PRAMOD SANKAR K.

Centre for Visual Information Technology
International Institute of Information Technology
Gachibowli, Hyderabad – 500032
Andhra Pradesh, INDIA.

+91 – 9704529444
pramod.kompalli@gmail.com
http://researchweb.iiit.ac.in/~pramod_sankar/

Research Interests

Computer Vision, Machine Learning, Information Retrieval, Image Processing

Education

Ph.D. in Computer Science, IIIT-Hyderabad (April 2004 – Present)
Thesis Title: *"Multimedia Retrieval through Automatic Alignment and Annotation"*
Thesis Advisor: Prof. C. V. Jawahar

B.Tech from VJIT (Affiliated to JNTU, Hyderabad) (May 1999 – April 2003)

Schooling at Padma Seshadri Bala Bhavan Junior College, Chennai and
V. P. Siddhartha Public School, Vijayawada

Work Experience

Organisation: University of Oxford, UK, Robotics Research Group
Designation: Visiting Researcher
Reporting to: Prof. Andrew Zisserman
Duration: Two 3-Month visits (2006 and 2008)
Role:
- Coordinated the TRECVID team (8 members) at Oxford and IIIT-Hyderabad.
- Participated in TRECVID challenge organised by NIST, USA.
- Research work related to Video Retrieval from TV shows and Movies.
- Ambassador of the UKIERI exchange program between UK and India.

Organisation: Centre for Visual Information Technology (CVIT), IIIT-Hyderabad, India.
Designation: Research Assistant
Reporting to: Prof. C. V. Jawahar
Duration: 6 Years
Role:
- Key role in the Digital Library of India project, aiming at digitisation of 1 Million books. Provided software for quality verification, and smooth functioning of the digitisation pipeline. Organised workshops for dissemination of standards and software.
- Contributed skew correction software to the OCR Consortium project (a DST project involving all IITs, IIITs, IISc, ISI, etc).
- Numerous presentations on the work at CVIT, to visiting dignitaries, heads of industrial organisations and faculty from other institutes
- Mentor for the News Video Retrieval project, coordinated three teams of 6 members each to build the fastest vision based news retrieval system.
- Co-advising the thesis of three M.S. by Research students along with Prof. C. V. Jawahar.

Organisation: IIIT-Hyderabad, India.
Designation: Teaching Assistant for Machine Learning
Reporting to: Prof. C. V. Jawahar
Duration: 3-Months (Monsoon 2007)

Role:
- Organising lectures.
- Setting and evaluating assignments and exams.
- Mentor and evaluator for the Semi-Supervised Learning Workshop, held as part of the course.

Organisation: Pentagram Research Centre, Hyderabad, India.
Designation: Research Engineer
Reporting to: Prof. E. G. Rajan
Duration: Dec 2002 – April 2004
Role:
- Head designer and developer for an image processing package (Logical Image Processing System –LIPS)
- Software programmer for Logical Pattern Generation System (LPGS)
- Built prototype software for "Fake Currency Detection using Optical Imaging"
- Gave talks at IDRBT, Gandhi Medical Hospital, NSRDC, etc. for marketing the developed software.
- Coordinated a team of 15 members towards organising the International Conference on Systemics, Cybernetics, Informatics (ICSCI), held in February 2004.

Publications

Large Scale Document Image Retrieval by Reverse Annotation
Pramod Sankar K., R. Manmatha, C. V. Jawahar
(under review) IEEE Transactions on Pattern Analysis and Machine Intelligence

An Indexing Approach for Speeding-Up Image Classification
Rahul Jain, Sudha Praveen, Pramod Sankar K., C. V. Jawahar
In Proceedings of 7th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), Chennai, India, 2010.

