AFRL-IF-RS-TR-2003-190
Final Technical Report
August 2003

JBI DISTRIBUTED TEST BED (JBI-DT) VIDEO

Northrop Grumman Information Technology, Inc.

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2003-190 has been reviewed and is approved for publication.

APPROVED: /s/

JONATHAN C. GREGORY
Project Engineer

FOR THE DIRECTOR: /s/

JOSEPH CAMERA
Chief, Information & Intelligence Exploitation Division
Information Directorate

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 074-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503			
1. AGENCY USE ONLY (Leave blank)	**2. REPORT DATE** August 2003	**3. REPORT TYPE AND DATES COVERED** Final Feb 02 – Feb 03	
4. TITLE AND SUBTITLE JBI DISTRIBUTED TEST BED (JBI-DT) VIDEO			**5. FUNDING NUMBERS** C - F30602-00-D-0159/0007 PE - 62702F PR - JBIT TA - PR WU - 07
6. AUTHOR(S) Thomas Heath			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Northrop Grumman Corporation Defense Mission Systems Beeches Technical Campus Rte 26N Rome NY 13440			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFRL/IFEC 32 Brooks Road Rome NY 13441-4114			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER** AFRL-IF-RS-TR-2003-190
11. SUPPLEMENTARY NOTES AFRL Project Engineer: Jonathan Gregory, IFEC, 315-330-4294, gregoryj@rl.af.mil			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			**12b. DISTRIBUTION CODE**

13. ABSTRACT *(Maximum 200 Words)*

JBI Distributed Test Bed (JBI-DT) Video final technical report details the efforts and accomplishments of utilizing information intrinsic to the MPEG video stream to augment and enhance mosaic generation capability.

14. SUBJECT TERMS MPEG Video Mosaic Generation			**15. NUMBER OF PAGES** 9 **16. PRICE CODE**
17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

INTRODUCTION

This final technical report (FTR) will describe the efforts and accomplishments performed under contract to the Air Force Research Laboratory (AFRL), Rome Research Site (RRS). The original terms of this contract specified the incorporation of Motion Imagery Exploitation capabilities, provided by AFRL's Exploitation Toolkit for Video (XTV), into the Joint Battlespace Infosphere (JBI). However, given the relatively small dollar amount of the contract, and the potential benefits anticipated for this purpose, it was decided to concentrate on enhancing existing exploitation capabilities.

Prior Motion Imagery Exploitation work performed by Northrop Grumman Information Technology (IT) under contract to AFRL in part consisted of a method to automatically determine significant scene changes in compressed MPEG video. This contract extended that effort to add segmentation capabilities to the scene change detection efforts. To make this tool user friendly, a graphical user interface (GUI) was built to support testing and utilization of this automatic segmentation tool. Additionally, this process was loosely integrated with an AFRL developed mosaicking capability to demonstrate a proof of concept for Phase 0 image exploitation.

OVERVIEW

This paper assumes that the reader has some general knowledge of the MPEG compression scheme. However, as a high level overview, Heath, Keller and Howlett (ref 1) described how macro-block distributions within an MPEG stream can be used to detect significant scene changes within an MPEG stream. Given that there is sufficient frame to frame spatial overlap, the MPEG process obtains compression by sharing spatial information among several frames and generates more P (forward predicted) and B (backward predicted) macro-blocks (and consequently fewer Intra-coded macro-blocks per frame.

If, on the other hand, there is not sufficient frame to frame spatial overlap (as in the case of significant scene changes), the number of P and B macro-blocks per frame will significantly decrease, with a comparable increase in Intra-coded macro-blocks.

By simply counting the numbers of each type of macro-block within each frame, and looking for significant statistical frame-to-frame deviations in those counts, scene changes can be easily identified. Using this information, an automatic video segmentation tool has been developed.

The ability to automatically detect and segment a video stream at points where there are significant scene changes will allow for the creation of homogeneous video clips as opposed to clips with dramatic focal length changes, or rapid sensor skewing. Among other things such noise in the video stream can significantly hinder a video exploitation process. On the other hand, that automatic nature of this process can significantly reduce the amount of work needed to be performed by an imagery or video analyst.

