Novel Methods for Forensic Multimedia Data Analysis: Part I

Petra Perner

Abstract

The increased usage of digital media in daily life has resulted in the demand for novel multimedia data analysis techniques that can help to use these data for forensic purposes. Processing of such data for police investigation and as evidence in a court of law, such that data interpretation is reliable, trustworthy, and efficient in terms of human time and other resources required, will help greatly to speed up investigation and make investigation more effective. If such data are to be used as evidence in a court of law, techniques that can confirm origin and integrity are necessary. In this chapter, we are proposing a new concept for new multimedia processing techniques for varied multimedia sources. We describe the background and motivation for our work. The overall system architecture is explained. We present the data to be used. After a review of the state of the art of related work of the multimedia data we consider in this work, we describe the method and techniques we are developing that go beyond the state of the art. The work will be continued in a Chapter Part II of this topic.

Keywords: multimedia forensic data analysis, standardization of forensic data analysis, video and image enhancement, video analysis, image analysis, speech analysis, case-based reasoning, multimedia feature extraction, handwriting, Twitter data analysis, novelty detection, legal aspects

1. Introduction

The objective of this work is to provide novel methods and techniques for the analysis of forensic multimedia data. These methods and techniques should form a novel toolkit for automatic forensic multimedia data. The data modalities the proposed work is considering are images and videos, text, handwriting, speech and audio signals, social media data, log data, and genetic data. The integration of methods for all these different data modalities in one tool kit should allow the cross-analysis of these data and the detection of events by interlinking between these data. The proposed methods will face on standard forensic tasks, for example, identification of events, persons, or groups and device recognition. Together with the end users and the police forces, new standard tasks will be worked out during the project and will give a new input to the standardization aspect of forensic data analysis.

The proposed novel methods and techniques will consider all aspects of multimedia data analysis such as device identification and trustworthiness of the data, signal enhancement, preprocessing, feature extraction, signal and data analysis, and interpretation.

Techniques for detecting artifacts in images and videos are of paramount importance. To trust the information extracted from images and videos, it is necessary to make sure that the image and video have been recorded by a camera, and that no artifact has been added. The detection of artifacts is a key element to use an image or a video in court. Thus, it should be clearly assessed the integrity of images and videos used as a proof of evidence.

In most image applications, the acquired images represent a degraded version of the original scene. Degradation in such images may appear in different forms. These types of degradations must be removed before the images are used for classification or decision making.

Novelty detection for the identification of novel situation and tasks will be another task that will be important in forensic applications, where the victims or events are very flexible. It will allow to identify new tasks, and by doing so, it will be an automatic method to improve standardization of the analysis of forensic data.

We will also develop learning methods to include new data into the existing cases and summarization of new and old cases into more general cases applicable to a wider range of tasks for further law purposes. For that, novel case-based reasoning methods will be developed that can keep the cases based on their multimedia features and specific event features in a case base, so that they can be easily retrieved and applied for new situations. The case-based reasoning system will consist of novel probabilistic and similarity-based methods. It will provide a wide range of novel similarity-based reasoning methods for the different feature types for identification and similarity determination. A special taxonomy for similarity determination and measures will be worked out and implemented in the CBR system. It will provide explanation capabilities for similarity and as those it will help a forensic data analyst to identify the right reasoning method for his particular problem. This aspect goes along with the training and education aspect for forensic data analysis. Part of this will be self-contained in the chosen methods and realized by the system.

In Section 2, the background and the motivation of our work will be described. Taking into account the special needs for multimedia forensic analysis, identification, and recognition system, we develop a novel architecture based on case-based reasoning. The data used are described in Section 3. Related work and the progress we want to make with our work are described in Section 4. This work does not only take into account to develop novel methods and techniques for multimedia content processing and reasoning, but we are also taking into account the legal aspect that is going along with processing sensible data. Finally, we given conclusions in Section 5. This chapter is continued in the Chapter Part II of Novel Methods for Forensic Multimedia Data Analysis.

2. Background, motivation, and overall system architecture

The analysis of multimedia data has to consider different aspects of the modalities of the data. We want to deal with images and videos, text, handwriting, speech and audio signals, social media data, log data, and genetic data. The idea is to come up with an automatic system that should cover all aspects of data analysis for the different modalities from the signal enhancement, preprocessing, feature extraction to the analysis, and interpretation. This includes image enhancement in order to eliminate the degradation in an image that might appear because of a known or an unknown blurring function, which leads to the

consideration of deconvolution and blind deconvolution problems or because of very low resolution devices, which lead to the combination of several low resolution images to obtain a high resolution one, the so called, super-resolution problem or to the utilization of highly compressed images, which suffer from compression artifacts.

Techniques for detecting artifacts in images and videos will be developed to trust the information extracted from images and videos. They should allow to make sure that the image and video have been recorded by a camera, and that no artifact has been added.

Feature extraction will be the selection of a set of sufficiently low- and high-level features in order to complement the existing standards for image, video, and audio data, with the aim at enabling novel and robust classification and recognition methods. They should allow modeling the standard tasks for forensic data analysis known so far but should be flexible enough to cover the needs of newly arising task.

Twitter was actively used by rivaling gang members to plan their assaults. Twitter data are hard to analyze because the text fragments are very short, multiple persons can be involved in a conversation about various topics, and the data are rapidly changing. Novel methods are necessary, which can be used to monitor in real-time Twitter and identify potential threats including individuals and communities of users who are planning illegal activities.

Furthermore, we plan to build a dynamic model on Twitter text to forecast the upcoming significant events and emotions of the crowd associated with these events. While there can be many events with strong presence in Social Media, some of them would have stronger negative emotions associated with them. These events are candidates that may have criminal nature or significant social consequences.

The huge amount of CCTV systems has increased the importance of video and image evidence in forensic labs. An automatic system should be able to select heads, vehicles, license plates, guns, dresses, and all other objects that can link a person to the event.

An important main focus of police work is the identification of people for which a decision of the public prosecutor's office or a judge to the observation or an arrest warrant was issued. Within the scope of this arrangement, the use of video supervised places and facilities, or at before not known places, the application of mobile video technology should occur for this purpose. The aim is to develop methods and procedures for an automatic system for identification of one or several target people in mobile video recordings based on passport photos or other available pictures.

A significant portion of data collected by Law Enforcement Agencies consists of speech and audio files. They form an important part of legal cases. Speech recognition systems (such as dictation systems) are now available in many languages. However, continuous spontaneous speech recognition is still an unsolved problem. Novel methods for the recognition of continuous spontaneous speech and other audio signals are necessary.

While the commercially available optical character recognition systems are very successful for printed documents, recognition of words in unconstrained settings or "in the wild" still is an open problem, and recognition of handwritten text continues to be a challenge. We propose to develop novel Handwriting Recognition Methods for unconstrained settings.

Novel Case-Based Reasoning (CBR) methods will be developed for the recognition, interpretation, and identification task. Case-based reasoning explicitly

uses past cases from the domain expert's successful or failing experiences. CBR is very useful in applications, where generalized knowledge is lacking. Therefore, case-based reasoning can be seen as a method for problem solving as well as a
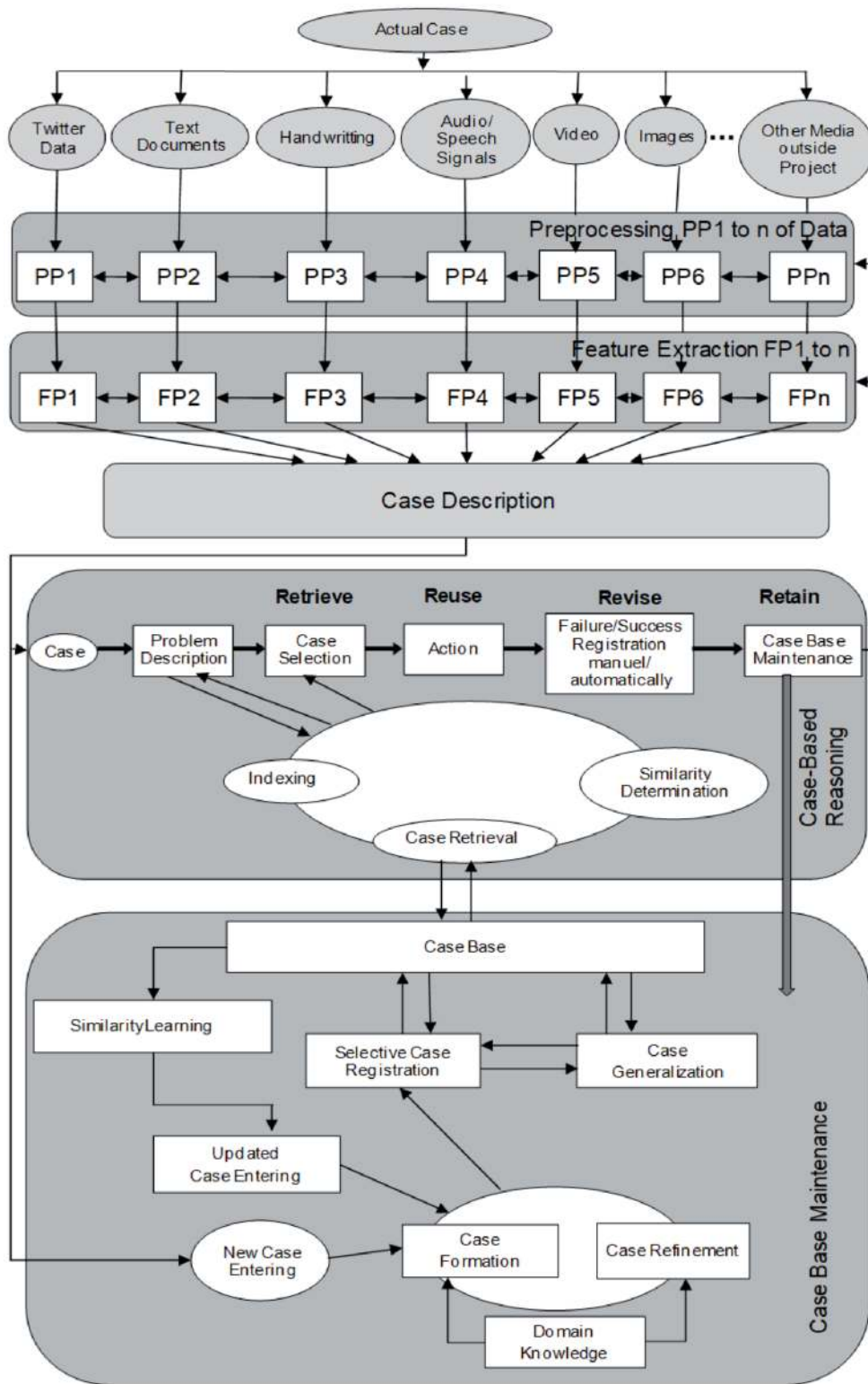

Figure 1.
System overview.

method to capture new experiences and make them immediately available for problem solving. It can be seen as a learning and knowledge discovery approach since it can capture from new experiences some general knowledge such as case classes, prototypes, and some higher-level concept. All these points make a CBR system very useful for the analyses of forensic data. The method is able to capture new cases and store new and old cases in a summarized way, so that they can be easily retrieved or used for reasoning. The reasoning methods are based on similarity that makes it very useful to detect and identify similar and identical cases without having generalized knowledge. Different similarity measures have to be developed that can deal with the different modalities of data and their case representation. A taxonomy of similarity will be developed that explains the relation, usefulness, and application of the different similarity measures to the data that will help a forensic data analyst to efficiently apply these reasoning methods to his problem.

