## Contents

1 VALUE PROPOSITION

1.1 Portrait .................................................. 1
1.2 Main research areas .................................... 1
1.3 Expertise directory .................................... 3
1.4 Dedicated development group ......................... 5

2 TECHNOLOGY TRANSFER .............................. 6

2.1 Project examples ........................................ 7
2.2 IdeArk: Idiap startup incubator ...................... 9

3 TECHNOLOGY COMPONENTS ....................... 10

3.1 Signal Processing and Machine Learning Toolbox (BOB) ........................................ 10
3.2 Speech-to-Text Transcription ............................ 11
3.3 Multilingual and accented speech recognition ........................................ 12
3.4 Grapheme-based and Instance-based Methods for Speech Processing ........................................ 13
3.5 Keyword Spotting for Searching and Indexing Spoken Documents ........................................ 14
3.6 Speech Synthesis ........................................ 15
3.7 Speaker Segmentation and Linking ....................... 16
3.8 Speaker recognition library and award winning technology ........................................ 17
3.9 Speaker Verification ....................................... 19
3.10 Large Scale Speaker Identification ..................... 20
3.11 Acoustic Simulator ...................................... 21
3.12 Very Low Bit-Rate Speech Coding ..................... 22
3.13 Intonation Modeling ..................................... 23
3.14 Automatic Speech Assessment .......................... 24
3.15 Ad Hoc Microphone Arrays .............................. 25
3.16 Speaker Localization and Separation .................... 26
3.17 Forensic Voice Comparison ............................. 27
3.18 Sparse Phonological Vocoding ......................... 28
3.19 Hybrid Recommender System for Multimedia .................. 29
3.20 Speech-Based Just-In-Time Information Retrieval .......................... 30
3.21 Machine Translation at the Document Level ........................................... 31
3.22 Face recognition library and awarded ISV technology .......................... 32
3.23 Fingerprint recognition ................................................................. 33
3.24 Palmvein recognition library ......................................................... 34
3.25 A method of detecting a falsified presentation to a vascular recognition system .................................................................................. 35
3.26 Scalable linear discriminant analysis for large scale pattern recognition .... 36
3.27 Fast Object Detector ......................................................................... 37
3.28 Large-scale Image Retrieval ............................................................ 38
3.29 Multi-camera Detection and Tracking ............................................... 39
3.30 Real-time Multiple face and head pose tracking .............................. 40
3.31 Gaze Estimation and Non-Verbal Behavior Extraction from consumer cameras ........................................................................ 41
3.32 Static Luggage Detection and Intrusion Detection ............................... 42
3.33 Unsupervised Activity Discovery, Monitoring, and Abnormality Detection ......................................................................................... 43
3.34 Motion Detection, Segmentation, and Estimation. Motion Stabilization ...... 44
3.35 Text detection and recognition in images and videos ..................... 45
3.36 Soft-skill analytics from online video job interviews ....................... 46
3.37 Activity Monitoring from Ambient Sensors ......................................... 47
3.38 Social context inference from proximity sensors ............................... 48
3.39 Robot Skills Transfer Toolbox ........................................................ 49
3.40 Cardiac Imaging .............................................................................. 50

4 CORPORATE SPONSORSHIP PROGRAM ........................................ 51
4.1 Corporate Sponsoring Program ......................................................... 51
4.2 Idiap Support Members ..................................................................... 52
4.3 Idiap Affiliated Members ................................................................... 53
4.4 Idiap Sponsor Members ................................................................... 54
4.5 Research Partners ........................................................................... 55
4.6 Swiss Center for Biometrics Research and Testing ............................... 56
1 VALUE PROPOSITION

1.1 Portrait

The Idiap Research Institute, based in Martigny, Valais - Switzerland, is a non-profit research foundation specialized in multimedia information management and multimodal man-machine interaction. The institution was founded in 1991 by the Town of Martigny, State of Valais, the Ecole Polytechnique Fédérale de Lausanne (EPFL), the University of Geneva and Swisscom. Idiap is connected to the EPFL through a Joint Development Plan.

1.2 Main research areas

**Speech & Audio Processing:** Speech processing has been one of the mainstays of Idiap’s research portfolio for many years. Today it is still the largest group within the institute, and Idiap continues to be recognized as a leading proponent in the field. The expertise of the group encompasses statistical automatic speech recognition (based on hidden Markov models, or hybrid systems exploiting connectionist approaches), text-to-speech, speaker recognition and generic audio processing (covering sound source localization, microphone arrays, speaker diarization, audio indexing, very low bit-rate speech coding, and perceptual background noise analysis for telecommunication systems).

**Computer Vision & Learning:** The main objective of the Computer Vision and Learning group is the development of novel machine-learning and statistical methods, with a strong emphasis on their algorithmic efficiency. The application domain we focus on encompasses any processing of images and video streams, for instance for object detection and scene analysis, tracking of people and biological structures, and image recognition in general. Our current research follows two main axes. The first is the study of novel generic learning methods to cope with feature spaces of large dimension, or large training sets. The second is the development of new algorithms for scene understanding, which involves practical problems such as object detection in natural scenes, or pedestrian tracking with a multi-camera setup.

**Social Computing:** Social computing is an interdisciplinary domain that integrates theory and models from mobile and ubiquitous computing, social media, machine learning, and social sciences to analyze and interpret human and social behavior in everyday life, and to create devices and systems that support interaction and communication. The group has unique expertise in human-centered approaches, and has conducted multidisciplinary research with experts in social psychology, human geography, and epidemiology. The current research in the group spans three areas: understanding user behavior in social video sites using online crowdsourcing and audio-visual behavioral analytics; inferring large-scale social and urban phenomena from mobile phone data, mobile crowdsourcing, and mobile social media; and understanding emerging social phenomena in face-to-face interaction through mobile and wearable sensors.

**Perception & Activity Understanding:** The Perception and Activity Understanding group conducts research in human activities analysis from multi-modal data (mainly vision and audio). This entails the investigation of fundamental tasks like the representation, detection, segmentation and tracking of people, the characterization of their state, and the modeling of sequential data and their interpretation in forms of gestures, activities, behavior or social relationships. These tasks are addressed through the design of principled algorithms extending models from computer vision, statistical learning, or multimodal signal processing. In particular, the group has focused on the recent years on the exploitation of probabilistic graphical models as one of its main modeling tools. Another objective of the group is the investigation, whenever possible, of unsupervised adaptation methods to leverage either the availability of potentially large amounts of test data from which exploitable statistics can be extracted, or of prior models that can help the soft tagging of test data, or of both. Surveillance, traffic analysis, analysis of behavior, human-robot interfaces, and multimedia content analysis are the main application domains.
Robot Learning & Interaction: The robot learning and interaction group focuses on human-centric robot applications, with the objective of developing probabilistic approaches to encode movements and behaviors in robots evolving in unconstrained human environments. In these applications, the models serve several purposes (recognition, prediction, online synthesis), and are shared by different learning strategies (imitation, emulation, incremental refinement or exploration). The aim is to facilitate the transfer of skills from end-users to robots, or in-between robots, by exploiting multimodal sensory information and by developing intuitive teaching interfaces. The group also takes the perspective that the robots actuation capability must be enriched with force and impedance information for haptic communication, as well as to anticipate the users’ behaviors and generate safe and natural movements.

Computational Bioimaging: Research in the Computational Bioimaging Group focuses on developing image acquisition, reconstruction and analysis algorithms to study dynamic processes and, in particular, live biological systems. Practical tools aim at (i) extending the physical limits of imaging hardware via techniques including super-resolution and multi-view, space variant deconvolution, and (ii) quantitative analysis of complex biological systems, including motion-based image analysis, cell tracking, microscopic fluid flow estimation, and integration of multi-modality images.

Biometric Person Recognition: The Biometrics group investigates and develops novel image-processing and pattern-recognition algorithms for face recognition (2D, 3D, and near-infrared), speaker recognition, anti-spoofing (presentation attack detection), and emerging biometric modes (EEG and vein). The group is geared toward reproducible research and technology transfer, using its own signal-processing and machine-learning toolbox.

Natural Language Processing: The Natural Language Processing (NLP) group studies how the analysis of texts at the semantic and pragmatic levels can contribute to two applications: machine translation and information retrieval. The group aims at improving the state of the art on core semantic and pragmatic analysis problems, such as semantic relatedness, disambiguation of pronouns or connectives, keyword extraction, or sentiment analysis. For machine translation, the NLP group combines text-level processing techniques with phrase-based statistical MT systems to improve translation quality. For information retrieval, the NLP group exploits content and sentiment analysis to improve multimedia recommendation over networked repositories; moreover, new keyword extraction methods are developed to improve the relevance of just-in-time document recommendation for conversations.