PRIOR WORK

Earlier work by Northrop Grumman IT relating to this effort was primarily basic research funded by AFRL/RRS, IFEC. This research proved that stable macro-block distributions across an MPEG stream (MPEG is used in this report to refer to either MPEG-1 or MPEG-2) are highly correlated with frame-to-frame spatial consistency. Conversely, highly erratic macro-block counts across frames are highly correlated with significant scene changes within the video stream. This current effort builds upon this prior work and has resulted in a user-friendly tool allowing the user to automatically or manually segment a video stream based upon scene changes, resulting in homogeneous video clips.

It should be noted that this process is performed entirely in the compressed domain alleviating the need to decompress and spatially compare each frame in the stream.

IMPLEMENTATION

Development of the auto-segmentation tool included the development of a graphical user interface. This GUI was developed using Microsoft Visual C++, Version 6.0 and consequently will only run in a windows environment.

To begin the process, the user left clicks on the **OPEN FILE** button. This causes a typical WINDOWS dialogue box to open in which a user can select an MPEG file for analysis in the traditional way. Once the file has been selected, the segmentation process begins the process of counting all macro-blocks within each from in the MPEG file. Once all frames have been processed, a plot of the raw Intra-coded macro-block counts is displayed. If desired, the user has the option of displaying any combination of the intra-coded, P and B macro-blocks. Also, the user can selectively display the raw, median of 5 or mean of 5 macro-blocks.

Statistical information, (such as File Name, Frame Size, etc.), about the MPEG file are also displayed. These data points are fixed and the user cannot modify their display. Since these entries are fairly self-explanatory, they will not be listed here, with the exception of GOP Count and Frames/Group.

A GOP is a Group Of Pictures. This is a component of the MPEG stream. Each MPEG stream is composed of some number of GOPs. Each GOP contains some number of Frames. The number of frames in a GOP is a function of the MPEG encoder. In this case, the Frames/Group is given as 9.

Once the macro-blocks are displayed, the user has the option of automatically determining where the scene changes have occurred. This is done by selecting the **AUTOMATIC** button, also shown in Figure 1. Once this button is clicked, a series of green vertical lines will appear on the screen indicating where the process has determined scene changes have occurred in the video.

Areas of segmentation are selected by first smoothing the raw data. A median of 5 filter has been determined to provide the easiest data with which to work by providing a good smoothing filter and retaining the general local trends within the data.

A sliding window of 10 elements is then applied to the macro-block counts and significant changes in the 10 element mean are sought. It is this change in slope, either up or down, that is indicative of a significant scene change in the video stream.

