UNT TRECvid: A Brighton Image Searcher Application [1]

Mark Rorvig [2], Ki-Tai Jeong, Anup Pachlag,
Ramprasad Anusuri, Diane Jenkins, Sara Oyarce
University of North Texas
Texas Center for Digital Knowledge

Abstract

The results, architecture and processing steps used by the University of North Texas team in the 2001 TRECvid video retrieval trials are described. Only a limited number of questions were selected by the team due to resource limitations described in the paper. However, the average precision of the team results over thirteen questions from the General search category were reasonable at 0.59.

1 Background

The Brighton Image Searcher was constructed following the 3rd International Conference on the Challenge of Image Retrieval sponsored by the Institute for Image Data Research of the University of Northumbria in Brighton, United Kingdom (Rorvig et al., 2000; Goodrum et al, 2001). Details of the system architecture and performance are documented in two papers presented in 2001 at the 10th World Wide Web Conference in Hong Kong, Special Administrative Region (Jeong et al., 2001). The system may be used at <http://archive4.lis.unt.edu/tdt/www/> for a collection of still images from the NASA Hubble Space Telescope Repair Mission.

This system is intended for use in the retrieval of still images. However, it was adapted for the TRECvid task by sampling salient, or key frames from the video segments comprising the NIST TRECvid test collection. The key frame extraction algorithm was developed at NASA and is described in Rorvig (1993). This algorithm is presently used to summarize International Space Station video downlink at the Johnson Space Center in Houston, Texas, USA.

2 Procedures

TRECvid Test Collection files were processed at equal intervals of five seconds to extract key frame candidates. The key frame candidates were then processed to extract a number of measures corresponding to primitive image features first proposed in 1980 by Marr in his posthumously published work *Vision*. These measures are typically rendered

as histograms but are treated in our system as Lorenz Information Measures This technique is more fully described in Jeong, et al. (2001). The image processing was performed using the relatively new Java Image Processing Libraries.