Uncertainty Quantification and Optimal Design: The Uncertainty Quantification and Optimal Design group focuses on quantifying and reducing uncertainties in the context of hi-fidelity models, with a main expertise on Gaussian Process methods and sequential design of computer experiments for optimization, inversion, and related problems. Application domains notably include energy and geosciences, with collaborations ranging from safety engineering to hydrology and climate sciences.
1.3 Expertise directory

<table>
<thead>
<tr>
<th>Expertise</th>
<th>Prof. Hervé Bourlard</th>
<th>Dr. Sylvain Calinon</th>
<th>Dr. Barbara Caputo</th>
<th>Dr. Ranan Collobert</th>
<th>Dr. François Fleuret</th>
<th>Dr. Philip N. Garner</th>
<th>Prof. Daniel Gatius-Perez</th>
<th>Dr. David Ginsbourger</th>
<th>Dr. Michael Liebling</th>
<th>Dr. Mathew Madhurai Doss</th>
<th>Dr. Sébastien Marcel</th>
<th>Dr. Petr Molick</th>
<th>Dr. Jean-Marc Odobez</th>
<th>Dr. Andrei Popescu-Răileanu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Converting spoken words into (searchable) text</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech synthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Converting written words into speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low bit rate speech coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enabling communication in difficult environments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech segmentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distinguishing speech from silence and other sound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prosody analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding meaning of speech by emotion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microphone array processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locating sound sources in a room</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker segmentation and linking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who spoke when? Linking speakers across heterogeneous recordings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noise robustness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making systems work in noisy environments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural language understanding and translation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Towards an increased level of abstraction of linguistic content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Document and text processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making knowledge accessible in the new information age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activities and attention recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding people behaviors and interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimodal processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expanding possibilities. Building future man-machine interfaces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computational cognitive science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enabling autonomy. Developing intelligence in systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal and non-verbal communication analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaching ever higher communication effectiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online and Mobile Crowdsourcing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tapping into the power of crowds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimedia information systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increasing the potential value for users in information systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User interfaces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making information useful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making systems work better</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expertise</td>
<td>Prof. Hervé Bourlard</td>
<td>Dr. Sylvain Calinon</td>
<td>Dr. Barbara Caputo</td>
<td>Dr. Ronan Collobert</td>
<td>Dr. François Fleuret</td>
<td>Dr. Philip N. Garner</td>
<td>Prof. Daniela Gatica-Perez</td>
<td>Dr. David Ginsbourger</td>
<td>Dr. Michael Leblang</td>
<td>Dr. Mathew Maginli-Doss</td>
<td>Dr. Sébastien Marcel</td>
<td>Dr. Petr Motlicek</td>
<td>Dr. Jean-Marc Odobez</td>
<td>Dr. Andrei Popescu-Belis</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>--------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>------------------------</td>
<td>---------------------</td>
<td>--------------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Face recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vein biometrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-spoofing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimodal person recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimodal person recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical machine learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leveraging unlabeled data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human-robot interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human-robot skills transfer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interfaces for robot programming</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptation of movements to new situations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of heterogeneous motion data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bio-microscopy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertainty Quantification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design and Metamodelling of Computer Experiments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1.4 Dedicated development group

Composed of a dozen highly skilled engineers, who understand our research domains, but are more motivated in bringing this into working prototypes or products (on multiple hardware platforms and OSs), this group has two main components: (1) the Technology Transfer Office (TTO) being the actual entry door to the development group, as well as the official contact point for development service/industrial requests and (2) the development group itself.

**TTO – Technology Transfer Office (Dr. Florent Monay):** Idiap is very active in multiple, national and international, technology transfer activities, and is also involved in numerous projects with industries, ranging from large institutions to SMEs and startups.

The TTO group represents the key link between researchers, development engineers, and industry. Besides maintaining Idiap’s technology portfolio (discussed below), and responding to industrial needs and contacts, the TTO also pro-actively investigates new opportunities. In this context, it assists Idiap researchers and industries to develop joint projects. As also discussed below, and illustrated in the present document, the TTO is also responsible to maintain a clear IPR strategy and track IP status across licenses. The TTO is thus a key entry link between the scientists, the industries, and the Development Group discussed just below.

**Development Group (Mr. Olivier Bornet):** Since the end of 2006, Idiap also maintains an active group of a dozen of highly talented developers, all working in the same office, and sharing multiple responsibilities, including:

- **Industrial projects:** Development of prototypes, and adaptation of research software to the needs of industries, with the goal to demonstrate feasibility of new products and systems, and in direct response to a contract with the industry or the needs of one of the Idiap’ spin-offs (always in the framework of well defined projects and IPR contracts). This also includes the development part of CTI projects, where responsibilities are then shared by the PI researchers and one of the developers.

- **Academic projects:** The development group is also responsible for the development of demonstration systems, either of general interest to Idiap (in which case efforts are funded by Idiap) or in the context of projects, e.g., EU project involving specific development efforts going beyond what researchers are able or willing to do.

- **Internal support:** The development group is also responsible to respond to in-house requests, e.g., helping out colleagues in resolving complex code issues, contributing in research or collaborative platform development, etc. In this context, the group is also responsible for the quality control and maintenance of the open source libraries made publicly available. As available for the Idiap web site, 31 open-source libraries are currently available from [http://www.idiap.ch/technology-transfer/open-source-software](http://www.idiap.ch/technology-transfer/open-source-software). Each library has to be thoroughly tested and validated before this release, and nobody at Idiap is allowed to release open source software without prior permission of the management and without doing this through the development group, who will make sure of the quality of what is being delivered, perform version control, etc.

- **Stable software and platforms:** This is not only part of the internal support discussed above but also for “close source” software which is being licensed to industries. There also, the development group is responsible for quality control, version control, and maintenance (if necessary). In collaboration with the TTO, they also have to make sure that we all follow the same IPR policy and keep track of the licensing and IPR status.
2 TECHNOLOGY TRANSFER

Idiap is committed to the transfer of knowledge and technology to the business community through an active engagement and partnership. The Technology Transfer Unit works in close collaboration with businesses to realize joint research and sponsored development programs along the research areas of Idiap. From small startups to large corporations, Idiap Research Institute is an ideal partner in “avant garde” technology and research!

Technology transfer is one of Idiap Research Institute’s three core mission including fundamental research and education. With “Technology transfer” we mean the process that leads the technology (software, algorithm, knowledge, expertise) to be transferred from Idiap to an industrial partner or a spin-off for economic development purposes.

This transfer is usually done by giving the grant of rights on the commercial exploitation of this technology (through license). We can therefore consider the process of technology transfer as a creation of economic value out of scientific discoveries. If the concept sounds simple, the path that leads from a technology demonstrator to an industrial product is highly complex. One reason is that researchers and industrialists do not speak the “same language” and their aim could be totally different.

One of the fundamental challenges in “technology transfer” is to facilitate the interface between the knowledge and skills of the researcher and the needs of the industrial partner. Idiap has resolved this by providing a pluri-disciplinary team of developers, programers dedicated to technology transfer. Also, the incubator IdeArk (http://www.ideark.ch/en/, see Section 2.2), a spin-off from Idiap has the mission to enhance Idiap’s technologies and promote the emergence of new startups by providing workspace, financing and mentorship.

![Figure 1: Value Creation](image-url)
2.1 Project examples

**Six-month project to design a novel alignment approach for a manufacturing process**

The industrial partner was a Swiss-based supplier of solutions for manufacturing plastic components. The project was initiated through pre-existing contacts with Idiap employees.

The problem at hand was to add into a pre-existing manufacturing process a procedure to align automatically plastic elements as they were manipulated before being assembled.

The preliminary discussions between Idiap and the company lasted for three months, after which a proposal was drafted collaboratively and submitted to TheArk to fund one developer at 50% for six months at Idiap, with a commitment from the industrial partner for four months of work on their side. The funding for the project was accepted in less than a month.

The company then provided Idiap with a test machine to record measurements from a reference sensor. From there at Idiap a software developer funded through the project, the project manager and a scientific supervisor met weekly. They designed and benchmarked a novel algorithm able to estimate automatically deviations from a desired alignment in spite of heavy structured noise in the measurements. This development was done with a monthly face-to-face meeting with the industrial partner to discuss progresses.

The outcome of the project was a novel algorithm, a performance report, and a reference implementation in C that could be directly transferred to the industrial partner for integration into its manufacturing robots. There were a few additional technical exchanges with Idiap for minor software changes.

In the following months the company filed a patent to protect the method.

**Resources:** 4 PM Engineer

**Funding:** 48'100 CHF

**One year project to develop activity analysis methods for a start-up company**

The industrial partner is a Swiss startup in the domain of health. The objectives of the collaboration are confidential; the domain of work is activity analysis.

The project was initiated through contacts between TTO and the company. After few months of interaction and concept definition, a CTI proposal was prepared, submitted, and accepted. Idiap acted as lead research partner.

The CTI project supported one postdoctoral researcher at full-time for one year and a software developer at part-time during one year.

Interaction included weekly interactions between research partners and monthly-to-bimonthly team meetings.

The main goals of the project were achieved, but there was a need to re-orient some of the research due to limitations related to availability of labeled data for machine learning.

This project encouraged the partners to apply for a follow-up CTI project, who was ultimately granted.

**Resources:** 12 PM Post-Doc, 8 PM Engineer

**Funding:** 270'000 CHF
**One year project to develop security solutions in the domain of mobile biometrics**

The industrial partner is a large US-based Internet company. The contact was initiated following a short research stay within the company premises. After the research stay the company was enthusiastic by the skills of our researchers and decided to fund a follow-up research project.

The precise objective of the collaboration is confidential and we can only mention that it is related to biometrics security on mobile devices.

In only one month, Idiap and the company prepared a research and an IP-agreement, specifying milestones and deliverables. The industrial partner agreed to fund through direct payment two researchers (scientific collaborators).

Over the duration of the project, the researchers will develop and evaluate algorithms to address specific vulnerability aspects in mobile biometrics. Idiap remains in contact with the industrial partner through remote meetings with the project manager.

The deliverables are composed of python/C++ software, in a form of an independent archive based on our BOB library ([https://www.idiap.ch/software/bob](https://www.idiap.ch/software/bob)), which was provided under an exclusive license of use in a restrictive field.

**Resources:** 24 PM Engineer  
**Funding:** 275’000 USD

---

**Fourteen months project to develop a novel tool for prediction in the domain of biometrics**

The industrial partner was a Russian Security company, and the contact was initiated by them by searching for Research Institutions with expertise in biometrics on the web.

The objective of the collaboration was to design a novel prediction algorithm for biometrics.

During two months, Idiap and the company conducted preliminary discussions and drafted a research and an IP-agreement, the former specifying milestones and deliverable. The industrial partner agreed to fund through direct payment two researchers at the post-doctoral level and a part-time software developer. It also provided Idiap with a test database for benchmarking the developed methods.

Over the duration of the project, the researchers funded on the project developed a new algorithm and a prototype implementation, and the software developer wrote a industrial-grade implementation packaged into a library, and developed an accompanying graphic interface for testing. Idiap remained in contact with the industrial partner through meetings with the CEO directly every two months.

