Steinebach, Kremer, Croce-Ferri: "Automatic Image Theft Detection in eBay by Digital Watermarking", Axmedis 2005: Proceedings of the 1st International Conference on Automated Production of Cross Media Content for Multi-channel Distribution, ISBN 88-8453-355-4, University Press Firenze, 2005

Automatic Image Theft Detection in eBay
by Digital Watermarking

Martin Steinebach, Ellen Kremer, Lucilla Croce Ferri
Fraunhofer Integrated Publication and
Information Systems Institute (IPSI),
Dolivostr. 15, D-64293 Darmstadt
{Martin.Steinebach}, {Ellen.Kremer}, {Lucilla.Croce-Ferri}@ipsi.fraunhofer.de

Abstract

Digital images are used in the Internet for a broad range of applications. One well known example for image usage are product photographs in eBay auctions. In recent times the misuse of these images is often discussed and reported. Either images are re-used from third party auctions or they are copied from other web sites like for example online catalogues. As eBay offers a high number of auctions, image theft often stays unnoticed. We introduce an automated method for image theft detection using digital watermarking and an eBay online interface. Images are scanned based on product description filters, downloaded and scanned for embedded watermarks.

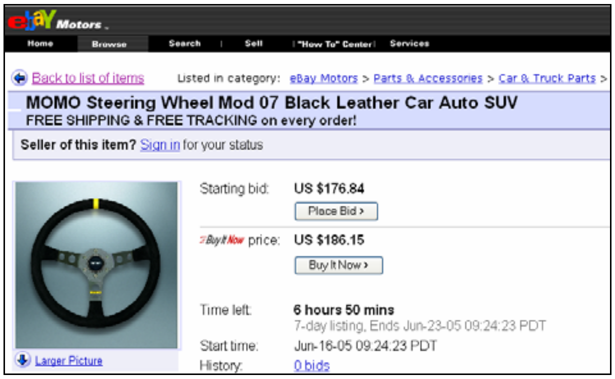
1. Motivation

Online auctions are very popular today. eBay is the best known representative for this business type. Many people use digital images to show potential bidders how the products they offer look like. Figure 1 provides an example.

Various reasons may lead a seller to a situation where he would like to provide an image, but has no access to one. Examples are lack of a digital camera, difficulties in providing professional images or a fraud where the seller actually does not own the product he is offering. In all cases sellers tend to use images already existing on the Internet, perpetrating a misrepresentation fraud. They can originate from other eBay auctions or from online shop catalogues where the products are sold. Either the image is copied by the seller to a web storage to which he refers or he simply uses the URL of the original copy of the image and places it into his offer. We call a seller using copied images "pirate seller" for the rest of this article. Both methods are copyright violations against the original creators of the digital images, leading to complains like this:

"someone named as […] stole my pictures and now he is selling the same items as mine. what will i do for him to stop copying my picture?"[1]

But it can be assumed that many copyright violations pass unnoticed by the original owners of the images due to the vast number of auctions hosted at eBay. Still owners of web shops using high quality image material to advertise for their products are looking for an efficient way to identify copyright violations. An automated system would be necessary to be able to scan the huge amount of images in an acceptable amount of time.



MOMO Steering Wheel Mod 07 Black Leather Car Auto SUV
FREE SHIPPING & FREE TRACKING on every order!

Seller of this item? Sign in for your status

Starting bid: US $176.84
Buy It Now price: US $186.15
Time left: 6 hours 50 mins
7-day listing, Ends Jun-23-05 09:24:23 PDT
Start time: Jun-16-05 09:24:23 PDT
History: 0 bids

Figure 1: example eBay auction with image

When an image theft is noticed, various reactions are possible, depending on the method of theft.

When the pirate seller uses an URL pointing to the original image, known countermeasures include (a) changing the image or (b) changing the image location.

