SOME STUDIES ON AUGMENTED PAPER SYSTEMS

KAVITA BHARDWAJ

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

OCTOBER 2014

SOME STUDIES ON AUGMENTED PAPER SYSTEMS

by
KAVITA BHARDWAJ

Department of
Electrical Engineering

Submitted

in fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JANUARY, 2014

Certificate

This is to certify that the thesis titled "**Some Studies on Augmented Paper Systems**" being submitted by Kavita Bhardwaj to the Department of Electrical Engineering, Indian Institute of Technology, Delhi, for the award of the degree of Doctor of Philosophy, is a record of bona-fide research work carried out by her under guidance and supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted to any other university or institute for the award of the degree or diploma.

Prof. S. Chaudhury
Professor
Department of
Electrical Engineering
Indian Institute of Technology Delhi
New Delhi - 110 016

Dr. S. Dutta Roy
Associate Professor
Department of
Electrical Engineering
Indian Institute of Technology Delhi
New Delhi - 110 016

Acknowledgements

I would like to express my deepest regards and sincere gratitude to my PhD supervisors Professor Santanu Chaudhury and Dr. Sumantra Dutta Roy, for supporting me throughout my research. Santanu Sir's vision and guidance to pursue quality research in my field of interest is the prime reason that my thesis has seen the light of day. During my years of working with him, I have observed and learned how hard-work combined with diligence leads to innovation based on creative ideas. He always inspired me to strive for excellence and put in my hundred percent to the job at hand. Santanu Sir, I really appreciate you for what you are. Whatever I write about you that seems very less. I only say that you are a visionary and talented with brain, intelligence and your presentation is wonderful. Sumantra Sir has been supportive and has given me the freedom to pursue various projects without objection. He has also provided insightful discussions and suggestions about the research. I really thank him for the last time push-up in helping and submitting the thesis. His presentation skills and writting skills are excellent. I hope, I could learn and also can work energetically like him.

I am grateful to both of my guides for their forbearance with my ignorance and for being patient with my struggles.

I am grateful to the members of SRC – Prof. Prem Kalra, Prof. S. D. Joshi and Prof. Brejesh Lall for their valuable suggestions and advice regarding the directions for my work. I remember my struggle with returning to academics and the travails

associated with achieving its milestones during my months of course-work. I have deep gratitude for all the faculty in Electrical Engineering, Computer Science and the Mathematics departments at IIT Delhi, who were instrumental in adding value to my learning in the areas important to my research.

I have had a long association with the Multimedia Lab and the administrative staff here, and have special regard for Mrs. Usha Bhola, who ensured that the equipment required for my work was always in place. I will remember Sushila ji, for her help. Mrs. Narender Kaur in the Electrical Engineering Department office has always helped in whatever way possible and I will remember her pleasing attitude.

At the end, I acknowledge the most important contribution to these long years of research work, which is of my family. They have stood by me patiently through all my efforts and given me complete leeway to pursue my doctorate even if it meant sometimes giving up my family responsibilities. I express my thanks to my close friends – Sonam Bhardwaj, Nisha Pahal, Deepti Wadhwa, Vandana Bhalla, Geetanjali Soni, Himanshu Govil and Anukriti Bansal who have waited forever for my perpetually busy status to end.

Kavita Bhardwaj
New Delhi

Abstract

Augmented Paper Systems involve a more natural interaction with both physical paper, coupled with the advantages of working with a networked digital medium. The thesis proposes an architecture for such a requirement. The user works on an input panel on a table-top where there are physical documents which can be seen by a camera, as well as a projector whose output is on the same table. A desirable requirement is for a user to have a personalised work-space. Natural interactions with such a mixed-media system typically involves the following important sub-tasks.

A user should be able to retrieve a document from a database by showing to the camera an image of possibly, a part of the document in question. The illumination in the same could be far from proper, there could be skew and other deformations in the exemplar being shown to the camera, plus there could be occlusions from the hand holding the document, or from other objects as well. The thesis proposes a robust strategy to fuse possibly noisy information from more than one source of measurement, for robust matching, and a geometric hashing-based strategy to retrieve the matching document. We have experimented with simple script-independent features for the same.

Identification of the script itself is an important problem for two reasons. First, to extract text from a given document needs an OCR (Optical Character Recognition engine), which works best when it is told the script of the document in question.

Second, a general search procedure could be simplified if a database organises documents according to the script: this prunes out a large part of the search space. We propose an Empirical Mode Decomposition (EMD)-based method to learn characteristics of a given script, and use it for automatic script identification. This is particularly important for the wide plethora of Indian scripts which have diversely differing characteristics.

A user working with a technical paper often desires a particular reading order (for different parts of the paper), which is often different from the order in which topics are presented in a document. A user may need to search the Internet for other related papers with the same keywords, or other publications of the same author, or simply search for a particular reference mentioned in the paper. For a structured document, we propose a hierarchical Conditional Random Field (CRF)-based strategy to learn the structure of different logical components of a research paper (which could in themselves, be quite different e.g., one column, two column, and even among these, the IEEE format, the Elsevier format and so on), and use this to automatically tag these entities in a given document page. The final contribution of the thesis is an overall architecture for a seamless paper-digital world interaction, with a camera-projector system, as mentioned initially. The system uses this information in conjunction with an Internet browser to search for the required entity, as mentioned above. The system enables any surface to be used as a simple low-cost touch-pad, with equivalent capabilities. The system can be used for annotating a document, or a learning a new language, based on automatic suggestions corresponding to a word for instance, and the system giving the relevant information either from itself, or from the Internet.

Contents