The deliverables were composed of an evaluation report and an archive of the library, which was provided under an exclusive license of use in a restrictive field.

**Resources:** 24 PM Post-Doc  
**Funding:** 250’000 EUR
Three year project on state of the art large vocabulary speech recognition

The industrial partner was Samsung in Korea. We were known to their research group mainly via some of our software open-sourced previously; we were invited to apply for their annual call for research partnerships.

Funding was granted under the Samsung GRO program. The object of the collaboration was to develop and evaluate state of the art speech recognition for distributed platforms. The project was renewed once with the same funding scheme and a slightly different focus. It was then renewed for a third year, with funding coming directly from the research lab, rather than the GRO program.

The Samsung GRO has a standard IP sharing agreement.

The collaboration was eased by use of open-source solutions, and resulted in a patent being filed (EP2736042A1).

Resources: 20 PM Researchers, 8 PM Post-Doc
Funding: 307’425 USD

2.2 IdeArk: Idiap startup incubator

As part of the Valais initiative TheArk, IdeArk relays the concerns of companies to the scientists, identifies market trends and facilitates the conversion of research results into innovative solutions. IdeArks task is also to support start-ups and SMEs active in Idiap-related fields.

The IdeArk Incubator provides multiple types of support and services to startups, including: free office spaces with furnitures, internet connection, coaching support, and financial support for collaborative projects with Idiap.

Many technologies from Idiap were developed into startups through the IdeArk support (KeyLemon, Klewel, AudioSearch, Koemei, recapp, . . .) and naturally benefit from Idiaps proximity and expertise.

Once a year, an International Create Challenge (ICC, http://www.createchallenge.org/) is also being held within the Idiap premises. Initiated in 2012, ICC is a free of charge, 3-week immersive super accelerator program that gives entrepreneurs and potential entrepreneurs the opportunity to drive their projects to a “Minimum Viable Product” (MVP). It is a unique program combining state-of-art technologies, cutting edge research, mentors-led coaching, and micro-seed investment (in collaboration with VCs and KickVenture).

Finally, in 2014, Idiap set up a crowdfunding portal (https://funding.idiap.ch/) to financially support the selected projects of the ICC accelerator program in their quest to become a viable start-up. Idiap/ICC is proud to be on the map of the crowdfunding industry in Switzerland in 2015 according to the recent study of Lucerne University of Applied Sciences and Arts1.

1http://blog.hslu.ch/retailbanking/crowdfunding/?sourceurl=/crowdfunding
3 TECHNOLOGY COMPONENTS

3.1 Signal Processing and Machine Learning Toolbox (BOB)

Keywords

Signal processing; machine learning; toolbox

Key contact researcher(s)

Dr. Sébastien Marcel
marcel@idiap.ch
Tel.: +41 27 721 77 27

Innovative aspects

– Multi-dimensional arrays Blitz arrays (tensors up to 11 dimensions) integration
– LAPACK integration
– HDF5 scientific file format integration
– Signal processing (FFT, DCT)
– Image processing (SIFT bridge, LBP, DCT block, Gabor, Smoothing, Optical Flow, illumination normalization)
– Biometric database and protocol support
– PCA/LDA, Linear Machines, MLP, SVM bridge, k-Means, Gaussian Mixture Models, Joint Factor Analysis, Inter-Session Variability Modeling, Probabilistic Linear Discriminant Analysis
– Video support
– Performance evaluation toolkit

Commercial application examples

– Face recognition, speaker recognition, vein recognition, multi-modal processing
– Biometrics-enabled identity management systems (Automated Border Control, Access Control, . . .)
– Multi-factor authentication security systems (Critical Infrastructures, e-Banking, . . .)
– Forensic Science, Video surveillance, Entertainment, Robotics, Man-Machine interaction

More information


Software & IPR status

The library is available under the BSD license:

– BOB: http://www.idiap.ch/software/bob/
3.2 Speech-to-Text Transcription

Functional description

Idiap has a growing collection of state of the art software to develop leading-edge speech-to-text systems and to perform speech recognition in various scenarios. Packages include Juicer, Tracter, ISS, and Kaldi. As one of the early pioneers in the development of hybrid systems using Hidden Markov Models (HMM) and Artificial Neural networks (ANN, now referred to DNN), we have leading-edge HMM/DNN technologies and speech recognizers available in several languages. Idiap can provide customized solutions of speech transcription systems for new languages (depending on availability of at least 100 hours of well transcribed audio data). Very large vocabulary (> 100 Kwords), continuous speech, speaker independent, HMM/GMM and HMM/DNN-based speech recognition available.

Innovative aspects

- Neural network based acoustic modeling, technology allowing for speaker/channel/sentence-based adaptation.
- Language model – domain customization using data crawling from Internet.
- Offline, near real-time, and real-time solutions, depending on algorithmic delay allowed by customers.

Commercial application examples

- Dictation system customized for specific domain
- Search for specific information in large audio archives
- Audio content data-mining and indexing

More information


Software & IPR status

IP lies in the copyright of the code, plus in expertise allowing us to create recognisers from suitable third party databases, dictionaries and tools. Some databases associated with certain languages may require separate agreements.
- Kaldi: http://kaldi.sourceforge.net/about.html
3.3 Multilingual and accented speech recognition

**Functional description**

Idiap is one of the world’s leading providers of multilingual speech recognition technology. Multilingual speech recognition involves numerous research challenges, including adaptation on limited amount of training data and mixed language recognition where it is not even possible to assume that one knows in advance the language being spoken. Idiap’s system can rapidly adapt to under-resourced languages, i.e. languages that have only a small amount of data available to train a speech recognition system, or to accented speech. Furthermore, Idiap’s system allows to perform mixed language speech recognition for Swiss German, Swiss French and English.

**Innovative aspects**

- Parsimonious use of language adaptation data (adaptation with very little amount of data)
- Fast adaptation to unseen languages
- Generic probabilistic multilingual framework

**Commercial application examples**

- Robust non-native speech recognition
- Mixed language speech recognition
- Language identification

**More information**


**Software & IPR status**

Most of the underlying software is open source. Some databases associated with certain languages may require separate agreements.

The Idiap spin-off recapp IT ([www.recapp.ch](http://www.recapp.ch)) commercialises parts of the technology and can provide multilingual speech recognition as a webservice.
3.4 Grapheme-based and Instance-based Methods for Speech Processing

Keywords
Speech processing; Markov models; computational linguistics; graphemes; phonemes

Key contact researcher(s)
Dr. Mathew Magimai Doss
Mathew.MagimaiDoss@idiap.ch
Tel.: +41 27 721 77 88

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number:
Software disclosure 8700
Software disclosure 11457

Functional description
Phonetic lexicon and acoustic resources (transcribed speech data) are the primary resources needed for development of speech technologies, such as automatic speech recognition, text-to-speech synthesis, spoken term detection. Not all languages or domains may have such well developed resources. To overcome these limitations, Idiap is actively engaged in research on alternate ways to model the words in the language through acoustics. More specifically, (a) by modeling units of written form i.e. graphemes and (b) by modeling spoken instances.

Innovative aspects
– Speech recognition without the need of a phonetically transcribed lexicon
– Development of linguistic resources and speech technology for under-resourced languages and domains

Commercial application examples
– Speech and language technology applications where lexicon with phonetic transcription is not readily available
– Front-end of text-to-speech systems
– Search for specific information in large audio archives
– Audio content data-mining and indexing

More information

Software & IPR status
The complete software consists of open source softwares such as HTK, Quicknet, SSP (SSP: https://github.com/idiap/ssp) and Idiap proprietary code that interfaces the open source tools, which can be customized to the customers need. Speech databases and lexical resources used for development may require commercial license.
3.5 Keyword Spotting for Searching and Indexing Spoken Documents

Functional description

Idiap offers solutions to search and index spoken documents (audio/multimedia archives) through keyword spotting (KWS) technology. The solutions support both text-based query input and query-by-example input i.e. spoken query. The KWS systems are in-parts or fully based on state-of-the-art automatic speech recognition technology. Consequently, Idiap can provide customized solutions of keyword spotting systems for different languages and domains.

Innovative aspects

- Neural network based acoustic modeling, technology allowing for speaker/channel/sentence-based adaptation
- Support of several speech transmission channels (landline telephone, mobile phone and voice over IP)
- Several times faster than real-time solution, enabling search in very large audio archives
- Robust solutions also with noisy recordings
- Unified framework for text query-based KWS and query-by-example based KWS

Commercial application examples

- Search for specific information in large audio archives
- Audio content data-mining and indexing

More information


Software & IPR status

Most of the underlying software is open source. Some databases associated with certain languages may require separate agreements.

The Idiap spin-off AudioSearch SARL is commercializing a part of the keyword spotting technology.
3.6 Speech Synthesis

Functional description

Idiap’s speech synthesis technology is based on years of experience in speech recognition, and in cross fertilisation of the two technologies. We are able to provide specific voices in several languages, including English and the main Swiss languages. We can also provide average voice based TTS, that is, generic voices that can be adapted quickly to more personal voices. We are active in further merging the technologies for recognition and synthesis allowing advances in either domain to benefit the other.

Innovative aspects

- Convergence of recognition and synthesis technology
- Cross-lingual adaptation

Commercial application examples

- Speech to speech translation
- Hands free feedback
- Prosthetic voices
- Assistance for blind and partially sighted

More information


Software & IPR status

Most of the underlying software is open source. Some databases associated with certain languages may require separate agreements.