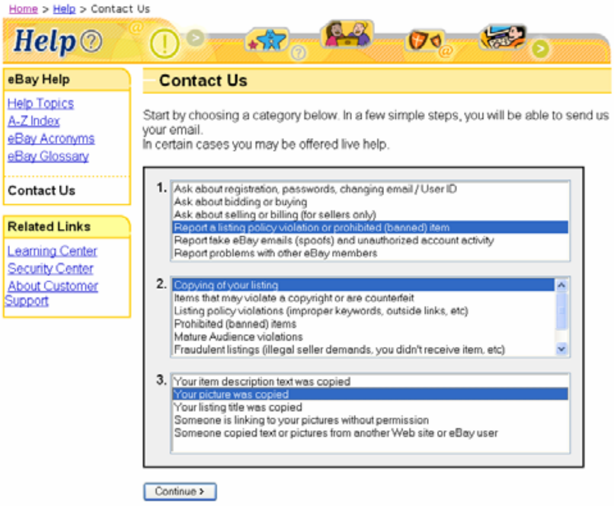
[1] http://forums.ebay.ph/thread.jspa?threadID=300000585&tstart=0&mod=1117000770127

In (a) the image could be changed into a sign indicating a copyright violation. Of course, the original owner of the referred image must change the URL of his image and his references to display the correct image. Method (b) is simpler and only leads to a broken link in the auction of the pirate seller.

But if a pirate seller copies the image to a web space the original owner has no access to, a third party must be called to handle the problem. In the case of eBay, a web form (see figure 2) for complains about pirate sellers is available. This leads into stopping the auctions of the pirate seller. Of course this method can also be chosen when the pirate seller is using only an URL to the original image[2].

As one can see, an owner of an image is not helpless against copyright violations of his images.

In this article, we address the part of automatically detecting image copyright violations in eBay using digital watermarking algorithms. In section 2 we introduce digital watermarking. In section 3 we introduce a concept using an eBay scanner and a watermarking detector. In section 4 we describe our implemented prototype with respect of reliability and performance.



Contact Us

Start by choosing a category below. In a few simple steps, you will be able to send us your email.
In certain cases you may be offered live help.

1. Ask about registration, passwords, changing email / User ID
   Ask about bidding or buying
   Ask about selling or billing (for sellers only)
   Report a listing policy violation or prohibited (banned) item
   Report fake eBay emails (spoofs) and unauthorized account activity
   Report problems with other eBay members

2. Copying of your listing
   Items that may violate a copyright or are counterfeit
   Listing policy violations (improper keywords, outside links, etc)
   Prohibited (banned) items
   Mature Audience violations
   Fraudulent listings (illegal seller demands, you didn't receive item, etc)

3. Your item description text was copied
   Your picture was copied
   Your listing title was copied
   Someone is linking to your pictures without permission
   Someone copied text or pictures from another Web site or eBay user

Figure 2: eBay image piracy notification form[3]

2. Digital Image Watermarking

Digital watermarking ([CoMB02], [Dit00]) invisibly embeds information into a cover with the help of a secret key. This information refers to the cover, and provides additional information about it accessible only by those who own the watermarking algorithm and the secret key. The most common application of this technology is copyright protection or customer tracing. Both can be seen as a simple stand-alone alternative to complex and more restrictive digital rights management (DRM) or easily inserted into existing security concepts.

One of the most important advantages of watermarking is the fact that content can escape from DRM environments based only on cryptography and access control. This would render it unprotected. One example for this is when the content is transmitted via an analogue channel. A well-designed digital watermark, on the other hand, stays in the content even after printing and scanning or manipulations like strong JPEG compression. So when watermarked content is found in an illegal environment, copyright claims can be proven or original customers can be identified.

For this reason, the challenge in DRM is mainly to keep material in a protected environment, while in watermarking one needs to find watermarked content which is used illegally. This makes efficient search strategies an important aspect of digital watermarking.

2.1 Image watermarking state of the art

A large number of watermarking algorithms has been proposed in the recent years. Most of them deal explicitly with still images and share common approaches.