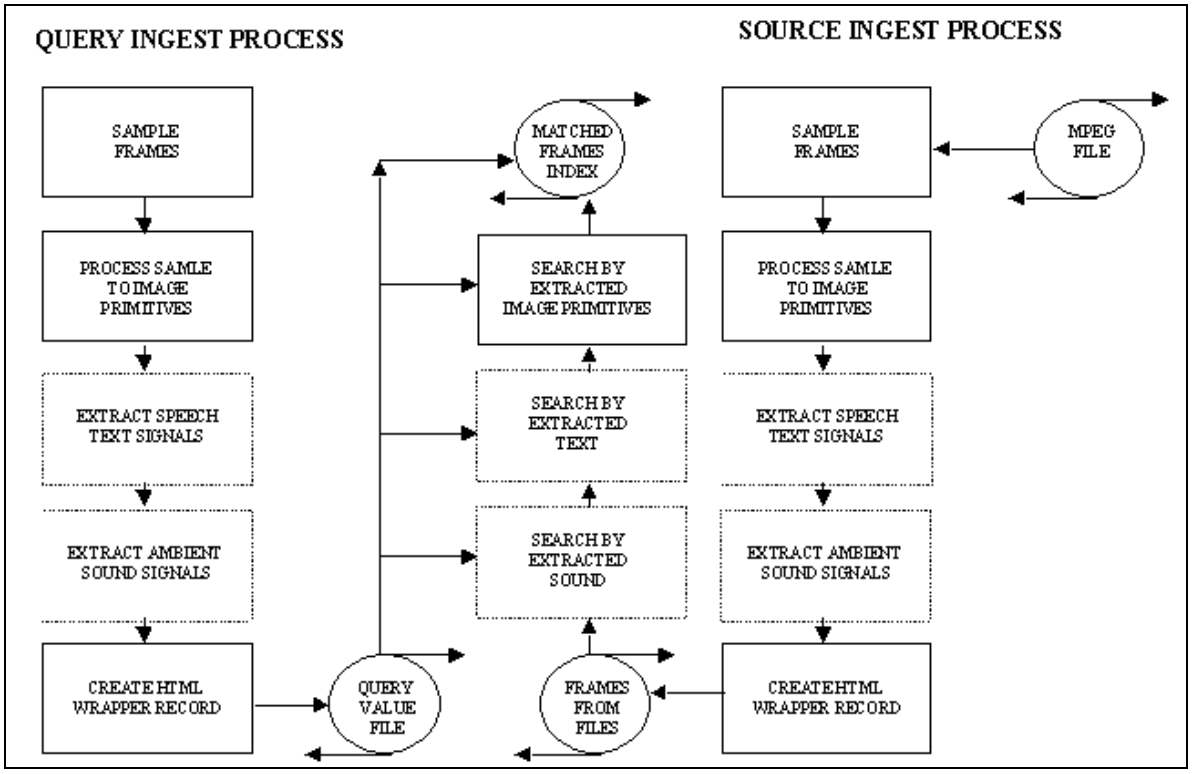

Illustration 1: System layout of the components of the Brighton Image Searcher adapted for use for the TRECvid trials. The dashed squares represent system functionality under construction and unavailable at the time of the TRECvid submission.

In future versions of this system, a query preprocessor will be built using blocks of this software to extract query representations from moving image clips. The extracted key frame measurements would then be matched by city block metric to key frames in the batch extracted collection. For this trial, however, only those test questions that provided either a known item or a search example that had already been processed in the key frame selection procedure were addressed. Further, use of the test question set even from this subset was limited to those examples or known item segments for which a key frame had been selected. (Salient or key frames are selected based on the frequency of the appearance of their features in the moving image document, and some examples were simply too short to have produced any frames in the prior batch process.)

Two of us (D.J. and S.O.) performed the searches by selecting a key frame from the collection that appeared in the example or known item intervals and then using that frame as a search exemplar for the Brighton Image Searcher. The key frame selection was based on the searcher's assessment that it was coextensive with the question semantics. In most cases the number of frames available for selection was limited,

however, precision scores better than expected may have been due to this introduction of human judgment into the search process lending the final results an "interactive" status.