All the above-mentioned facts result in the overall system architecture given in **Figure 1**. The architecture consists of the three main processing units: media preprocessing, feature extraction, and decision unit based on case-based reasoning. The input is the different media data. The architecture is open, so that new input media data can be considered when the necessary processing modules are available. The outcome of the preprocessing and the feature extraction unit is a description of the different media data by sufficiently low- and high-level features that will be combined to the case representation. The reasoning will be done by the case-based reasoning unit based on formerly calculated case representation. The reasoning will be the identification and recognition of the objects or scenario's as well as the detection of novel events. The CBR unit will be criticized based on the result of the action, and the decision of the CBR unit has been proposed. Depending on that outcome, case-based maintenance will be done. New case will be stored in the case base, the similarity measure will be updated or changed, or case generalization will be done.

Besides the development of novel processing and reasoning methods, it is necessary to develop a legal framework regulating the process of gathering, processing, analyzing, and integrating multimedia data.

3. Data used

Different types of security-related data will be used for the work provided by the end users:

- Passive millimeter-wave (PMMW) images and video are used for security screening as many materials, including clothing, are transparent to millimeter-waves. The imagers that use this technology, such as those developed by ALFA [1], are therefore installed at security checkpoints to screen people for hidden weapons (including powders, liquids, and gels) and contraband. They are characterized by a low resolution compared to visible images, due to the wavelength used. ALFA's current software automatically detects objects within the spatial and thermal resolution of the system and draws a red box around them. Some examples of this image type are given in **Figures 2–4**. These are then represented at the approximate locations on a generic silhouette to preserve the subject's privacy. However, object classification to automatically distinguish between a threat and a nonthreat object is not currently performed. A new system will be developed to make a classification based on the shape and size of

the objects detected in the raw millimeter-wave image. This would reduce the number of false alarms.

- Anonymous Data from Text will be collected. These data are freely available on the Web. We propose to perform initial experiments on anonymized data to validate the feasibility of our approach. After authorization of the responsible superiors of the cybercrime unit is obtained, we will use the developed system for real-life investigations.

- A Telekom company will prepare a speech database obtained under various conditions and under various speech coders and encoders to test the new algorithms.
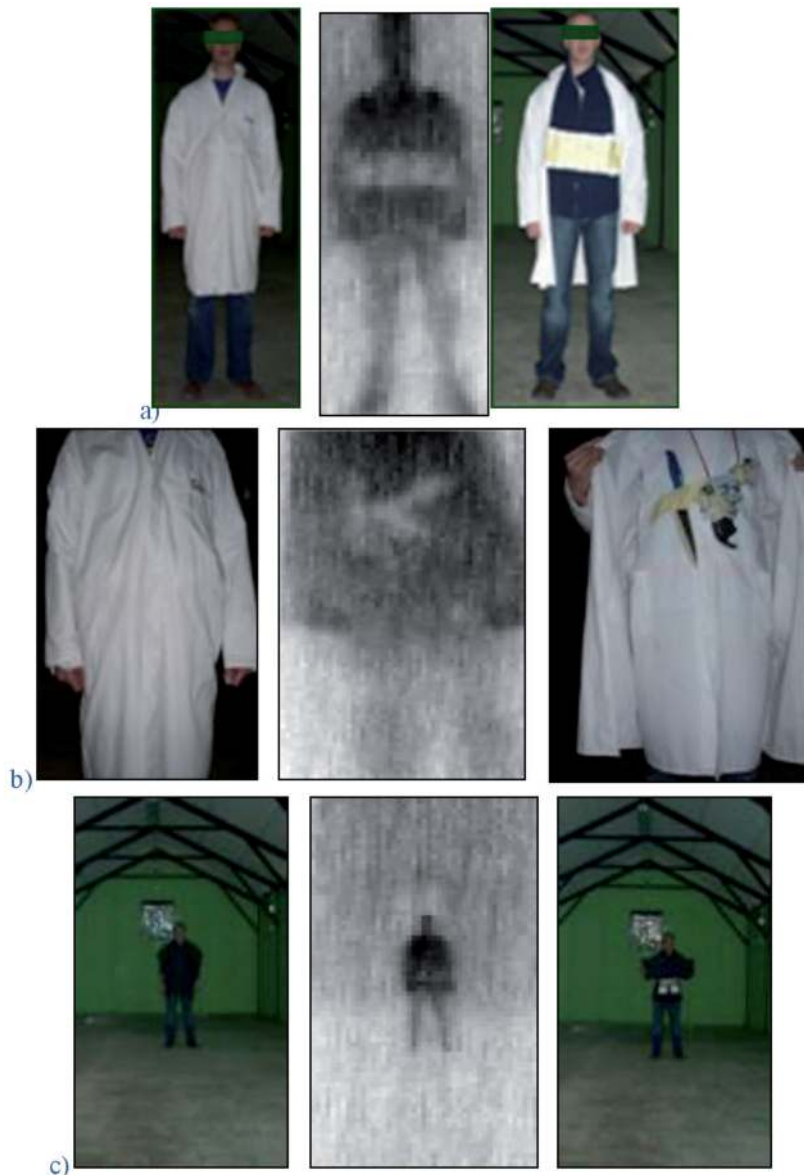

Figure 2.
(a) Left: clothed subject; center: raw millimeter-wave image of subject; right: subject showing hidden suicide bomber belt; (b) left: clothed subject; center: raw millimeter-wave image of subject; right: subject showing hidden gun and knife; (c) left: clothed subject at 10 m; center: millimeter-wave image of subject at 10 m; right: subject showing two hidden bags of powder explosives. Subject with gel pack hidden between the legs and automatic millimeter-wave detection marked + raw millimeter-wave image of subject; right: subject with gel pack hidden under the arm and automatic millimeter-wave detection marked + raw millimeter-wave image of subject.

Figure 3.
Automatic object and potential threat detection (ATD) on processed millimeter-wave image on the left and privacy protection output to operator on the right.
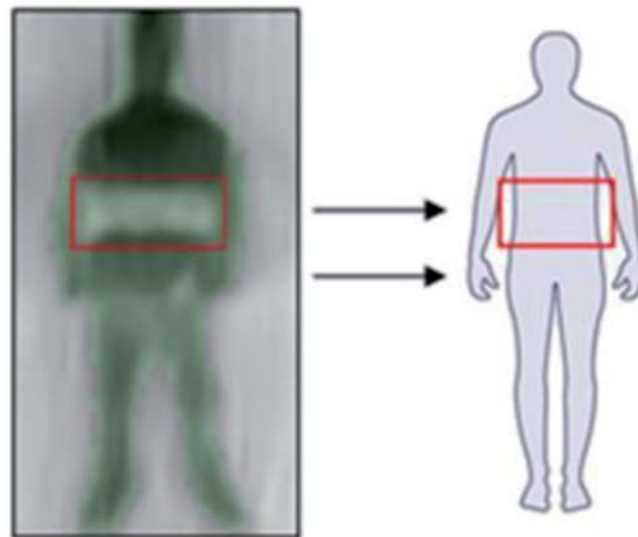

Figure 4.
Person with hidden object around the hip.

- Video and Image databases with case scenarios will be provided by police forces.

- Handwriting documents will be collected through the involvement of graduate and undergraduate students. We also plan to use the following benchmark data set: IAM Database for Off-line Cursive Handwritten Text http://www.iam.unibe.ch/~zimmerma/iamdb/iamdb.html. The database contains the forms of unconstrained western handwritten text. It includes 27,000 isolated words (400 pages).

4. Related work and progress

4.1 Video and image enhancement, filtering, and assessment

4.1.1 State of the art

In most image applications, the acquired images represent a degraded version of the original scene. These applications include astronomical imaging [2] (e.g., using

ground-based imaging systems or extraterrestrial observations of the earth and the planets), commercial photography [3, 4], surveillance and forensics [5, 6], medical imaging [7] (e.g., X-rays, digital angiograms, autoradiographs, MRI, and SPECT), and security tasks where commercial photography and other image modalities like Synthetic Aperture Radar (SAR) [8] and Passive Millimeter (PMMW) [9] are frequently used.

Degradations in such images may appear in different forms. They may be due to a known or an unknown blurring function that leads to the consideration of deconvolution [9–13] and blind deconvolution [3, 14] problems. They may also be due to the use of very low-resolution devices, which lead to the combination of several low-resolution images to obtain a high-resolution one, the so called, super-resolution problem [15, 16] or to the utilization of highly compressed images, which suffer from compression artifacts [17]. These types of degradations must be removed before the images or video sequences are used for classification or decision making. Interestingly, all the problems described above can be formulated within the Bayesian framework [18–20]. A fundamental principle of the Bayesian philosophy is to regard all parameters and unobservable variables as unknown stochastic quantities, assigning probability distributions based on subjective beliefs. Thus, the original image(s), the observation noise, and even the function(s) defining the acquisition process are all treated as samples of random fields, with corresponding prior probability density functions that model our knowledge about the imaging process and the nature of images.

4.1.2 Beyond the state of the art

Once the problem is modeled, inference is then needed. The recently developed variational Bayesian methods have attracted a lot of interest in Bayesian statistics, machine learning, and related areas [18–20]. A major disadvantage of traditional methods (such as expectation maximization (EM)) is that they generally require exact knowledge of the posterior distributions of the unknowns, or poor approximations of them are used. Variational Bayesian methods overcome this limitation by approximating the unknown posterior distributions with simpler, analytically tractable distributions, which allow for the computation of the needed expectations and therefore extend the applicability of Bayesian inference to a much wider range of modeling options: more complex priors (which are very much needed in applications involving images) modeling the unknowns can be utilized with ease, resulting in improved estimation accuracy.