Some basic watermarking requirements can be identified independently from the various applications. These requirements are related mainly to the perceptual transparency after the watermarking embedding, to the watermarking capacity, i.e. the quantity of information that can be embedded into the data, the security of the watermarking technique and the watermarking robustness against common processing techniques or intentional manipulations of the data. Even if transparency, capacity and robustness are trade-off parameters, for images the robustness represents the most challenging parameter. The geometrical transformations of images caused by printing and scanning are very challenging processes for watermarks search mechanisms. They can cause serious robustness problems to many watermarking algorithms not explicitly designed to survive them. General methods to achieve high robustness against these transformations are based on resynchronization techniques, such as the usage of the original image for non-blind detection methods, registration patterns or extraction of characteristic feature points found in the original image for blind ones.

[2] http://cgi3.ebay.com/ws/eBayISAPI.dll?ViewUserPage&userid=net_enforcers_inc

[3] http://pages.ebay.com/help/contact_us/_base/index.html

A local exhaustive search mechanism can be necessary in combination of the previous mentioned methods. Other possibities are invariant watermarks, which remain unchanged under the considered geometrical transformation and autocorrelation techniques for period watermarks. [3]

Geometrical transformations are not the only manipulations that can corrupt the embedded watermarks. Also lossy compression, noise and luminance changes can modify the images in such a way that the watermarks cannot be detected anymore or only partially. The approached used in this case to achieve the required robustness are based mainly on the redundant embedding into perceptually significant parts of the image. Spread spectrum techniques [4] are used for this purpose in the spatial and frequency domain.

2.2 Searching Strategies

Different commercial applications and services based on web-crawling already exist, that look for watermarked data in Internet. Perhaps the most known one is the service offered by Digimark[4], but also other companies are providing commercial watermarking and searching services, promising secure online distribution of images with usage tracking[5]. Not only images need to be protected. Audio mp3 data are particularly vulnerably and new business models based on their individual fingerprinting are adopted by always more publishing companies [7].

3. Concept

In this section we describe the general idea of our watermarking approach for eBay. A prototype for auctions crawling, based on the Gnutella architecture was already implemented [6]. This search system looked for watermarked data.

The section discusses the different stages and the necessary components. The actual implementation of these is described in the next section.

Our approach is to watermark images used by original owners and search for misuse of these marked images in eBay using keywords to identify fitting auctions. A complete process including fraud detection would feature the following steps:

1. **Original image is watermarked by its owner.** Only the watermarked copy is used in the public. The unmarked copy is stored in a secure place or deleted. The embedded message identifies the original owner. The secret key used in the embedding process is stored.

2. **Marked image is put in the public.** This can either be an eBay auction, an online catalogue or even a printed catalogue if the watermark is robust against printing and scanning, Thereby the image is made accessible for potential pirates.

3. **Image is included in scanner list.** The original owner informs the online eBay scanner that it should search for misuse of this image. He provides the secret key which is necessary to retrieve the watermark and a list of keywords which describe the product to be seen on the image.

4. **Scanning starts.** The scanner is now continuously looking for auctions with fitting keywords. If such an auction is found and it features an image, it is downloaded and the secret key is used to check if the original owners watermark is present. If the original owner uses his image in an eBay auction, this process will also find his own auction showing the successful operation of the scanner.

5. **Pirate seller places auction using stolen image.** Now the pirate seller copies the marked image into his eBay auction. He describes the product he wants to sell and uses some of the keywords the original owner submitted to the scanner.

6. **Scanner detects misuse.** The auction of the pirate seller is found by the scanner using the keywords. The image is downloaded and the watermark of the original owner is detected.

7. **Alert.** The scanner now informs the original owner about the misuse, sending a copy of the auction including the image to him. The original owner can now ask eBay to shut down the auction. As an alternative, the scanner could also inform eBay automatically about the misuse.

3.1 Example

A seller wants to start an auction of a steering wheel. He takes a photograph of the product to sell and embeds a watermark consisting of his name and using his secret key into it. The marked copy is then used in the eBay auction illustrating the product to be sold. Figure 3 shows the usage of the image.