- VTLN: http://www.idiap.ch/software/hts-vtln/
- SSP: https://github.com/idiap/ssp
- HTS: http://hts.sp.nitech.ac.jp/
- Kaldi: http://kaldi.sourceforge.net/about.html

Keywords

Speech synthesis; text-to-speech systems; Markov models; speech signal processing

Key contact researcher(s)

Dr. Philip N. Garner
Phil.Garner@idiap.ch
Tel.: +41 27 721 77 68

Technology Transfer Office

Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program

See Section 4 of the present document

File reference & version number:

Software disclosure 5879
Software disclosure 8700
Software disclosure 11457
3.7 Speaker Segmentation and Linking

**Functional description**

The speaker segmentation and linking systems automatically discover speaker boundaries in an audio and link similar speakers across audio recordings in a database. Idiap’s speaker segmentation system, also referred to as the speaker diarization system, employs the Information Bottleneck approach for faster-than-real-time, unsupervised segmentation of speakers in an audio recording. The speakers discovered are then linked across audio recordings using the state-of-the-art speaker modelling and comparison approaches. Our systems have been demonstrated to work on a variety of recording conditions such as TV shows, talks, meetings, etc.

**Innovative aspects**

- Rapid speaker segmentation of audio with the Information Bottleneck approach
- Longitudinal linking of speakers across audio recordings with i-vectors
- Speaker overlap detection to reduce segmentation errors

**Commercial application examples**

- Pre-processing for speech processing systems (e.g., for speech recognition, speaker recognition, etc.)
- Mining speakers in a large audio collection

**More information**


**Software & IPR status**

The Information Bottleneck based speaker diarization system is available as a free and open source software.

- [http://www.idiap.ch/scientific-research/resources/speaker-diarization-toolkit](http://www.idiap.ch/scientific-research/resources/speaker-diarization-toolkit)
3.8 Speaker recognition library and award winning technology

**Keywords**

Speaker recognition library

**Key contact researcher(s)**

Dr. Sébastien Marcel  
marcel@idiap.ch  
Tel.: +41 27 721 77 27

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document

**File reference & version number:**

Software disclosure 9142

---

**Functional description**

This library is designed to run speaker verification experiments.

**Innovative aspects**

Different state-of-the-art algorithms for voice activity detection, feature extraction (MFCC, LFCC), modelling techniques (Gaussian Mixture Models, Session Variability Modelling, Total Variability Modelling – iVectors) and database interfaces are implemented.

**Commercial application examples**

– Biometrics-enabled identity management systems (Automated Border Control, Access Control, …)
– Multi-factor authentication security systems (Critical Infrastructures, e-Banking, …)
– Forensic Science, Video surveillance, Entertainment, Robotics, Man-Machine interaction

**More information**


NIST Speaker i-vector Machine Learning Challenge 2014

**Software & IPR status**

Open sourced at https://pypi.python.org/pypi/bob.bio.spear but re-licensing possible for commercial purposes.
3.9 Speaker Verification

Functional description

Idiap’s speaker verification/authentication technology focuses on datasets with large speaker populations. Besides state-of-the-art approaches based on iVectors and SGMM modeling generating low-dimensional speaker models, followed by channel compensation techniques, a hybrid approach successfully integrates our state-of-the-art speech recognition (or key-word spotting) technology into a speaker recognition system that benefits from linguistic resources.

Our environment simulator provides massive resources for the development of speaker verification systems that are more robust to acoustic mismatch.

Innovative aspects

– Combining state-of-the-art speech and speaker recognition technologies
– Exploiting the content of the spoken message to enhance verification capabilities
– Simulation of a large variety of acoustic environments for improved robustness
– Multiple enrollment solutions offering enhanced scoring capabilities

Commercial application examples

– Verification/authentication of caller against large database of speaker models
– Speaker verification of multi-styled and noisy speech
– Security: ID verification, forensics, tracking of criminal activities
– Gatekeeper for accessing secured systems (“voice-as-password” applications exploiting users’ knowledge)

More information


Software & IPR status

The above mentioned approaches to speaker verification are developed for the open-source Kaldi toolkit. The acoustic simulator is available as an open-source package from publicly available resources.
3.10 Large Scale Speaker Identification

Functional description

Idiap’s speaker identification technology focuses on datasets with large speaker populations. It exploits several complementary technologies (iVectors, Subspace GMMs) to generate low-dimensional but representative speaker models, followed by channel compensation technique. Our environment simulator provides massive resources for the development of speaker identification systems that are more robust to acoustic mismatch.

Innovative aspects

- Combining state-of-the-art speech and speaker recognition technologies
- Simulation of a large variety of acoustic environments for improved robustness
- The technology is language-, text- and channel-independent
- Possibility to detect phone exchange within the same speaker
- New speaker can be easily added by end-users
- The technology supports all speaker identification alternatives: 1:1, 1:N, N:M (group of speakers against group of speaker models)

Commercial application examples

- Search and retrieval of speaker identities (in large audio archives)
- Speaker identification of multi-styled and noisy speech
- The technology can be easily combined with gender and language identification

More information


Software & IPR status

The above mentioned approaches to speaker identification are developed within the open-source Kaldi toolkit. The acoustic simulator is available as an open-source package from publicly available resources.
3.11 Acoustic Simulator

Functional description

Idiap maintains an open-source acoustic simulator that provides a large number of speech signal degradation processes. Based on a large database of 80h of real noise recordings, over a hundred impulse responses of speakers, cabinets, smartphones and rooms along with 15 different speech and audio codecs, this simulator stands out as a valuable pool of resources to develop detailed and robust models for speaker and speech recognition technology. It currently covers landline, cellular, satellite, VOIP, interview and playback scenarios.

Innovative aspects

- Large number of speech signal degradation processes
- Open-source data and code
- Expandable functionality

Commercial application examples

- Development of robust speech processing systems
- Low-resource speech applications

More information


Software & IPR status

IP lies in the copyright of the code. Noise data and impulse responses are online resources released under a Creative Commons Attribution license. Codec code has more restrictive licenses.
3.12 Very Low Bit-Rate Speech Coding

Keywords
Speech processing; speech coding; very low bit rate real-time speech communication

Key contact researcher(s)
Dr. Philip N. Garner
Phil.Garner@idiap.ch
and
Dr. Milos Cernak
milos.cernak@idiap.ch
Tel.: +41 27 721 77 68/99

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number:
N/A

Functional description
Idiap’s very low bit-rate (VLBR) speech coding system is a conventional speech coder with an acceptable communication delay, designed for real-time speech communication. Cascaded phone-based speech recognition and synthesis systems transmit phonetic and syllabic information, encoded independently, i.e., asynchronously. It operates at an uncompressed bit rate of 213 bits/sec and achieves an average communication delay of 243 ms. A more recent version also uses phonological speech representation instead of phonetic, and is purely based on artificial neural networks, and no hidden Markov models are then used.

Innovative aspects
– Speech coding inspired by the human speech signal processing, viz., has explicit simultaneous phonetic and syllabic components, which are asynchronously related.
– The recent version transmits binary phonological speech representation, and thus is more ready for a multi-lingual use.

Commercial application examples
End-to-end speech transmission suitable for military and tactical communication systems, and for environments with highly restricted bandwidth, such as under-see communication systems.

More information

Software & IPR status
Most of the underlying software is open source. Some databases associated with certain languages may require separate agreements.
– https://gitlab.idiap.ch/milos.cernak/sct-ext
– https://gitlab.idiap.ch/milos.cernak/phonovoc
### 3.13 Intonation Modeling

**Keywords**

Intonation modelling; text-to-speech systems; emotion recognition; speech signal processing

**Key contact researcher(s)**

Dr. Philip N. Garner  
Phil.Garner@idiap.ch  
Tel.: +41 27 721 77 68

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document

**File reference & version number:**

Software disclosure 8700  
Software disclosure 11457

---

**Functional description**

Intonation refers to the way pitch and energy of a speech signal vary in order to convey meaning, intent or emotion. It is normally ignored in speech recognition, but is essential for natural sounding speech synthesis. Idiap has novel intonation models with a physiological basis, closely matching the human production mechanism. Designed for cross-lingual emotion transfer in speech to speech translation, the models have potential application in emotion recognition and speech modification.

**Innovative aspects**

- Biologically plausible physiological model
- Cross-lingual adaptation

**Commercial application examples**

- Speech to speech translation
- Emotion recognition
- Emotional speech synthesis

**More information**


**Software & IPR status**

Much of the underlying software is or will soon be open-source.

- SSP: [https://github.com/idiap/ssp](https://github.com/idiap/ssp)
3.14 Automatic Speech Assessment

Keywords
Speech assessment; speech signal processing; Markov models

Key contact researcher(s)
Dr. Mathew Magimai Doss
Mathew.MagimaiDoss@idiap.ch
Tel.: +41 27 721 77 88

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number:
N/A

Functional description
Idiap conducts research on the development of automatic speech assessment methods. The R&D activities range from objective assessment of speech intelligibility to automatic accentedness evaluation of non-native speech. Automatic speech assessment reduces the need for time-consuming listening tests with human subjects and aids in developing speech communication and assistive speech technologies, such as speech coding, speech synthesis, computer aided language learning (CALL).

Innovative aspects
– Unified framework to assess speech at various levels, such as phone/phoneme, word
– Quantitative results that can be interpreted like traditional scores from human subjects

Commercial application examples
– Benchmarking of speech telecommunication or synthesis systems
– Development of assistive technologies such as CALL, pathological speech processing systems

More information

Software & IPR status
The complete software consists of open source softwares such as HTK, Quicknet, SSP (SSP: https://github.com/idiap/ssp) and Idiap proprietary code that interfaces the open source tools. Speech databases and lexical resources used for the development of the assessment system may require commercial license. Application of automatic accent evaluation for CALL is presently licensed to an industrial partner.
3.15 Ad Hoc Microphone Arrays

**Keywords**

Microphone position calibration, Synchronization, Distributed source localization

**Key contact researcher(s)**

Prof. Hervé Bourlard  
Dr. Afsaneh Asaei  
herve.bourlard@idiap.ch  
afsaneh.asaei@idiap.ch  
Tel.: +41 27 721 77 11

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document

**File reference & version number**

Software disclosure 11610

**Functional description**

Ubiquitous sensing exploits the abundance of microphone-embedded devices, such as notebooks and smart phones, in developing innovative technologies capable of aggregating widely sensed spatial audio. The key challenges pertain to the unknown geometry of the microphones and asynchronous recordings to enable higher level speech applications. Idiap has the prototype of a technology for calibration and synchronization of ad hoc microphones relying on properties of the structured matrices as well as the acoustic reverberation.