Techniques for detecting artifacts in images and videos are of paramount importance. In order to trust the information extracted from images and videos, it is necessary to make sure that the image and video have been recorded by a camera, and that no artifact has been added. The detection of artifacts is a key element to use an image or a video in court. Thus, the integrity of images and videos used as a proof of evidence should be clearly assessed. The trustworthiness of images and videos has clearly an essential role in many security areas, including forensic investigation, criminal investigation, surveillance systems, and intelligence services.

As stated by Mahdian and Saic [21], verifying the integrity of digital images and detecting the traces of tampering without using any protecting pre-extracted or pre-embedded information have become an important research field of image processing. We will utilize and develop blind methods for detecting image forgery, that is, methods that use the image function to perform the forgery detection task. These methods are based on the fact that forgeries bring into the image-specific detectable changes (e.g., statistical changes). In high-quality forgeries, these changes cannot be found by visual inspection. Existing methods mostly try to identify various traces

of tampering and detect them separately. The final decision about the forgery can be carried out by fusion of results of separate detectors.

Blind methods can be classified into several categories. In detection of near-duplicated image regions, a part of the image is copied and pasted into another part of the same image with the intention to hide an object or a region. There are methods capable of detecting near duplicated parts of the image that usually require a human interpretation of the results, see Refs. [21–23]. A different category includes interpolation and geometric transformation that are typically based on the resampling of a portion of an image onto a new sampling lattice, see, for example, Ref. [24]. In the photomontage detection problem, one of the fundamental tasks is the detection of image splicing, which can sometimes be based on analyzing the lighting conditions. Another category is related to compression method. In order to alter an image, typically the image is loaded to photoediting software, and once the changes are done, the digital image is resaved. Methods capable of finding the image compression history can be helpful in forgery detection. Another important category is the study of the noise characteristics and the chromatic aberrations [25, 26]. In the same line, blur and sharpening can also be analyzed to detect the concealment of traces of tampering. When two or more images are spliced together, it is often difficult to keep the appearance of the image correct perspective. Applying the principles from projective geometry to problems in image forgery detection can be also a proper way to detect traces of tampering. There are also other groups of forensic methods effective in forgery detection, see, for instance, single-view recaptured image detection, aliveness detection for face authentication, and device identification in digital image forensics, Refs. [27–30].

4.2 Case-based reasoning

4.2.1 State of the art

Case-Based Reasoning has been shown a successful problem-solving method in different applications were generalized knowledge is lacking. CBR has been used to interpret images [31, 32], 1-D signals [31, 33, 34], and text cases [35]. It also has been used for meta-learning of the best parameter of image segmentation [36] and classification methods [37], so that the best processing and classification results can be achieved, although domain knowledge is lacking. The success of these systems is because cases can be more easily collected than rules or other domain data and because of the flexibility of the systems based on their learning and maintenance mechanisms that allow incrementally improvement of their system performance during usage of the system.

4.2.2 Beyond the state of the art

The necessity to study the taxonomy of similarity measures and a first attempt to construct a taxonomy over similarity measures has been given by Perner [38] and has been further studied by Cunningham [39]. More work is necessary especially when not only one feature type and representation is used in a CBR system, as it is the case for multimedia data. These multimedia cases will be more complex as the cases used in the system described above that only face on one specific data type. To understand the similarity between these multimedia cases will require more complex knowledge of similarity by the police investigator for the different types of multimedia data. To develop novel similarity measures for text, videos, images, and audio and speech signals and to construct a taxonomy that allows understanding the relation between the different similarity measures will be a challenging task. Similarity aggregation

of the different types of similarity measures is another challenging topic. Specific knowledge for the different types of data such as text [40, 41], images [42–44], video [45], 1-D signals, and meta-learning [36] is required in this work. The development of new similarity measures for multimedia data types and new data representations and ontologies will be done. A complex CBR system that can handle so many different data types, similarities, and data sources is a novelty.

Retrieval of multimedia data from a case base can be refined by relevance feedback mechanisms [46–52]. The user is asked to mark retrieved results as being "relevant" or not with respect to his/her interests. Then, feature weights and the similarity measures are suitably adapted to reflect user's interests. Relevance feedback can be implemented in a number of ways, for example, as the solution of an optimization problem, or as a classification problem. According to the problem at hand, the most suited formulation has to be devised. Thus, the main challenge will be to formulate the relevance feedback problem for forensic applications, so that the search is driven toward the cases more relevant to the case at hand.

Research has been described for learning of feature weights and similarity measures [53–55]. Case mining from raw data in order to get more generalized cases has been described by Jaenichen and Perner [56]. Learning of generalized cases and the hierarchy over the case base has been presented by the authors of Refs. [45, 57]. These works demonstrate that the system performance can be significantly improved by these functions of a CBR system.

New techniques for learning of feature weights and similarity measures and case generalization for different multimedia types are necessary and will be developed for these tasks.

The question of the Life Cycle of a CBR system goes along with the learning capabilities, case base organization and maintenance mechanism, standardization, and software engineering for which new concepts should be developed. As the result, we should come up with generic components for a CBR system for multimedia data analysis and interpretation that form a set of modules that can be easily integrated and updated into the CBR architecture. The CBR system architecture should easily allow configuring modules for new arising task.

The partner IBAI has a number of national and international patents that protect their work on CBR for images and signals. It is to expect that new methods will be developed that can be protected by patents and can ensure the international competition of European entities on CBR systems.

4.3 Multimedia feature extraction

4.3.1 State of the art

Most of computer vision algorithms rely on the extraction of meaningful features that transform raw data values into a more significant representation, better suited for classification and recognition. Although considered often not a central problem, the quality of feature representation can have critically important implications for the performance of the subsequent recognition methods.

Features are usually defined and selected according to a problem-oriented strategy, that is, ad hoc in light of the information considered relevant for the task at hand. In forensics, a plethora of features have been defined for the automated solutions to different problems, such as face detection, retrieval and recognition in video and images [58–60], individual people tracking over video sequences [61, 62], recognition of different biometric parameters (ear, gait, and iris) in images or videos [63, 64], speaker identification in audio signals, suspicious word detection, and handwriting recognition in text document.

Main challenges in forensics scenarios regard the unconstrained conditions in which multimedia data are collected. For audio signals, this is usually in the form of channel distortion and/or ambient noise. For videos and images, problems arise from changes in the illumination direction and/or in the pose of the subjects, occlusions, aging, and so on.

For images and videos, according to the problem at hand, the features selected can be based on specific morphologic parameters of individuals, such as face characteristics (e.g., nose width and eye distance) [65], posture and gesture, ear details, and so on or on general appearance features computed with low-level descriptors. These descriptors can be either global or local and can exhibit different degrees of invariance. Global descriptor category includes features based on Principal Component Analysis (PCA) [66] and Linear Discriminant Analysis (LDA) [67]. The local descriptor category is currently spreading and comprises features based on local values of color, intensity, or texture. To this category belong Scale-Invariant Feature Transform (SIFT) [68], Local Binary Pattern (LBP) [69], Histograms of Oriented Gradients (HOG) [70], or Gabor wavelets [71]. LBP is a well-known texture descriptor and a successful local descriptor robust to local illumination variations [72]. LBP descriptors are compact and easy to compare by various histogram metrics. In addition, there are many LBP variants that improve the description performance; among these, the most popular is Multi-Scale LBP (MSLBP) [73]. HOG has been successfully applied to tasks such as human detection [70] and face recognition [74]. Similar to LBP, edge information captured by gradients within blocks is packed into a histogram. Discarding pixel location information by block-based histogram binning, LBP and HOG gain invariance to local changes such as small facial expressions and pose variations in pedestrian images. The Gabor wavelets are also successful descriptors that capture global shape information centered at a pixel [75]. The convolution of multiple Gaussian-like kernels with different scales and orientations captures information insensitive to expression variation and blur at a pixel's location. Recently, a generalization of the Pairs of Pixels (POP) descriptor, called Centre Symmetric-Pairs of Pixels (CCS-POP), has been presented for face identification [76]. Another line of research currently gaining attention regards the computation of biologically inspired descriptors that result from the attempt to mimic natural visual systems. Several works have shown interesting results in a variety of different face and object recognition contexts [77–79].

The approach based on local descriptors has recently gained popularity, especially in relation to the spreading of the bag-of-feature representation. Indeed, in this frame, local feature descriptors, which can achieve high robustness with respect to appearance variations, are employed to develop a bag of descriptors that represent image content. All such descriptors are, then, quantized using learned visual words to facilitate the retrieval or classification [80–83]. The approach seems promising in forensic scenarios to fit the high variation of object appearance across different views since some very informative local features can accommodate to bad localizations or part visibility [62].

4.3.2 Beyond the state of the art

The problem of automatically extracting relevant information out of the enormous and steadily growing amount of electronic text data is becoming much more pressing. To overcome this problem, various technologies for information management systems have been explored within the Natural Language Processing (NLP) community. Two promising lines of research are represented by the investigation and development of technologies for (a) ontology learning from document collections and (b) feature extraction from texts.

Ontology learning is concerned with knowledge acquisition from texts as a basis for the construction of ontologies, that is, an explicit and formal specification of the concepts of a given domain and of the relations holding between them; the learning process is typically carried out by combining NLP technologies with machine learning techniques. Buitelaar [84] organized the knowledge acquisition process into a "layer cake" of increasingly complex subtasks, ranging from terminology extraction and synonym acquisition to the bootstrapping of concepts and of the relations linking them. Term extraction is a prerequisite for all aspects of ontology learning from text: measures for termhood assessment range from raw frequency to Information Retrieval measures such as TF-IDF, up to more sophisticated measures [85–88]. The dynamic acquisition of synonyms from texts is typically carried out through clustering techniques and lexical association measures [89, 90]. The most challenging research area in this domain is represented by the identification and extraction of relationships between concepts (taxonomical ones but not only); this research area presents strong connections with the extraction of relational information from texts, both relations and events (see below).

With feature extraction, we refer to the task of automatically identifying in texts instances of semantic classes defined in an ontology. This task includes recognition and semantic classification of items representing the domain referential entities ("Named Entity Recognition" or NER), either "named entities" or any kind of word or expression that refers to a domain-specific entity. Recently, extraction of inter-entity relational information is becoming a crucial task: relations to be extracted range from "place_of", "author_of," etc. to specific events, where entities take part in with usually predefined roles ("Relation Extraction"). Currently, there exist several feature extraction approaches, addressing different requirements, operating in different domains and on different text types, and extracting different information bits. If we look at the type of the underlying extraction methodology, systems can be classified into the following classes:

- rule-based systems, using hand-crafted rules. Rule-based systems are particularly appropriate for dealing with documents showing very regular patterns, such as standard tables of data, Web pages with HTML markup, or highly structured text documents;

- systems incorporating supervised machine learning: an alternative to the time-consuming process of hand coding of detailed and specific rules is represented by supervised semantic annotation systems, which learn feature extraction rules from a collection of previously annotated documents; and

- systems using unsupervised machine learning: they represent a viable alternative, currently being explored in different systems, to supervised machine learning approaches, as they dispense with the need for training data whose production may be as time consuming as rule hand coding.

