MPEG compressed videos”, Journal of Network and Computer Applications, 2002, vol. 5, no. 2, s. 109-127.
19. Hanjalić A.: Content-based Analysis of Digital Video. Kluwer Academic Publishers, Boston 2004.
20. Hori C., Furui S.: „A New Approach to Automatic Speech Summarization”, IEEE Transactions on Multimedia, 2003, vol. 5, no. 3, pp. 368-378.
21. ITU-R Recommendation BT.500: Methodology for the Subjective Assessment of the Quality of Television Pictures, 06/2002.
22. Kender J. R., Yeo B. L.: „Video Scene Segmentation Via Continuous Video Coherence”, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR’1998, Santa Barbara, USA, 1998, pp. 367-377.
23. Kranz M., Tripathi S. K.: „Scene-Based Characterization of VBR MPEG-Compressed Video Traffic”, Proc. ACM Sigmetrics 97, Seattle, USA, 1997.
24. Lee H. Y., Lee, H. K., Ha Y. H.: „Spatial Color Descriptor for Image Retrieval and Video Segmentation “, IEEE Transactions on Multimedia, 2003, vol. 5, no. 3, pp. 358-367.
25. Leszczuk M.: „Usage of Bisection Method for Shot Detection in Video Content for Digital Video Library”, Proc. 11th Open European Summer School: „Networked Applications” EUNICE 2005, Colmenarejo (Madrid), Hispania, 2005, pp. 218-221.
26. Lombardo A., Morabito G., Schembra G.: „Modeling Intra media and Intermedia Relationships in Multimedia Network Analysis Through Multiple Timescale Statistics”, IEEE Transactions on Multimedia, 2004, vol. 6, no. 1, pp. 142-157.
27. Manjunath B. S., Salembier P., Sikora T.: Introduction to MPEG-7 Multimedia Content Description Interface. John Wiley & Sons, Chichester, 2002.
28. Mashat A., Kara M.: „Performance evaluation of a scene-based model for VBR MPEG traffic”, [in]: System Performance Evaluation: Methodologies and Application, CRC Press, 2000, pp. 123-142.
29. Mitchell J., Pennebaker W., Fogg C., LeGall D. J.: MPEG video compression standard, International Thomson Publishing, New York 1996, pp. 58.
30. Naci U., Hanjalić A.: „A Unified Framework for Fast and Effective Shot Transition Detection Based on Analysis of Spatiotemporal Video Data Blocks”, Proc. Fourth International Workshop on Content-Based Multimedia Indexing CBMI’2005, 2005, Riga, Latvia.
31. Po L. M., Wong K. M.: „A New Palette Histogram Similarity Measure for MPEG-7 Dominant Color Descriptor”, Proc. IEEE International Conference on Image Processing 2004 ICIP’2004, Singapore, 2004, vol. 3, pp. 1533-1536.
32. Roberts J., Mocci U., Virtamo J.: Broadband network tele-traffic. Springer-Verlag, Berlin, 1996, pp. 20-25.
33. Rose O.: „Simple and efficient models for variable bit rate MPEG video traffic”, Performance Evaluation, 1997, vol. 30, pp. 69-85.
34. Sang-Jo Y., Seong-Dae K.: „Traffic modelling and QoS prediction for MPEG-coded video services over ATM networks using scene level statistical characteristics”, Journal of High-Speed Networks, 1999, vol. 8, no. 3, pp. 211-224.
35. Skarbek W., Galiński G., Wnukowicz K.: „Tree Based Multimedia Indexing – a Survey”, Proc. Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji KKRRiT 2004, Warszawa 2004, pp. 77-85.
36. Skarbek W.: „MPEG-7”, Proc. IX Konferencja PLOUG’2003, Kościelisko, 2003, pp. 102-119.
37. Smith J. R., Chang S. F.: „Searching for Images and Videos on the World-Wide Web”, CU/CTR Technical Report 459-96-25, 1996.
38. Smith J. R., Chang S.-F.: „An Image and Video Search Engine for the World-Wide Web”, Proc. Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases V EI’1997, San Jose, USA, 1997, pp. 84-95.
39. Smith J. R.: „Digital Video Libraries and the Internet”, IEEE Communications Magazine, 1999, vol. 37, no. 1, pp. 92-97.
40. Sulkowski B., Sulkowski A.: „Application of a Monte Carlo Method to Calculate some Functions of Images”, Proc. International Conference on E-he@lth in Common Europe, Kraków 2003, pp. 439-446.
41. Taskiran C., Chen J. Y., Albion A., Torres L., Bouman C., A., Delp E. J.: „Vibe: A Compressed Video Database Structures for Active Browsing and Search”, IEEE Transactions on Multimedia, 2004, vol. 6, no. 1, pp. 103-118.
42. Wong K. M., Cheung C. H., Liu T. S., Po L. M.: „Dominant color image retrieval using Merged Histogram”, Proc. of IEEE International Symposium on Circuit and Systems 2003 ISCIT’2003, Bangkok, Thailand, 2003, vol. 2, pp. 908-911.
43. Wong K. M., Cheung C. H., Po L. M.: „Merged-Color Histogram for color image retrieval”, Proc. IEEE International Conference on Image Processing 2002 ICIP’2002, Rochester, USA, 2002, vol. 3, pp. 949-952.
44. Wong K. M., Po L. M.: „MPEG-7 Dominant Color Descriptor based relevance feedback using Merged Palette Histogram”, Proc. the IEEE International Conference on Acoustics, Speech and Signal Processing 2004 ICASSP’2004, Montreal, Canada, 2004, vol. 3, pp. 433-436.
45. Yeo B. L., Liu B.: „A unified approach to temporal segmentation of motion JPEG and MPEG compressed video”, Proc. IEEE International Conference on Multimedia Computing and Systems, 1995, pp. 81-88.
46. Yeo B. L., Liu B.: „Rapid scene analysis on compressed video”, IEEE Transactions on Circuits and Systems for Video Technology, 1995, vol. 5, no. 6, pp. 533-544.
47. Auster berry D.: The Technology of Video and Audio Streaming. Focal Press, Oxford 2002.
48. Halsall F.: Multimedia Communications Applications, Networks, Protocols and Standards, Addison-Wesley, Essex 2001.
49. Yeo B. L., Liu B.: „On the extraction of DC sequence from MPEG compressed video”, Proc. IEEE International Conference on Image Processing, vol. 2, 1995, pp. 260-263.

1. Save workspace variables to file: <http://www.mathworks.com/access/helpdesk/help/techdoc/ref/save.shtml>, <http://www.mathworks.com/access/helpdesk/help/techdoc/ref/uisave.html> [↑](#footnote-ref-1)