**Innovative aspects**

- Calibration of microphones from partial distances
- Synchronization of microphone recordings
- Distributed speaker localization
- Ability to handle very large network of ad hoc microphones
- Ability to deal with malfunctioning microphones

**Commercial application examples**

- Hands-free high-quality speech recording
- 3D audio technologies
- Sound field reproduction
- Surveillance

**More information**


**Software & IPR status**

The research software is available upon agreement.
3.16 Speaker Localization and Separation

**Functional description**

Multi-party speaker localization and separation play a key role in many applications involving hands-free meeting recordings and scene analysis. Idiap’s multiparty speech processing technology provides a flexible framework for multi-speaker localization and speech separation based on the recently emerged techniques of the structured sparse coding. The principles rely on the models of the auditory system in structural grouping and spatial analysis.

**Innovative aspects**

- Ability to pinpoint the exact source location
- Ability to separate individuals from overlapping speech
- Requirement for a few number of microphones
- Flexible in ad-hoc and constrained microphone placement

**Commercial application examples**

- Distant speech recognition
- Hands-free video conferencing
- Sound field reproduction
- Surveillance
- Entertainment

**More information**


**Software & IPR status**

- The reference implementation is available upon agreement.
- The rights are partly covered by a patent application: US20130096922: Method, apparatus and computer program product for determining the location of a plurality of speech sources.
3.17 Forensic Voice Comparison

Keywords
Voice comparison; speaker recognition; security; forensics; tracking of criminal activities

Key contact researcher(s)
Dr. Petr Motlicek
pmotlic@idiap.ch
Tel.: +41 27 721 77 49

Technology Transfer Office
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number: 141

Functional description
Forensic voice comparison involves the expert assessment of voice recordings to determine the identity of a speaker (offender). The speakers involved are termed as known speakers and questioned speakers. The former is a set of one or more speakers whose identity is known while the latter is the speaker whose identity we wish to confirm with respect to the known speakers. A fully-automatic (or semi-automatic) approach available at Idiap exploits an automatic speaker recognition system to obtain, through a series of carefully designed steps, a score signifying the proximity of the voice characteristics of the offender to that of the known speaker.

Innovative aspects
– Language-independent technology based on a distribuable service that takes into account relevant population data matching the questioned-speaker data
– Score evaluation and quality assurance provided in a final report
– Client-server interaction using the standard REST API

Commercial application examples
– Pay-Per-Use licensing scheme or fully commercial system can be offered
– Service specifically targeted for forensic voice comparison that exposes latest technologies in speaker recognition
– Latest technology with large interest for government security bureaus and law-enforcement agencies

Software & IPR status
The above mentioned approach to forensic voice comparison is being developed internally and for the partners in the relevant projects.
### 3.18 Sparse Phonological Vocodering

#### Functional description

Phonological representation of speech signal and multimedia in general, are sparse and their coefficients are highly structured. The underlying structured sparsity patterns are related to segmental and supra-segmental speech parameters, associated to the production and perception of speech at different time scales correlated with physiology of articulation as well as linguistics. The structured sparsity can thus be utilized for innovative technologies in the context of speech production and ineligibility assessment. They also enable highly efficient speech representation. Our developed computational platform can achieve ultra low bit speech coding and it is applicable for automatic speech segmentation or linguistic parsing as well as assessment of speech production and ineligibility.

#### Innovative aspects

- Deep learning based phonological vocoding
- Structured compressive sensing of phonological representations
- Structured sparse coding of phonological representation
- Class-specific codebook of diverse linguistic structures

#### Commercial application examples

- Low bit rate speech coding
- Automatic speech segmentation and alignment
- Speech production and ineligibility assessment
- Robust recognition of articulatory distorted speech

#### More information

“A. Asaei, M. Cernak, H. Bourlard, On Compressibility of Neural Network phonological Features for Low Bit Rate Speech Coding, proceeding of Interspeech, Dresden, Germany, September 2015.

#### Software & IPR status

Some databases associated with certain languages may require separate agreements. The rights are covered by a patent application: US2015846036: Signal processing method and apparatus based on structured sparsity of phonological features.
3.19 Hybrid Recommender System for Multimedia

**Keywords**

Recommender system; content-based recommendation; collaborative filtering; hybrid recommender system; sentiment analysis; multimedia recommendation

**Key contact researcher**

Dr. Andrei Popescu-Belis  
andrei.popescu-belis@idiap.ch  
Tel.: +41 27 721 77 29

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document

**File reference & version number:**

Software disclosure 10782

---

**Functional description**

This system performs generic and personalized recommendation of multimedia content: typically A/V lectures, but also images or videos. We have identified the metadata most relevant for content-based (CB) recommendation. We have defined a method to combine CB and collaborative filtering (CF), which is applicable to both cold-start and non-cold start settings. In one-class CF problems (when users only mark items as favorites, or ignore them), we can complement user action information with user sentiment extracted from user-generated texts such as comments. Furthermore, we can perform fine-grained sentiment analysis in terms of aspects, learning to detect which sentences of reviews or comments are likely to refer to given aspects, for review segmentation and summarization.

**Innovative aspects**

– Hybrid content-based / collaborative filtering recommender  
– Complement one-class ratings with sentiment of comments  
– Aspect-based analysis of reviews, extracting most representative sentences

**Commercial application examples**

– Recommender systems for large or small collections  
– Explanation of recommendations based on user reviews

**More information**


**Software & IPR status**

Open sources:  
– CBRec v1.0: https://github.com/idiap/cbrec  
– EMORec v1.0: https://github.com/idiap/cbrec
3.20 Speech-Based Just-In-Time Information Retrieval

**Functional description**

We have designed a system that uses words from a conversation between users to create implicit queries, submits them to a search and retrieval system, clusters the results and displays suggestions of documents that are potentially relevant to users at a given moment in the conversation. At the core of the system lies an algorithm that extracts a relevant and diverse set of keywords from a conversation fragment, avoiding ASR noise. These are then clustered according to conversation topics. Retrieval results from several implicit queries are merged and ranked by rewarding diversity. Moreover, we can also allow users to ask explicit queries in speech, and expand them based on the conversational context.

**Innovative aspects**

- Extract relevant and diverse keywords, avoiding noise
- Use topical clustering to formulate implicit queries
- Rank results by rewarding diversity
- Expand explicit queries using context

**Commercial application examples**

- Real-time document recommender system in meetings
- Document recommender for lecture recordings

**More information**


**Software & IPR status**

Research software, demonstrator (Idiap Showroom).
3.21 Machine Translation at the Document Level

**Keywords**

Statistical machine translation; discourse structure; connectives; pronouns; verbs

**Key contact researcher**

Dr. Andrei Popescu-Belis  
andrei.popescu-belis@idiap.ch  
Tel.: +41 27 721 77 29

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document

**File reference & version number:**

Software disclosure 8452  
Software disclosure 9011

---

**Functional description**

This technology improves current systems for automatic translation, by using text-level information. Several pre-processors detect inter-sentence dependencies and use them to label different types of words, prior to translation. The labels are used by a third-party statistical phrase-based machine translation system (open source) which otherwise translates sentence-by-sentence. We currently deal with discourse connectives (e.g. ‘since’ or ‘while’), pronouns (‘it’ vs. ‘il’ or ‘elle’) and verb tenses in English, and with Chinese and German compounds. Our method improves the translation of entire texts.

**Innovative aspects**

- Coupling of discourse-level classifiers with SMT  
- Learning to detect semantic features for connectives, pronouns, compounds, and verbs from parallel corpora

**Commercial application examples**

- Improved MT systems for long coherent texts  
- Post-editing tools for existing MT engines

**More information**


**Software & IPR status**

- Open source software for translation of connectives  
  [https://github.com/idiap/DiscoConn-Classifier](https://github.com/idiap/DiscoConn-Classifier)  
- and their evaluation [https://github.com/idiap/act](https://github.com/idiap/act)

Research software and know-how for other types of words (verbs, pronouns, compounds).
3.22 Face recognition library and awarded ISV technology

Functional description

The Face Recognition Library (FaceRecLib) is designed to perform a fair comparison of face recognition algorithms. It contains scripts to execute various kinds of face recognition experiments on a variety of facial image databases, and running baseline algorithms.

Innovative aspects

The library contains interfaces to many publicly available facial image databases, and default evaluation protocols. Default implementations of face recognition algorithms, which rely on BOB, are provided, e.g., for: Eigenfaces, Linear Discriminant Analysis, Probabilistic Linear Discriminant Analysis (PLDA), Local Gabor Binary Pattern Histogram Sequences, Gabor Graph Matching, Gaussian Mixture Modeling, Inter-Session Variability Modeling (ISV), and the Bayesian Intrapersonal/Extrapersonal Classifier.

Commercial application examples

- Biometrics-enabled identity management systems (Automated Border Control, Access Control, …)
- Multi-factor authentication security systems (Critical Infrastructures, e-Banking, …)
- Forensic Science, Video surveillance, Entertainment, Robotics, Man-Machine interaction

More information


Software & IPR status

Open sourced at https://pypi.python.org/pypi/bob.bio.face but re-licensing possible for commercial purposes.
3.23 Fingervein recognition

Keywords
Fingervein recognition library

Key contact researcher(s)
Dr. Sébastien Marcel
marcel@idiap.ch
Tel.: +41 27 721 77 27

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number:
Software disclosure 10264

Functional description
This library is designed to perform a fair comparison of finger vein recognition algorithms.

Innovative aspects
The library contains interfaces to publicly available fingervein image databases, and default evaluation protocols. Default implementations of vein recognition algorithms, which rely on BOB, are provided, e.g., for: Maximum Curvature, Repeated Line Tracking and Wide Line Detector.

