

Improving Image to Video Matching to Support Entity Resolution with Motion Detection and Feature Extraction

H. S. Senevirathna

149232V

Faculty of Information Technology

University of Moratuwa

December 2017

Improving Image to Video Matching to Support Entity Resolution with Motion Detection and Feature Extraction

H. S. Senevirathna

149232V

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa,
Sri Lanka for the partial fulfilment of the requirements of the MSc in Information
Technology

December 2017

Declaration

We declare that this thesis is our own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student (s)

Signature of Student (s)

Date:

Supervised by

Name of Supervisor(s)

Signature of Supervisor(s)

Date:

Dedication

To my parents.

Acknowledgements

From the numerous people who helped me in this project, I must first thank my supervisor Mr. Saminda Premarathne who suggested this research area to me and guided me throughout with his advice, encouragement, expertise and wisdom. Then I must thank my parents, husband and parents-in-law for encouraging me to extend effort and for supporting me through the numerous difficulties I faced during the course of the project, with their wisdom. Last, but not the least I must thank my fellow students, the course coordinator Mr. Sudantha B. H. and coordinating assistant Achala Subhashini for their cooperation to make this project a reality.

Abstract

The need for image based video search is increasing rapidly as today with the expansion of big data and the increasing power of hardware. But there are only a few highly successful implementations in existence. In this project I have developed a search method combining motion detection and Different Feature Detection algorithms, then evaluated the method's effectiveness and compared the two approaches of Real-time Video Search and searching against a Database of feature data taken from videos. Key frames of videos are extracted using motion detection, by the difference of consecutive key frames and the Otsu's threshold. Speeded Up Robust Features (SURF), Harris-Stephens corners with Fast Retina Keypoint (FREAK) descriptor and color features are the feature detection/description methods used for extracting features. The features extracted from key frames are matched with those of the given image and M-estimator SAmple and Consensus (MSAC) algorithm is used to find 'Affine transformations' from the matching points. Different thresholds are taken by combining the feature extraction methods for filtering the results. Two prototypes were produced for comparing searching normally and searching against a database of features. Images of cats are being used to search videos where, some of which have cats throughout, some which have intermediate intervals of cats and while others have no cats. After evaluating against sets of images of incrementing size, the search method produced an intermediate level accuracy (48.89%) of search results. Furthermore, comparing the two prototypes for 5 images and 3 sets of videos, the stored summary prototype is seen slower than the real-time video search, and a trivial difference in result statistics is found.

Table of Contents

Dedication	iii
Acknowledgements	iv
Abstract	v
Table of Contents	vi
List of Figures	vii
List of Tables	viii
Introduction	1
Developments and Challenges in Image Matching, Image Based Video Search and Video Segmentation	4
Technologies Used For Improving Image to Video Matching	15
Method of Employing Motion Detection and Feature Matching in the Prototypes	21
Designs of the Prototypes	25
Implementation	30
Evaluation	39
Conclusion and Further Enhancements	48
References	52
Appendix A - Approach	54
Appendix B - Design	55
Appendix C - Implementation	57
Appendix D - Evaluation	71
Glossary of Terms	76

List of Figures

Figure 2.1 - Architecture of an early CBVR	13
Figure 3.1 - Video components	15
Figure 3.2 - Video segmentation hierarchy	16
Figure 4.1 - Level 1 DFD of Real-time searching prototype	22
Figure 4.2 - Level 1 DFD of Stored summary searching prototype	23
Figure 5.1 - Architecture of Real-time searching prototype	25
Figure 5.2 – Activity diagram of Real-time searching prototype	26
Figure 5.3 - Architecture of Stored summary searching prototype	27
Figure 5.4 - ER diagram	28
Figure 6.1 - Screenshot of the database design (relationships)	31
Figure 6.2 - List of code components	32
Figure 6.3 - Screenshot of execution of the run-time prototype	37
Figure 7.1 - Result of searching more than 90% appearance videos	43
Figure 7.2 - Result of searching Medium maching videos	44
Figure 7.3 - Result of searching 0% maching videos	44
Figure 7.4 - Total result of searching all maching videos	45
Figure 7.5 - Total result for video sets	45
Figure 7.6 - Comparison between prototypes' runing times	46
Figure 7.7 - Comparison between prototypes' percentages of results	46
Figure B.1 – Activity diagram of Stored summary searching prototype search senario	55
Figure B.2 - Stored summary searching prototype's Add Video scenario	56

List of Tables

Table 2.1 - Comparison of researches	12
Table 6.1 - Detailed data dictionary	32
Table 7.1 - Matching results for sample images with all video sets using real-time prototype	41
Table 7.2 - Matching results for different sets using real-time prototype	42
Table 7.3 - Comparison of search times for different sets using both prototypes	42
Table 7.4 - Comparison of matching video percentages for different video sets using both prototypes	43
Table A.1 - DFD data dictionary	54
Table D.1 - Evaluation details of prototype comparison	72
Table D.2 – Evaluation details of video results for different image sets	74
Table D.3 – Evaluation details of video results for different video sets	75