Reverse Annotation Based Retrieval from Large Document Image Collections
Pramod Sankar K.
SIGIR Doctoral Consortium 2010

Nearest Neighbor based Collection OCR
Pramod Sankar K., C. V. Jawahar, and R. Manmatha
In Proceedings of Document Analysis Systems (DAS), Boston, USA, 2010

Subtitle-free Movie to Script Alignment
Pramod Sankar K., C. V. Jawahar, and Andrew Zisserman
In Proceedings of British Machine Visions Conference (BMVC), London, UK, 2009

Oxford TRECVID 2008 – Notebook paper
James Philbin, Manuel Marin-Jimenez, Siddharth Srinivasan and Andrew Zisserman
Mihir Jain, Sreekanth Vempati, Pramod Sankar and C. V. Jawahar,
In Proceedings of NIST TRECVID 2008

Probabilistic Reverse Annotation for Large Scale Image Retrieval
Pramod Sankar K and C. V. Jawahar
In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, MN, USA, 2007

Text Driven Temporal Segmentation of Cricket Videos.
Pramod Sankar K., Saurabh Pandey and C.V.Jawahar
In Proceedings of 5th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), Madurai, India, LNCS 4338 pp.433-444, 2006.

Enabling Search over Large Collections of Telugu Document Images - An Automatic Annotation Based Approach.
Pramod Sankar K. and C.V.Jawahar
In Proceedings of 5th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), Madurai, India, LNCS 4338 pp.837-848, 2006.

Annotation of Images and Videos based on Textual Content without OCR.
Pramod Sankar K., Million Meshesha, C. V. Jawahar,
Workshop on Computation Intensive Methods for Computer Vision (in conjunction with ECCV 2006), Graz, Austria, 2006

Digitizing a Million Books: Challenges for Document Analysis.
Pramod Sankar K., Vamshi Ambati, Lakshmi Pratha, C. V. Jawahar,
In Proceedings of 7th IAPR Workshop on Document Analysis Systems (DAS), 2006, Nelson, New Zealand, LNCS Springer-Verlag, 425-436

Quality Management in Digital Libraries.
Vamshi Ambati, Pramod Sankar K., Lakshmi Pratha, C. V. Jawahar,
In Proceedings of the 1st International Conference on Universal Digital Library (ICUDL), Hangzhou, P. R. China, Zhejiang University Press, 2005, pp. 314-321

Content Development for Digital Library of India: Regional Mega Scanning Centre Hyderabad
Lakshmi Pratha, Vamshi Ambati, Pramod Sankar
In Digital Libraries Special issue, VishwaBharati Magazine, TDIL, Government of India. (http://www.tdil.mit.gov.in/apr_2005.htm)

Hexagonal Pixel Grid Modeling and Processing of Digital Images Using CLAP Algorithms.
K. Pramod Sankar, T. Sanjay, E.G.Rajan,
In Proceedings of the International Conference on Systemics, Cybernetics and Informatics, ICSCI-2004, Hyderabad, Vol.1, pp. 327-336.

Cellular Logic Array Processing of Digital Images on a Hexagonal Grid
K. Pramod Sankar, E. G. Rajan
Book Chapter. Symbolic Computing - Signal and Image Processing,
B. S. Publications, India, 2003.

Geometric Filters over Hexagonal Pixel Grids and Their Use in Machine Vision
K. Pramod Sankar
Undergraduate Thesis, submitted to JNTU-Hyderabad

Awards

- Travel grant from SIGIR to attend SIGIR-2010 and present at the Doctoral Consortium
- Bursary for attending BMVC, sponsored by Microsoft
- Best Poster award at ICVGIP-2006, for the paper titled, "Text Driven Temporal Segmentation of Cricket Videos"
- Adjudged among Top-5 demos for "Cricket Video Retrieval System", at MSR Techvista-2007
- Awarded Gold Medal at VJIT for highest overall percentage in B. Tech (CS&E)

Course Work (at IIIT-H)

Core courses: Pattern Recognition, Computer Vision, Machine Learning, Computer Graphics
Breadth courses: Algorithms, Operating Systems, Probability and Random Processes, Advanced Computer Architecture