Commercial application examples
– Biometrics-enabled identity management systems (Automated Border Control, Access Control, …)
– Multi-factor authentication security systems (Critical Infrastructures, e-Banking, …)

More information

Software & IPR status
Open sourced at https://pypi.python.org/pypi/xbob.fingervein but re-licensing possible for commercial purposes.
### Palmvein recognition library

#### Keywords
Palmvein recognition library

#### Key contact researcher(s)
Dr. Sébastien Marcel  
marcel@idiap.ch  
Tel.: +41 27 721 77 27

#### Technology Transfer Office
Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

#### Corporate Sponsorship Program
See Section 4 of the present document

#### Functional description
This library is designed to perform a fair comparison of palm vein recognition algorithms.

#### Innovative aspects
The library contains interfaces to publicly available palm vein image databases, and default evaluation protocols. Default implementations of vein recognition algorithms, which rely on BOB, are provided, e.g., for: Maximum Curvature, Repeated Line Tracking, Wide Line Detector and Local Binary Pattern Histogram Sequences.

#### Commercial application examples
- Biometrics-enabled identity management systems (Automated Border Control, Access Control, …)
- Multi-factor authentication security systems (Critical Infrastructures, e-Banking, …)

#### More information

#### Software & IPR status
Open sourced at [https://pypi.python.org/pypi/bob.palmvein](https://pypi.python.org/pypi/bob.palmvein) but re-licensing possible for commercial purposes.
3.25 A method of detecting a falsified presentation to a vascular recognition system

**Functional description**

Biometric vascular recognition is the recognition of an individual through the vascular (veins and arteries) pattern of their hand/finger. If the biometric is stolen it has been shown that it is possible to create a fake. Presenting a fake biometric on the sensor of a biometric recognition system is denoted as spoofing or presentation attack. Anti-spoofing or presentation attack detection, also termed as liveness detection, is the process of detecting a fake biometric. This invention disclosure describes processes to detect a presentation attack with a fake vascular biometric.

**Innovative aspects**

The invention is based on a multi-algorithm and multi-spectrum framework processing both Visual spectra and Near-InfraRed spectra static and dynamic images.

**Commercial application examples**

- Biometrics-enabled identity management systems (Automated Border Control, Access Control, . . . )
- Multi-factor authentication security systems (Critical Infrastructures, e-Banking, . . . )

**More information**

on request (patent pending)
3.26 Scalable linear discriminant analysis for large scale pattern recognition

Keywords

An efficient and scalable version of probabilistic linear discriminant analysis (PLDA).

Key contact researcher(s)

Dr. Sébastien Marcel
marcel@idiap.ch
Tel.: +41 27 721 77 27

Technology Transfer Office

Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program

See Section 4 of the present document

File reference & version number:

Software disclosure 7088

Functional description

This implements probabilistic linear discriminant analysis (PLDA). The implementation is a scalable and efficient implementation.

Innovative aspects

The commercial application is potentially widespread as it provides an exact and scalable way of training and computing the likelihood for PLDA. Thus, it is applicable to many pattern recognition tasks involving discriminant analysis and it has already been shown to be applicable to face and speaker recognition tasks to obtain state-of-the-art performance. Other general pattern recognition and object recognition applications are also foreseen.

Commercial application examples

- Object recognition
- Biometrics-enabled identity management systems (Automated Border Control, Access Control, ...)
- Multi-factor authentication security systems (Critical Infrastructures, e-Banking, ...)
- Forensic Science, Video surveillance, Entertainment, Robotics, Man-Machine interaction

More information


Software & IPR status

Open sourced at https://pypi.python.org/pypi/xbob.paper.tpami2013 but re-licensing possible for commercial purposes.
3.27 Fast Object Detector

Functional description

Object detection in images is a key component of many systems able to process images automatically. This can be for semantic analysis, or as a pre-processing for identification or robotic control. In many practical situations, this operation has to be done extremely rapidly, to allow for instance the processing of video streams in real-time. We have developed a novel strategy to speed up a large class of such methods, which is close to one order of magnitude faster than the best pre-existing algorithms.

Innovative aspects

- One order of magnitude speed-up with respect to pre-existing approaches
- Technique applicable to a large class of detection methods

Commercial application examples

- Pedestrian detection for surveillance in video stream
- Low-power object detection systems

More information


Software & IPR status

The reference implementation is available under the GPL3 license.

- FFLD: https://www.idiap.ch/scientific-research/resources/exact-acceleration-of-linear-object-detectors
- The algorithm is covered by US patent (US20140089365) on “Object detection method, object detector and object detection computer program”.

Keywords

Object detection

Key contact researcher(s)

Dr. François Fleuret
francois.fleuret@idiap.ch
Tel.: +41 27 721 77 39

Technology Transfer Office

Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program

See Section 4 of the present document

File reference & version number:

Software disclosure 9496
3.28 Large-scale Image Retrieval

**Keywords**

Content-based image retrieval; large scale

**Key contact researcher(s)**

Dr. François Fleuret  
francois.fleuret@idiap.ch  
Tel.: +41 27 721 77 39

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document

**File reference & version number:**

Software disclosure 6006

---

**Functional description**

Iterative Relevance Feedback is a software technology to facilitate search in very large image databases. The core idea is to provide the user with an interactive interface which displays candidate images among which he has to pick the most relevant. From that choice, the system updates the scores attached to every image, and proposes a new selection. The approach we develop relies on a sophisticated statistical model to maximize the information gathered from the user at every step.

Our technology can cope with very large databases. Concurrent methods with the same statistical models are limited to a few tens of thousands of images. Our technique allows to scale this up by two orders of magnitude on a single PC, and has the potential to scale by several other orders of magnitude on a dedicated cluster of machines.

**Innovative aspects**

- Use an adaptive granularity for the image set representation
- Scales gracefully to very large data-sets

**Commercial application examples**

- On-line product search
- Interactive person identification
- Image bank browsing

**More information**


**Software & IPR status**

The reference implementation is available under the GPL3 license:

- HEAT system: [http://www.idiap.ch/scientific-research/resources/heat-image-retrieval-system](http://www.idiap.ch/scientific-research/resources/heat-image-retrieval-system)
3.29 Multi-camera Detection and Tracking

Keywords
Multi-camera; pedestrian detection; surveillance; behavior analysis

Key contact researcher(s)
Dr. François Fleuret
francois.fleuret@idiap.ch
Tel.: +41 27 721 77 39

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number:
Software disclosure 7632

Functional description
This technology allows the detection of pedestrians or vehicles from multiple synchronized video streams taken from calibrated cameras looking at a common area of interest. It can either performs the detection in real-time in separate frames, or process a batch of frames to exploit temporal consistency to increase the accuracy.

Innovative aspects
– Robust system with a limited number of parameters
– Can handle very a degraded signal
– Consistent estimates in case of occlusion
– Meaningful probabilistic prediction

Commercial application examples
– Video-surveillance
– Behavioral analysis
– Crowd counting

More information

Software & IPR status
– An implementation of the tracker ("Multi-Tracked Paths") is available under the GPL3 license: http://www.idiap.ch/scientific-research/resources/mtp
– The algorithms were developed in collaboration with the CVLab at EPFL and are covered by two (EPFL) patents:
  • International patent WO2013072401 on “Tracklet-based Multi-Commodity Network Flow for Tracking Multiple People”.
  • US patent US20130177200 on “A method and apparatus for multiple object tracking with k-shortest paths”.
3.30 Real-time Multiple face and head pose tracking

Keywords
Computer Vision, Face tracking, Head orientation

Key contact researcher(s)
Dr. Jean-Marc Odobez
odobez@idiap.ch
Tel.: +41 27 721 77 26

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

Functional description
Idiap has developed a real-time system that can track continuously from 1 to 5 faces, which in addition can extract for each detected face the head orientation, providing some hint on where the person is looking at. The technology does not only rely on face detection but also tracking, since due to the large variety of head poses that people can take (e.g. when looking down at paper, or in some profile views) face detection is rapidly insufficient when a person does not look straight at the camera.

Innovative aspects
– Robust to background noise and illumination variability
– Can tolerate camera motion
– Track faces even when they are not frontal
– Provide head pose cue, and visual focus of attention cues (e.g. are people looking at the camera)
– Robust track creation and deletion mechanisms for lower false alarm rates.

Commercial application examples
Tracking people faces is a prerequisite in many applications before further behavior analysis, e.g. when interacting implicitly (e.g. in the remote conferencing systems) or explicitly (e.g with a robot) with a device.
– HCl application with one-to-many interactions
– Communication systems (e.g. webcam) that can exploit behavior cues for communication improvement (e.g. zooming on individuals)
– Behavior analysis study systems.

More information


Software & IPR status
A demonstration software is available.
3.31 Gaze Estimation and Non-Verbal Behavior Extraction from consumer cameras

Functional description

We have developed softwares for and a long-standing experience with the automatic analysis of a variety of real-life human (non-verbal) behaviors in interaction modeling (attention, head gestures, addressee, engagement,...).

Using RGBD sensors, we have developed precise eye gaze estimation systems comprising the fitting of personalized 3D mesh face models, a real-time estimation head pose tracker, and gaze estimation modules. When relying on RGB cameras only (typically at further distance from the sensor), multiperson real-time head pose trackers have been developed.

In both domains, software have been developed to infer the attention of people towards visual targets (people, objects or scene regions), and recognize head gestures.

Innovative aspects

- No user restriction (non-intrusive system)
- Generic or personalized head and eye gaze models
- Handles gaze tracking under head pose changes
- 3D reasoning (not only for screen-based applications)
- Works under low-resolution eye imaging

Commercial application examples

- Attention modeling (e.g. consumer preference analysis)
- User studies, eg in sociology (NVB and gaze coding; dyadic and group interactions analysis)

More information


Software and IPR status

- RGBD simplified version - http://www.idiap.ch/scientific-research/resources/hg3d
- Gaze tracking under patenting
Functional description

Detecting changes in videos is an essential component of surveillance applications. Idiap has expertise in this domain, with efficient background subtraction schemes and further exploitation of this technology for automatic activity analysis from outdoor or indoor scenes. Several modules have been developed in this direction. The first one detects and analyses changes in perimeters and space surveillance application, allowing for detecting intruders. The second detects static abandoned luggage left by passengers, including the detection of potential people remaining close to these objects.