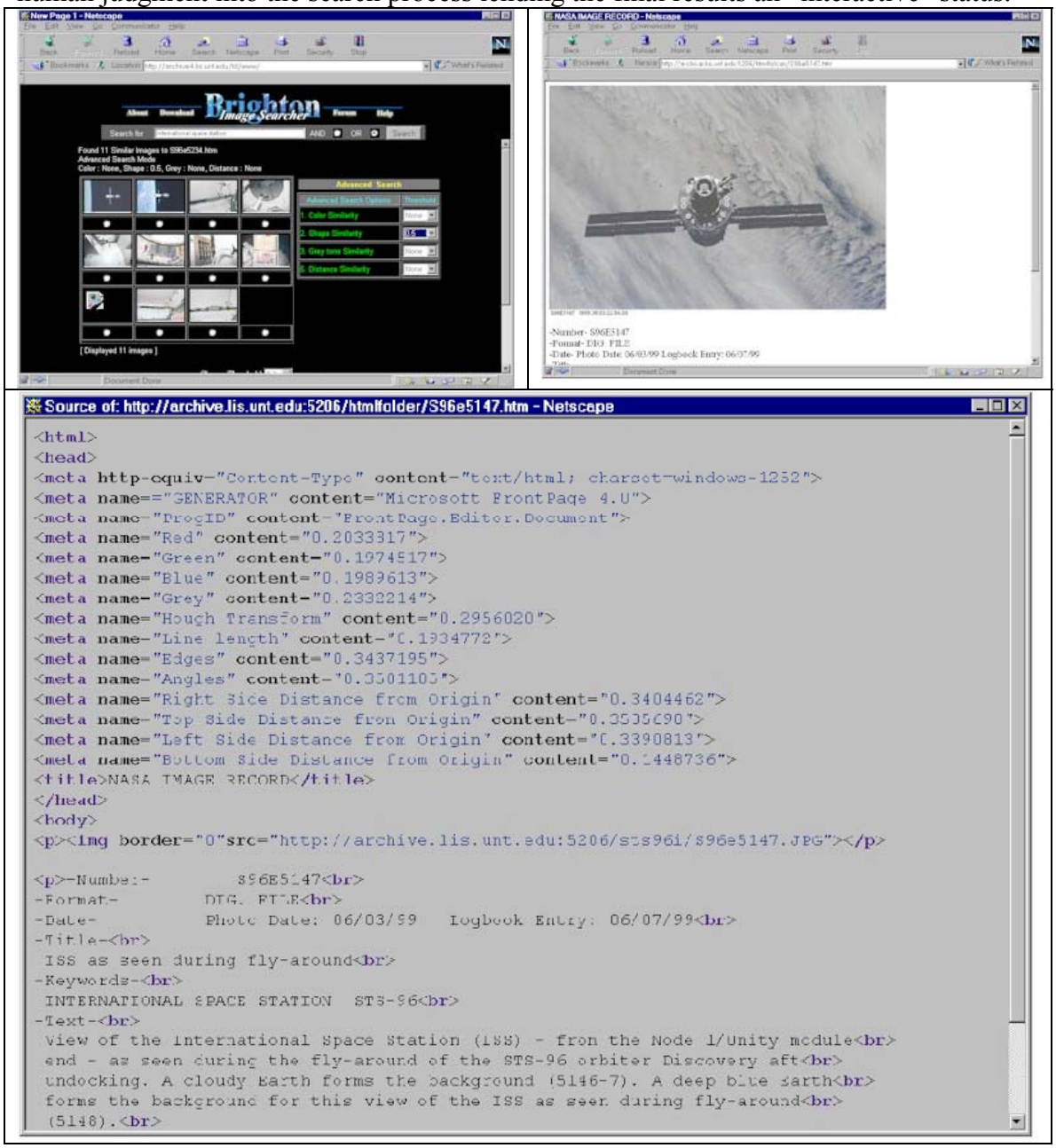

Illustration 2: The three frames above record (a-upper left) a retrieval session page for the image searching system, (b-upper right) a single image chosen by double clicking on an image from (a), and (c-below) the metadata tags for the image represented by the twelve measures used in this system.

In response to an image selected as an exemplar for searching, the Brighton Image Searcher can potentially retrieve all the images in the collection ranked by the order of metric correspondence to the features of the candidate images. However, for this trial,

only the images retrieved from the first twelve matches in the system were included in the results.

3 Results

For the thirteen questions addressed by our team, the overall precision score was 0.59 with extremes ranging from 0.92 to 0.00. The table below records these scores from the TRECvid evaluation sponsors.

TRECVID QUESTION	UNT SYSTEM DESIGNATION	RELEVANT SEGMENTS	UNT RLEVANT	UNT SELECTED	PRECISION SCORE
002	Sys21	7	4	9	0.78
024	Sys21	1	1	2	0.50
041	Sys21	9	8	12	0.75
043	Sys21	2	2	5	0.40
048	Sys21	3	0	4	0.75
049	Sys21	3	2	5	0.60
050	Sys21	11	9	12	0.92
051	Sys21	6	5	7	0.86
052	Sys21	0	0	2	0.00
055	Sys21	0	0	1	0.00
056	Sys21	1	1	2	0.50
057	Sys21	1	0	2	0.50
059	Sys21	1	1	4	0.25
061	Sys21	8	7	9	0.89

Table 1: Average precision over 13 questions for the UNT system was 0.59.

4 References

Goodrum, A., Rorvig, M., Jeong, K., Suresh, C. (2001) "An Open Source Agenda for Research Linking Text and Image Content Features," *Journal of the American Society for Information Science*, [in press].

Jeong, K.T., Rorvig, M., Jeon, J., Weng, N. (2001) "Image Retrieval by Content Measure Metadata Coding," *10th International World Wide Web Conference, May 1-5, 2001 Hong Kong, International World Wide Web Conference Committee:*
<http://www.iw3c2.org/Conferences/Welcome.html>,
<http://www.www10.org.hk/cdrom/posters/p1142/index.htm>.

Rorvig, M. E., "A Method for Automatically Abstracting Visual Documents," *Journal of the American Society for Information Science*, 44(1):040-056, 1993.

Rorvig, M., Jeong, K., Suresh, C., Goodrum, A. (2000) "Exploiting Image Primitives for Effective Retrieval," *CIR2000 - Third UK Conference on Image Retrieval, 4 - 5 May 2000, Old Ship Hotel, Brighton, United Kingdom* (J.P. Eakins, P.G.B. Enser, Eds.), [unnumbered manuscript] 7pps., University of Brighton: School of Information Management.

Notes

(1) This work was supported by Intel Corporation, the Special Libraries Association, and the National Aeronautics and Space Administration.

(2) All correspondence should be addressed to Mark Rorvig, Ph.D., UNT/SLIS/TxCDK, P.O. Box 311068, Denton, Texas, USA 76203, Tel: 940-300-5344, Fax: 940-565-3101, <mrorvig@unt.edu>.