1.1 The REWIND Newsletter

REWIND is a FET (Future and Emerging Technology) research project funded by the European Commission within the ICT area of the Seventh Framework Program (FP7). The project is devoted to the analysis of multimedia content in order to trace back its processing history with potential impact in several areas ranging from forensics and law enforcement to quality assessment. The project started on May 1 2011 and closed on April 30 2014.

This newsletter is addressed to subjects that might be potentially interested in the outcomes of the project, including companies, public and private research institutions or individual researchers working in this field. This is the fourth issue of the REWIND newsletter. All the published issues can be downloaded at www.rewindproject.eu.

Even if the EU funded activities are now concluded the REWIND partners are strongly committed to continue the activities started during the project along different research lines. Further issues of the newsletter will be published and the project web-site will remain open and continuously updated with new results.

1.2 The REWIND has reached its objectives

In the following a brief description of the activities carried out during the lifetime of the project.

The work started with the definition of a wide and comprehensive set of application scenarios and relevant use cases, in which the possibility of revealing the processing history of multimedia objects is highly relevant. These scenarios embrace several fields, including multimedia forensics, quality assessment, restoration of digital content, etc. This served as background for devising accurate mathematical models representing the most common processing operators, a necessary prerequisite to design practical footprint detectors. The basic idea is that complex processing chains can be described as the composition of smaller elementary blocks. For this reason, a taxonomy of simple operators was proposed, distinguishing those operators that act on the domain (e.g., upsampling, downsampling, scaling, rotation, affine warping, etc.), from those that act on the co-domain (e.g., pointwise functions, filtering, transforms, etc.). A set of processing chains was identified, ranging from simple chains (e.g., linear filtering, resampling, etc.) to more complex examples (e.g., blockwise and predictive coding, analog-to-digital and digital-to-analog conversion, etc.).
Due to the extremely large number of processing chains that arise from plausible combinations of simple operators, the research focused on a relevant subset of the identified chains. At the same time, different chains might lead to similar outputs. Hence, the problem of distinguishing among different chains (also composed by the same elementary blocks but with different ordering) was systematically addressed, based on well-established detection and information theoretic measures. The optimality and the links between those distinguishability measures were analyzed, considering the different prior knowledge that could be available to the content analyst. In particular, the proposed method can also be used to determine the minimum number of observations (e.g., number of samples, number of DCT coefficients, etc.), which are required to reliably determine the properties of the underlying processing chain. Although general in its nature, this framework was applied to a broad set of scenarios, which involve quantization, possibly re-iterated, combined with other operators, e.g., non-linear point-wise functions.

Mathematical modeling focused on the operators that are often part of common processing chains, so as to devise suitable detectors. Special attention was given to signal acquisition, resampling as well as to chains involving the composition of quantization with other operators, e.g., transform coding, possible re-iterated multiple times or followed by linear filtering. In all cases, the results achieved express the statistics of the signal produced as output of the processing chain, as a function of the source statistics and the parameters of the operators.

The research on mathematical modeling led to the definition of a novel theoretical framework for the reverse engineering of signal acquisition chains. It was investigated how signals are transformed through a chain of signal acquisition and reconstruction stages, relying on the theory of sampling signals with finite rate of innovation. The model allows one to determine the chain structure and corresponding acquisition history from unknown query signals.

In the case of resampling, the work provided, for the first time, a solid theoretical justification for the use of derivative pre-filters, which are commonly used in resampling detectors based on empirical evidence. In addition, an optimum pre-filter was determined for a given resampling factor and source statistics, so as to maximize the detection accuracy.

Block-wise transform coding is able to accurately model the underlying processing of lossy coding schemes. In this context, the research led to the precise characterization of the statistics of the DCT coefficients when the input signal undergoes double compression. Unlike previous methods, which assumed alignment between block boundaries of the first and second compression, the proposed model works also in the general case of non-aligned double compression. When transform coding is cascaded with full-frame linear filtering, it is widely acknowledged that the detection of quantization footprints is significantly more
difficult. This research addressed an explicit mathematical modeling of this chain, which is particularly challenging due to the fact that combines together block-wise and frame-wise processing operations. The result achieved so far reveal, somewhat surprisingly, that the common assumptions on the statistical independency between DCT coefficients in different blocks do not hold. Indeed, it is exactly this residual redundancy that gives rise to an unexplored class of footprints, which might lead, e.g., to the identification of the properties of the linear filter applied. The work also covered the particularly challenging case in which the type of transform being used is unknown. The proposed solution relies on lattice theory and it successfully recovers the transform matrix when the number of observations exceeds, by a small margin, the dimensionality of the space (e.g., the number of transform coefficients in a block).