Depending on nature and depth of the features to be extracted, different amounts of linguistic knowledge must be resorted to. This means that type and role of the linguistic analysis differ from one system to another. The condition part of feature extraction rules may check the presence of a given lexical item, the syntactic category of words in context, and their syntactic dependencies. Different clues such as typographical features, relative position of words, or even coreference relations can also be exploited. Most feature extraction systems therefore involve linguistic text processing and semantic knowledge: segmentation into words, morphosyntactic tagging, (either shallow or full) syntactic analysis, and sometimes even lexical disambiguation, semantic tagging, or anaphora resolution.

Text analysis can be carried out either at the preprocessing stage or as part of the feature extraction process. In the former case, the whole text is first analyzed. The analysis is global in the sense that items that are spread all over the document can contribute to build the normalized and enriched representation of the text. Then, the feature extraction process operates on the enriched representation of the text. In the latter case, text analysis is driven by the process of verifying a specific condition. The linguistic analysis is local, focuses on the context of the triggering item associated with a specific feature, and fully depends on the conditions to be checked for that feature.

Different approaches to feature extraction will be investigated to assess their strength and effectiveness to detect and describe the multimedia data content relevant to forensic activities. Both biometric features and local informative descriptors will be studied and collected to create a range of different opportunities to describe multimedia data content. More precisely, low level, local, invariant descriptors will be explored to assure a good performance of detection algorithms, especially for recognition in the wild, whereas global biometric features and properties will be considered as high-level information that is better understandable by end users.

A formal model will be adopted to define the features of different kinds. This will result into an ontological model that will organize different classes of features and foster their sharing and reuse. This will be a very innovative result since the ontology will be general and will approach the domain of multimedia data analysis. It will go further current metadata standards such as MPEG 7 or 21 and will be much more comprehensive and specific than other existing ontologies, which are only partially focused on feature extraction and always aimed at other problems such as multimedia data annotation. Additionally, the ontology will be enriched with algorithms to compute the features included, resulting into a toolbox for feature extraction. This will be another very innovative result.

As far as feature extraction from texts is concerned, the main challenge is represented by the typology of texts to be dealt with, testifying noncanonical language usages.

4.4 Text mining

4.4.1 State of the art

Twitter is a new multimedia communication channel that is rapidly gaining popularity and users, yet police forces do not dispose of adequate methods to analyze the large amounts of textual data that are generated each day. Recently, several retrospective investigations concerning football riots revealed that Twitter was actively used by rivaling gang members to plan their assaults. Twitter data are hard to analyze because the text fragments are very short, multiple persons can be involved in a conversation about various topics, and the data are rapidly changing.

Twitter is a recently introduced microblogging and information sharing platform [91] with over 140 million users and 340 million tweets per day. In the past, several studies have been dedicated to analyzing twitter feeds, for example, in the field of opinion mining and sentiment analysis. For example, in Ref. [92], the authors analyzed the text content of daily Twitter feeds by two mood tracking tools: OpinionFinder, which measures positive versus negative mood, and Google-Profile of Mood States (GPOMS), which measures mood in terms of six dimensions (Calm, Alert, Sure, Vital, Kind, and Happy). They cross-validated the resulting mood time series by comparing their ability to detect the public's response to the presidential election and thanksgiving day in 2008. Ratkiewicz et al. [93] used machine learning for analyzing politically motivated individuals and organizations that use multiple

centrally controlled twitter accounts to create the appearance of widespread support for a candidate or opinion and to support the dissemination of political misinformation.

4.4.2 Beyond the state of the art

We propose to develop and use an integrated data visualization environment based on formal concept analysis, temporal concept analysis, temporal relational semantic systems, and self-organizing maps to identify suspicious tweets.

Formal concept analysis (FCA) is a mathematical technique that was introduced in 1982 by Rudolf Wille [94] and takes its roots in earlier work of Birkhoff [95] and the early work on applying lattice-theoretical ideas in information science, like it was done by Barbut et al. [96]. FCA was used in several security text mining projects. The goal in each of these papers was to make an overload of information available in an intuitive visual format that may speed up and improve decision making by police investigators on where and when to act. In the first case study, with the Amsterdam-Amstelland police (RPAA), which started in 2007, FCA was used to analyze statements made by victims to the police. The concept of domestic violence was iteratively enriched and refined, resulting in an improved definition and highly accurate automated labeling of new incoming cases [97]. Later on, the authors made a shift to the millions of observational and very short police reports from which persons involved in human trafficking and terrorism were extracted. Concept lattices allowed for the detection of several suspects involved in human trafficking or showing radicalizing behavior [98, 99].

Temporal concept analysis (TCA) was introduced by Wolff [100] and offers a framework for representing and analyzing data containing a temporal dimension. In previously discussed security applications, suspects were mentioned in multiple reports, and a detailed profile of one suspect (and persons in his social network) depicted as a lattice, with timestamps of the observations as objects and indications as attributes helped to gain an insight into his (their) threat to society [101]. Recently, TCA and its relational counterpart temporal relational semantic systems (TRSS, [100]) were successfully applied to the analysis of chat conversations [102].

Self-organizing maps [103] have been used in many applications, where high-dimensional unsupervised data spaces had to be visualized in a two-dimensional plane to make the data accessible for human experts. For example, Ramadas et al. [104] used self-organizing maps for identifying suspicious network activity. In a previous security case study, a special type named emergent self-organizing maps was used to identify domestic violence in police reports [105, 106]. They were found to be more suitable than multidimensional scaling for text mining. Claster et al. [107] used self-organizing maps to mine over 80 million twitter micro logs in order to explore whether these data can be used to identify sentiment about tourism and Thailand amid the unrest in that country during the early part of 2010 and further whether analysis of tweets can be used to discern the effect of that unrest on Phuket's tourism environment.

Nevertheless, there are several differences between analyzing twitter feeds and traditional police reports. Whereas individual tweets may not be so interesting, a lot of information can be distilled from conversations consisting of many tweets that emerged between different users concerning a certain topic. Such feeds do not contain a summary of facts; rather several topics emerge between two or more persons. We should judge the interestingness of the feed from a security enforcement perspective and distinguish between several types of twitter users in a relevant conversation, for example, is this person someone who contributed only marginally or did he or she actually contribute to or promote criminal behavior. Ebner et al. [108]

used Formal Concept Analysis (FCA) to categorize twitter users who write tweets about the same topics in the context of a conference event. Cuvelier et al. [109] used FCA as an e-reputation monitoring system in combination with tag clouds. Also, the Natural Language Processing of tweets is nowadays a challenging task since Twitter is characterized by a so-called noncanonical language. It is widely acknowledged that NLP systems have a drop of accuracy when tested against text characterized by this kind of language. This negatively affects different levels of text analysis ranging from the linguistic annotation to the information extraction process. It follows that the analysis of noncanonical languages is one of the main topics of the most recent NLP conferences, for example, the First Workshop on Syntactic Analysis of Noncanonical Language (SANCL-2012) (https://sites.google.com/site/sancl2012/), the workshop series on Scritture brevi (lit.: short writings) organized by the University of Rome Tor Vergata (https://sites.google.com/site/scritturebrevi/atti-dei-workshop), and the First Shared Task on Dependency Parsing of Legal Texts at SPLET-2012 (https://sites.google.com/site/splet2012workshop/shared-task). The main challenges in analyzing noncanonical languages, as tweet language, result from the fact that they have different linguistic characteristics with respect to the data from which the tools are trained, typically newswire texts. Among the others, punctuation and capitalization are often inconsistent; slang, technical jargon is widely exploited; and noncanonical syntactic structures frequently occur [110–112]. Accordingly, several domain adaptation methods and different strategies of analysis have been investigated to improve the accuracies of the NLP tools, among the most recent ones the self-training method used by Le Roux et al. [113], the active-learning method used by Attardi et al. [114], and the term-extraction method proposed by Bonin et al. [88].

Event detection in Twitter has been recently an area of active research and successfully applied to detecting earthquakes [115] and sport events [116]. For events of interest to legal forces, one can utilize the generic features, such as emerging common terms, location, date, and also potentially the participants of the event. Hence, we extract the date/time information and time-event phrases that are learnt from tweets and set the presence of them as a feature. Participant information is also captured via the presence of the '@' character followed by a username within tweets. Specific to the events of legal interest, one can also utilize the overall sentiment of the tweets as a potential feature. According to a recent research by Leetaru [117] at the University of Illinois at Chicago, strong negative emotions in news can suggest upcoming of a significant event. A sentiment analysis in a long period of news revealed that the textual sentiments before the revolutions in Libya and Egypt have shown significant negative signals. The strength of this negativity is found comparable to the signals in 1991 news, right before the United States entered Kuwait; and also in 2003, when the United States-Iraq was about to start.

While the current approaches, such as Ref. [117], have been shown to work on static data and static models, more research is needed to enhance these methods for the dynamic case. Also, the news text is highly structured and formal, while Twitter consists of informal short text. Based on our prior work on classifying short tweets [118], and sentiment analysis on large-scale data [119], we will categorize the tweets for event detection and identify tweets with strong sentimentality. Our initial hypothesis is that strong sentiment increases the probability of event being of interest to legal forces. Recently, distributional semantic models (DSMs) have been applied to affective text analysis with good results across languages [120]. In this WP, we will also apply DSMs to sentiment analysis of multilingual tweets. The more interesting problem is the forecasting problem, where the events can be predicted beforehand. This would be of high value for preventive law enforcement. Besides the prediction problem, one can also use this approach to get feedback from

the crowd on actions taken by the law officers. Such approaches have already been deployed for finance and marketing applications to understand the mood of financial markets and consumer opinions [92, 121, 122]. Similar concepts can be adapted for forensic applications. In fact, FBI and Pentagon have already started to utilize these methods to predict criminal and terrorist activities and monitor persons and regions of high interest [AP Exclusive].

The innovativeness of tool in this area lays in the fact that the combination of the discussed methods has never been proposed for visualizing and clustering data, nor integrated in a software system. It will be the first integrated human centered data discovery environment that combines both statistical methods from machine learning with order-theoretic methods such as concept lattices. The self-organizing map that can handle high-dimensional data spaces and, as a consequence, is an ideal tool for an initial preprocessing is at the start of the human centered discovery process. FCA can then be used to explore dependencies and information links in a smaller subset. TCA and TRSS are used for in-depth profiling of identified individuals and communities. In particular, we focus on the niche of twitter user and feed mining in the broader text-mining field. State-of-the-art domain adaptation methods will be tested to improve the accuracies of the linguistic annotation tools on Twitter data, and customized term-extraction methods will be devised in order to reliably extract relevant keywords from tweets. Needless to say that the proposed system can be easily expanded to other text mining applications.