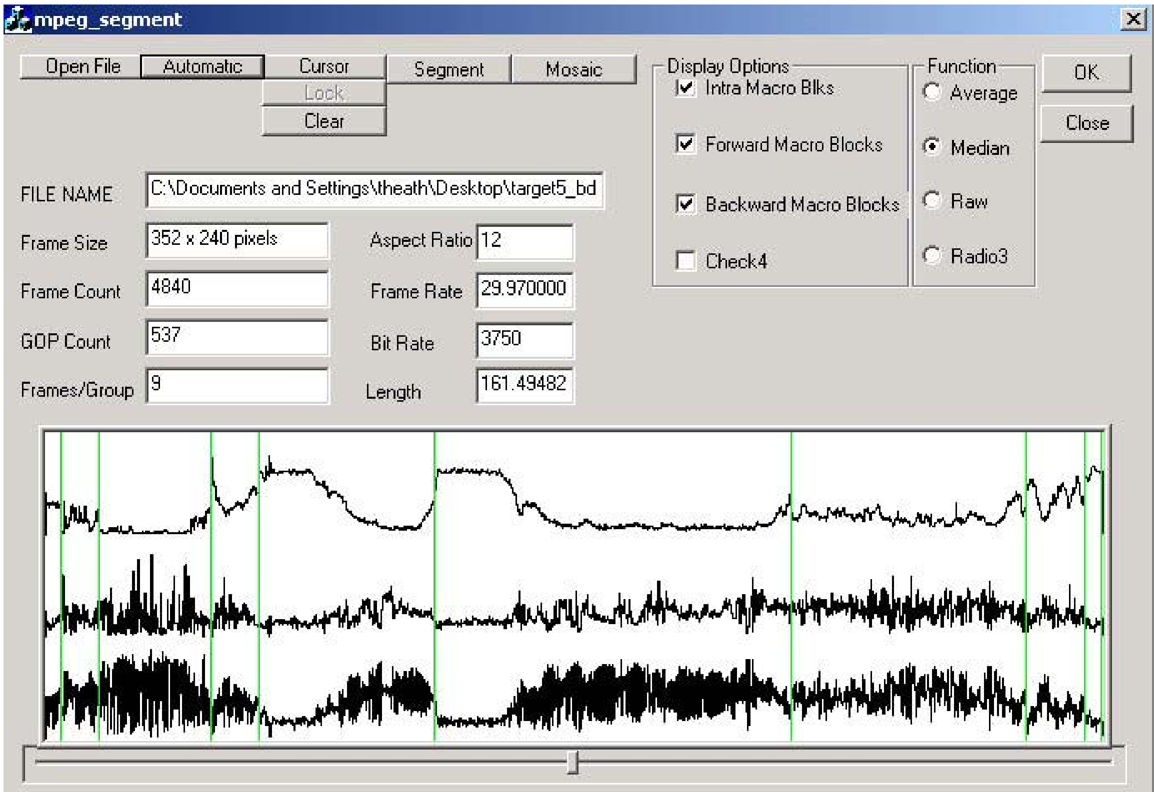

Figure1: Shows a Median of 5, Macro-Block Distribution Spanning 4840 Frames

If the user is satisfied with the automatic segmentation selections, the **SEGMENT** button can then be selected. However, the user can override the automatically selected segmentation locations by clicking on the **CLEAR** and then on the **CURSOR** buttons. This will cause the automatically selected green vertical lines to disappear, and display a single red vertical cursor. This red cursor can be positioned by clicking on the **SCROLL** button at the bottom of the display and dragging it to the desired position.

Once the cursor is moved to the desired position, the user should click on **LOCK**. This will lock the cursor into position and display a green vertical line at that point. This process should be repeated until all segmentation positions have been selected by the user.

Upon completion of the segmentation selection process, the user clicks on the **SEGMENT** button. This causes the process to parse through the selected MPEG file and output one file for each of the selected segmentation areas. For example, the segmentation areas shown in Figure 1 will generate 10 mpeg files.

Output file names are generated by using the original MPEG file name and appending the start and stop frame indices. For example, the first file output from the example displayed in Figure 1 would be named target5_bda_0_to_63.mpg. In this case, target5_bda.mpg was the original file name. Frame 0 was the first frame in this segment and frame 63 was the last frame. This process is repeated until the entire mpeg file is segmented as defined by the cursors. This process is performed in a non-destructive manner, leaving the original mpeg file intact.

If desired, the user can begin the early phase exploitation of the video. The output file names are saved in program memory, and can be passed to the AFRL mosaicking process mentioned above by clicking on the **MOSAIC** button. Upon clicking that button, all file segments are sent to the mosaicking process. All files output from the mosaic process are saved in Pix Map (PPM) format. Figure 2 shows a composite picture of 8 of the mosaics generated from segmentation/mosaicking process.