In some cases, signals are being modified by people in a malicious way. Hence, the research considered mathematical modeling of processing chains characterized by the presence of adversaries. This work started from the observation that state-of-the-art forensic, counter-forensic and counter-counter-forensic techniques address the problem from either the perspective of the forensic analyst or that of the adversary, thus being entangled in a sort of chicken-and-egg dilemma. Conversely, the problem was addressed in a rigorous way adopting a game theoretic approach. This was initially applied to the problem of source identification, which is one of the most relevant in the field of forensics, leading to the derivation of the ultimate achievable performance of the forensic analysis in the presence of an adversary aiming at deceiving it. The work was then extended in two ways: first, a theoretical analysis considered the problem of understanding the fundamental limits of the source identification problem, by investigating the distinguishability of two sources under adversarial conditions; second, a general information-theoretic framework was proposed for fusing multiple observations, also considering the presence of an adversary who might corrupt part, or all, of them.

The investigation of mathematical models describing single operators and processing chains was paralleled by the design and implementation of several practical footprint detectors applicable to multimedia objects, advancing the current state-of-the-art in the field. In the case of acquisition-based footprints, the main contributions were in the context of re-acquisition of image and video signals. Image recapture might be applied to disguise non-original images and was detected through the analysis of visualized edges. Indeed, a professionally recaptured image from an LCD monitor can be visually very difficult to distinguish from its original counterpart. The proposed method uses elements of the theory of sampling signals with finite rate of innovation to characterize the way edges are blurred and distorted by the sampling kernel modeling camera optics during acquisition. Based on these findings, two alternative dictionaries were trained to provide a sparse representation for single captured edges and recaptured edges, respectively. In addition, another technique was proposed that considers image recapture as a series of local resampling operations. In the case of video, motion-related footprints were studied, to reveal scene jitter introduced in the
recapturing process due to either physiological tremor or environmental factors. Moreover the footprints left by geometric transformations such as forward 3D projection during video recapture; by temporal synchronization between projector and recapture device as well as lens distortion cues were considered.

Coding-based footprints were thoroughly investigated in the case of images, specifically targeting the case of multiple JPEG compression due to its widespread use. The aforementioned mathematical modeling of DCT coefficients undergoing double quantization led to design of practical state-of-the-art detectors, which work also in the challenging case of non-aligned double JPEG compression. This was also extended to the challenging case in which double JPEG compression is interleaved by a linear contrast enhancement, by estimating both the JPEG quality factor and the amount of contrast enhancement. The study of multiple compressions, i.e., involving up to four consecutive compression stages with different parameters, successfully led to the design of a detector based on the analysis of the first digit’s statistics, which is able to discriminate, for the first time, the number of compression stages.

The research addressed also the challenging case of video signals. The problem of video codec identification was addressed for the first time, leading to several strategies for the identification of the class of video coding architecture (e.g., MPEG-2, MPEG-4 or H.264/AVC) used to originally encode a video sequence, when the latter is re-encoded in a possibly different format. The same techniques are also able to identify coding parameters, GOP size and also the characteristics of non-normative tools (e.g., the motion estimation algorithm) used at the encoder side. This enabled to address local tampering detection by exploiting coding-based footprints. Similarly, it was possible to detect the insertion and deletion of whole frames in compressed video sequences. In addition, a method was developed to automatically classify the video streaming source, to detect illegal distribution of content. The work also considered footprints left in video sequences by packet losses, leading to a no-reference video quality monitoring scheme, which relies exclusively on decoded data in the pixel domain, being applicable in cases when the bit-stream is unavailable.

Regarding audio data, fake quality detection was addressed, leading to methods for both lossless and MP3 data, and then extended to the case of double MP3 compression. Specifically, the effect of cross-codec framing grid in audio files was analyzed, in order to detect possible forgeries.