[4] http://www.digimarc.com/

[5] http://www.bluespike.com/giovanni.html, http://www.alphatecltd.com/watermarking/eikonamark/eikonamark.html

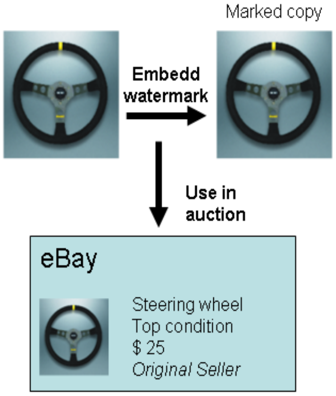


Figure 3: Image of steering wheel is marked and used in an auction

Now a pirate seller also wants to sell a similar steering wheel. Instead of taking an own photograph, he is looking for other auctions selling steering wheels in eBay. He chooses the image of the original seller and copies the image to his own web space. Then he uses the marked copy in his own auction. As seen in figure 4, he offers a "top steering wheel" which is "used, as new".

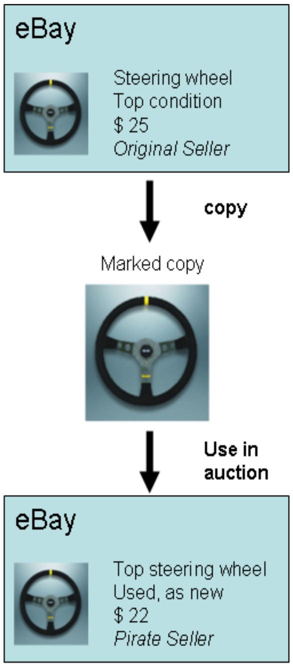


Figure 4: The image is copied and misused by a pirate seller

But the original seller is aware of potential image theft and therefore sends his secret key and a list of keywords to an eBay scanner. The scanner now checks all images in auctions of the transportation domain featuring the keywords "car wheel" and "steering wheel" as shown in figure 5. Over time, it finds three images. One is from the auction of the original seller and a watermark is detected in it. The auction is then listed in a report sent from the scanner to the original seller. The next image is the stolen image of the pirate seller where the watermark is also found in. The occurrence of the image is also noted in the report. The third image is a different image of a steering wheel. No watermark of the original seller is detected here. The scanning is repeated for a given amount of time specified by the original owner. Images or auctions already checked are stored in a data base to prevent repeated downloading and detection.

The original seller regularly receives the report of the scanner. If any other than his own auctions use his images, he can either react on his own or address eBay to stop the pirate seller auction. In either way, he does not need to scan for auctions misusing his images on his own. Any auction in the transportation domain selling steering wheels or car wheels will be scanned for his image automatically for a period defined by him.

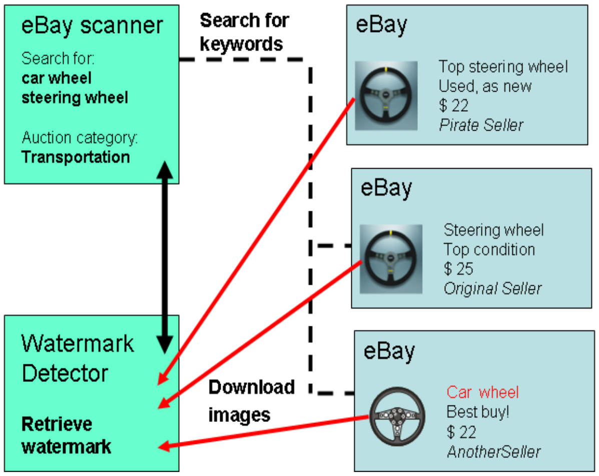


Figure 5: The eBay scanner is searching for fitting keywords and finds the auction of the pirate seller

4. Implementation and test results

In this section we describe our prototypic implementation based on the concept introduced in the previous section. We also provide test results and identify possible bottlenecks in a commercial application.

4.1 System design

Our scanner uses the eBay product search feature to search image files. It accesses the eBay platform via Internet, utilizing an API provided by eBay. This API enables to access eBay via web services (see Figure 6).

