

3:05 20

CINF Symposia

ACS Chemical Information Division (CINF)
Fall, 2011 ACS National Meeting
Denver, CO (August 28 - September 1)

MONDAY AFTERNOON

Section A
Colorado Convention Center, DENVER

Computer Vision based chemical information extraction from digital images and streaming videos

Muthukumarasamy Karthikeyan*,
m.karthikeyan@ncl.res.in.
Digital Information Res Centre (Chemoinformatics Division)
& Centre of Excellence in Scientific Computing,
National Chemical Laboratory (CSIR),
Pune Maharashtra 411 008, India

Chemical structures are the perfect examples to test the power of robotic vision to recognize and translate them into truly three dimensional chemical structures. Here we present a open-source based computer program to use digital video devices to capture and analyze rapidly hand drawn or computer generated molecular structures from plain papers. The computer program is capable of extracting molecular images from live streaming digital video signals and pre-recorded chemistry oriented educational videos. The images captured from these sources are further transformed into vector graphics for edge detection, node detection, Optical Character Recognition (OCR) and interpret them into bonds, atoms in molecular context. The molecular information generated is further transformed into re-usable data formats (MOL, SMILES, InCHI, SDF) for modeling and simulation studies. The connection table and atomic co-ordinates (2D) generated through this automatic process can be further used for generation of IUPAC names of the molecules and also for searching the chemical data from public and commercial chemical databases. Applying this software the digital webcams, camcorders can be used for recognition of molecular structure from hand-drawn or computer generated chemical images. The method and algorithms can be further used to harvest chemical structures from other digital documents or images such as PDF, JPEG formats. Effective implementation of this program can be further used for automatic translation of chemical images into common names or IUPAC names for chemical education and research. The performance and efficiency of this workflow can be extended to mobile devices (smart phones) with wi-fi and camera capabilities.