A web crawler will be designed to collect the feeds from the twitter website. This is a technically challenging yet known task to the scientific community (see e.g., [107]). The data collection can be done by an employee hired by the police who received a type P screening. The type of data is fragments of texts. Concerning languages, we will first focus on Dutch tweets. This may later be extended to Hungarian and Bulgarian since most organized crime in areas such as human trafficking is committed by these nationalities in Amsterdam. Since a tweet consists of among others a user name, his twitter ID and the posted text, as well as potentially ID and name of other users, we will first replace these user-identifiable information items by numeric values using regular expressions. In the second step, we will use available Named Entity Recognition methods for removing person names from the tweets themselves.

4.5 Video analysis

4.5.1 State of the art

Video retrieval has a long history [123–125]. According to the type of video at hand (e.g., film, news, CCTV recording, etc.), different retrieval tasks can be defined both in terms of the type of query and in terms of the processing techniques that are suited for extracting meaningful concepts. For example, it is easy to see that the making of a film comprises the use of techniques whose goal is to provoke sentiments in the watcher. Thus, in order to retrieve concepts from videos, automatic techniques must take into account not only the characteristic of the scene but also the movements of the camera and video editing techniques. On the other hand, still cameras used for video-surveillance purposes allow for the detection of persons and objects moving within the monitored area, as the characteristic of the scene is well known in advance. On these topics, a vast corpus of research has been carried out in the past years, and a number of automatic analysis techniques are embedded into commercial products [126].

One of the first steps in video analysis is the detection of shots, that is, video sequences that contain a continuous camera action in time and space [127, 128].

In the case of films, broadcasted news, and sport videos, shot detection is performed by looking at well-known separators, such as fading and black frames. Each shot is then characterized by one or more key frames, that is, those frames that can be used to characterize the shot. Shot classification can be performed by extracting suitable features and using machine-learning techniques for concept classification. Features can be either extracted from key frames, as well as by looking at global characteristics of the video sequence. They can represent low-level information of such as color and textures as well as characteristics of the shot such as temporal features.

A number of techniques for carrying out these steps have been developed for TV broadcasters, in particular for sport as well as news programs [123, 124]. In these areas, the knowledge of the rules of the game and the rules of video shooting allowed for building a reliable ground truth that allows to make objective comparisons of different algorithms. The classification of video shots can be used for retrieval purposes, as soon as the goal is to retrieve all videos related to a particular class. On the other hand, the use of these techniques for forensic applications still needs more investigation due to the low resolution of the cameras, the variability of the recorded scenes, and the presence of person and objects typically in nonfrontal positions and with many occlusions.

Today, it is of particular interest the reidentification of people in videos [129, 130]. This problem can be formulated as follows. In many real scenarios, an area is monitored by a number of cameras. When persons move in the monitored environment, they can be identified by their face only if they appear in the video in some pose. After they have been identified in one of the videos, they can be tracked (i.e., reidentified) according to their global appearance (e.g., their clothes) rather than by their face.

Speech and sound files constitute an important part of the data collected by Law Enforcement Agencies. For the last 35 years, practical speech recognition systems have been based on Hidden Markov Models (HMMs), which model the training data using the Baum-Welch algorithm in a global manner. Markov state probability distributions are also represented using Gaussian Mixture Models (GMMs). HMMs try to represent the time-varying speech and sound files [131, 132]. This approach is successful to some extent in controlled environments and dictation systems in which people clearly speak to the machines [133].

HMMs and GMMs use features extracted from temporal speech windows. Current speech and sound feature extraction schemes are based on Fourier analysis [131, 134, 135]. Temporal information is only incorporated to the automatic speech recognition systems by only dividing speech into temporal analysis windows. Unfortunately, this global approach loses keyword or speaker-specific features, which are needed in forensic applications. For example, a person cannot modify his or her own average temporal zero crossing rate, even if he or she tries to change his or her own voice by mumbling, or talking with a mouth full of food or cotton balls, etc. [136]. This kind of temporal and person specific information is not used in today's systems, which are globally trained using all the available data.

Global approaches provide good speech and speaker recognition and identification results as long as it is possible to have a good description of the unobserved data. However, continuous spontaneous speech recognition is still an unsolved problem [133, 137]. Unfortunately, most of the speech data in legal cases are spontaneous speech data. In many applications, it is required to retrieve keywords, phrases, names, and speakers from spontaneous speech in real time. Therefore, it is necessary to develop not only new feature extraction and speech and sound representation schemes but also exemplar type case and similarity-based reasoning methods to improve the current speech and sound processing systems.

4.5.2 Beyond the state of the art

The analysis of videos for forensic applications can be carried out by relying on some of the above techniques, provided they are tailored to the scenario at hand. It is easy to see that in the case of surveillance videos, we cannot define a shot according to the paradigm used to segment a film or a sport video [126, 138]. Rather, the definition of "shot" can be driven by the event that is looked for in the video. In particular, the video analyst should be able to query the system, so that the video is first segmented according to the particular event, and then, the shots that can contain the event of interest with high probability are further analyzed by more sophisticated technique in order to detect the object of interest [139]. The development of such a system is beyond the current state of the art, and it will be carried out within this project.

The development of reidentification techniques may allow tracking a person in videos collected by multiple cameras at different locations and in different periods. Detecting people can be carried out by face detection. Many of the existing facial recognition systems are sensitive to variations in the enrolment phase [140–145]. Often these systems have been trained by a huge number of pictures of the same person to estimate reliable values of the parameters for statistical classifiers. The current state of the art does not include a suitable system for the generation of a prototype picture of a person nor a suited prototype-based classifier [146, 147]. Some automatic prototype generation developed in the area of pattern recognition could be used for face recognition [148–150].

Prototype-based system could effectively handle changes in illumination, as they can perform recognition by part resemblance [151, 152]. For the above reasons, most of the facial recognition systems available today assume a standardized enrolment procedure to be performed in a controlled environment (e.g., a cabin), where a number of pictures of the face in a frontal position (2-D) with respect to the camera are taken. In addition, the picture is renewed whenever the recognition accuracy decreases.

Many different methods have been used so far for face recognition and cover a wide spectrum of methods in the pattern recognition field: geometrical representation of the face [153], templates [154, 155], hidden Markov models [156], principal component analysis [157], independently component analysis [158], elastic graph matching [159], trace transform [160], and SVM [161]. None of the methods can be seen as the most promising method because the performance depends on the scenario at hand, and the assumption behind the proposed theoretical models might not be met in real scenarios. Thus, new techniques based on the exploitation of different picture representations, such as shape, texture, signs for skin, eyes and spatial, sign-based connections, and the prototype-based system, have to be investigated.

Case and similarity-based recognition and sensing methods for speech, sound, and audio recognition using both temporal and frequency domain information will be developed. Development of "query by example," keyword, and phrase-based retrieval schemes using exemplar-based schemes, which will be capable of part and whole similarity matching, will be a significant contribution to the existing speech recognition systems.

Current methods for speech and audio analysis emphasize spectral methods. For example, well-known Shazam music recognition method uses only spectral peaks [162]. Commonly used mel-cepstral coefficients, line spectral frequencies, and RASTA features [134, 135] do not have any temporal information, either. We believe that temporal information is not fully utilized in current methods. Temporal information will provide critical information for speaker recognition and keyword spotting

applications. We are developing temporal speech representation methods based on delta modulation [163, 164], zero-crossing, and wavelet scattering [165, 166] information will be incorporated into content based audio and sound retrieval and speech and audio recognition applications.

As pointed above, another important avenue, which is not explored by current methods, is compressive recognition, similarity-based reasoning, and case-based reasoning. Current data modeling methods assume a global representation. On the other hand, case and similarity-based reasoning methods will be able to incorporate fine details of the test case and will likely to provide better recognition results, especially in spontaneous speech. Temporal representation methods such as delta modulation and zero-crossing information are ideal for exemplar and similarity-based reasoning approaches. It is also possible to combine the differential representation of temporal data with the spectral data using compressive sensing [167], which extends this differential data processing concept by using random weights adding to zero to linearly combine the data and/or features. In this way, similarity learning, case generalization and case storage, and compressive learning and sensing will allow the handling of very large amount (terabytes) of data. Once the keyword and phrases are detected, analysts can manually process the proposed retrieval results.

Cut-and-paste locations in speech can be also detected using delta modulation and wavelet scattering, providing a differential representation of speech, sound, and audio data. Fragile watermarking schemes based on wavelet scattering and delta modulation will be developed to prevent tampering. Resulting representation can be easily stored, and it will be ideal for different forensic purposes.

5. Conclusions

Forensic investigations on multimedia evidence usually develop along four different steps: analysis, selection, evaluation, and comparison. During the analysis step, technicians typically look at huge amounts of different multimedia data (e.g., hours of video or audio recordings, pages and pages of text, and hundreds and hundreds of pictures) to reconstruct the dynamic of the event and collect any piece of relevant information. This step obviously requires a lot of time, and many factors can make it difficult, among which data heterogeneity, quality, and quantity are the most relevant. Afterward, during the selection step, technicians select and acquire the most meaningful pieces of information from the different multimedia data (e.g., frames from videos, audio fragments, and documents). Then, in the evaluation step, they look for relevant elements in the selected data, which will be further investigated in the comparison step. They can select heads, vehicles, license plates, guns, sentences, sounds, and all other elements that can link a person to the event. The main problems are the low quality of media data due to high compression, adverse environmental conditions (e.g., noise, bad lighting condition), camera/object position, and facial expressions. Finally, during the comparison step, technicians place the extracted elements side by side with a known element of comparison. From the comparison of general and particular characteristics, the operators give a level of similarity. In forensic application, the use of automatic pattern recognition system gives poor performance because of the high variability of data recording. On the other hand, human perception is a great pattern recognition system but is characterized by high subjectivity and unknown reproducibility and performance.

In this chapter, we propose to develop a toolkit of methods and instruments that will be able to support analysts along all these steps, strongly reducing human intervention. First of all, it will include instruments to process different kinds of

media data and, possibly, correlate them. This will obviously reduce the time spent to find the correct instruments for processing the medium at hand. Furthermore, it comprises preprocessing tools that alleviate, by filtering and enhancement, the problem of low-data quality. In particular, for image and video data, a great help will come from super-resolution methods that will maximize the information contained in low-resolution images or videos (e.g., foster the process of face reconstruction and recognition from blurred images). This feature will greatly support all the subsequent steps.