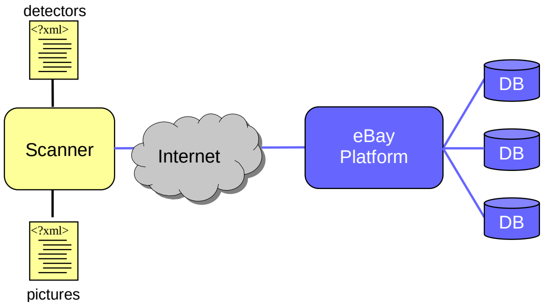
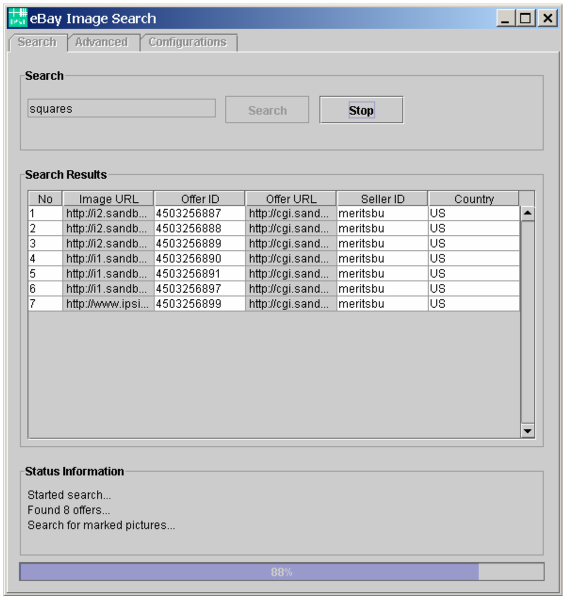


Figure 6: eBay Image Scanner communicates via internet with the eBay platform

The information whether an already scanned image is watermarked or not, is stored in an xml file in order to avoid multiple scanning of images.
Following steps are performed by the system to process a search query:

1. Submit keywords given by the user to the product search offered by eBay.
2. Identify the product offers containing an image.
3. Download the images from this selected product list and calculate their hash value
4. Check if the hash value already exists in the xml file. In this case, check if the image was watermarked.
5. Otherwise, check if the image is watermarked and store its hash value.
6. Display the results of these steps to the user as shown in Figure 7.



squares

No	Image URL	Offer ID	Offer URL	Seller ID	Country
1	http://i2.sandb...	4503256887	http://cgi.sand...	meritsbu	US
2	http://i2.sandb...	4503256888	http://cgi.sand...	meritsbu	US
3	http://i2.sandb...	4503256889	http://cgi.sand...	meritsbu	US
4	http://i1.sandb...	4503256890	http://cgi.sand...	meritsbu	US
5	http://i1.sandb...	4503256891	http://cgi.sand...	meritsbu	US
6	http://i1.sandb...	4503256897	http://cgi.sand...	meritsbu	US
7	http://www.ipsi...	4503256899	http://cgi.sand...	meritsbu	US

```
Started search...
Found 8 offers...
Search for marked pictures...
```

Figure 7: Example search in the eBay image scanner

The result of the search query contains the URLs of the image and the offer, together with the offer and seller IDs. Furthermore the country, where the eBay offer originates, is also displayed.

4.2 Implemented functionalities

The graphical interface for the scanner (Figure 8) offers the possibility to use advanced search functionalities for the keywords in the offer title. The user can specify or exclude selected words or all the listed words as criteria for the search. Also portions of sentences can be used and the search can be performed not only in the title but also in the complete offer text. Furthermore the user can decide to search only selected eBay categories, the eBay stores or offers from a specific date.