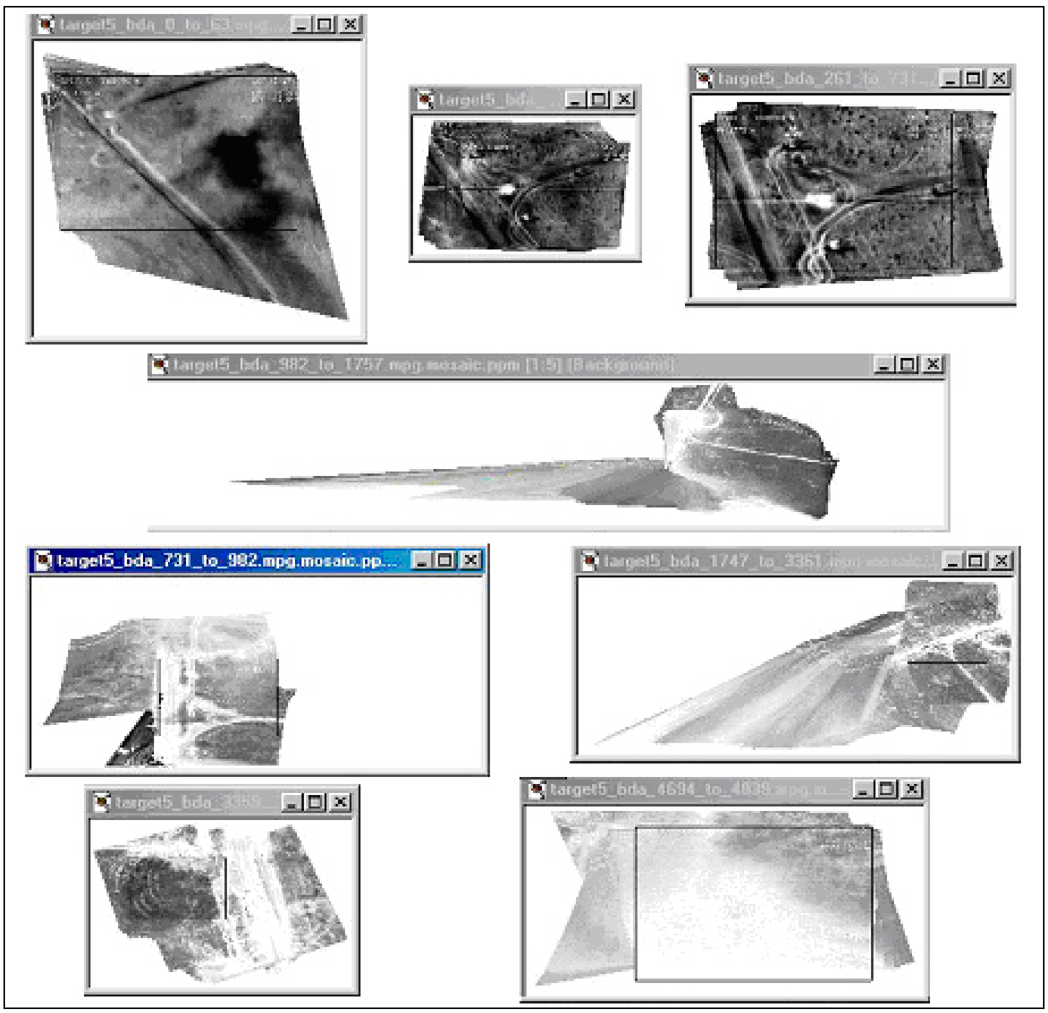

Figure 2: Shows a Composite of 8 Mosaics Generated by the Autosegmentation/Mosaic Processes

To gracefully terminate program execution, the user can simply click on either the **OK** or **CLOSE** buttons. Either of these buttons will cause all information extracted from the MPEG file to be lost, and the program removed from memory.

FINDINGS AND CONCLUSIONS

Surveillance video can be of very long duration, often longer than 24 hours, thus making effective use of this video difficult. To facilitate the exploitation of this video and reduce required manpower, tools such as automatic segmentation become necessary. This effort has shown that automatic video segmentation can be an effective tool for the purpose of partitioning video into homogeneous segments and will assist in the effective exploitation of aerial surveillance video.

Integration of the automatic segmentation tool with the mosaicking capability provides the exploiter with a completely hands-free analysis tool which will allow the user to quickly view the mosaics and identify which video clips may be of interest, and therefore provide a more robust exploitation tool.

This effort has proven that with a relatively small investment, a user-friendly, fairly robust aid to motion imagery exploitation can be developed. This capability should be beneficial to both military and commercial applications.

REFERENCES

Prior efforts of others that were of help in developing this algorithm follow.

1. Thomas Heath, John Keller and Todd Howlett, “Automatic Video Segmentation In The Compressed Domain”, SPIE- 2001, A.

2. Thomas Heath, Mark Robertson, John Keller, Todd Howlett, “Segmentation of MPEG-2 Motion Imagery Within The Compressed Domain,” IEEE-2002.

3. V. Kobla, D. Doermann, and A. Rosenfeld, “Compressed Domain Video Segmentation.”

4. ISO/IEC 11172-2, “MPEG-1-Information Technology–Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s, Part 2, Video,” 1993.

5. ISO/IEC 13818-2, “MPEG-2–Information Technology, Generic Coding of Moving Pictures and Associated Audio, Part 2: Video,” 1996.