In this chapter, we focused on the background and motivation for our work. The overall system architecture is explained. We present the data to be used. After a review of the state of the art of related work of the multimedia data we consider in this work, we describe the method and techniques we are developing that go beyond the state of the art. The work will be continued in the Chapter Part II of Forensic Multimedia Data Analysis.

Author details

Petra Perner
Institute of Computer Vision and Applied Computer Sciences, Leipzig, Germany

*Address all correspondence to: pperner@ibai-institut.de

IntechOpen

References

[1] Passive millimeter wave images, copyright. Alfa Imaging S.A. ALFA, Spain

[2] Molina R, Murtagh F, editors. DSP soars into space. IEEE Signal Processing Magazine. 2001;**18**(2):1-3

[3] Babacan SD, Molina R, Do MN, Katsaggelos AK. Blind deconvolution with general sparse image priors. In: European Conference on Computer Vision (ECCV), Berlin, Heidelberg, Florence, Italy: Springer; 2011. pp. 984-999

[4] Starck JL, Murtagh F, Candès EJ, Donoho DL. Gray and color image contrast enhancement by the curvelet transform. IEEE Transactions on Image Processing. 2003;**12**:706-717

[5] Daubos T, Murtagh F. High-quality still images from video frame sequences. In: Geradts Z, Rudin LI, editors. Investigative Image Processing II, Proceedings of the SPIE. Vol. 4709. 2002. pp. 49-59

[6] Daubos T, Geradts Z, Starck JL, Campbell J, Murtagh F. Automated wavelet-based image addition: application to surveillance video. In: Whelan PF, editor. IMVIP'99—Irish Machine Vision and Image Processing Conference 1999, Dublin City University. 1999. pp. 15-25

[7] Luessi M, Babacan SD, Molina R, Booth JR, Katsaggelos AK. Bayesian symmetrical EEG/fMRI fusion with spatially adaptive priors. NeuroImage. 2011;**55**(1):113-132

[8] Vega M, Mateos J, Molina R, Katsaggelos AK. Bayesian TV Denoising of SAR images. In: IEEE International Conference on Image Processing ICIP 2011. Bruselas (Bélgica); 2011. pp. 169-172

[9] Amizic B, Spinoulas L, Molina R, Katsaggelos AK. Compressive sampling with unknown blurring function: Application to passive millimiterwave imaging. In: IEEE International Conference on Image Processing. Orlando, Florida; 2007. pp. 321-329

[10] Babacan SD, Molina R, Katsaggelos AK. Total variation image restoration and parameter estimation using variational posterior distribution approximation. In: International Conference on Image Processing (ICIP)—IBM Student Paper Award for ICIP 2007. Vol. I. San Antonio, Texas (USA); 2007. pp. 97-100

[11] Starck J-L, Murtagh F, Fadili J. Sparse Image and Signal Processing: Wavelets. Curvelets: Morphological Diversity. Cambridge University Press; 2010

[12] Starck J-L, Murtagh F. Astronomical Image and Data Analysis. 2nd ed. Springer-Verlag; 2006

[13] Chantas G, Galatsanos N, Molina R, Katsaggelos AK. Variational Bayesian image restoration with a spatially adaptive product of total variation image priors. IEEE Transactions on Image Processing. 2010;**19**(2):351-362

[14] Babacan D, Molina R, Katsaggelos AK. Bayesian blind deconvolution from differently exposed image pairs. IEEE Transactions on Image Processing. 2010;**19**(11):2874-2888

[15] Katsaggelos AK, Molina R, Mateos J. Super resolution of images and video. In: Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan & Claypool; 2007

[16] Babacan SD, Molina R, Katsaggelos AK. Variational Bayesian super resolution. IEEE

Transactions on Image Processing. 2011;**20**(4):984-999

[17] Mateos J, Katsaggelos AK, Molina R. A Bayesian approach to estimate and transmit regularization parameters for reducing blocking artifacts. IEEE Transactions on Image Processing. 2000;**9**(7):1200-1215

[18] Bishop C. Pattern Recognition and Machine Learning. Springer; 2006

[19] Barber D. Bayesian Reasoning and Machine Learning. Cambridge University Press; 2012

[20] Murphy KP. Machine Learning: A Probabilistic Perspective. MIT; 2012

[21] Mahdian B, Saic S. A bibliography on blind methods for identifying image forgery. Signal Processing: Image Communication. 2010;**25**(6):389-399

[22] Poisel R, Tjoa S. Forensics investigations of multimedia data: A review of the state-of-the-art. In: Proc. Sixth Int. IT Security Incident Management and IT Forensics (IMF) Conf. 2011. pp. 48-61

[23] Farid H. Exposing digital forgeries in scientific images. In: ACM MM&Sec. 2006

[24] Popescu AC, Farid H. Exposing digital forgeries by detecting traces of resampling. IEEE Transactions on Signal Processing. 2005;**53**(2):758-767

[25] Chen M et al. Determining image origin and integrity using sensor noise. IEEE Transactions on Information Forensics and Security. 2008;**3**(1):74-90

[26] Li C-T, Satta R. An empirical investigation into the correlation between vignetting effect and the quality of sensor pattern noise. IET Computer Vision. 2012;**6**(6):560-566

[27] Gao X, Ng TT, Qiu B, Chang S-F. Single-view recaptured image detection based on physics-based features. In: IEEE International Conference on Multimedia & Expo (ICME). 2010

[28] Ng T-T. Camera response function signature for digital forensics Part II: Signature extraction. In: IEEE Workshop on Information Forensics and Security (WIFS). 2009

[29] Khanna N, Delp EJ. Source scanner identification scanned documents. In: IEEE Workshop on Information Forensics and Security (WIFS). 2009

[30] Khanna N et al. Forensic techniques for classifying scanner, computer generated and digital camera images. In: IEEE ICASSP. 2008

[31] Perner P, editor. Case-Based Reasoning for Image and Signals, Series Computational Intellignece. Berlin: Springer Verlag; 2007

[32] Perner P, Holt A, Richter M. Image processing in case-based reasoning. The Knowledge Engineering Review. 2005;**20**(3):311-314

[33] Ahmed MU, Begum S, Funk P. An overview of three medical applications using hybrid case-based reasoning. In: Perner P, editor. ICDM 2012, Workshop Proceedings, Workshop on Case-Based Reasoning CBR-MD 2012. Fockendorf: IBAI-Publishing; 2012. pp. 79-94 ISBN 978-3-942952-16-3

[34] Perner P, Attig A, Machno O. Novel method for the interpretation of spectrometer signals based on delta-modulation and similarity determination. Transactions on Mass-Data Analysis of Images and Signals. 2011;**3**(1):3-14

[35] Weber RO, Ashley KD, Breueninghaus S. Textual case-based reasoning. The Knowledge Engineering Review. 2006;**20**(3):255-260

[36] Attig A, Perner P. Model building in image processing by meta-learning based on case-based reasoning. In: Wang PS-P, editor. Pattern Recognition and Machine Vision-In Honor and Memory of Late Prof. King-Sun Fu, River Publishers' Series in Information Science and Technology. River Publishers; 2010. pp. 149-164

[37] Murtagh F, Starck JL. Wavelet and curvelet moments for image classification: Application to aggregate mixture grading. Pattern Recognition Letters. 2008;**29**:1557-1564

[38] Perner P. Why case-based reasoning is attractive for image interpretation. In: Perner P, Aha D, Watson I, editors. Case-Bases Reasoning Research and Developments, LNAI. Vol. 2080. Heidelberg: Springer; 2001. pp. 27-44 (invited paper)

[39] Cunningham P. A taxonomy of similarity mechanisms for case-based reasoning. IEEE Transactions on Knowledge and Data Engineering. 2009;**21**(11):1532-1543

[40] Iosif E, Potamianos A. Similarity computation using semantic networks created from web-harvested data. Natural Language Engineering. 2015:49-79

[41] Iosif E, Potamianos A. Unsupervised semantic similarity computation between terms using web documents. IEEE Transactions on Knowledge and Data Engineering. 2010;**22**(11):1637-1647

[42] Perner P, Perner H, Jänichen S. Recognition of airborne fungi spores in digital microscopic images. Journal Artificial Intelligence in Medicine, Special Issue on CBR. 2006;**36**(2):137-157

[43] Geradts Z, Bijhold J, Hermsen R, Murtagh F. Image matching algorithms for breech marks and firing pins in a database of spent cartridge cases of

firearms. Forensic Science International. 2001;**119**:97-106

[44] Geradts Z, Bijhold J, Hermsen R, Murtagh F. Matching algorithms using wavelet transforms for a database of spent cartridge cases of firearms. In: Proceedings of SPIE. Vol. 4232. 2001. pp. 545-552

[45] Contreras P, Murtagh F. Fast, linear time hierarchical clustering using the Baire metric. Journal of Classification. 2012;**29**:118-143

[46] Thomee B, Lew M. Interactive search in image retrieval: a survey. Journal of Multimedia Information Retrieval. 2012;**1**(2):71-86

[47] Piras L, Giacinto G, Paredes R. Enhancing image retrieval by an exploration-exploitation approach. In: Perner P, editor. Machine Learning and Data Mining in Pattern Recognition, LNCS. Vol. 7376. Berlin: Springer; 2012. pp. 355-365

[48] Datta R, Joshi D, Li J, Wang JZ. Image retrieval: ideas, influences, and trends of the new age. ACM Computing Surveys. 2008;**40**:1-60

[49] Giacinto G, Roli F. Instance-based relevance feedback in image retrieval using dissimilarity spaces. In: Perner P, editor. Case-Based Reasoning for Signals and Images. Berlin: Springer-Verlag; 2007. pp. 419-430

[50] Giacinto G. A nearest-neighbor approach to relevance feedback in content based image retrieval. In: Proceedings of the 6th ACM International conference on Image and video retrieval (CIVR'07). ACM Press; 2007. pp. 456-463

[51] Tronci R, Murgia G, Pili M, Piras L, Giacinto G. ImageHunter: A novel tool for relevance feedback in content based image retrieval. In: Loi C, Semeraro G, Vargiu E, editors. New Challenges in

Distributed Information Filtering and Retrieval, SCI. Vol. 439. Heidelberg: Springer; 2013. pp. 53-70

[52] Lew MS, Sebe N, Djeraba C, Jain R. Content-based multimedia information retrieval: state of the art and challenges. ACM Transactions on Multimedia Computing, Communications, and Applications. 2006;**2**:1-19

[53] Craw S. Introspective learning to build case-based reasoning (CBR) knowledge containers. In: Perner P, Rosenfeld A, editors. Machine Learning and Data Mining in Pattern Recognition, LNCS. Vol. 2734. Heidelberg: Springer; 2003. pp. 1-6