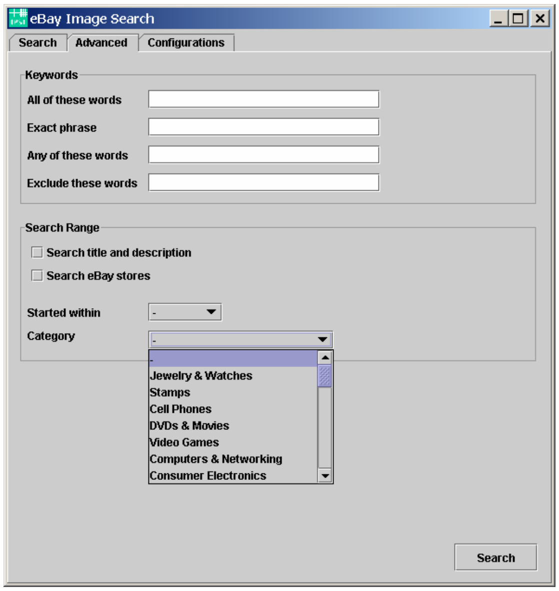


Figure 8: Advances search options of the eBay image scanner

The eBay Image Scanner allows using different watermarking detection algorithms. They are shown in the "configuration" page. The user can decide which one is used for the current search.

The current advanced prototype can be enhanced in the future particularly in relation to the configuration and documentation facilities and its user friendliness.

A possible extension of the searching filters is considered and a more structurable display of the results. Other new possible functionalities are related to an automatic notification of the results to the user or directly to the eBay supporters. In this last case, mechanisms should be added to definitively ensure the fraudulence of the founded watermarked images.

Another important issue is the amount of information stored in the result list. At the moment, the user receives only a list of possible illegal usages of a specified image. For forensic applications, it would be necessary to collect all information about the suspected offer and a screenshot that can be used as fraud evidence.

To enhance user friendliness, user profiles can be supported, containing the different configuration settings.

4.3 Performance

Two types of performances are important for our eBay Image Scanner. The first one is related to the false and positive detection errors of the watermarking algorithm and the second one to the time performance of the whole system.

Compared to other existing commercial crawling tools, for example the Digimarc image tracking service[6], the Image Scanner is not limited to a specific watermarking algorithm. Different methods can be linked to the Image Scanner thought a well defined interface specification.

Following requirements have to be satisfied by the watermarking methods, to optimize the system performance:

- Blind detection, i.e. the detection algorithm does not utilize the original image to extract the watermark. This is, of course, a mandatory requirement.
- High robustness against scaling and cropping, since pirate images can be a slightly different version of the legal ones.
- High robustness against luminance and compression changes for the same reason.
- Low complexity, at least for the detection process, in order to reduce the detection time. This will play a fundamental role in minimizing the processing time of the whole scanning process.

4.4 Possible attacks

If the watermarking algorithm is available to everyone, attacks trying to overwrite or destroy the watermark are possible [5]. The robustness and security of the watermarking algorithm have to be proved critically, before the method is registered by the scanner.

It has to be pointed out, that the knowledge necessary to perform complicated attacks are normally behind the possibilities of most of normal eBay users.

Also secure protocols for the transmission of the secret key needed during the detection process have to be utilized. If a pirate has access to the secret key, he could try to generate his own watermark with the purpose to replace the original one.

It could be also necessary to make the scanner anonymous and to mask its IP address, to avoid misleading it. A professional attacker could monitor the searching activities of the scanner, temporarily substitutes the illegal used images with some other images thereby hide his illegal usage of the images.

Possible technical interferences of the scanner with the auction functionalities are not an issue, since the scanner acts like a normal user, looking for a specific product. Still an explicit cooperation with eBay would

[6] http://www.digimarc.com/products/imagebridge/MarcSpider/default.asp

be desirable, may in the form of an additional security service offered by eBay.

5. Discussion and Future Work

In this section we briefly describe the advantages of our approach for different potential groups of users and the planned extensions to the prototype, in order to enhance its usability.

5.1 Potential users

The main benefits of our automatic eBay Image Scanner are related to the possibility of monitoring copyright infringements for images. This is an open issue in particular for professional online catalogues, online image archives or professional eBay sellers, whose images are often stolen and somewhere else illegally used. With the eBay Image Scanner, the catalogue operators would be able to track these abuses, having a mean to demonstrate their legal position. Since the watermarking algorithm has to be designed to be robust against geometrical attacks, images need to be watermarked only once. The legal user can publish them in different contexts, in different Web pages with different formats, without the necessity to mark all the different versions of the images.

