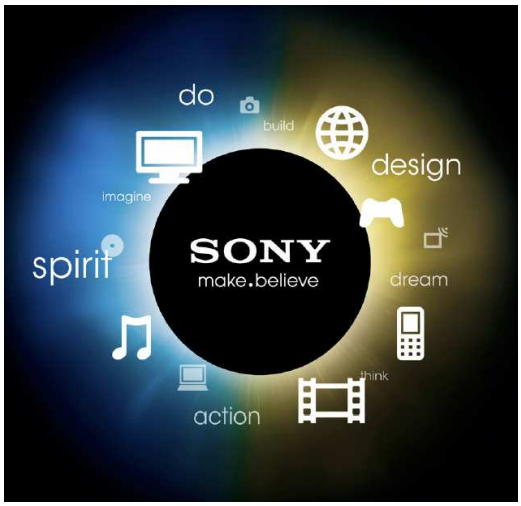
10.04.2012



Sony manufactures audio, video, communications, and information technology products for the global consumer and professional markets. With its music, pictures, games and online businesses Sony is uniquely positioned to become a leading personal broadband network company in the 21st century. Sony's European electronics operations cover a full range of activities from R&D, design, engineering and manufacturing to marketing, sales and service.

Our **Stuttgart Technology Center** (STC) is where innovations and future generations of products are being developed to meet the requirements and needs of the worldwide markets and customers.

In our STC **European Technology Center** (EuTEC) our engineers are providing advanced European Audio and Video Technology for worldwide and European specific CE devices. The R&D is ranging from architecture and algorithm to software and standardization in the areas of optical technology (illumination and projection systems), digital transmission technology (digital radio and TV, near field communication and power-line communication), TV picture improvement (HDTV and MPEG on flat-screen displays) as well as standardization (DRM & Media Standards, DVB, NFC, PLC).

We are offering a

Research Internship: Alternative OCR system for text in videos

Inside the Sony Stuttgart Technology Center, we are working on algorithms for the detection and classification of text in video signals. Our current system is using morphologic operators for text detection. For the character recognition the Tesseract OCR system is used which is optimized to work on scanned book pages. In order to address the special characteristics of a movie such as background noise and the variation of fonts and color, we developed various algorithms and filters in order to improve the recognition accuracy.

While Tesseract itself is regarded as a quite mature technology that provides high recognition accuracy, it is somehow limited with regard to input from videos. Since the development of Tesseract (1985-1995) new algorithms such as Support Vector Machines (SVM) and Scale Invariant Feature Transform (SIFT) have become available. Such new technologies have the potential to enable a more precise, and especially a more robust OCR technology which is suitable for the application in videos.

The task is to implement an alternative OCR system based on such new technologies. The resulting performance will be evaluated against the Tesseract system on both, scanned text pages and text segments extracted from video frames.

We are looking for a dedicated student with a background in either signal/video processing and/or machine learning. The candidate should know how to use Matlab and should be familiar with Linux/Unix. If you are interested in this position, please send a CV including your grades and your availability to the address below. The duration of the internship is 6 months, starting around October 2012. A compensation of 1000 Euro per month is paid.

Sony Deutschland GmbH
Speech and Sound Group -- ssg@sony.de