Several approaches were developed towards detecting editing-based footprints in images, videos and audio. For images, the footprints left by editing operations were analyzed from multiple angles based on, e.g., the effect on DCT coefficients, resampling traces, edited/duplicated regions. As for video, algorithms were devised to detect copy-move attacks. In the case of audio editing footprints, a toolbox was developed to detect music plagiarism.
Besides considering detectors for operators applied to individual media, the research considered the very challenging problem involving multimodal signals. An information theoretic framework was proposed for quantifying the operator estimation improvement that can be achieved by considering the joint pdf of the involved signals, i.e., by exploiting the correlation of the different input signals; specifically, closed formulas were derived for several noiseless and memoryless scenarios where the input signals are Gaussian distributed and the operators are linear. From a different perspective, audio-visual cues were jointly studied to reveal inconsistencies between a video sequence and the accompanying audio track, which might be the result of tampering. On the one hand, a multi-trace analysis on audio files was proposed, based on the combination of blind microphone analysis and phase analysis of stable tones. On the other hand, the problem of identifying the demosaicing algorithm used during the acquisition phase was addressed. The combination of audio and visual analysis was then applied for obtaining a more robust identification of the acquisition device. A multimodal analysis of audio-visual content was also proposed to detect the footprints left by the environment (e.g., indoor vs. outdoor), rather than by the acquisition device. In order to deal systematically with multimodal analysis, a decision fusion framework was explored. Two different methodologies were considered to decline this framework in the context of audio-visual content analysis: Dempster-Shafer theory of evidence and Fuzzy theory. The work was also extended to the case in which background knowledge about the performance of the individual detectors is available to the analyst.

The presence of adversaries was explicitly considered for different concrete footprint detectors, thus producing several strategies that the adversary may use to fool state-of-the-art detectors and, on the other hand, corresponding countermeasures against specific counter-forensic attacks.

From the adversary side, different methods were developed to attack methods based on local features. Specifically, copy-move detection based on matching of SIFT local features was addressed not only by removing the SIFT features present in the to-be-attacked image, but also by reintroducing fake SIFT keypoints into the previously cleaned image. A similar strategy was pursued to fool automatic image classifiers based on Bag-of-Visual-Words, which rely on the extraction of local features in the first place. In some cases, the specific details of the detector are unknown to the adversary. Yet, the adversary looks at the detector as a black-box, computing the feature he wants to attack and then injecting within a digital image a specific feature that tells the history the adversary wants to be believed. Particular emphasis was given to attacks related to the JPEG compression history. On the adversary side, a tool able to mislead a double compression detector based on the Benford’s law was proposed, which relies on the analysis of the histogram of First Significant Digits (FSD). This was generalized by devising a strategy that is able to attack any detector based on FSD first-order histograms, regardless of the specific method adopted, thus being universal. In the case of video, a counter-forensic strategy was proposed for attacking the video camera.
identification algorithm, based on the recognition of the motion estimation strategy used by the video encoder. From the analyst side, tools for revealing the traces left by counter-forensic strategies were developed, addressing attacks towards single/double compressed based on denoising and full-frame filtering.

In some cases, the analysis of a single multimedia object might not be sufficient to reveal its processing history. As such, it is interesting to consider sets of multimedia objects simultaneously. In this direction, REWIND has addressed the challenging goal of multimedia phylogeny, which aims at revealing the dependencies among sets of multimedia objects (e.g., audio or image near-duplicates, video composites referring to the same event, etc.). The research started with the definition of suitable dissimilarity measures considering either pairs of near-duplicates (including audio, in addition to images and videos), or objects obtained combining parts of other objects. This is a necessary prerequisite to populate a dissimilarity matrix from which phylogenetic trees or forests might be reconstructed. To this end, a number of significant contributions were made: i) the reconstruction of large-scale trees from incomplete dissimilarity matrices, by exploring the phylogeny problem for near-duplicate images in large-scale scenarios, and devising solutions that have straightforward extension to other media such as videos; ii) the reconstruction phylogenetic forests, in which the identification of the structure of relationships underlying the images was addressed, correctly reconstructing their past history and ancestry information, and grouping them in distinct trees of processing history, also in the challenging case in which the number of trees is not known a-priori; iii) the proposal of a state-of-the-art algorithms to reconstruct phylogenetic trees, based on either heuristic criteria or optimum branching; iv) the exploration of dependency graph algorithms for finding multiple parenting relationships for both images and videos (e.g., one object that is a result of a combination of two or more objects through splicing/blending algorithms); v) the fusion of complementary phylogeny algorithms to improve the robustness of the solution to the phylogeny problem.