Innovative aspects

- Robust multi-cue background subtraction scheme
- Light robust multi-tracker scheme
- Combination of long-term cues (background) and dynamic cues (tracklets)

Commercial application examples

- Public space monitoring (luggage detection)
- Perimeter or indoor monitoring (intrusion detection)

More information


Software status

Demonstration software validated in real conditions.
3.33 Unsupervised Activity Discovery, Monitoring, and Abnormality Detection

Keywords
Activity Discovery, Domotic, Surveillance, Abnormality detection.

Key contact researcher(s)
Dr. Jean-Marc Odobez
odobez@idiap.ch
Tel.: +41 27 721 77 26

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number:
Software disclosure 11637

Functional description
Idiap had developed methods and software for the unsupervised discovery of recurrent temporal pattern in multivariate time series, where observed values are caused by the superposition of multiple phenomena. This typical situation occurs when multiple sensors record the activities of multiple objects/people, like multimodal sensors (proximity, water, light sensors, etc) in domotics, multiple cameras -or even microphone-arrays- in indoor/outdoor complex scenes. Discovered patterns correspond to recurrent activities like trajectories of different object classes in cameras, or sequences of activated sensors in domotics. In addition, the method allows the identification of abnormal situations, and can be exploited to select streams to be displayed in control rooms of large public spaces (to direct the attention of operators towards relevant information).

Innovative aspects
– Capture sub-activities temporal ordering
– No Prior on scene needed
– Small number of parameter to set
– Estimate automatically the numer of activities

Commercial application examples
– Activity discovery from large multi-sensor datasets
– Surveillance (anomaly detection, stream selection)
– Video content filtering (summary generation)
– Adaptation to user’s behaviors (heating,...)

More information

Software status
– Recurrent pattern learning http://www.idiap.ch/resource/research/probamod-v1/
– More complete prototype for surveillance videos
3.34 Motion Detection, Segmentation, and Estimation. Motion Stabilization

**Functional description**

Idiap has expertise and software on video motion analysis, which is important in both the surveillance and video content analysis domains. Robust motion estimation is often a way to stabilize shaky videos, and can be exploited as such or prior to further video content analysis tasks.

**Innovative aspects**

- Robust automatic estimation of dominant motion (works even in presence of secondary motion sources)
- Accurate image motion estimation
- Proven efficiency in both consumer video application (video stabilization) and industrial background (Alcatel, Ifremer,...)
- Real-time

**Commercial application examples**

- Video Stabilization (e.g. from mobile or table video acquisition)
- Surveillance

**More information**


**Software status**

- Commercial motion estimation and stabilization software available.
- Non-exclusive license granted to Cinetis SA.
**3.35 Text detection and recognition in images and videos**

**Keywords**

Computer Vision, text detection, text recognition, video content analysis.

**Key contact researcher(s)**

Dr. Jean-Marc Odobez  
odospez@idiap.ch  
Tel.: +41 27 721 77 26

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document

**File reference & version number:**

Licence 20121015-A01

---

**Functional description**

Consumer pictures and broadcast videos often contain text that provide semantic information about the media content (e.g. name of a person appearing in the video). Idiap has developed a set of robust tools for detecting and transcribing the text present in such media, which have been validated in several European research projects and Swiss technology transfer projects.

**Innovative aspects**

- Works with non black/white content (colored text, text superimposed on textured background)
- Multihypothesis text segmentation and transcription
- High text recall

**Commercial application examples**

- Broadcast data analysis
- Scanned document analysis

**More information**


**Software status**

A demonstration software is available.

**IPR status**

A non-exclusive license has been granted to Klewel SA.
3.36 Soft-skill analytics from online video job interviews

**Functional description**

A system to profile, filter, screen, compare, search, and rank job candidates on the basis of soft-skills analytics. Soft-skills is an umbrella term to refer to a cluster of personality traits, communication, language, interpersonal skills, teamwork, leadership, etc. that characterize people in the working place, that are as relevant as technical or hard skills, and are typically not shown in a resume. We integrate online-video recordings, a crowdsourcing system that uses a crowd to collect a ground-truth of soft-skills, a set of machine learning algorithms to estimate soft-skills based on annotations and automatic methods that extract nonverbal and verbal behavioral cues, and a suit of machine machine learning algorithms to compare, search, and rank job candidates on the basis of soft-skills.

**Innovative aspects**

- Crowdsourcing of high-agreement soft-skill annotations
- Consistent measures of soft skills on the basis of behavioral video data
- Scales the understanding of soft-skills from personal video collections

**Commercial application examples**

- Online video job interviewing processes

**More information**


**Software & IPR status**

- Internal Invention Disclosure
3.37 Activity Monitoring from Ambient Sensors

Functional description

We have developed several algorithms related to the analysis of activities from ambient sensors. These algorithms allow the identification of recurrent activity patterns and the detection of changes in these activity patterns. Ambient sensing represents a challenge in terms of sensor types and configurations: ambient sensors can measure different types of information and work at different scales and precisions. This heterogeneity has been addressed by developing a sensor abstraction layer that allows for transversal comparisons between activity levels estimated in different environments.

Innovative aspects

– Detection of abnormal activity patterns
– Sensor abstraction layer
– Visualisation of activity patterns

Commercial application examples

– This solution is currently used by our industrial partner to remotely monitor elderly people

More information

Ambient sensing has been addressed during the CTI funded project: DomoCare - A new Home Care Preventive Protocol.

Software & IPR status

The Software is fully owned by our industrial partner. We have acquired a significant expertise in working and mining ambient sensor data that can be applied to any scenario involving ambient activity sensing.
3.38 Social context inference from proximity sensors

**Keywords**

proximity sensors, mobile phones, wearables

**Key contact researcher**

Prof. Daniel Gatica-Perez  
gatica@idiap.ch  
Tel.: +41 27 721 77 33

**Technology Transfer Office**

Dr. Florent Monay  
tto@idiap.ch  
Tel.: +41 27 721 77 72

**Corporate Sponsorship Program**

See Section 4 of the present document.

**File reference & version number:**  
N/A

**Functional description**

A method to infer social context from proximity sensors like bluetooth, for example embedded on mobile phones but also in wearable devices. The method uses longitudinal proximity data generated by a population of mobile device users and generates, for each member of the population, the groups of people with whom each user is in proximity, and the typical times when this occurs. The method can be used for group discovery or for event mining applications.

**Innovative aspects**

– Statistical machine learning approach that discovers groups and their temporal context from raw observations of physical proximity.
– The method is applicable to various types of sensor data related to physical proximity.

**Commercial application examples**

– Group discovery and event mining in large gatherings.
– Organizational mapping.

**More information**


**Software & IPR status**

– Matlab code
3.39 Robot Skills Transfer Toolbox

Keywords
Human-robot skills transfer; robot programming interfaces; optimal control; adaptive movements; gesture synthesis

Key contact researcher(s)
Dr. Sylvain Calinon
sylvain.calinon@idiap.ch
Tel.: +41 27 721 77 61

Technology Transfer Office
Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program
See Section 4 of the present document

File reference & version number:
Software disclosure 11619

Functional description
Development of technologies to facilitate the re-programming of robots. We can provide tools and expertise for the automatic analysis and generation of movements, gestures and manipulation skills. The developed software components rise from the cross-fertilization of statistical learning, dynamical systems and optimal control, enabling skill transfer techniques that rely on the observation of human demonstrations and on intuitive interaction with the robot.

Innovative aspects
– Layman interface for robot programming
– Automatic adaptation of movements to new situations
– Handling of task variations and options
– Safe robot controlled by minimal intervention principle

Commercial application examples
– Human-robot skill transfer
– Optimal control based on human demonstrations
– Motion analysis and synthesis

More information

Software & IPR status
The underlying software comes in two distinct and independent versions in Matlab and C++ to facilitate both analysis and integration aspects. The Matlab version is fully compatible with the GNU Octave open source software. The C++ version is a library with minimal dependencies to facilitate its inclusion in other softwares. An independent and optional frontend GUI is available for monitoring and fast prototyping purposes.
3.40 Cardiac Imaging

Keywords

Biological microscopy, heart, cardiac, image processing, image reconstruction, dynamic imaging, movie synchronization, 3D imaging

Key contact researcher(s)

Dr. Michael Liebling
michael.liebling@idiap.ch
Tel.: +41 27 721 77 82

Technology Transfer Office

Dr. Florent Monay
tto@idiap.ch
Tel.: +41 27 721 77 72

Corporate Sponsorship Program

See Section 4 of the present document

File reference & version number:

N/A

Functional description

Development of imaging protocols and reconstruction algorithms for 3D cardiac imaging.

Innovative aspects

– Volumetric and dynamic imaging of the heart
– Extraction of cardiac function parameters
– Minimal user interaction

Commercial application examples

– Combined microscopy-software tools for live biological imaging
– 3D image reconstruction from multiple views

More information

Software & IPR status

The underlying software consists of a set of Matlab, C, and Java tools to perform problem-specific microscopy acquisition protocols, digital image reconstruction of dynamic 3D volumes, analysis and display routines.