[54] Wettschereck D, Aha DW, Mohri T. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. Artificial Intelligence Review. 1997;**11**:273-314

[55] Zhang L, Coenen F, Leng P. Formalising optimal feature weight settings in case-based diagnosis as linear programming problems. Knowledge-Based Systems. 2002;**15**:391-298

[56] Jaenichen S, Perner P. Conceptual clustering and case generalization of two-dimensional forms. Computational Intelligence. 2006;**22**(3/4):178-193

[57] Perner P. Case-base maintenance by conceptual clustering of graphs. Engineering Applications of Artificial Intelligence. 2006;**19**(4):381-393

[58] Schwartz W, Guo H, Choi J, Davis L. Face identification using large feature sets. IEEE Transactions on Image Processing (TIP). 2012;**21**(4):2245-2255

[59] Tolba AS, El-baz AH, El-Harby AA. Face Recognition: A literature review. International Journal of Signal Processing. 2006;**2**(2):88-103

[60] Jain AK, Klare B, Park U. Face matching and retrieval in forensics applications. IEEE MultiMedia. 2012;**19**(1):20-28

[61] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In: Proc. ECCV 2008. 2008. pp. 262-275

[62] Liu K, Yang J. Recognition of people reoccurrences using bag-of-features representation and support vector machine. In: Chinese Conference on Pattern Recognition, Nanjing, 2009. pp. 1-5. DOI: 10.1109/CCPR.2009.5344034

[63] Sanderson C. Biometric Person Recognition: Face, Speech and Fusion. VDM Verlag; 2008

[64] Ali H, Salami MJE. Wahyudi: Iris recognition system by using support vector machines. In: International Conference on Computer and Communication Engineering, ICCCE; 2008. pp. 516-521

[65] Vezzetti E, Marcolin F. 3D human face description: Landmarks measures and geometrical features. Image and Vision Computing. 2012;**30**:698-712

[66] Turk M, Pentland A. Face recognition using eigenfaces. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. 1991. pp. 586-591

[67] Belhumeur P, Hespanha J, Kriegman D. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997;**19**(1):11-720

[68] Lowe DG. Object recognition from local scale-invariant features. In: Proceedings of the International Conference on Computer Vision. Vol. 2. 1999. pp. 1150-1157

[69] Ahonen T, Hadid A, Pietikainen M. Face description with local binary patterns: Application to face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006;**28**(12):2037-2041

[70] Dalal N, Triggs B. Histograms of oriented gradients for human detection. CVPR. 2005

[71] Shen L, Bai L. A review on Gabor wavelets for face recognition. Pattern Analysis and Applications. 2006;**9**:273-292

[72] Heikkilä M, Pietikäinen M, Schmid C. Description of interest regions with local binary patterns. Pattern Recognition. 2009;**42**:425-436

[73] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002;**24**(7):971-987

[74] Schwartz WR, Guo H, Davis LS. A robust and scalable approach to face identification. ECCV. 2010

[75] Zhu ZF, Tang M, Lu HQ. A new robust circular Gabor based object matching by using weighted Hausdorff distance. Pattern Recognition Letters. 2004;**25**(4):515-523

[76] Choi J, Schwartz WR, Guo H, Davis LS. A complementary local feature descriptor for face identification. In: IEEE Workshop on the Applications of Computer Vision (WACV). 2012

[77] Serre T, Wolf L, Bileschi S, Riesenhuber M, Poggio T. Robust object recognition with cortex-like mechanisms. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007;**29**(3):411-426

[78] Mutch J, Lowe DG. Object class recognition and localization using sparse features with limited receptive fields. IJCV. 2008

[79] Cox D, Pinto N. Beyond simple features: A large-scale feature search approach to unconstrained face recognition. In: IEEE Int. Conference on Automatic Face & Gesture Recognition. 2011. pp. 8-15

[80] Sivic J, Zisserman A. Video Google: A text retrieval approach to object matching in videos. In: Proc. ICCV 2003. Nice, France; 2003. pp. 11-17

[81] Moosmann F, Nowak E, Jurie F. Randomized clustering forests for image classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2008;**30**(9):1632-1646

[82] Colantonio S, Martinelli M, Salvetti O. Ontology and algorithms integration for image analysis. In: Salerno E, Çetin AE, Salvetti O, editors. Muscle 2011, LNCS. Vol. 7252. Berlin Heidelberg: Springer-Verlag; 2012. pp. 17-29

[83] Perner P. Image mining: Issues, framework, a generic tool and its application to medical-image diagnosis. Journal Engineering Applications of Artificial Intelligence. 2002;**15**(2):105-216

[84] Buitelaar P, Cimiano P, Magnini B, editors. Ontology Learning from Text: Methods, Evaluation and Applications Frontiers in Artificial Intelligence and Applications Series. Vol. 123. IOS Press; 2005

[85] Frantzi K, Ananiadou S. The C–value/NC value domain independent method for multi-word term extraction. Journal of Natural Language Processing. 1999;**6**(3):145-179

[86] Dell'Orletta F, Lenci A, Marchi S, Montemagni S, Pirrelli V, Venturi G. Dal

testo alla conoscenza e ritorno: estrazione terminologica e annotazione semantica di basi documentali di dominio. In: AIDA Informazioni, Atti del Convegno Nazionale Ass.I.Term "I–TerAnDo", Università della Calabria, 5-7 giugno 2008. Roma: AIDA, n. 1-2/2008, ISSN 1121-0095; 2008. pp. 185-206

[87] Lenci A, Montemagni S, Pirrelli V, Venturi G. Ontology learning from Italian legal texts. In: Breuker J et al., editors. Law, Ontologies and the Semantic Web—Channelling the Legal Information Flood, Frontiers in Artificial Intelligence and Applications. Vol. 188. Heidelberg: Springer; 2009. pp. 75-94

[88] Bonin F, Dell'Orletta F, Venturi G, Montemagni S. A contrastive approach to multi–word extraction from domain–specific corpora. In: Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010). La Valletta, Malta; 2010

[89] Lin D. Automatic retrieval and clustering of similar words. In: Proceedings of COLING/ACL98. Montreal, Canada; 1998

[90] Allegrini P, Montemagni S, Pirrelli V. Example-based automatic induction of semantic classes through entropic scores. In: Linguistica Computazionale. Vol. XVI–XVII. 2003. pp. 1-45

[91] Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In: Proceedings of the 19th international conference on World wide web (WWW '10). New York, NY, USA: ACM; 2010. pp. 591-600

[92] Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. Journal of Computational Science. 2011;**2**(1):1-8

[93] Ratkiewicz J et al. Detecting and tracking political abuse in social media. In: Proc. of ICWSM. 2011

[94] Wille R. Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival I, editor. Ordered Sets. Dordrecht, Boston: Reidel; 1982. pp. 445-470

[95] Birkhoff G. Lattice Theory. 3rd ed. Vol. 25. Providence, RI: American Mathematical Society Coll. Publ; 1973

[96] Kimberly Dozier AP exclusive: CIA following Twitter, Facebook. Available from: http://www.guardian.co.uk/world/feedarticle/9929898 [Accessed: 08 October 2012]

[97] Poelmans J, Elzinga P, Viaene S, Dedene G. Formally analyzing the concepts of domestic violence. Expert Systems with Applications. 2011;**38**(4):3116-3130. DOI: 10.1016/j.eswa.2010.08.103

[98] Elzinga P, Poelmans J, Viaene S, Dedene G, Morsing S. Terrorist threat assessment with formal concept analysis. In: Proc. 8th IEEE International Conference on Intelligence and Security Informatics. 23-26 May. Vancouver, Canada; 2010. pp. 77-82 ISBN: 978-1-42446460-9/10

[99] Poelmans J, Elzinga P, Viaene S, Dedene G, Kuznetsov S. Semi-automated knowledge discovery in unstructured text: Identifying and profiling human trafficking. International Journal of General Systems. 2012;**41**(8):774-804

[100] Wolff KE. Temporal concept analysis. In: Nguifo EM et al, editors. ICCS-2001 International Workshop on Concept Lattices-Based Theory, Methods and Tools for Knowledge Discovery in Databases, Stanford University. Palo Alto, CA; 2001. pp. 91-107

[101] Poelmans J, Elzinga P, Viaene S, Dedene G, Kuznetsov S. A concept discovery approach for fighting human

trafficking and forced prostitution. In: 19th International Conference on Conceptual Structures, July 25-29, Derby, England, LNCS. Vol. 6828. Heidelberg: Springer; 2011. pp. 201-214

[102] Elzinga P, Wolff KE, Poelmans J. Analyzing chat conversations of pedophiles with temporal relational semantic systems. In: 1st IEEE European Conference on Intelligence and Security Informatics. Odense, Denmark; 22-24 August 2012. 2012. pp. 242-249

[103] Kohonen T. Self-organized formation of topologically correct feature maps. Biological Cybernetics. 1982;**43**:59-69

[104] Ramadas M, Ostermann S, Tjaden B, Vigna G, Kruegel C, Jonsson E. Detecting anomalous network traffic with self-organizing maps. In: Recent Advances in Intrusion Detection, LNCS. Vol. 2820. Heidelberg: Springer; 2003. pp. 36-54

[105] Poelmans J, Elzinga P, Viaene S, Van Hulle M, Dedene G. Text mining with emergent self organizing maps and multi-dimensional scaling: A comparative study on domestic violence. Applied Soft Computing. 2011;**11**(4):3870-3876. DOI: 10.1016/j.asoc.2011.02.026

[106] Poelmans J, Elzinga P, Viaene S, Van Hulle M, Dedene G. Gaining insight in domestic violence with emergent self-organizing maps. Expert Systems with Applications. 2009;**36**(9):11864-11874

[107] Cha M, Haddadi H, Benevenuto F, Gummad KP. Measuring user influence on twitter: The million follower fallacy. In: 4th Int'l AAAI Conference on Weblogs and Social Media. Washington, DC; 2010

[108] Ebner M, Mühlburger H, Schaffert S, Schiefner M, Reinhardt W, Wheeler S. Getting granular on Twitter: Tweets from a conference and their limited usefulness for non-participants. In: Key Competencies in the Knowledge Society. Vol. 324. Boston: Springer; 2010. pp. 102-113

[109] Cuvelier E, Aufaure M-A. A buzz and e-reputation monitoring tool for twitter based on galois lattices. In: Andrews S, Polovina S, Hill R, Akhgar B, editors. Conceptual Structures for Discovering Knowledge, LNCS. Vol. 6828. Berlin: Springer; 2011. pp. 91-103

[110] Bonin F, Dell'Orletta F, Venturi G, Montemagni S. Contrastive filtering of domain-specific multi-word terms from different types of corpora. In: Proceedings of the workshop Multiword Expressions: from Theory to Applications (MWE 2010), 23rd International Conference on Computational Linguistics (COLING2010), Beijing, China, August 28. 2010. pp. 76-79

[111] Dell'Orletta F, Marchi S, Montemagni S, Plank B, Venturi G. The SPLeT-2012 shared task on dependency parsing of legal texts. In: Proceedings of the 4th Workshop on "Semantic Processing of Legal Texts" at LREC 2012. Istanbul, Turkey; 2012

[112] Petrov S, McDonald R. Overview of the 2012 shared task on parsing the web. In: Shared Task on Domain Adaptation for Parsing the Web At the First Workshop on Syntactic Analysis of Non-Canonical Language. At HLT-NAACL 2012 in Montreal on June 8, 2012. 2012

[113] Le Roux J, Foster J, Wagner J, Zadeh Kaljahi RS, Bryl A. DCUParis13 systems for the SANCL 2012 shared task. In: Notes of the First Workshop on Syntactic Analysis of Non-Canonical Language (SANCL). 2012

[114] Attardi G, Sartiano D, Simi M. Active learning for domain adaptation of dependency parsing on legal texts.