The scientific activities of the partners involved in the project have generated more than 90 papers published on scientific journal or on the proceedings of international conferences. All these paper are listed and can be downloaded from the project web-site (www.rewindproject.eu).

The REWIND Consortium recognizes that, for a fruitful and comprehensive dissemination, it is necessary that the results obtained by the project are reproducible. In order to achieve this goal, particular attention was given to the identification of specific application scenarios and corresponding use cases. For each of them, carefully annotated data with very fine-grain information about single actions performed upon content. Many of the datasets used are now available to the public on the project website, thus representing a significant contribution to the scientific community working in this field.
1.3 Dissemination of the project results

Dissemination is a key issue for the REWIND project. It is pursued through the publication of the research results on top-ranked journals and conferences, but also by means of demos and special sessions organized at important conferences.

In particular, during the last few months, REWIND participated to the following events:

**The 13th ENFSI DIWG Meeting**
A REWIND representative (Alessia De Rosa) participated to the 13th Meeting of the European Network of Forensic Sciences, Expert Working Group for Digital Imaging (ENFSI-DIWG) that took place in Athens on October 1-4 2013. More than 60 European government forensic laboratories belong to ENFSI ([http://www.enfsi.eu](http://www.enfsi.eu)), and the DIWG meetings are dedicated to present and analyze the innovation and trends on image/video forensics. At the meeting in Athens REWIND presented the results of the project activities in the field of image and video analysis. These results were extremely appreciated by the attendees that also expressed their interest in testing the tools that are available on the REWIND website. The ENFSI-DIWG Theme Leader (Patrick De Smert) wrote an appreciation letter to the REWIND project (see Appendix 1).

**The ICT-Event (Vilnius, November 2013)**
REWIND participated to ICT-Event that was organized by the European Commission in Vilnius on November 6-8 2013. REWIND showcased its approach and results in a booth at the exhibition organized during the event. The booth was visited by a large number of people and several contacts were established for possible exploitations of the project technologies.
The REWIND booth at the ICT Event in Vilnius.

The 2014 GTTI Thematic Meeting on MMPS
The Thematic Meeting on Multimedia Signal Processing gathers researchers and professionals focused on multimedia signal processing and its applications. It provides a venue for people to interact and exchange ideas about state-of-the-art scientific achievements, as well as to promote initiatives and encourage the development of the signal processing community. The fourth edition of the meeting took place in Forni di Sopra (Udine, Italy) on February 16-18, 2014 and it was supported by the IEEE Signal Processing Society - Italian Chapter and by the National Telecommunications and Information Technologies Group (GTTI). During the event the REWIND Italian partners (PoliMi and CNIT) presented the recent results and demos developed by the project.

REWIND Audio Demos presented at CEBIT 2014
The Fraunhofer IDMT Institute, as a REWIND partner, presented at its CEBIT 2014 (Hannover, March 10-14 2014) booth, the demos developed within the REWIND project related to the analysis and processing of audio data. In particular demos on Audio manipulation detection, Sampling Plagiarism detection, Melody plagiarism were showcased.
IEEE IFS-TC Image Forensics Challenge – supported by REWIND

In the last years, the field of Digital Image Forensics has emerged with several tools and algorithms presented to help researchers and forensics experts to interpret the authenticity of digital documents. However, the lack of a clear benchmark and common comparison protocol of such algorithms has limited us to gauge them under real-world conditions.

In this context, the IEEE Information Forensics and Security Technical Committee (IFS-TC) has worked on the idea of launching a detection and localization forensics challenge. The REWIND Consortium decided to join their effort and significantly contribute to the organization and success of the IEEE IFS-TC Image Forensic Challenge. This was possible due to the fact that several researchers involved in the REWIND project are also part of the IFS-TC. Specifically, the REWIND Consortium is represented in the Steering Committee of the Challenge (3 persons) by A. Piva (CNIT) and A. Rocha (Unicampinas).

To guarantee the maximum visibility and attractiveness of the challenge, it appeared as an IEEE initiative while REWIND was a major sponsor.