Note on IPR: Ongoing efforts, part of software suite was developed while at Caltech and UCSB. Some of the methods are covered by patent:

4 CORPORATE SPONSORSHIP PROGRAM

4.1 Corporate Sponsoring Program

While Idiap is very active in fundamental research and in the development of large collaborative research programs, we are also proud to be very much system driven, i.e., driven by shared long-term visions and continuous by-product applications. We are also very well aware of the need of continuous feedback from (and interaction with) industry, through (1) collaborative projects, (2) creation of spin-offs or incubation of young start-ups (3) exchange programs, (4) corporate sponsorship, and (5) partnership or sponsorship with the Swiss Center for Biometrics Research and Testing:

1. Collaborative projects: Idiap currently has several (usually well focused) research projects directly funded by industries. In this case, the research project is precisely defined in collaboration with the two parties and resulting IPR is discussed on a case-by-case basis.

2. Creation of spin-offs or incubation of young start-up: In collaboration with the State of Valais, Idiap has set up its own technology transfer and start-up incubation instrument through an independent company IdeArk. For more information, see http://www.ideark.ch.

3. Exchange program: Although also part of the current Corporate Sponsoring Program (limited to industries), Idiap has set a large scale international visitor program.

4. Corporate sponsorship program: In addition to the above, Idiap also have a specific corporate sponsorship program to boost long-term interaction with industries through training of industrial visitors, joint research programs, etc. While depending on the level of sponsoring and the goals being sought, the main mission of the present sponsoring programs can be summarized as follows:

- To ensure complementary funding for Idiap, while fully benefiting our corporate sponsors and optimizing technology transfer opportunities.
- To allow the industry to be involved with current research.
- To offer a platform for information, meeting, cooperation and interaction between industrial and research partners.
- To encourage meeting and interaction between different industrial partners.

5. The Swiss Center for Biometrics Research and Testing (http://www.biometrics-center.ch) is a center for biometric excellence established to advance biometric security technologies, mainly by facilitating collaboration between industry and academia.

Depending on the importance of their financial commitment and the strength of the established relationship, Idiap offers companies the opportunity to be involved at different levels, as discussed next.

[Note that all amounts mentioned below are exclusive of VAT.]
4.2 Idiap Support Members

(In principle, this status is reserved for small businesses)

Advantages

Regular information on the activities of Idiap:

1. Provision of publications and research reports.
2. Invitation to participate in the annual Idiap Symposium. This provides a forum for the demonstration of recent developments, as well as the interaction between researchers and the different partners.
3. Facilitated access to software developed by Idiap (in so far as the software in question is free of rights). In the event a significant level of support is requested, a specific agreement will have to be established.
4. Access to Idiap for advice and information on technologies within the domain of Idiap’s research. If consultation requiring a significant study is requested, a specific agreement will have to be established.
5. Mention as a support member in the annual report of Idiap.

Contribution

Payment of a minimum amount of CHF 10,000.- per year.

Duration

Automatic renewal subject to cancellation with 3 months advanced notice.
4.3 Idiap Affiliated Members

**Advantages**

1. Regular information on the activities of Idiap.
2. Provision of publications and research reports.
3. Invitation to participate in the annual Idiap Symposium. This provides a forum for the demonstration of recent developments, as well as the interaction between researchers and the different partners.
4. Facilitated access to software developed by Idiap (in so far as the software in question is free of rights). In the event a significant level of support is requested, a specific agreement will have to be established.
5. Access to Idiap for punctual advice and information on technologies within the domain of Idiap’s research. If consultation requests requiring a significant study of work is or if requests are repeatedly placed, a specific agreement will have to be established.
6. Ability to invite Idiap researchers for conferences.
7. On request, and subject to the agreement of those involved, access to the CVs of doctoral students.
8. A position for one or more scientific visitors, corresponding to a maximum of one person per year. The visitor(s) will have access to a work-station and administrative support, as well as the computer resources and databases of Idiap. They will be integrated into the research groups of Idiap and the eventual result of their contribution will belong to Idiap. The institute will be consulted on the choice of visitors, and will have the right of veto. The visitor will receive an allowance from Idiap (CHF 2,500.- per month). The remaining expenses and remuneration of the visitor are the responsibility of the sponsoring company.
9. Mention as an affiliated member in the annual report of Idiap.

**Contribution**

Payment of an amount of CHF 60’000.- /year (including allowance paid to visitor(s)).

**Duration**

To take account of the stability required for research, in principle a minimum 2 year term applies, with automatic renewal subject to cancellation with 12 months advanced notice.
4.4 Idiap Sponsor Members

Advantages

1. Regular information on the activities of Idiap.
2. Provision of publications and research reports.
3. Invitation to participate in the annual Idiap Symposium. This provides a forum for the demonstration of recent developments, as well as the interaction between researchers and the different partners.
4. Facilitated access to software developed by Idiap (in so far as the software in question is free of rights). In the event a significant level of support is requested, a specific agreement will have to be established.
5. Access to Idiap for advice and information on technologies within the domain of Idiap’s research. If consultation requiring a significant study is requested, a specific agreement will have to be established.
6. Ability to invite Idiap researchers for conferences.
7. On request, and subject to the agreement of those involved, access to the CVs of doctoral students.
8. Positions for scientific visitors, corresponding to a maximum of two full positions per year. The visitors will have access to a work-station and administrative support, as well as the computer resources and databases of Idiap. They will be integrated into the research groups of Idiap and the eventual result of their contribution will belong to Idiap. The Institute will be consulted on the choice of visitors, and will have the right of veto. The visitor will receive an allowance from Idiap (CHF 2,500.- per month). The remaining expenses and remuneration of the visitor are the responsibility of the sponsoring company.
9. Ability to request that the visitors be associated with particular research directions, at the complete discretion of Idiap.
10. Mention as a sponsor member in the annual report of Idiap.

Contribution

Payment of an amount of CHF 120'000.-/year (including allowance paid to visitor(s)).

Duration

To take account of the stability required for research, in principle a minimum 2 year term applies, with automatic renewal subject to cancellation with 12 months advanced notice.
4.5 Research Partners

Research Partners are companies involved with particular research projects which finance a minimum of two researcher positions with a related contribution of at least CHF 350,000.-/year during a minimum term of 2 years.

As the main partners of Idiap, Research Partners receive all the general benefits offered to other partners. Research Partners have an observers seat on the Board of Management (except for closed sessions).

The exact benefits granted, as well as the amount of the contribution, will be determined in the specific agreements established with each Research Partner.

In particular, the agreement will determine the definition of the research project and, to the extent appropriate, the terms and conditions in connection with the exploitation of the results. The agreement also provide the way the association of the Research Partner with the project is presented and manifested.
4.6 Swiss Center for Biometrics Research and Testing

As part of Idiap Research Institute, the Swiss Center for Biometrics Research and Testing is a center for biometric excellence established to advance biometric security technologies, mainly by facilitating collaboration between industry and academia. Its core missions are to carry out high quality research, train talented engineers and researchers, foster technology transfer from academia to industry, and to propose evaluation and testing services in the domain of biometric security.

The center promotes two instruments: a cooperative research consortium and an evaluation platform for biometrics.

1. The center operates the European Cooperative Identification Technology Research Consortium (EAB-CITER). EAB-CITER is split between partners and affiliates to ensure an environment of mutual benefit. While the partners create research proposals for funding consideration according to the affiliates’ needs, the affiliates fund and drive the research carried out among the partners.

   This structure helps industry partners to gain unparalleled access to cutting-edge researchers and their expertise and supports researchers to acquire the opportunity to bid for funding from driven and relevant organizations. Through this multidisciplinary environment, it is expected that organizations will enjoy a golden age of biometric development, leading to rapid improvements in current biometric technology and the realization of novel technologies.

2. The center initiated and maintains the Biometrics Evaluation and Testing (BEAT) platform. The BEAT platform ([www.beat-eu.org/platform](http://www.beat-eu.org/platform)) provides easy online access to experimentation and testing in biometrics and computational science in general.

   BEAT also provides a mechanism to generate online attestations (certifications) to reproducible results or evaluations in biometrics. Attestations (e.g., ISO/IEC 19795-1) can be used to support formal certification processes or improve evidence on published academic articles.

EAB-CITER – Affiliates Members

An “Affiliate” member (Privileged or Regular) is a company, public research and development organization or agency, or any government-owned contractor operated laboratory affiliated to EAB-CITER via the affiliate agreement.

Advantages

1. Advisory board: right to assign one representative for regular affiliates and two representatives for privileged affiliates to the advisory body, designated to decide on the allocation of resources to research projects.
2. Licensing rights for regular affiliates: right to obtain a non-exclusive, worldwide, royalty-free, perpetual commercial license to funded-only results.
3. Licensing rights for privilege affiliates: right to obtain a non-exclusive, worldwide, royalty-free, perpetual commercial license to any results, sub-licensable to wholly affiliated entities.

Contribution

- Regular affiliates: Payment of an amount of CHF 5’000.- /year.
- Privileged affiliates: Payment of an amount of CHF 20’000.- /year.

Duration

2-year term, with automatic renewal subject to cancellation with 3 months’ advanced notice.

---

2[https://www.beat-eu.org/platform/attestations/1721510690](https://www.beat-eu.org/platform/attestations/1721510690)
EAB-CITER – Partners Members

A “Partner” member is an academic institution, a non-for-profit research institution or an individual researcher or scientist associated with the Center by means of the partnership agreement.

Advantages

1. Research funding: Right to propose research projects to affiliates for funding
2. Intellectual property: Ownership of all intellectual property on results
3. Possibility to propose activities to be sponsored

Contribution

- CHF 0.-/year: bring two affiliates
- CHF 800.-/year: bring no affiliates

Duration

2-year term, with automatic renewal subject to cancellation with 3 months’ advanced notice.

BEAT platform

The Swiss Center for Biometrics Research and Testing will rely on a sponsorship program:

http://www.biometrics-center.ch/sponsors/development-of-beat-platform

to fund development activities related to the platform but also to offers services such as:

1. Allocating dedicated computing resources for evaluation
2. Hosting of biometric databases
3. Running International evaluations
4. Licensing the platform (including support) to third parties that are interested to host the platform on their facilities.
Contact:
  Prof. Hervé Bourlard, bourlard@idiap.ch
Technology Transfer Office: tto@idiap.ch
  Dr. Florent Monay