In: Proceedings of the 4th Workshop on "Semantic Processing of Legal Texts" at LREC 2012. Istanbul, Turkey; 2012

[115] Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international conference on World wide web (WWW '10). New York, NY, USA: ACM; 2010. pp. 851-860

[116] Lanagan J, Smeaton AF. Using Twitter to detect and tag important events in sports media. In: Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona. Catalonia, Spain; 2011

[117] Leetaru KH. Culturomics 2.0: Forecasting large-scale human behavior using global news media tone in time and space. First Monday. 2011;**16**(9)

[118] Demir E, Fuhry D, Sriram B, Demirbas M, Ferhatosmanoglu H. Short text classification in twitter to improve information filtering. In: Proceedings of the ACM SIGIR 2010 Posters and Demos. Vol. 2010. Geneva, Switzerland

[119] Kucuktunc O, Cambazoglu BB, Weber I, Ferhatosmanoglu H. A large-scale sentiment analysis for Yahoo! answers. In: Proceedings of the fifth ACM international conference on Web search and data mining (WSDM '12). New York, NY, USA: ACM; 2012. pp. 633-642

[120] Zhai CX. Statistical Language Models for Information Retrieval (Synthesis Lectures Series on Human Language Technologies). Morgan & Claypool Publishers; 2008

[121] Archak N, Ghose A, Ipeirotis PG. Deriving the pricing power of product features by mining consumer reviews. Management Science. 2011;**57**(8):1485-1509

[122] Liu B, Hu M, Cheng J. Opinion observer: analyzing and comparing opinions on the Web. In: Proceedings of the 14th international conference on World Wide Web. 10-14 May 2005. Chiba, Japan; 2005

[123] Yuan J, Wang H, Xiao L, Zheng W, Li J, Lin F, et al. A formal study of shot boundary detection. IEEE Transactions on Circuits and Systems for Video Technology. 2007;**17**:168-186

[124] Xu C, Wang J, Lu H, Zhang Y. A novel framework for semantic annotation and personalized retrieval of sports video. IEEE Transactions on Multimedia. 2008;**10**:421-436

[125] Hauptmann AG, Yan R, Lin W-H, Christel MG, Wactlar H. Can high-level concepts fill the semantic gap in video retrieval? A case study with broadcast news. IEEE Transactions on Multimedia. 2007;**9**:958-966

[126] Stringa E, Regazzoni CS. Real-time video-shot detection for scene surveillance applications. IEEE Trans. on Image Processing. 2000;**9**(1):69-79

[127] Snoek CGM, Worring M. Concept-based video retrieval. Foundations and Trends in Information Retrieval. 2009;**4**(2):215-322

[128] Fan J, Elmagarmid AK, Zhu X, Aref WG, Wu L. ClassView: Hierarchical video shot classification, indexing and accessing. IEEE Transactions on Multimedia. 2004;**6**:70-86

[129] Tian Y, Hampapur A, Brown L, Feris R, Lu M, Senior A. Event detection, query, and retrieval for video surveillance. In: Ma Z, editor. Artificial Intelligence for Maximizing Content Based Image Retrieval. 2009. pp. 342-370

[130] Doretto G, Sebastian T, Tu P, Rittscher J. Appearance-based person reidentification in camera networks: problem overview and

current approaches. Journal of Ambient Intelligence and Humanized Computing. 2011;**2**:127-151

[131] Heisele B, Ho P, Poggio T. Face recognition with support vector machines: Global versus component-based approach. In: Proc. of the Eighth IEEE International Conference on Computer Vision. Vancouver, Canada; Vol. 2. 2001. pp. 688-694

[132] Candès EJ, Wakin MB. An introduction to compressive sampling. IEEE Signal Processing Magazine. 2008;**25**(2):21-30

[133] Baker JM, Deng L, Glass J, Khudanpur S, Lee C-H, Morgan N, et al. Research developments and directions in speech recognition and understanding, Part 1. IEEE Signal Processing Magazine. 2009;**26**(3):75-80

[134] Walker MA, Rudnicky A, Aberdeen J, Bratt EO, Garofolo J, Hastie H, et al. DARPA communicator evaluation: Progress from 2000 to 2001. In: ICSLP 2002. Vol. 1. 2002. pp. 273-276

[135] Hermansky H, Morgan N. RASTA processing of speech. IEEE Transactions on Speech and Audio Processing. 1994;**2**(4):578-589

[136] Saon G, Chien J-T. Special issue on fundamental technologies in modern speech recognition. IEEE Signal Processing Magazine. 2012:18-33

[137] Erzin E, Cetin AE. Interframe differential coding of line spectrum frequencies. IEEE Transactions on Speech and Audio Processing. 1994;**2**(2):350-352

[138] NIST: TRECVID video retrieval evaluation—Online proceedings 2002-2018. Available from: http://www-nlpir.nist.gov/projects/tvpubs/tv.pubs.org.html

[139] Satta R, Fumera G, Roli F. Fast person re-identification based on dissimilarity representations. Pattern Recognition Letters. 2012;**33**(14):1838-1848

[140] Dee HM, Cohn AG, Hogg DC. Building semantic scene models from unconstrained video. Computer Vision and Image Understanding. 2012;**116**(3):446-456

[141] Abate A, Riccio MND, Tortora G. An ifs based approach for face recognition. In: Proc. IEEE International Conference on Image Processing. Vol. II. 2005. pp. 938-941

[142] Arandjelovi O, Cipolla R. An information-theoretic approach to face recognition from face motion manifolds. Image Vision Comput. 2006;**24**(6):639-647

[143] Beymer D, Poggio T. Face recognition from one example view. Tech. Rep. 1536. MIT AI Lab.; 1995

[144] Distasi R, Nappi M, Tucci M. Fire: Fractal indexing with robust extensions for image databases. IEEE Transactions on Image Processing. 2003;**12**(3):373-384

[145] Gao Y, Leung M. Face recognition using line edge map. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002;**24**(6):764-779

[146] Gao Y, Leung M, Hui S, Tananda M. Facial expression recognition from line-based caricatures. IEEE Transactions on Systems, Man, and Cybernetics Part A. 2003;**33**(3):407-412

[147] Perner P. Prototype-based classification. Applied Intelligence. 2008;**28**(3):238-246

[148] Perner P, Attig A. Prototype-based classification for automatic knowledge acquisition of pathological processes at the cellular level. Transactions on Mass-Data Analysis of Images and Signals. 2010;**2**(1):41-54

[149] Blanz V, Vetter T. A morphable model for the synthesis of 3D faces. In: Computer Graphics Proceedings SIGGRAPH'99. 1999. pp. 187-194

[150] Chowdhury AKR, Chellappa R. Face reconstruction from monocular video using uncertainty analysis and a generic model. Computer Vision and Image Understanding. 2003;**91**(1-2):188-213

[151] Cristinacce D, Cootes TF. Feature detection and tracking with constrained local models. In: Proceedings IEEE British Machine Vision Conference. 2006

[152] Georghiades AS, Belhumeur PN, Kriegman DJ. From few to many: Illumination cone models for face recognition under variable lighting. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001;**23**(6):643-660

[153] Tan T, Yan H. Face recognition using the weighted fractal neighbor distance. IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews. 2005;**35**(4):576-582

[154] Brunelli R, Poggio T. Face recognition through geometrical features. In: LNCS. Vol. 588. Springer; 1992. pp. 792-800

[155] Fishler M, Elschlager R. The representation and matching of pictorial structures. IEEE Transactions on Computers. 1973;**C-22**(1):67-92

[156] Brunelli R, Poggio T. Face recognition: Features versus templates. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993;**15**(10):1042-1052

[157] Nefian AV, Hayes MH. Hidden Markov models for face recognition. In: Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Seattle, Washington, USA; 1998. pp. 2721-2724

[158] Pentland A, Moghadam B, Starner T. View-based and modular eigenspaces for face recognition. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 1994. pp. 84-91

[159] Bartlett MS, Movellan JR, Sejnowski TJ. Face recognition by independent component analysis. IEEE Transactions on Neural Networks. 2002;**13**(6):1450-1464

[160] Wiskott L, Fellous J, Kruger N, von der Malsburg C. Face recognition by elastic bunch graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997;**19**:775-779

[161] Srisuk S, Petrou M, Kurutach W, Kadyrov A. Face authentication using the trace transform. In: Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, Wisconsin, USA; 2003. pp. 305-312

[162] Saunders J. One of the most indicative and robust measures to discern voiced speech is the average zero-crossing rate (ZCR) of the time domain waveform. In: Real-time Discrimination of Broadcast Speech/Music. IEEE International Conf. On Acoustics, Speech, and Signal Processing (ICASSP). 1996

[163] Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE. 1989;**77**(2):257-286

[164] Perner P. Data reduction methods for technological industrial robots with direct teach-in-programming. Dissertation IH Mittweida 1985. 2nd edn. Fockendorf: IBAI Publishing; 2010. ISBN 978-3-940501-16-5

[165] Perner P, Attig A, Machno O. Novel method for the interpretation of spectrometer signals based on delta-modulation and similarity determination. Transactions on Mass-Data Analysis of Images and Signals. 2011;**3**(1):3-14 and The Patent: P. Perner "Method and Device for Automatically Determining a Substance Based on Spectroscopic Examinations," US020110153227A1

[166] Andén J, Mallat S. Deep scattering spectrum. IEEE Transactions on Signal Processing. 2014;**62**(16):4114-4128

[167] Jabloun F, Cetin AE, Erzin E. Teager energy based feature parameters for speech recognition in car noise. IEEE Signal Processing Letters. 1999;**6**(10):259-261