Software Skills

Operating Systems:	Windows, Linux
Programming Languages:	C, C++, Matlab, Java
Programming Tools & APIs:	VC++, VB, QT Designer, OpenGL
Others:	HTML, MySQL, Gimp, LaTeX, Microsoft Office

Thesis Overview

Title: *"Multimedia Retrieval through Automatic Alignment and Annotation"*
Focus: Building text-retrieval systems for large multimedia collections, hitherto un-searchable.
Challenge: i) Visual recognition is a hard problem to solve, ii) variety and scale of data is huge.
Novelty: Utilise the availability of parallel text to annotate multimedia.

<u>**Subtitle-free Movie to Script Alignment**</u>

Problem: Alignment of videos with their (text) scripts.
Challenge: Matching semantic textual information with raw visual information.
Solution: A multiple assignment problem that assigns multiple sentences of the script to video-shots. Sentences and shots are matched based on three cues: <Location, Person, Dialogues>.
Achievement: One can now search from episodes of Seinfeld, Charlie Chaplin and Indian films. User can get video results for queries such as "Kramer's entrance", "Charlie Chaplin making shoe-soup", "Utpal Datt standing up", etc.

<u>**Text-driven Temporal Segmentation of Cricket Videos**</u>

Problem: Temporal segmentation of Cricket-match videos into individual events + Annotation with commentary.
Challenge: Simultaneous segmentation and annotation is a chicken-and-egg problem. If one is available the other can be estimated to a certain extent; estimating them together is the challenge.
Solution: The idea is that the scene categories could be differentiated from each other. Each scene category is modelled using visual-temporal descriptors. A *video model* is generated using information from text-commentary. This model is now aligned with the observed data from the video. Once aligned, we could segment and annotate in one step.
Achievement: Retreival enabled on matches from one entire tournament. Users can retrieve all videos where Sachin Tendulkar hits a Four, Harbhajan Singh claims a wicket, etc. Users can also generate their own *personalised highlights,* by specifying their scenes of interest.

<u>**Large-Scale Annotation of Document Images**</u>

Problem: Annotation of scanned document images in Indian scripts, with their text-equivalent.

Challenge: i) OCRs are not robust for non-Latin scripts, ii) heavy degradations in document-images and iii) large scale of data to be processed

Solution: Firstly, pose the problem as one of word-recognition, instead of character recognition; this eliminates problem with character segmentation and classification. Secondly, make the word-recognition step computationally efficient using carefully engineered indexing schemes.

Achievement: 1000 Telugu books were made searchable, consisting of 120K pages and 36M words. Search results are delivered instantly. Annotation performance defines the state-of-the-art.

Professional Activities

- Reviewer for CVPR, SIGIR, DAS, ICDAR, IJDAR, ICVGIP, COIN, etc.
- Student Volunteer for Indo-Israeli Workshop on Computer Vision (WCV), International Joint Conference on Artificial Intelligence (IJCAI), Asian Conference on Computer Vision (ACCV) and the Workshop on Computer Vision, Graphics, and Image Processing (WCVGIP).
- Was the *"Conference Secretary"* of the International Conference on Systemics, Cybernetics and Informatics, ICSCI-2004. Was responsible for much of the organizational effort over a six-month span. Coordinated a team of 20 volunteers, before and during the event. MC for the conference.
- Student member of IEEE
- Volunteer for "Lend Your Voice" campaign of RadioMirchi 98.3FM, to convert text-books to audio-books for the visually challenged.

References

Prof. C. V. Jawahar
Centre for Visual Information Technology
International Institute of Information Technology – Hyderabad.
email: jawahar@iiit.ac.in

Dr. R. Manmatha
Centre for Intelligent Information Retrieval
University of Massachusetts, Amherst.
email: manmatha@cs.umass.edu

Prof. Andrew Zisserman
Visual Geometry Group,
University of Oxford.
email: az@robots.ox.ac.uk