The major objectives of this Challenge are:

- To provide the community an open data set and protocol for evaluation of the latest forensics techniques to identify forgeries in digital images.
- To evaluate the current state-of-the-art techniques with respect to their ability to detect image forgeries (content manipulation).
To set forth for standardization protocol as a common comparison ground truth for new techniques.

The challenge was organized in 2 phases: Phase 1 - deciding whether or not an image is manipulated, and Phase 2 - localizing the manipulated regions. The challenge started on July 201 and ended before the Workshop WIFS’13 (November 2013), where results of the whole challenge were presented. Although REWIND did not participate to the challenge (since the RFEWIND partners had access to the ground-truth material which have been disclosed to the public only at the end of the challenge), the Consortium used the challenge datasets to validate the techniques developed within the project. Moreover, to increase the attractiveness of the challenge (and to disseminate REWIND as much as possible among the forensics community) the REWIND partners contributed to the money prize granted to the challenge winners.

1.4 REWIND on newspapers and on the web

Even if the project was strictly oriented to develop new approaches and techniques (as typical for a FET-Open project) and therefore without specific and immediate applications, thanks to the REWIND researchers proactivity, the results of the project have been showcased several times on newspapers and on news websites. A complete list of them is reported on the REWIND website, and it appears in the following:

- March 18, 2014 – A presentation of the REWIND audio tampering detection activities presented by Fraunhofer IDMT at CEBIT 2014 on www.schattenblick.de
• March 16, 2014 – A presentation of the REWIND audio tampering detection activities presented by Fraunhofer IDMT at CEBIT 2014 on www.radio-satisfaction.de
• February 24, 2014 – An article on the REWIND tool on Audio Plagiarism Detector developed Fraunhofer IDMT on www.silicon-saxony.de
• February 10, 2014 - REWIND mentioned in an article appeared on Ilsole24ore
• February 4, 2014 – An article on the REWIND tool on Audio Plagiarism Detector developed Fraunhofer IDMT on www.ingenieur.de
• November 13, 2013 - The workshop organized by REWIND on "Digital doctoring: How to tell the real from the fake" is covered in wired.it
• November 11, 2013 – An article on How detect fakes on Internet Contents mentioning REWIND appeared on the Marco Bosco’s blog
• November 8, 2013. An interview to the REWIND coordinator on SoundCloud
• November 8, 2013. An interview to the REWIND coordinator on Radio Città del Capo
• May, 17 2013 - REWIND mentioned in a wired.it article on digital doctoring.
• February 13, 2013 - Early in 2013, the software described in the paper “Audio Forensics Meets Music Information Retrieval - A Toolbox for Inspection of Music Plagiarism” (Dittmar et al. EUSIPCO 2012) was successfully applied to disprove allegation of music plagiarism with the German entry to the Eurovision Song Contest (ESC). BILD Zeitung
• March 9, 2013 - Anderson Rocha - University of Campinas - interviewed in a featured article on image forensics on the Economist
• April 11, 2012 – An article with an interview to Prof. Fernando Perez Gonzales (University of Vigo): "O ponemos coto al photoshop o las fotografías perderán su valor testimonial", Faro de Vigo

1.5 NewDisclaimer Note for the REWIND tools
A newDisclaimer notehas been added to the REWIND Tools available for download on the project website, please take a look.

The REWIND researchers invites all the newsletter readers to test the tools and datasets available on the website, any comment and suggestion is welcome.

Further information is available at: www.rewindproject.eu.
<table>
<thead>
<tr>
<th>Project Public WebSite</th>
<th><a href="http://www.rewindproject.eu/">http://www.rewindproject.eu/</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Contact Point</td>
<td>Prof. Stefano Tubaro</td>
</tr>
<tr>
<td></td>
<td>Dipartimento di Elettronica, Informazione e Bioingegneria</td>
</tr>
<tr>
<td></td>
<td>Politecnico di Milano</td>
</tr>
<tr>
<td></td>
<td>Piazza Leonardo da Vinci 32</td>
</tr>
<tr>
<td></td>
<td>20133 Milano</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
</tr>
<tr>
<td></td>
<td>Tel/Fax: +39 02 2399 3647/3413</td>
</tr>
<tr>
<td></td>
<td>Email: <a href="mailto:stefano.tubaro@polimi.it">stefano.tubaro@polimi.it</a></td>
</tr>
</tbody>
</table>