Chapter 7

Conclusions

7.1 Concluding remarks

In this thesis, we have covered different aspects of the field of protection of data that have to be made available to non completely trusted users. Data must be protected against unauthorized uses while preserving their utility. Therefore, protection must stay imperceptible, *i.e.* transparent.

The primary concern has been to offer a broad overview of current techniques for transparent data protection. Such techniques depend on the kind of data and the kind of risks we wish to protect data against. We have focused on transparent protection for two kinds of data: multimedia contents and statistical microdata.

Regarding multimedia contents, we first have studied the use of steganography for protecting data against unauthorized redistribution. More precisely, we have presented proposals corresponding to the two main subfields of copyright protection: watermarking and fingerprinting. A second achievement has been to develop steganographic techniques for providing lossless authentication of multimedia contents. In both cases, the multimedia

contents being considered consist of digital images.

Statistical disclosure control methods have been studied for microdata protection together with measures to compare performance of different methods in terms of information loss and disclosure risk.

7.2 Results of this thesis

Several results have been presented in this thesis.

In Chapter 3, our contributions to watermarking for digital images have been presented. First, a new visual components algorithm to guide image watermarking in achieving imperceptibility has been proposed. The algorithm provides information on the maximum alteration each pixel of an image can suffer without damaging the visual quality of the image. This algorithm can be plugged into watermarking algorithms that operate in the spatial domain and embed the information by directly incrementing/decrementing the color level of pixels.

Next, two new image watermarking schemes have been proposed. Both of them achieve imperceptibility using the aforementioned visual components algorithm. The first one is semi-public and offers high embedding capacity and robustness against compression, filtering and scaling attacks. The second one is oblivious and offers medium embedding capacity; its mark recovery algorithm does not require previous knowledge on the embedded watermark and offers robustness against compression, filtering, scaling and moderate geometric distortion attacks.

Mixture of watermarked digital objects has been presented as a way to increase robustness of current proposals. Prior mixture has been shown to be effective as a technique to combine the robustness properties of a set of image

watermarking algorithms. Posterior mixture has been proposed as a techique which aims at making the mark recoverable again when several attacked versions of the same content are found and the mark is not recoverable from any of them.

In the field of watermarking for image authentication, we have studied the invertibility of the well known spatial-domain spread-spectrum algorithm. Our study shows the suitability of this algorithm for lossless image authentication.

In the field of fingerprinting, Chapter 4 proposes a new construction for obtaining collusion-secure fingerprinting codes robust against collusions of up to three buyers. This construction provides, for a moderate number of possible buyers, shorter codewords than those offered by the general construction of [BS95] for $c = 3$.

In the protection of statistical microdata, a modification to a current performance metric to evaluate masking methods has been proposed in Chapter 5. The modified metric is more general in that it can deal with methods resulting in masked files whose number of records is not the same of the original file.

Also in statistical data protection, a procedure has been presented which enhances the performance of masking methods in terms of the previously mentioned metric. It is a post-masking optimization procedure which postprocesses masked files in order to decrease information loss while leaving disclosure risk practically unchanged. In this way, a better tradeoff between information loss and disclosure risk is obtained. This procedure has been shown to enhance the performance of the two best-performing masking methods: rankswapping and multivariate microaggregation.

The last chapter of this thesis has presented the novel concept of

multilevel access to precision-critical data. In this way, protected data are made available to different users, who, depending on their clearance, can remove part of the noise introduced by protection, thus obtaining better data quality. Two methods have been described to provide multilevel access to masked microdata files and watermarked digital images.

7.3 Future research

We sketch here some open problems that remain to be solved and possible extensions to some of the presented proposals that will be addressed in the future.

In the field of watermarking, our future research will be directed to:

- Achieving tamper-proofness in watermarking. That is, design watermarks that cannot be removed even if the intruder knows the particular watermarking algorithm used to embed them.
- Study the invertibility of other robust watermarking schemes for lossless image authentication.

Research on short binary fingerprinting codes robust against collusions of size greater than three should be done. Our future research in this field will be directed to the construction of codes robust against collusions of size greater than 3.

In statistical microdata protection, the presented post-masking optimization procedure can be extended in at least two directions:

- Study its applicability as a generator of *synthetic microdata*.
- Extend the procedure to preserve all moments up to m-th order.