The deterrent effect would produce also an advantage for eBay itself, since it would reduce the number of reclamations that they have to process, as answer to stolen victims' protests. But, of course, also users and new potential sellers would appreciate the enhanced security against image misrepresentations. This could contribute to increase the transactions' volume and to open new online markets for high valued images and photographs.

Another application for the eBay Image Scanner is the online brand monitoring. Services based on the search of illegal usages of marked images in eBay could be offered to assist marketing and branding professionals to protect to their brands.

5.2 Extension

The current advanced prototype can be enhanced in the future particularly in relation to the configuration and documentation facilities and its user friendliness.

A possible extension of the searching filters is considered and a more structured display of the results.

Other new possible functionalities are related to an automatic notification of the results to the user or directly to the eBay supporters. In this last case, mechanisms should be added to definitively ensure the fraudulence of the founded watermarked images.

Another important point is about the protocols of the results. At the moment, the user receives only a list of possible illegal usages of a specified image. For forensic applications, it would be necessary to collect all information about the suspected offer and a screenshot that can be used as fraud evidence.

About enhanced user friendliness, user profiles can be supported, containing the different configuration settings.

6. Summary and Conclusion

We proposed an automated method for image theft detection based on digital watermarking and providing an eBay online interface. Images are scanned using product description filters, downloaded and scanned for embedded watermarks. The listed results can be analysed by the legal image copyright holder, who decides the countermeasures against the pirates. It is important to point out that, in order to ensure the optimal image scanner performance, all online existing images have to be watermarked before their publication, otherwise the attacker could utilize image versions which are not protected and these would not be found by the scanner.

Another strategy used to discourage image theft is the embedding of a visible watermark, such as a company logo, or the website URL, or any other copyright text into the images.[7]

The most important advantage of our approach is that the high quality of the published images is not damaged by the visible mark. Furthermore, in most cases, the visible mark could also be easily cropped out, while the invisible watermark can be detected by the scanner also after cropping transformations.

Not only the sellers of eBay actions would benefit from the image scanner services, but also eBay itself should have interest in supporting and offering them to its users.

Acknowledgements

The work described in this paper has been supported in part by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT. The information in this document reflects only the author's views, is provided as is and no guarantee or warranty is given that the information

[7] http://www.vendio.com/my/ihost/promo_wm.html

is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

References

[1] I. Cox, M. Miller, and J. Bloom, ”Digital Watermarking”, San Francisco, CA: Morgan Kaufmann, ISBN 1-55860-714-5, (2002)

[2] J. Dittmann, "Digitale Wasserzeichen" Springer Verlag, Berlin/Heidelberg, ISBN 3-540-66661-3, (2000)

[3] P. Bas, JM. Chassery and B. Macq, “Geometrically Invariant Watermarking Using Feature Points”, IEEE
Trans. Image Processing, Vol. 11, Nr. 9, pp. 1014-1028, (2002)

[4] Cox, I., Kilian, J., Leighton, T., Shamoon, T., "Secure Spread Spectrum Watermarking for Multimedia", IEEE Transactions on Image Processing, Vol.6, Nr.12, pp.1673-1687,(1997)

[5] Johnson, Duric, Jajodia, “Information Hiding: Stenography and Watermarking Attacks and Countermeasures”, Kluwer Academic Publishers, ISBN: 0-7923-7204-2, (2001)

[6] Steinebach, Dittmann, Lang: "Konzepte zur Vermeidung oder Verfolgung von Urheberrechtsverletzungen in Netzwerken auf der Basis digitaler Wasserzeichen", Competence in Content, Ralf Schmidt (Hrsg.), Tagungsband 25. Online-Tagung der DGI, S. 113 – 125, ISBN 3-925474-58-x, 2003

[7] Steinebach, Dittmann: "Design Principles for Active Audio and Video Fingerprinting, Multimedia Security", Chapter V, Idea Group Publishing, Chun-Shien Lu (Hrsg.), ISBN 1-59140-275-1,pp 157-172, 2004